Antenna Theory in Resonating Systems derived from Fundamental Electromagnetism

Habilitationsschrift

von Dr. rer. nat. Frank Gronwald

geb. am 15. Juli 1968   in Aachen

zur Verleihung des akademischen Grades

**Doctor rerum naturalium habilitatus
(Dr. rer. nat. habil.)**

genehmigt von der Fakultät für Elektrotechnik und Informationstechnik
der Otto-von-Guericke-Universität Magdeburg am 02. Oktober 2006

Gutachter:

       Herr Prof. Dr. J. Nitsch, Magdeburg
       Herr Prof. Dr. G. Wollenberg, Magdeburg
       Herr Prof. Dr. A. Enders, Braunschweig
       Herr Prof. Dr. Dr. C.E. Baum, Albuquerque, USA

Contents

List of Symbols

a_n	coefficient of nonlinear current-voltage power series relation
$\boldsymbol{a}$	(i) two-dimensional surface, (ii) three-component vector function
a	normal mode of electromagnetic field
da_i	two-dimensional surface element
$A_i, \mathbf{A}, \boldsymbol{A}, \mathsf{A}$	vector gauge potential
$A_{jk}, [A]$	method of moment matrix
$A_{ij}{}^k$	interaction field
$A_i^{xx}, A_i^{yy}, A_i^{zz}$	sign coefficient of mirror source
$\boldsymbol{b}$	three-component vector function
b_n	coefficient of nonlinear voltage-current power series relation
b	normal mode variable of electromagnetic field
$B^i, \mathbf{B}, \boldsymbol{B}, \mathsf{B}$	magnetic field strength
c	velocity of light
$\boldsymbol{c}$	(i) pilot vector, (ii) curve in three-dimensional space (iii) three-component vector function
c	normal mode variable of electromagnetic field
cond	condition number
C	conductance
C'	per-unit-length capacitance
C_1, C_2	integration constant
$\mathbb{C}$	set of complex numbers
dc^i	one-dimensional line element
d	(i) metric, (ii) distance
$\boldsymbol{d}$	three-component vector function
$D^i, \mathbf{D}, \boldsymbol{D}, \mathsf{D}$	electric excitation
$D_{\mathcal{L}}$	domain of operator $\mathcal{L}$
e	(i) elementary charge, (ii) measure of error in function space
$\boldsymbol{e}_i$	vector frame

e_n	orthonormal basis of a function space
$\boldsymbol{e}_n$	normal unit vector
$\boldsymbol{e}_t, \boldsymbol{e}_\tau$	unit tangent vector
$\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$	cartesian unit vector
$\boldsymbol{e}_{I_A}, \boldsymbol{e}_{I_{TL}}$	unit vector of antenna and transmission line mode current
$\boldsymbol{e}_{\boldsymbol{r}', \boldsymbol{r}}$	unit vector that points from $\boldsymbol{r}'$ to $\boldsymbol{r}$
e_R	reference frame of wave function
erfc	complementary error function
E	(i) energy, (ii) Ewald parameter
E^{inc}	incident electric field component
$E_i, \mathbf{E}, \boldsymbol{E}, \mathsf{E}$	electric field strength
$\mathcal{E}$	amplitude of electric field strength
f	(i) general scalar function, (ii) element of function space, (iii) frequency
f_n	basis of a function space
$\boldsymbol{f}$	general function with values in $\mathbb{C}^m$
$\boldsymbol{f}_n$	eigenfunction with values in $\mathbb{C}^m$
$F^i, \mathbf{F}, \mathsf{F}$	general vector field
F^i_n	component in function space
F_i	Lorentz force
$\boldsymbol{F}_n$	transverse eigenfunctions
g	(i) determinant of metric field, (ii) element of function space, (iii) general scalar function
$\boldsymbol{g}$	general function with values in $\mathbb{C}^m$
g_{ij}, g^{ij}	metric field
G	Green's function
G'	per-unit-length conductance
$G^A_{xx}, G^A_{yy}, G^A_{zz}$	cartesian component of the dyadic Green's function of the vector gauge potential $\boldsymbol{A}$
$G^E_{xx}, G^E_{yy}, G^E_{zz}$	cartesian component of the dyadic Green's function of the electric field strength $\boldsymbol{E}$
G_0	free space Green's function of Helmholtz equation
G^ϕ	Green's function of the scalar gauge potential ϕ
$\overline{\boldsymbol{G}}^A$	dyadic Green's function of the vector gauge potential $\boldsymbol{A}$
$\overline{\boldsymbol{G}}^E$	dyadic Green's function of the electric field strength $\boldsymbol{E}$
$\overline{\boldsymbol{G}}^B$	dyadic Green's function of the magnetic field strength $\boldsymbol{B}$

$\overline{\boldsymbol{G}}_0$	dyadic free space Green's function of Helmholtz equation
h	(i) Planck constant, (ii) element of function space
$\hbar$	Planck constant divided by 2π
H	(i) Hilbert space, (ii) linear transfer function
H_n	nth order nonlinear transfer function
$H_i, \mathbf{H}, \boldsymbol{H}, \mathsf{H}$	magnetic excitation
$\mathcal{H}$	amplitude of magnetic excitation
i	current in time domain
$\hat{i}$	current phasor
I	(i) electric current, (ii) identity
I_k	expansion function for electric current
I'_s	distributed current source
$\boldsymbol{I}_{\mathrm{A}}, I_{\mathrm{A}}$	antenna mode current
$\boldsymbol{I}_{\mathrm{TL}}, I_{\mathrm{TL}}$	transmission line mode current
$\overline{\boldsymbol{I}}$	unit dyad
$\mathcal{I}$	mathematical functional
j	imaginary unit
$J^i, \mathbf{J}, \boldsymbol{J}, \mathsf{J}$	electric current density
J_i^Φ	magnetic flux current
$\boldsymbol{J}_s$	electric surface current
k	(i) wavenumber, (ii) integral kernel
k', k''	real and imaginary part of wavenumber
k_n, k_{mnp}	wavenumber corresponding to an eigenmode
k_x, k_y, k_z	cartesian wavenumber component of an eigenmode
d^3k	volume element of reciprocal space
$\boldsymbol{k}, \hat{\boldsymbol{k}}$	wave vector and normalized wave vector
l_u	Lie derivative with respect to velocity field u^i
l_x, l_y, l_z	dimension of rectangular cavity
L	(i) Lagrangian, (ii) inductance (iii) antenna length
L'	per-unit-length inductance
$\boldsymbol{L}_n, \boldsymbol{L}_{mnp}$	longitudinal eigenfunction
$\boldsymbol{L}^p(\Omega)^m$	space of Lebesgue integrable functions
$\mathcal{L}$	(i) Lagrangian density, (ii) linear operator
$\mathcal{L}_D$	linear differential operator
$\mathcal{L}_\lambda$	the operator $\mathcal{L} - \lambda I$

$\mathcal{L}^{\star}$	adjoint operator
m	mass
M	subspace of a Hilbert space
$M^{\perp}$	orthogonal complement of M
$\boldsymbol{M}_n, \boldsymbol{M}_{mnp}$	transverse eigenfunction
$\boldsymbol{N}_n, \boldsymbol{N}_{mnp}$	transverse eigenfunction
$\mathcal{M}_{\xi_0}^{\xi}\{\overline{\boldsymbol{P}}\}$	product integral of $\overline{\boldsymbol{P}}(\xi)$
N	normalization function
p_i, p^i	momentum
P	probability
P_k	pulse function
$\overline{\boldsymbol{P}}$	per-unit-length parameter matrix
q	electric point charge
q'	per-unit-length electric charge
Q	(i) electric charge, (ii) quality factor
$\mathbb{Q}$	set of rational numbers
$\boldsymbol{r}, \boldsymbol{r}', \boldsymbol{r}''$	position vectors
d^2r	area element
d^3r	volume element
R	resistance
R'	per-unit-length resistance
$R_{i,mnp}$	distance between mirror source and observation point
$\boldsymbol{R}$	position vector
$R_{\mathcal{L}}$	range of operator $\mathcal{L}$
$\mathbb{R}$	set of real numbers
S	(i) two-dimensional surface, (ii) action, (iii) topological space
S_k	sinusoidal basis function
∂S	closed boundary of two-dimensional surface
t	(i) time, (ii) line parameter
$u^i, \boldsymbol{u}$	velocity field
v	voltage in time domain
$\hat{v}$	voltage phasor
v_1, v_2	voltage amplitude
$v^i, \boldsymbol{v}$	velocity field
v_i	covariant vector

$\boldsymbol{v}$	three-dimensional volume
dv	three-dimensional volume element
V	(i) three-dimensional volume, (ii) voltage in frequency domain
V_s'	distributed voltage source
∂V	closed surface of three-dimensional voltage
w	electromagnetic energy density
w^i	contravariant vector density
w_j	weighting function of the method of moments
W	electromagnetic energy
x^i	spatial coordinate
X_i	cartesian coordinate of mirror source
$\boldsymbol{X}$	voltage-current vector
y^i	spatial coordinate
Y	admittance
Y_i	cartesian coordinate of mirror source
Z	impedance
Z_{mn}	self ($m = n$) and mutual ($m \neq n$) impedance
Z_i	cartesian coordinate of mirror source
$\mathcal{Z}_{\text{int}}$	intrinsic impedance
$\mathcal{Z}_{\text{s}}$	surface impedance
α	real or complex scalar
α_i	general covariant vector
$\alpha_n, [\alpha]$	set of real or complex scalars
β	(i) velocity divided by velocity of light, (ii) real or complex scalar
β^j	general contravariant vector density
$\boldsymbol{\beta}$	velocity field divided by velocity of light
$\beta_n, [\beta]$	set of real or complex scalars
β_R	reference phase
γ	scalar density
γ^0, γ^i	Pauli matrix
$\gamma^i{}_j$	magnetoelectric material tensor
Γ	boundary of volume Ω
δ	(i) delta function, (ii) infinitesimal parameter
δ_i^n	unit tensor
δ_{mn}	Kronecker symbol

$\delta_{\mathrm{t}}, \delta_{x^i}$	translation in time and space
δ_ϵ	gauge transformation
$\delta_{\omega_i{}^j}$	rotation in space
ϵ	gauge parameter
ϵ_{0N}	Neumann's factor, equal to 1 for $N = 0$, equal to 2 for $N = 1, 2, \dots$
$\epsilon_{ijk}, \epsilon^{ijk}$	Levi-Civita symbol
ε	(i) absolute permittivity, (ii) infinitesimal parameter
ε_0	vacuum permittivity
ε^{ij}	permittivity tensor
θ, θ^R	phase of wave function and its component
λ	(i) complex number, (ii) wavelength
λ_n	set of eigenvalues
Λ_{mn}	abbreviation, denotes $\int_\Gamma \big(\boldsymbol{e}_n \times \boldsymbol{H}_m(\boldsymbol{r})\big) \cdot \big(\boldsymbol{e}_n \times \boldsymbol{H}_n(\boldsymbol{r})\big)\, d^2 r$
μ_0, μ	vacuum and absolute permeability
μ_{ij}^{-1}	impermeability tensor
ξ	position variable along transmission line
ρ	electric charge density
ρ_{mag}	magnetic charge density
ρ_s	electric surface charge density
$\rho(\mathcal{L})$	resolvent set
σ	conductivity
$\sigma(\mathcal{L})$	spectrum
τ_1, τ_2	position variable along wire
ϕ	(i) scalar gauge potential, (ii) phase of a propagating field
$\phi_j{}^k$	interaction field
ϕ_n	function basis
Φ	magnetic flux
φ	electrostatic scalar potential
φ_n, φ_{mnp}	eigenfunction corresponding to Poisson equation
χ_n, χ_{mnp}	eigenfunction of scalar Helmholtz equation
ψ	general scalar function
ψ_k	basis function of the method of moments
ψ_n, ψ_{mnp}	eigenfunction of scalar Helmholtz equation
Ψ, Ψ_0, Ψ^R	wave function, its amplitude, and its component
ω	angular frequency

ω', ω''	real and imaginary part of angular frequency
ω_{p}	angular eigenfrequency of perfectly conducting cavity
$\omega_{m,n}$	intermodulation frequency
Ω	subset of $\mathbb{R}^n$
$\partial_i, \partial_{\mathrm{t}}$	partial derivative
D_u/Dt	material derivative with respect to velocity field u^i
$D_i^A, D_{\mathrm{t}}^\phi$	gauge covariant derivative
∇	nabla operator
$\|\ \ \|$	norm
$[\ \ ]$	matrix
$\langle\ ,\ \rangle$	inner product
$\langle\ ,\ \rangle_p$	pseudo inner product
$\langle a, b \rangle_p$	reaction between field strength $\boldsymbol{E}^a$ and current density $\boldsymbol{J}^b$
$\parallel$	longitudinal component
$\perp$	transverse component
$*$	complex conjugate

Zusammenfassung

In der vorliegenden Arbeit werden Konzepte der Antennentheorie mit denen der Mikrowellentheorie verknüpft, um eine "Antennentheorie innerhalb resonierender Systeme" zu formulieren. Resonierende Systeme sind in diesem Zusammenhang als räumliche Umgebungen definiert, innerhalb derer sich elektromagnetische Resonanzen ("stehende Wellen") ausbilden können. Eine Antennentheorie innerhalb resonierender Systeme bietet einen geeigneten Rahmen zur Modellierung innerer Probleme der elektromagnetischen Verträglichkeit. Diese Modellierung beinhaltet hauptsächlich die Untersuchung der Wechselwirkung von Antennen, wobei elektromagnetische Störquellen durch sendende Antennen und elektromagnetische Störsenken durch empfangende Antennen repräsentiert werden.

Unsere Vorgehensweise orientiert sich an den folgenden drei Fragestellungen:

1. Welche Gleichungen bestimmen das Verhalten von Antennen innerhalb resonierender Systeme (physikalische Modellbildung und mathematische Formulierung)?

2. Welche mathematischen Methoden sind anzuwenden, um diese Gleichungen hinreichend genau und schnell auswerten zu können (analytische und numerische Lösungsverfahren)?

3. Welche Schlussfolgerungen lassen sich aus den gewonnenen Lösungen ziehen (physikalische Interpretation und technische Anwendung)?

Um auf diese Fragestellungen adäquat eingehen zu können, ist die Kenntnis der grundlegenden Konzepte der klassischen Elektrodynamik unabdingbar. Diese Konzepte werden in *Kapitel 1* vollständig eingeführt und interpretiert. Die gewählte Darstellung ist als Kombination von Maxwellscher Axiomatik und eichtheoretischer Beschreibung originär. Eine wichtige Konsequenz ist die Identifikation der zwei komplementären Arten von Singularitäten des elektromagnetischen Feldes, welche durch Coulomb-Singularitäten und elektromagnetische Resonanzen gegeben sind. Entsprechend lassen sich elektromagnetische Felder in Coulomb-Felder und Strahlungsfelder unterteilen. Für die in praktischen Anwendungen auftretenden elektromagnetischen Felder ist eine exakte Aufspaltung in diese beiden Feldanteile in der Regel nicht möglich. Diese untrennbare Verknüpfung von Coulomb-Anteilen und Strahlungsanteilen ist der hauptsächliche Grund für die bei der Formulierung und Anwendung einer Antennentheorie in resonierenden Systemen auftretenden Schwierigkeiten.

Das für die weitere mathematische Formulierung notwendige Rüstzeug wird maßgeblich durch die Funktionalanalysis geliefert. Daher beginnt *Kapitel 2* mit einer Bereitstellung funktionalanalytischer Begriffe und Methoden. Wesentlich sind die Methode der Entwicklung nach Eigenfunktionen eines selbstadjungierten Differentialoperators und die Methode der Greenschen Funktion als Basis der Antennentheorie. Auch numerische Lösungsmethoden finden im funktionalanalytischen Rahmen eine natürliche und einheitliche Darstellung.

In *Kapitel 3* werden zunächst elektromagnetische Begriffe für die Antennentheorie eingeführt. Hierzu gehören Definitionen elektromagnetischer Kopplung, das Prinzip der Reziprozität, und die Darstellung von Antennenimpedanzen in Umgebungen mit diskretem elektromagnetischen Spektrum. Es zeigt sich, dass die in der Antennentheorie zu lösenden Feldintegralgleichungen innerhalb von resonierenden Systemen Greensche Funktionen als Integralkerne aufweisen, die sowohl durch Coulomb-Singularitäten als auch durch elektromagnetische Resonanzen gekennzeichnet sind. Eine praktikable Auswertung solcher Integralgleichungen erfordert eine getrennte Berechnung beider Arten von Singularitäten. Hierfür eignen sich hybride Strahlen-Moden Darstellungen Greenscher Funktionen, die sich mit Interpolationsverfahren und der Methode der analytischen Regularisierung kombinieren lassen. Mit diesen Hilfsmitteln werden für kanonische Beispiele Antennenimpedanzen innerhalb von Resonatoren berechnet. Diese Berechnungen repräsentieren vollständige Lösungen von Antennenproblemen innerhalb resonierender Systeme und liefern Erkenntnisse für die Elektromagnetische Verträglichkeit.

Nichtlinear belastete Antennen innerhalb resonierender Ungebungen werden in *Kapitel 4* betrachtet. An solchen Antennen treten Intermodulationseffekte auf, die komplexe Frequenzspektren generieren. Zur Berechnung solcher Frequenzspektren ist es vorteilhaft, ein gegebenes Antennenproblem zunächst auf ein Netzwerkproblem zu reduzieren. Anschließend können Methoden der nichtlinearen Netzwerktheorie angewendet werden. Ein für die Elektromagnetische Verträglichkeit wichtiger Aspekt ist das Phänomen der Unwandlung von hohen Frequenzen zu niedrigen Frequenzen innerhalb von Resonatoren. Dieser Effekt wird qualitativ beschrieben und anhand von Beispielen quantitativ berechnet.

Die Integralgleichungen der Antennentheorie bilden auch die Grundlage der herkömmlichen Leitungstheorie und ihrer Verallgemeinerungen. Daher bietet es sich an, abschließend in *Kapitel 5* auf elektromagnetische Leitungen einzugehen. Für gleichförmige Leitungen ist eine Aufspaltung des Leitungsstromes in Gleichtaktstrom und Gegentaktstrom auch in resonierenden Umgebungen sinnvoll. Der Gleichtaktstrom kann dann mit den in Kapitel 3 vorgestellten Methoden der Antennentheorie berechnet werden, während sich der Gegentaktstrom näherungsweise durch klassische Leitungstheorie berechnen läßt. Zur Beschreibung allgemeiner Leitungskonfigurationen ist die klassische Leitungstheorie aber nicht mehr geeignet und es sind in diesem Fall die Integralgleichungen der Antennentheorie ohne einschränkende Annahmen zu lösen.

Introduction and Summary

Antennas are engineering devices that are built to transmit and receive electromagnetic signals. In everyday life they are known as essential components of radio, television, mobile communication, and radar systems. The interaction between a transmitting and a receiving antenna is represented by the sequence

$$\text{transmitting antenna} \longrightarrow \text{electromagnetic field} \longrightarrow \text{receiving antenna} \,. \qquad (1)$$

In *antenna theory* we are mainly concerned with the modeling of this sequence. The corresponding physical framework is provided by classical electrodynamics. In classical electrodynamics it is appropriate to view the sequence (1) as electromagnetic interaction between electromagnetic sources, where electromagnetic sources are represented by electric charges and their currents,

$$\text{electromagnetic sources} \longleftrightarrow \text{electromagnetic field} \longleftrightarrow \text{electromagnetic sources} \,. \quad (2)$$

The coupling between electromagnetic sources and electromagnetic fields is governed by the Maxwell equations. If we want to explicitly calculate the interaction sequence (2) we have to solve the Maxwell equations. In *free space* the general solution of the Maxwell equations with respect to elementary electromagnetic sources is known. It is fortunate that this case is the one which is of primary interest in antenna theory: Usually, we require electromagnetic propagation between a transmitting and a receiving antenna to take place in a free space environment where it is not too much influenced by obstacles and boundaries. This is why antennas often are mounted on top of mountains, towers, roofs, or poles. Consequently, most results of antenna theory are based on the assumptions that the electromagnetic coupling and propagation takes place in free space [5, 211].

However, in antenna theory we also are confronted with situations where the assumption of a free space environment is no longer justified. Examples are *cavity backed antennas* that are located within a semi-open cavity [52, 26] or *periodic antenna arrays* that are modeled by equivalent configurations of single antennas which are enclosed by a number of mirrors [132, 5]. The necessity to formulate antenna coupling not only in free space becomes particularly evident if we consider the increasingly important and expanding subject of *Electromagnetic Compatibility (EMC) analysis* [217, 117, 168, 197]: In EMC analysis we encounter a large class of situations where Electromagnetic Interference (EMI) sources act as transmitting antennas and produce unwanted electromagnetic

fields. These fields may eventually couple to EMI victims that behave like receiving antennas. Therefore, it is natural to view in this context both EMI sources and EMI victims as unintentionally coupled antennas. Since EMI sources and victims often are located within metallic enclosures it is then required to adapt the concepts of antenna theory to this situation.

Probably one of the most important features of a free space environment is the existence of a continuous set of possible states of the electromagnetic field. An antenna in free space may couple, depending on its excitation, to all electromagnetic states of this continuous set. This set mathematically is described by a simple function, the Green's function of the Helmholtz equation in free space. It follows from the properties of this function that in free (three-dimensional) space electromagnetic wave propagation can be characterized by outgoing spherical waves in a rather simple way. But once we introduce boundaries these can impose severe conditions on the possible states of the electromagnetic field. In particular, boundary conditions may single out discrete and bound states, the so-called *modes* or *resonances*. We call an environment which supports resonances a *resonating environment* or *resonating system*. If an antenna is placed within a resonating system the coupling between antenna and electromagnetic field at a particular frequency can be greatly inhibited or enhanced. Also the solution of the Maxwell equations becomes more complex and involved if compared to free space [146].

These considerations lead us to the subject of this thesis: We will investigate the interaction between transmitting and receiving antennas in resonating systems and, accordingly, call the corresponding framework *antenna theory in resonating systems*. This framework emerges from the combination of usual antenna theory and *microwave theory*. In microwave theory, which includes the *theory of guided waves*, the Maxwell equations are solved within resonating systems in order to model electromagnetic propagation [136, 30, 172]. Therefore, many concepts that are familiar from microwave theory can be applied to antenna theory in resonating systems as well. But microwave theory usually assumes given electromagnetic sources which excite electromagnetic fields, that is, the electric currents that generate the propagating electromagnetic fields usually are assumed to be known. This is in contrast to antenna theory where antenna currents *a priori* are unknown and also need to be determined on the basis of the Maxwell equations. It follows that in the framework of an antenna theory of resonating systems we always have to consider not only freely propagating electromagnetic fields but also their interaction with transmitting and receiving antennas which do represent extended electromagnetic sources. As a result, both antenna currents and electromagnetic fields, which both interact, need to be determined.

Our approach is based on the following chapters:

Chapter 1 provides an account of the foundations of classical electrodynamics. It yields the basic equations of electromagnetics that we will need to solve together with their physical interpretation. Our presentation is original, unique, and believed to be of appreciable pedagogical value. It combines an *axiomatic approach* with

the *classical gauge field approach*. The axiomatic approach has been outlined in
a recent monograph [84] which is based on a careful study of relevant literature,
among we especially mention the work of Truesdell & Toupin [225] and Post [171].
The classical gauge field approach has been established by Weyl [236] and nowa-
days constitutes a cornerstone of elementary particle physics [21]. We will use a
formulation of the gauge field approach which stresses the physical importance of
the gauge potentials and has been found useful both in gravity and electromag-
netics [62, 66, 67]. After a comparison of the axiomatic and the classical gauge
field approach we will pay special attention to the *dynamical properties of the elec-
tromagnetic field*. In the course of this it will be essential to introduce the split
of the electromagnetic field into an irrotational and a rotational part. This split
will also lead to a distinction between *Coulomb fields* and *radiation fields*. As a
further consequence it will be seen that Coulomb fields and radiation fields are
inseparably intertwined. It is this circumstance that leads to many conceptual and
practical problems in electrical engineering applications such as antenna theory or
transmission line theory. Therefore, our presentation of the foundations of classical
electrodynamics is beneficial in order to recognize the link between basic electro-
magnetic field properties and fundamental difficulties that are encountered during
the solution of actual engineering problems.

Chapter 2 begins with a summary of important results of *functional analysis*. This
mathematical subject developed from the necessity to solve boundary value prob-
lems that are of importance in physics. These boundary value problems include
vector wave equations that form the dynamical basis of many electromagnetic prob-
lems. We review the relevant concepts that have been introduced by Hilbert [32]
and later were applied to problems of theoretical physics [146]. A main goal of
this review is to introduce the *method of eigenfunction expansion for self-adjoint
differential operators*. The functional analytic framework also yields an elegant ap-
proach to the *Green's function method* which is of utmost importance for a proper
formulation of antenna theory. Green's functions are essential to represent the elec-
tromagnetic interaction either in free space or in the presence of boundaries. As
an example we will outline the construction of electromagnetic Green's functions
in cavities. Finally, it will be stressed that the framework of functional analy-
sis naturally leads to the discretization of boundary value problems and, thus, to
numerical solution methods.

Chapter 3 first presents some basic concepts of antenna theory. In particular, *mea-
sures of electromagnetic coupling* will be introduced and related to electromagnetic
field properties. For the formulation of equations that are appropriate to deter-
mine antenna currents we will concentrate on the class of linear wire antennas and
review how to obtain relevant *electric field integral equations*. For the case of an-
tennas in resonating systems these integral equations are characterized by kernels

which are given by Green's functions that contain both Coulomb singularities and resonances, i.e., poles in the complex frequency plane. Efficient solution methods of these integral equations require to separate both types of electromagnetic singularities. Concepts that realize such a separation include the *method of analytical regularization* and the *representation of Green's functions in hybrid form*. They allow to explicitly calculate *antenna impedances in resonating systems* in an efficient way. This will be exemplified for linear wire antennas within rectangular cavities. The analytic-numerical results confirm the properties of antenna coupling within resonating systems that are qualitatively expected from physical arguments.

Chapter 4 serves to consider antennas in resonating systems that are *nonlinearly loaded*. A nonlinear load will cause *intermodulation effects* and drastically change the frequency spectrum that is excited by an antenna. Consequently, it will be of main concern to determine the spectrum that is generated by a nonlinearly loaded antenna within a resonating system. As a general solution strategy, the corresponding electromagnetic field problem will be reduced to an *equivalent circuit problem*. An important aspect is the phenomenon of *"high to low frequency conversion"* which will be discussed in the context of Electromagnetic Compatibility Analysis.

Chapter 5 specializes on linear wire structures that form *transmission lines*. The coupling of transmission lines to electromagnetic fields is modeled by the same electric field integral equations that occur in antenna theory and it follows that the methods of the previous Chapter 3 apply. It will be outlined how to put the Maxwell equations in the form of *generalized Telegrapher equations* and how to arrive at the *classical transmission line theory*. In the course of this it will turn out to be useful to introduce *antenna mode currents* and *transmission line currents*. For uniform transmission lines both types of currents decouple in free space and may also approximately decouple within a cavity. The benefit of such a decoupling is that the antenna mode may be calculated along the lines of Chapter 3 while the transmission line mode can, approximately, be calculated from the comparatively simple classical transmission line theory. However, the classical transmission line theory is not suitable to characterize general tranmission line structures. In this case it is necessary to solve the integral equations of antenna theory without restricting approximations.

Chapter 1

Fundamentals of Electromagnetics

In nature one has, up to now, identified four fundamental interactions: Gravity, electromagnetism, weak interaction, and strong interaction. Gravity and electromagnetism manifest themselves on a macroscopic level. The weak and the strong interactions are generically microscopic in nature and require a quantum field theoretical description right from the beginning. Most electrical engineering applications are modeled by means of the electromagnetic interaction and in this case it is not required to take into account the other three interactions. This is especially true for the development of the usual antenna theory and microwave theory.

The four interactions can be modeled individually. Thereby it is recognized that electromagnetism has the simplest structure amongst these interactions. This simplicity is reflected in the Maxwell equations. They, together with a few additional assumptions, explain the electromagnetic phenomena that we observe in nature or in laboratories. Authoritative accounts of classical electromagnetic field theory are provided by [199, 93, 182], e.g..

Rather then simply accepting the Maxwell equations and studying their consequences we have in mind to derive them from some deeper lying structures, using as few assumptions as possible. This is the main motivation for the development of the axiomatic approach and the gauge field approach. Both approaches to electromagnetism will be introduced below. It is known from elementary particle physics that the Maxwell equations rely on conservation laws and symmetry principles [21, 185]. These concepts will be the main ingredients of the axiomatic approach in the following Sec. 1.1 and also of the gauge field approach which is the subject of Sec. 1.2. The two different approaches will be related to each other in Sec. 1.3. In Sec. 1.4 we will mainly be concerned to rewrite Maxwell equations in terms of wave equations while the subject of Sec. 1.5 is to derive important electromagnetic field properties from the electromagnetic equations of motion.

1.1   Axiomatic approach

The axiomatic approach to classical electrodynamics is based on electric charge conservation, the Lorentz force, magnetic flux conservation, and the existence of local and linear constitutive relations [225, 84]. The *inhomogeneous* Maxwell equations, expressed in terms of D^i and H_i , turn out to be a consequence of electric charge conservation, whereas the *homogeneous* Maxwell equations, expressed in terms of E_i and B^i, are derived from magnetic flux conservation. The excitations D^i and H_i , by means of constitutive relations, are linked to the field strengths E_i and B^i. Whereas the axiomatic approach has been presented in a relativistic framework [225, 84] we will be able to mostly do without relativistic notions. This is quite remarkable and requires, in particular for the derivation of the homogeneous Maxwell equations from magnetic flux conservation, some steps that are not necessary if the complete framework of relativity is available.

The axiomatic approach is not only characterized by simplicity and beauty, but is also of appreciable pedagogical value. The more clearly a structure is presented, the easier it is to memorize. Moreover, an understanding of how the fundamental electromagnetic quantities D^i, H_i, E_i, B^i are related to each other may facilitate the formulation and solution of actual electromagnetic problems.

As it is appropriate for an axiomatic approach, we will start from as few prerequisites as possible. What we will need is some elementary mathematical background that comprises differentiation and integration in the framework of tensor analysis in three-dimensional space. In particular, the concept of integration is necessary for introducing electromagnetic objects as integrands in a natural way. To this end, we will use a tensor notation in which the components of mathematical quantities are explicitly indicated by means of upper (contravariant) or lower (covariant) indices [190]. The advantage of this notation is that it allows to represent geometric properties clearly. In this way, the electromagnetic objects become more transparent and can be discussed more easily. For the formalism of differential forms, which provides similar conceptual advantages, we refer to [128, 84].

Some mathematical material is compiled in Appendix A. It might be helpful in order to get comfortable with the tensor notation. For a quick start we introduce the following conventions:

- Partial derivatives with respect to a spatial coordinate x^i (with $i, j, \cdots = 1, 2, 3$) or with respect to time t are abbreviated according to

$$\frac{\partial}{\partial x^i} \longrightarrow \partial_i, \qquad \frac{\partial}{\partial t} \longrightarrow \partial_t. \tag{1.1}$$

- We use the "summation convention". It states that a summation sign can be omitted if the same index occurs both in a lower and an upper position. That is,

we have, for example, the correspondence

$$\sum_{i=1}^{3} \alpha_i\,\beta^i \;\longleftrightarrow\; \alpha_i\,\beta^i\,. \tag{1.2}$$

- We define the Levi-Civita symbols ϵ_{ijk} and ϵ^{ijk}. They are antisymmetric with respect to all of their indices. Therefore, they vanish if two of their indices are equal. Their remaining components assume the values $+1$ or -1, depending on whether ijk is an even or an odd permutation of 123:

$$\epsilon_{ijk} = \epsilon^{ijk} = \begin{cases} 1\,, & \text{for } ijk = 123,\ 312,\ 231, \\ -1\,, & \text{for } ijk = 213,\ 321,\ 132. \end{cases} \tag{1.3}$$

With these conventions we obtain for the *gradient* of a function f the expression $\partial_i f$. The *curl* of a (covariant) vector v_i is written according to $\epsilon^{ijk}\partial_j v_k$ and the *divergence* of a (contravariant) vector (density) w^i is given by $\partial_i w^i$.

Now we are prepared to move on to the Maxwell theory. In the next four subsections, we will establish classical electrodynamics from electric charge conservation (axiom 1), the Lorentz force (axiom 2), magnetic flux conservation (axiom 3), and the existence of constitutive relations (axiom 4). This represents the core of classical electrodynamics: It results in the Maxwell equations together with the constitutive relations and the Lorentz force law.

In order to complete electrodynamics, one can require two more axioms, which we only mention shortly (see [84] for a detailed discussion). One can specify the energy-momentum distribution of the electromagnetic field (axiom 5) by means of its so-called energy-momentum tensor. This tensor yields the energy density $(D^i E_i + H_i B^i)/2$ and the energy flux density $\epsilon^{ijk} E_j H_k$ (the Poynting vector), inter alia. Moreover, if one treats electromagnetic problems of materials in macrophysics, one needs a further axiom by means of which the total electric charge (and the current) is split (axiom 6) in a bound or material charge (and current), which is also conserved, and in a free or external charge (and current).

1.1.1 Electric charge conservation (axiom 1) and the inhomogeneous Maxwell equations

In classical electrodynamics, the electric charge is characterized by its density ρ. From a geometric point of view, the charge density ρ constitutes an integrand of a volume integral. This geometric identification is natural since, by definition, integration of ρ over a three-dimensional volume V yields the total charge Q enclosed in this volume

$$Q := \int_V \rho\,dv\,. \tag{1.4}$$

We note that, in the SI-system, electric charge is measured in units of "ampere times second" or coulomb, $[Q] = \mathrm{As} = \mathrm{C}$. Therefore, the SI-unit of charge density ρ is $[\rho] = \mathrm{As/m^3} = \mathrm{C/m^3}$.

It is instructive to invoke at this point the Poincaré lemma. There are different explicit versions of this lemma. We use the form (A.23) that is displayed in Appendix A. Then (if space fulfills suitable topological conditions) we can write the charge density ρ as the divergence of an integrand D^i of a surface integral. Thus,

$$\boxed{\partial_i D^i = \rho} \qquad\qquad (\boldsymbol{\nabla} \cdot \boldsymbol{D} = \rho)\,. \qquad\qquad (1.5)$$

This result already constitutes one inhomogeneous Maxwell equation, the Coulomb-Gauss law. In parenthesis we display for comparison the more conventional vector notation.

Electric charges often move. We represent this motion by a material velocity field u^i, that is, we assign locally a velocity to each portion of charge in space. The product of electric charge density ρ and material velocity u^i defines[1] the electric current density J^i,

$$J^i = \rho u^i\,. \qquad\qquad (1.6)$$

Geometrically, the electric current density constitutes an integrand of surface integrals since integration of J^i over a two-dimensional surface S yields the total electric current I that crosses this surface,

$$I = \int_S J^i \, da_i\,. \qquad\qquad (1.7)$$

We have, in SI-units, $[I] = \mathrm{A}$ and $[J^i] = \mathrm{A/m^2}$.

We now turn to electric charge conservation, the first axiom of our axiomatic approach. To this end we have to determine how individual packets of charge change in time as they move with velocity u^i through space. A convenient way to describe this change is provided by the material derivative D_u/Dt which also is often called convective derivative [225, 183]. It allows to calculate the change of a physical quantity as it appears to an observer or a probe that follows this quantity. Then electric charge conservation can be expressed as

$$\frac{D_u Q}{Dt} = 0\,, \qquad\qquad (1.8)$$

where the material derivative is taken with respect to the velocity field u^i. It can be rewritten by means of the Reynold's transport theorem in the following way [183, p.

---

[1]This definition is a microscopic one, since the movement of individual electric charges that constitute the electric charge density is considered. On a macroscopic "averaged" level it is possible that the effective charge density vanishes while an electric current is present. An example is a configuration of an electric current that flows within a wire and exhibits no net charge density since negative charges of moving electrons are compensated by positive charges of atoms that constitute the wire.

457],

$$\frac{D_u Q}{Dt} = \frac{D_u}{Dt} \int_{V(t)} \rho \, dv$$

$$= \int_{V(t)} \frac{\partial \rho}{\partial t} \, dv + \oint_{\partial V(t)} \rho u^i \, da_i$$

$$= \int_{V(t)} \left(\frac{\partial \rho}{\partial t} + \partial_i (\rho u^i) \right) dv \,. \tag{1.9}$$

Here we used in the last line the Stokes theorem in the form of (A.24). The volume $V(t)$ that is integrated over depends in general on time since it moves together with the electric charge that it contains. By means of (1.6), (1.8), and (1.9) we obtain the axiom of electric charge conservation in the local form as continuity equation,

$$\partial_t \rho + \partial_i J^i = 0 \,. \tag{1.10}$$

We mention that we can obtain this result also from the direct application of the material derivative D_u/Dt to the charge density. Noting that the material derivative is the sum of the partial time derivative and the Lie derivative with respect to the velocity field u^i [225],

$$\frac{D_u}{Dt} = \frac{\partial}{\partial t} + l_u \,, \tag{1.11}$$

we find with (A.39) for the material time derivative of the scalar density ρ

$$\frac{D_u \rho}{Dt} = \frac{\partial \rho}{\partial t} + l_u \rho \tag{1.12}$$

$$= \frac{\partial \rho}{\partial t} + \partial_i (\rho u^i) \tag{1.13}$$

$$= \frac{\partial \rho}{\partial t} + \partial_i J^i \,. \tag{1.14}$$

In this way the continuity equation (1.10) follows from $D_u \rho / Dt = 0$.

Now we use the inhomogeneous Maxwell equation (1.5) in order to replace within the continuity equation (1.10) the charge density by the divergence of D^i. This yields

$$\partial_i \left(\partial_t D^i + J^i \right) = 0 \,. \tag{1.15}$$

Again we invoke the Poincaré lemma, now in the form (A.22), and write the sum $\partial_t D^i + J^i$ as the curl of the integrand of a line integral which we denote by H_i. We obtain

$$\boxed{\epsilon^{ijk} \partial_j H_k - \partial_t D^i = J^i} \qquad \left(\boldsymbol{\nabla} \times \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} = \boldsymbol{J} \right) \,. \tag{1.16}$$

Equation (1.16) constitutes the remaining inhomogeneous Maxwell equation, the Ampère-Maxwell law, which, in this way, is derived from the axiom of electric charge

conservation. The fields D^i and H_i are called electric excitation (historically: electric displacement) and magnetic excitation (historically: magnetic field), respectively. From (1.5) and (1.16) it follows that their SI-units are $[D^i] = $ As/m^2 and $[H_i] = $ A/m.

Some remarks are appropriate now: We first note that we obtain the excitations D^i and H_i from the Poincaré lemma and charge conservation, respectively, without introducing the concept of force. This is in contrast to other approaches that rely on the Coulomb and the Lorentz force laws [43]. Furthermore, since electric charge conservation is valid not only on macroscopic scales but also in microphysics[2], the inhomogeneous Maxwell equations (1.5) and (1.16) are microphysical equations as long as the source terms ρ and J^i are microscopically formulated as well. The same is valid for the excitations D^i and H_i. They are microphysical quantities – in contrast to what is often stated in textbooks, see [93], for example. We finally remark that the inhomogeneous Maxwell equations (1.5) and (1.16) can be straightforwardly put into a relativistically invariant form. This is not self-evident but suggested by electric charge conservation in the form of the continuity equation (1.10) since this fundamental equation can also be shown to be relativistically invariant.

1.1.2 Lorentz force (axiom 2) and merging of electric and magnetic field strengths

During the discovery of the electromagnetic field, the concept of force has played a major role. Electric and magnetic forces are directly accessible to experimental observation. Experimental evidence shows that, in general, an electric charge is subject to a force if an electromagnetic field acts on it. For a point charge q at position $x_{\mathrm{q}}{}^i$ we have $\rho(x^i) = q\delta(x^i - x_{\mathrm{q}}{}^i)$. If it has the velocity u^i we postulate the Lorentz force

$$\boxed{F_i = q(E_i + \epsilon_{ijk}u^j B^k)} \tag{1.17}$$

as second axiom. It introduces the electric field strength E_i and the magnetic field strength B^i. The Lorentz force already yields a prescription of how to measure E_i and B^i by means of the force that is experienced by an infinitesimally small test charge q which is either at rest or moving with velocity u^i. Turning to the dimensions, we introduce voltage as "work per charge". In SI, it is measured in volt (V). Then $[F_i]=$VC/m and, according to (1.17), $[E_i] = $ V/m and $[B^i] = $ Vs/m$^2 = $ Wb/m$^2 = $ T, with Wb as abbreviation for Weber and T for Tesla.

From the axiom of the Lorentz force (1.17), we can draw the conclusion that the electric and the magnetic field strengths are not independent of each other. The corre-

---

[2]Microphysics commonly is understood as the physics on small scales that describes the interaction between single, elementary particles. The concept of an elementary particle not necessarily involves quantum effects and also is useful and important for classical electrodynamics [182]. On microphysical scales electric charges and their related currents often are represented by distributions that reflect the physical model of a point particle.

a) charge observed from its rest frame

b) charge observed from inertial frame
moving with respect to q
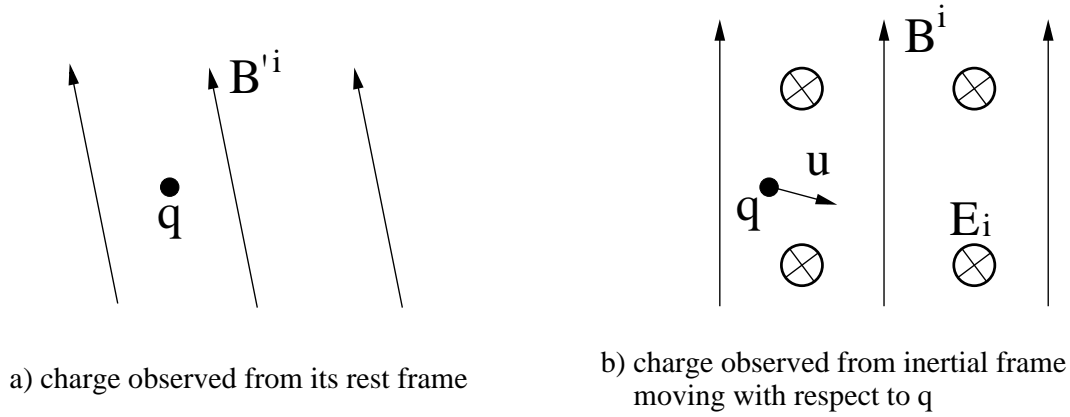
Figure 1.1: A charge that is, in some inertial frame, at rest and is immersed in a purely magnetic field experiences no Lorentz force, see Fig.1a. The fact that there is no Lorentz force should be independent of the choice of the inertial system that is used to observe the charge. Therefore, a compensating electric field accompanies the magnetic field if viewed from an inertial laboratory system which is in relative motion to the charge, see Fig.1b.

sponding argument is based on the special relativity principle: According to the special relativity principle, the laws of physics are independent of the choice of an inertial system [43]. Different inertial systems move with constant velocities v^i relative to each other. The outcome of a physical experiment, as expressed by an empirical law, has to be independent of the inertial system where the experiment takes place.

Let us suppose a point charge q with a certain mass moves with velocity u^i in an electromagnetic field E_i and B^i. The velocity and the electromagnetic field are measured in an inertial laboratory frame. The point charge can also be observed from its instantaneous inertial *rest frame.* If we denote quantities that are measured with respect to this rest frame by a prime, i.e., by u'^i, E'_i, and B'^i, then we have $u'^i = 0$. In the absence of an electric field in the rest frame, i.e., if additionally $E'_i = 0$, the charge experiences no Lorentz force and, therefore, no acceleration,

$$F'_i = q(E'_i + \epsilon_{ijk}u'^j B'^k) = 0 \,. \tag{1.18}$$

The fact that the charge experiences no acceleration is also true in the laboratory frame. This is a consequence of the special relativity principle or, more precisely, of the fact that the square of the acceleration can be shown to form a relativistic invariant. Consequently,

$$F_i = q(E_i + \epsilon_{ijk}u^j B^k) = 0 \,. \tag{1.19}$$

Thus, in the laboratory frame, electric and magnetic field are related by

$$E_i = -\epsilon_{ijk}u^j B^k \,. \tag{1.20}$$
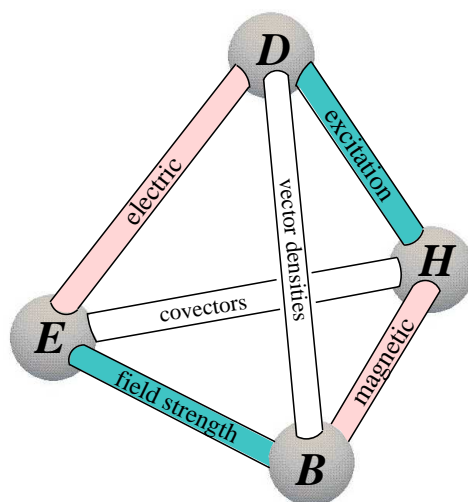
Figure 1.2: The tetrahedron of the electromagnetic field. The electric and the magnetic excitations D^i, H_i and the electric and the magnetic field strengths E_i, B^i build up 4-dimensional quantities in spacetime. These four fields describe the electromagnetic field completely. Of electric nature are D^i and E_i, of magnetic nature H_i and B^i.

This situation is depicted in Fig.1.1. Accordingly, we find that electric and magnetic field strength cannot be viewed as independent quantities. They are connected to each other by transformations between different inertial systems.

Let us pause for a moment and summarize: So far we have introduced the four electromagnetic field quantities D^i, H_i and E_i, B^i. These four quantities are interrelated by physical and mathematical properties. This is illustrated in Fig.1.2 by the "tetrahedron of the electromagnetic field"[3].

1.1.3 Magnetic flux conservation (axiom 3) and the homogeneous Maxwell equations

We digress for a moment and turn to hydrodynamics. Helmholtz was one of the first who studied rotational or "vortex" motion in hydrodynamics, see [115]. He derived theorems for vortex lines. An important consequence of his work was the conclusion that vortex lines are conserved. They may move or change orientation but they are never spontaneously created nor annihilated. The vortex lines that pierce through a two-dimensional surface can be integrated over and yield a scalar quantity that is called circulation. The circulation in a perfect fluid, which satisfies certain conditions, is constant provided the loop enclosing the surface moves with the fluid [115].

There are certainly fundamental differences between electromagnetism and hydro-

---

[3]Thanks are due to Dr. Christian Heinicke for providing this figure.

dynamics. But some suggestive analogies exist. A vortex line in hydrodynamics seems analogous to a magnetic flux line. The magnetic flux Φ is determined from magnetic flux lines, represented by the magnetic field strength B^i, that pierce through a two-dimensional surface S,

$$\Phi := \int_S B^i \, da_i \,. \tag{1.21}$$

As the circulation in a perfect fluid is conserved, we can guess that, in a similar way, the magnetic flux may be conserved. Of course, the consequences of such an axiom have to be borne out by experiment.

At first sight, one may find vortex lines of a fluid easier to visualize than magnetic flux lines. However, on a microscopic level, magnetic flux can occur in quanta. The corresponding magnetic flux unit is called flux quantum or fluxon and it carries $\Phi_0 = h/(2e) \approx 2,07 \cdot 10^{-15}$ Wb, with h the Planck constant and e the elementary charge. Single quantized magnetic flux lines have been observed in the interior of type II superconductors if exposed to a sufficiently strong magnetic field, see [84, p. 131]. They even can be counted. The corresponding experiments provide good evidence that magnetic flux is a conserved quantity.

But how can we formulate magnetic flux conservation mathematically? In Sec. 1.1.1 we applied the material derivative D_u/Dt with respect to a velocity field u^i to the total electric charge Q and, equivalently, to the electric charge density ρ. This yielded the continuity equation (1.10) which expresses electric charge conservation. We may follow the same pattern to express magnetic flux conservation and write down the conservation law

$$\frac{D_u \Phi}{Dt} = 0 \,. \tag{1.22}$$

This expression has to be examined and this, in turn, requires to clarify the following two points:

- How do we define a velocity field u^i with respect to a magnetic field B^i?

- What is a physically reasonable definition of the current of a magnetic flux?

To answer the first point we have to know how to observe, in general, a magnetic field. The only means that we have to our disposal is the Lorentz force law, the second axiom of our approach. With the Lorentz force law we may use electric test charges to measure the electric and magnetic field strength. We have already noted in the last subsection that electric and magnetic field strength are connected to each other by relativistic transformations. This makes it impossible to state in a relativistically invariant way which contribution to a Lorentz force is due to an electric field and which contribution is due to a magnetic field. *Any* observer who uses an electric test charge which is located in his rest frame might state that the Lorentz force on his test charge is "purely electric" since for his test charge $u^i = 0$ in (1.17) and, thus, $F_i = qE_i$. He might furthermore draw the conclusion that in his rest frame the velocity of the magnetic field

vanishes as well. This is, of course, a wrong conclusion since the velocity which appears in the Lorenz force law is the relative velocity between an observer and a test charge but it is, a priori, not the relative velocity between a test charge and the magnetic field.

To nevertheless associate a velocity field to a magnetic field we note that the *vanishing* of the Lorentz force on a test charge is relativistically invariant. If the Lorentz force on a test charge vanishes in one inertial systems it will vanish in all inertial systems. In this case there will be exactly one inertial system where this test charge is at rest. In this inertial system we have

$$F'_i = q(E'_i + \epsilon_{ijk} u'^j B'^k) = 0 \tag{1.23}$$

and it follows, since $u'^j = 0$, that $E'_i = 0$. We now *define* this distinguished inertial system to be, at the position considered, the rest frame of the magnetic field. This definition requires that we can always find an inertial system where $F'_i = 0$, i.e., that we can always find an inertial system where $E'_i = 0$. For an arbitrary electromagnetic field this will not be true, but we assign this property to a *purely* magnetic field. A purely magnetic field is an electromagnetic field where, at any point in space and time, we can make the electric field vanish in *one* inertial system. In the definition of magnetic flux conservation we will only consider electromagnetic fields which are purely magnetic. Otherwise we are not able to associate a velocity field to a magnetic field in a relativistically invariant way. Therefore, the answer to the first point is that the velocity u^i associated to a (purely) magnetic field is the velocity of a specified inertial system which moves with respect to a laboratory system with velocity u^i and where the Lorentz force on a test charge vanishes.

An answer to the second point requires to provide a physically meaningful definition of magnetic flux current. To this end we reconsider the notion of electric charge,

$$Q = \int_V \rho \, dv \,, \tag{1.24}$$

together with its corresponding conservation law

$$\partial_t Q + \int_{\partial V} J^i \, da_i = 0 \,. \tag{1.25}$$

We see from this representation that the rate of change of the electric charge within a specified volume V is balanced by the out- or inflowing charge across the surface ∂V. This charge transport is described by the electric charge current J^i that is integrated over the enveloping surface ∂V. By means of the Stokes theorem in the form (A.24), equation (1.25) yields the local continuity equation

$$\partial_t \rho + \partial_i J^i = 0 \,. \tag{1.26}$$

Let us follow the same pattern to define the current of a magnetic flux: Starting with the definition (1.21) of the magnetic flux, the corresponding geometric conservation law,

in analogy to (1.25), reads

$$\partial_t \Phi + \int_{\partial S} J_i^{\Phi} \, dc^i = 0 \,, \tag{1.27}$$

where we introduced the magnetic flux current J_i^{Φ}. This is a covariant vector that is integrated along a line ∂S, that is, along the curve bordering the 2-dimensional surface S. The conservation law (1.27) tells us that the rate of change of the magnetic flux within a specified area S is balanced by the magnetic flux current J_i^{Φ} that is integrated along the boundary ∂S. Then the Stokes theorem in the form (A.25) yields the local continuity equation

$$\partial_t B^i + \epsilon^{ijk} \partial_j J_k^{\Phi} = 0 \,. \tag{1.28}$$

One interesting consequence is the following: The divergence of (1.28) reads

$$\partial_i (\partial_t B^i) = 0 \qquad \Longrightarrow \qquad \partial_i B^i = \rho_{\text{mag}} \,, \quad \partial_t \rho_{\text{mag}} = 0 \,. \tag{1.29}$$

Thus, we find a time-*in*dependent term ρ_{mag} which acquires tentatively the meaning of a magnetic charge density. Let us choose a specific reference system in which ρ_{mag} is constant in time, i.e., $\partial_t \rho_{\text{mag}} = 0$. Now we go over to an arbitrary reference system with time coordinate t' and spatial coordinates $x^{i'}$. Clearly, in general $\partial_{t'} \rho_{\text{mag}} \neq 0$. The only way to evade a contradiction to (1.29) is to require $\rho_{\text{mag}} = 0$, that is, the magnetic field strength B^i has no sources, its divergence vanishes:

$$\boxed{\partial_i B^i = 0} \qquad (\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0) \,. \tag{1.30}$$

This is recognized as one of the homogeneous Maxwell equations.

To specify the magnetic flux current we finally explore magnetic flux conservation, expressed by means of the material time derivative D_u/Dt with respect to a velocity field u^i which is associated to a purely magnetic field. We have

$$\begin{aligned} \frac{D_u \Phi}{Dt} &= \frac{D_u}{Dt} \int_{S(t)} B^i \, da_i \\ &= \int_{S(t)} \left(\partial_t B^i - \epsilon^{ijk} \partial_j \epsilon_{klm} u^l B^m + u^i \partial_j B^j \right) \, da_i \,, \end{aligned} \tag{1.31}$$

where we applied the Helmholtz transport theorem [183, p. 456]. Alternatively, we can work with the local expression

$$\begin{aligned} \frac{D_u B^i}{Dt} &= \partial_t B^i + l_u B^i \\ &= \partial_t B^i + u^j \partial_j B^i - B^j \partial_j u^i + B^i \partial_j u^j \,, \end{aligned} \tag{1.32}$$

where we used the formula (A.36) for the Lie derivative of a contravariant vector density.

The Lie derivative of B^i can be rewritten according to[4]

$$u^j \partial_j B^i - B^j \partial_j u^i + B^i \partial_j u^j = -\epsilon^{ijk} \partial_j \epsilon_{klm} u^l B^m + u^i \partial_j B^j \,. \qquad (1.33)$$

Hence, it follows

$$\frac{D_u B^i}{Dt} = \partial_t B^i - \epsilon^{ijk} \partial_j \epsilon_{klm} u^l B^m + u^i \partial_j B^j \qquad (1.34)$$

and it is recognized that (1.34) is the local version of (1.31)

According to (1.30) the divergence of the magnetic field strength B^i vanishes, $\partial_i B^i = 0$. Also, by virtue of (1.20), we can locally identify the term $-\epsilon_{klm} u^l B^m$ with an electric field strength E_k. This is because we assumed that we work with a purely magnetic field, i.e., an electromagnetic field which does not exert a Lorentz force on an electric charge which moves with velocity u^i in the laboratory frame. Then magnetic flux conservation, $D_u B^i/Dt = 0$, yields

$$\boxed{\partial_t B^i + \epsilon^{ijk} \partial_j E_k = 0} \qquad \left(\frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{\nabla} \times \boldsymbol{E} = \boldsymbol{0} \right) \,. \qquad (1.35)$$

This equation reflects magnetic flux conservation, the third axiom of our axiomatic approach. It constitutes the remaining homogeneous Maxwell equation, that is, Faraday's induction law. We compare this result to the continuity equation (1.28) and deduce that the electric field that appears in the Faraday's induction law has to be interpreted as a magnetic flux current.

1.1.4 Constitutive relations (axiom 4) and the properties of spacetime

So far we have introduced $4 \times 3 = 12$ unknown electromagnetic field components D^i, H_i, E_i, and B^i. These components have to fulfill the Maxwell equations (1.5), (1.16), (1.30), and (1.35), which represent $1 + 3 + 1 + 3 = 8$ partial differential equations. In fact, among the Maxwell equations, only (1.16) and (1.35) contain time derivatives and are dynamical. The remaining equations, (1.5) and (1.30), are so-called "constraints". They are, by virtue of the dynamical Maxwell equations, fulfilled at all times if fulfilled at one time. It follows that they do not contain information on the time evolution of the electromagnetic field. Therefore, we arrive at only 6 dynamical equations for 12 unknown field components. To make the Maxwell equations a determined set of partial differential equations we still have to introduce additionally the so-called "constitutive relations" between the excitations D^i, H_i and the field strengths E_i, B^i.

---

[4]In a more conventional notation this identity reads, compare (B.11),

$$(\boldsymbol{u} \cdot \boldsymbol{\nabla})\boldsymbol{B} - \boldsymbol{B}(\boldsymbol{\nabla} \cdot \boldsymbol{u}) + (\boldsymbol{B} \cdot \boldsymbol{\nabla})\boldsymbol{u} = -\boldsymbol{\nabla} \times (\boldsymbol{u} \times \boldsymbol{B}) + \boldsymbol{u}(\boldsymbol{\nabla} \cdot \boldsymbol{B}) \,.$$

The simplest case to begin with is to find constitutive relations for the case of electromagnetic fields in vacuum. There are guiding principles that limit their structure. We demand that constitutive relations in vacuum are invariant under translation and rotation, furthermore they should be local and linear, i.e., they should connect fields at the same position and at the same time. Finally, in vacuum the constitutive relations should not mix electric and magnetic properties. These features characterize the vacuum and not the electromagnetic field itself. We will not be able to prove them but postulate them as fourth axiom.

If we want to relate the field strengths and the excitations we have to remind ourselves that E_i, H_i are natural integrands of *line* integrals and D^i, B^i are natural integrands of *surface* integrals. Therefore, E_i, H_i transform under a change of coordinates as covariant vectors while D^i, B^i transform as contravariant vector densities. To compensate these differences we will have to introduce a symmetric metric field $g_{ij} = g_{ji}$. The metric tensor determines spatial distances and introduces the notion of orthogonality. The determinant of the metric is denoted by g. It follows that $\sqrt{g}\, g^{ij}$ transforms like a density and maps a covariant vector into a contravariant vector density. We then take as fourth axiom the constitutive relations for vacuum,

$$\boxed{D^i = \varepsilon_0 \sqrt{g}\, g^{ij}\, E_j\,,} \tag{1.36}$$

$$\boxed{H_i = (\mu_0 \sqrt{g})^{-1} g_{ij}\, B^j\,.} \tag{1.37}$$

In flat spacetime and in Cartesian coordinates, we have $g = 1$, $g^{ii} = 1$, and $g^{ij} = 0$ for $i \neq j$. We recognize the familiar vacuum relations between field strengths and excitations. The electric constant ε_0 and the magnetic constant μ_0 characterize the vacuum. They acquire the SI-units $[\varepsilon_0] = \mathrm{As/Vm}$ and $[\mu_0] = \mathrm{Vs/Am}$.

What seems to be conceptually important about the constitutive equations (1.36), (1.37) is that they not only provide relations between the excitations D^i, H_i and the field strengths E_i, B^i, but also connect the electromagnetic field to the structure of spacetime, which here is represented by the metric tensor g_{ij}. The formulation of the first three axioms that were presented in the previous sections does not require information on this metric structure. The connection between the electromagnetic field and spacetime, as expressed by the constitutive relations, indicates that physical fields and spacetime are not independent of each other. The constitutive relations might suggest the point of view that the structure of spacetime determines the structure of the electromagnetic field. However, one should be aware that the opposite conclusion also has a truth value: It can be shown that the propagation properties of the electromagnetic field determine the metric structure of spacetime [84, 114].

Constitutive relations in matter usually assume a more complicated form than (1.36), (1.37). In this case it would be appropriate to derive the constitutive relations, after an averaging procedure, from a microscopic model of matter. Such procedures are the

subject of solid state or plasma physics, for example. A discussion of these subjects is out of the scope of this work but, without going into details, we quote the constitutive relations of a general linear *magnetoelectric* medium:

$$D^i = \left(\boldsymbol{\varepsilon^{ij}} - \epsilon^{ijk} n_k \right) E_j + \left(\boldsymbol{\gamma^i}_{\boldsymbol{j}} + \tilde{s}_j{}^i \right) B^j + (\boldsymbol{\alpha} - s) B^i , \qquad (1.38)$$

$$H_i = \left(\boldsymbol{\mu_{ij}^{-1}} - \epsilon_{ijk} m^k \right) B^j + \left(-\boldsymbol{\gamma^j}_{\boldsymbol{i}} + \tilde{s}_i{}^j \right) E_j - (\boldsymbol{\alpha} + s) E_i . \qquad (1.39)$$

This formulation is due to Hehl & Obukhov [84, 85, 163], an equivalent formulation was given by Lindell & Olyslager [173, 128]. Both matrices ε^{ij} and μ_{ij}^{-1} are symmetric and possess 6 independent components each, ε^{ij} is called *permittivity* tensor and μ_{ij}^{-1} *impermeability* tensor (reciprocal permeability tensor). The magnetoelectric cross-term $\gamma^i{}_j$, which is trace-free, $\gamma^k{}_k = 0$, has 8 independent components. It is related to the Fresnel-Fizeau effects. Accordingly, these pieces altogether, which we printed in (1.38) and (1.39) in boldface for better visibility, add up to $6 + 6 + 8 + 1 = 20 + 1 = 21$ independent components.

With the introduction of constitutive relations the axiomatic approach to classical electrodynamics is completed. We will see in the next Section 1.3 how this approach relates to the framework of gauge theory.

1.1.5 Remarks

We have presented an axiomatic approach to classical electrodynamics in which the Maxwell equations are derived from the conservation of electric charge and magnetic flux. In the context of the derivation of the inhomogeneous Maxwell equations, one introduces the electric and the magnetic excitation D^i and H_i, respectively. The explicit calculation is rather simple because the continuity equation for electric charge is already relativistically invariant such that for the derivation of the inhomogeneous Maxwell equations no additional ingredients from special relativity are necessary. The situation is more complicated for the derivation of the homogeneous Maxwell equations from magnetic flux conservation since it is not immediately clear of how to formulate magnetic flux conservation in a relativistic invariant way. It should be mentioned that if the complete framework of relativity were available, the derivation of the axiomatic approach could be done with considerable more ease and elegance [225, 84].

At this point we would like to comment on a question that sometimes leads to controversial discussions, as summarized in [183], for example. This is the question of how the quantities E_i, D^i, B^i, and H_i should be grouped in pairs, i.e., the question of "which quantities belong together?". Some people like to form the pairs (E_i, B^i), (D^i, H_i), while others prefer to build (E_i, H_i) , (D^i, B^i). Already from a dimensional point of view, the answer to this question is obvious. Both, E_i and B^i are *voltage*-related quantities, that is, related to the notions of force and work: In SI, we have $[E_i] = \mathrm{V/m}$, $[B^i] = \mathrm{T} = \mathrm{Vs/m}^2$, or $[B^i] = [E_i]/\text{velocity}$. Consequently, they belong

together. Analogously, D^i and H_i are *current*-related quantities: $[D^i] = \mathrm{C/m^2} = \mathrm{As/m^2}$, $[H_i] = \mathrm{A/m}$, or $[D^i] = [H_i]/\mathrm{velocity}$.

These conclusions are made irrefutable by relativity theory. Classical electrodynamics is a relativistic invariant theory and the implications of relativity have been proven to be correct on macro- and microscopic scales over and over again. And relativity tells us that the electromagnetic field strengths E_i, B^i are inseparably intertwined by relativistic transformations, and the same is true for the electromagnetic excitations D^i, H_i. In the spacetime of relativity theory, the pair (E_i, B^i) forms one single quantity, the tensor of electromagnetic field strength, while the pair (D^i, H_i) forms another single quantity, the tensor of electromagnetic excitations. If compared to these facts, arguments in favor of the pairs (E_i, H_i), namely that both are covectors, and (D^i, B^i), both are vector densities (see the tetrahedron in Fig.2), turn out to be of secondary nature.

1.2 The gauge field approach towards electromagnetism

Modern descriptions of the fundamental interactions heavily rely on symmetry principles. In particular, this is true for the electromagnetic interaction which can be formulated as a gauge field theory that is based on a corresponding gauge symmetry. In recent articles this approach towards electromagnetism has been explained in an original and descriptive way [64, 66, 67]. We want to put the gauge field approach next to the axiomatic approach since it furnishes further information that will complement our picture of classical electrodynamics. In particular, it allows to clarify the concept of gauge invariance which often accompanies explicit calculations in the solution of electrodynamic boundary value problems. It also shows that the electromagnetic potentials, which often are viewed as mathematical auxiliary variables, are of major physical relevance. Furthermore, in the gauge field approach the inhomogeneous Maxwell equations turn out to be true equations of motion while the homogeneous Maxwell equations become a mere mathematical identity.

While it is rewarding to gain the additional insights that are provided by the gauge field approach it should be admitted that this approach, at first sight, extends on a rather abstract level. But it only requires a small number of steps:

1. Accept the fact that physical matter fields (which represent electrons, for example) are described microscopically by complex wave functions.

2. Recognize that the absolute phase of these wave functions has no physical relevance. This arbitrariness of the absolute phase constitutes a one-dimensional rotational type symmetry $U(1)$ (the circle group). This is the gauge symmetry of electrodynamics.

3. To derive observable physical quantities from the wave functions requires to define derivatives of wave functions with respect to space and time. These derivatives need to be invariant under the gauge symmetry. The construction of such "gauge covariant" derivatives makes it necessary to introduce gauge fields. One gauge field, the vector potential A_i, defines gauge covariant derivatives D_i^A with respect to the three independent directions of space, while another gauge field, the scalar potential ϕ, defines a gauge covariant derivative D_t^ϕ with respect to time.

4. Finally, the values of the gauge fields ϕ and A_i are obtained from equations of motion that turn out to be the inhomogeneous Maxwell equations. The gauge fields are related to the gauge invariant electric and magnetic field strengths via

$$E_i = -\partial_i \phi - \partial_t A_i \,, \tag{1.40}$$

$$B^i = \epsilon^{ijk} \partial_j A_k \,. \tag{1.41}$$

1.2.1 Differences of physical fields that are described by reference systems

To introduce the concept of gauge symmetry we formally denote a physical field by $\mathbf{F}$, its components by F^i, and a reference frame with respects to these components by $\boldsymbol{e}_i$,

$$\mathbf{F} = F^i \boldsymbol{e}_i \,. \tag{1.42}$$

The field $\mathbf{F}$ might change in space or time. Its components F^i depend on the reference frame $\boldsymbol{e}_i$ and usually do not have an absolute significance since usually we have a certain freedom to choose $\boldsymbol{e}_i$. Therefore, a change of $\mathbf{F}$ involves both its components and the corresponding reference frame. Then we have

$$\partial_i \mathbf{F} = (\partial_i F^j) \boldsymbol{e}_j + F^j (\partial_i \boldsymbol{e}_j) \,, \tag{1.43}$$

$$\partial_t \mathbf{F} = (\partial_t F^j) \boldsymbol{e}_j + F^j (\partial_t \boldsymbol{e}_j) \,. \tag{1.44}$$

This simple looking "product rule for differentiation" poses severe difficulties: We might be able to determine the differences $\partial_i F^j$ and $\partial_t F^j$, e.g., by measurements, but how do we determine changes $\partial_i \boldsymbol{e}_j$, $\partial_t \boldsymbol{e}_j$ of reference frames? What is an *unchanged* reference frame? How do we *gauge* a reference frame? We have to face the fact that a priori changes of reference frames are not defined as long as they involve the change between different points in space or time.

At this point the concept of "interactions" enters the stage. If a physical field changes its value we might intuitively think that this is due to an interaction. The gauge principle states that the information on the change of reference frames is contained in interaction fields. Mathematically, this is formulated as follows: The changes $\partial_i \boldsymbol{e}_j$, $\partial_t \boldsymbol{e}_j$ of a reference frame are determined by interaction fields $A_{ij}{}^k$ and $\phi_j{}^k$ according to

$$\partial_i \boldsymbol{e}_j := A_{ij}{}^k \boldsymbol{e}_k \,, \tag{1.45}$$

$$\partial_t \boldsymbol{e}_j := \phi_j{}^k \boldsymbol{e}_k \,. \tag{1.46}$$

With these definitions the interaction fields provide a gauging of reference system at different positions in space and time. This is the reason why they are called *gauge fields*. A corresponding mathematical term is *connection*, because gauge fields connect reference systems at different positions in space and time.

With the definitions (1.45), (1.46) we may write the differences $\partial_i \mathbf{F}$, $\partial_t \mathbf{F}$ as

$$\partial_i \mathbf{F} = (\partial_i F^j)\boldsymbol{e}_j + F^j A_{ij}{}^k \boldsymbol{e}_k \,, \qquad (1.47)$$

$$\partial_t \mathbf{F} = (\partial_t F^j)\boldsymbol{e}_j + F^j \phi_j{}^k \boldsymbol{e}_k \,. \qquad (1.48)$$

So far it is not clear if the formal introduction of the gauge fields $A_{ij}{}^k$ and $\phi_j{}^k$ is physically meaningful. However, it turns out that these gauge fields correctly describe the fundamental interactions that we observe in nature. In the following this circumstance will be explained for the case of the electromagnetic interaction.

1.2.2 The phase of microscopic matter fields

We first identify the physical field and its corresponding reference system which leads to the introduction of the electromagnetic field as a gauge field. To this end we note that microscopic matter fields, like electrons, for example, are represented in the framework of quantum mechanics by *wave functions* $\Psi(x^i, t)$ [10, 58]. A wave function assumes complex values and depends on space coordinates x^i and a time coordinate t, $\Psi(x^i, t) \in \mathbb{C}$. It follows that a microscopic particle with specified momentum p_i and specified energy E is represented by a wave function of the form

$$\Psi(x^i, t) = \Psi_0 \, e^{-\frac{j}{\hbar}(p_i x^i - Et)} \,, \qquad (1.49)$$

with $\Psi_0 \in \mathbb{R}$, and $\hbar$ a fundamental constant which carries the dimension of an action, $\hbar = h/2\pi \approx 1.0546 \times 10^{-34} Js$.

The plane wave (1.49) is a very special case of a wave function since momentum and energy of a microscopic particle usually are not known exactly but affected by an *uncertainty*. In this more general case microscopic particles are represented by *wave packets* that are obtained by the superposition of plane waves of the form (1.49) with different momenta and energies. The value $\Psi(x^i, t)$ has no *direct* physical significance, but the square of its absolute value yields a probability density. This is, the real value

$$P(x^i, t) = |\Psi(x^i, t)|^2 \, dv \qquad (1.50)$$

is the probability to find a particle, which is described by $\Psi(x^i, t)$, at time t within a volume dv.

The wave function $\Psi(x^i, t)$ of (1.49) has the structure

$$\Psi(x^i, t) = \Psi_0 \, e^{-j\theta} \qquad (1.51)$$

with the phase

$$\theta(x^i, t) = (p_i x^i - Et)/\hbar. \tag{1.52}$$

It is seen that the characteristic quantities p_i and E of a free microscopic particle are included in its phase $\theta(x^i, t)$. They can explicitly be obtained by differentiation of the phase of the wave function,

$$p_i = \hbar \, \partial_i \theta, \tag{1.53}$$

$$E = -\hbar \, \partial_t \theta, \tag{1.54}$$

and these relations reveal that the momentum and the energy of a microscopic particle is determined by the phase differences $\partial_i \theta$ and $\partial_t \theta$. These phase differences depend on reference frames that, a priori, can arbitrarily be chosen, as we will see next.

1.2.3   The reference frame of a phase

In this subsection we consider a wave function Ψ at *one* space point x^i at a *fixed* time t. We wish to assign a specific phase θ to Ψ. The phase can be taken as a real number of the interval $[0, 2\pi[$, $0 \leq \theta < 2\pi$. In order to assign a fixed value to θ we need a reference system which determines a reference phase. We denote this reference phase by β_R. The specification of θ is explained in Fig. 1.3. There the wave function is displayed as a curly line. The arrow along this curly line indicates that the wave function is characterized by a certain phase. This phase can be thought of as a point on a circle and is given by an element of the interval $[0, 2\pi[$. The reference phase β_R is drawn as another arrow which also indicates a certain direction, i.e., represents an element of the interval $[0, 2\pi[$. We take this value as reference phase β_R. The choice of β_R a priori is arbitrary. A convenient choice would be $\beta_R = 0$, for example.
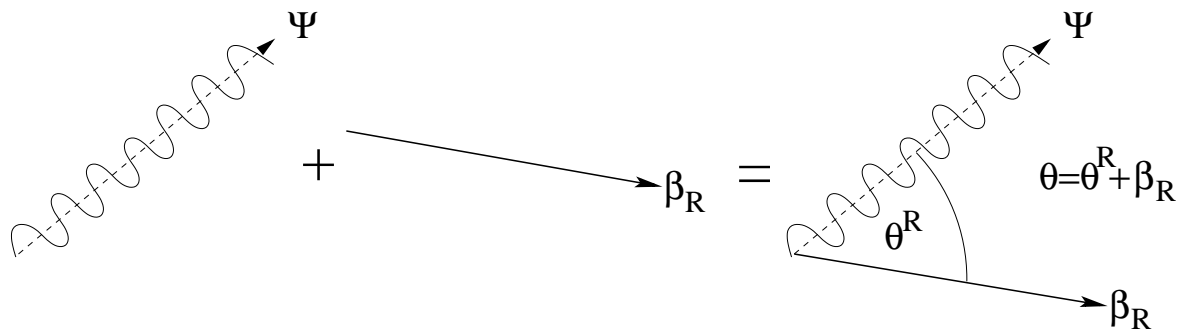

Figure 1.3: A reference system which determines a reference phase β_R allows to determine the phase of a wave function according to $\theta = \theta^R + \beta_R$.

To read off the value θ of the wave function we determine the angle between the wave function and the reference system. This angle is denoted by θ^R and refers to the

reference system β_R. Therefore, θ^R can be understood as a component with respect to the basis β_R. The component θ^R and the basis β_R yield the phase θ according to

$$\theta = \theta^R + \beta_R. \tag{1.55}$$

In terms of the wave function we write

$$\begin{aligned}\Psi =& \Psi_0\, e^{-j\theta} = \Psi_0\, e^{-j(\theta^R + \beta_R)} \\ =& \Psi_0\, e^{-j\theta^R} e^{-j\beta_R}.\end{aligned} \tag{1.56}$$

It follows that the wave function is of the form (1.42),

$$\Psi = \Psi^R e_R, \tag{1.57}$$

with the correspondences

$$\Psi^R = \Psi_0\, e^{-j\theta^R} \qquad \text{(component)} \tag{1.58}$$

and

$$e_R = e^{-j\beta_R} \qquad \text{(reference system)}. \tag{1.59}$$

The choice of a reference system of a phase is not unique. This is clarified in Fig. 1.4. We have the *gauge freedom* to choose between reference systems which differ from each other by a one-dimensional rotation. The choice of a fixed reference system is called the choice of a *gauge*. In this language the value θ^R is a *gauge dependent* quantity since it depends on the gauge, i.e., it depends on the choice of a reference system.

We denote a gauge transformation by δ_ϵ. Its effect on the component θ^R of the phase is

$$\delta_\epsilon \theta^R := \theta'^R - \theta^R = \frac{q}{\hbar}\epsilon. \tag{1.60}$$

Here we have assigned to the difference $\theta'^R - \theta^R$ the value $q\epsilon/\hbar$. This notation seems a bit awkward but it is in accordance with conventions that have their origin in quantum mechanics. The phase difference carries no physical dimension and q has the dimension of electric charge, $[q] = \text{As} = \text{C}$. Therefore, the parameter ϵ has the dimension of an action per charge, $[\epsilon] = \text{J}/(\text{As}^2)$. The interpretation of a gauge transformation (1.60) might appear to be rather trivial: If we shift between two reference systems that differ by an angle $(q/\hbar)\epsilon$ then the component θ^R of the phase changes by an angle $(q/\hbar)\epsilon$.

The reference system β_R can also be used to gauge the reference system β'_R. In order to show this we set the value of β'_R to

$$\beta'_R = \beta_R - \frac{q}{\hbar}\epsilon. \tag{1.61}$$

Then β_R transforms under a gauge transformation according to

$$\delta_\epsilon \beta_R := \beta'_R - \beta_R = -\frac{q}{\hbar}\epsilon. \tag{1.62}$$
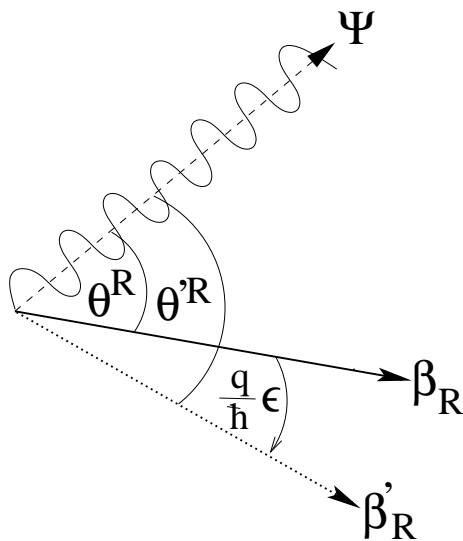
Figure 1.4: The phase θ^R of the wave function Ψ depends on the choice of a reference system. A gauge transformation corresponds to the transition from one reference system to another equivalent reference system. This transition is accomplished by a rotation about an angle $(q/\hbar)\epsilon$.

It follows that the phase θ is a gauge independent quantity,

$$
\begin{aligned}
\delta_\epsilon \theta &= \delta_\epsilon \theta^R + \delta_\epsilon \beta_R \\
&= \frac{q}{\hbar}\epsilon - \frac{q}{\hbar}\epsilon = 0 \,.
\end{aligned}
\tag{1.63}
$$

But in spite of this gauge independence the value of θ has no absolute significance since it depends, by virtue of $\theta = \theta^R + \beta_R$, on a choice of a reference frame β_R.

The situation is different if we consider the *difference* between two phases θ_1 and θ_2. Since both phases are defined at the same point x^i and to the same time t they can be characterized by a common reference system β_R,

$$
\theta_1 = \theta_1^R + \beta_R \,,
\tag{1.64}
$$

$$
\theta_2 = \theta_2^R + \beta_R \,.
\tag{1.65}
$$

Taking the difference yields

$$
\theta_1 - \theta_2 = \theta_1^R - \theta_2^R \,,
\tag{1.66}
$$

that is, the difference $\theta_1 - \theta_2$ is both independent of the reference system β_R and, due to

$$
\delta_\epsilon(\theta_1 - \theta_2) = \delta_\epsilon \theta_1 - \delta_\epsilon \theta_2 = 0 - 0 = 0 \,,
\tag{1.67}
$$

a gauge independent quantity. Therefore, the phase difference $\theta_1 - \theta_2$ at one point (x^i, t) in spacetime *has* an absolute significance.

1.2.4 The gauge fields of a phase

So far we considered the pase θ at one point (x^i, t) in spacetime. Now we turn to the change of the phase between *two different* points in spacetime and first concentrate on two points (x^i, t) and $(x^i + dx^i, t)$ that are separated by an infinitesimal spatial distance dx^i. Mathematically, we find from (1.55) for the difference $\partial_i \theta$ between two points x^i and $x^i + dx^i$ the expression

$$\partial_i \theta = \partial_i \theta^R + \partial_i \beta_R \,, \tag{1.68}$$

this is, the change of the phase θ is the sum of the change of its component and the change of its reference frame.

We will now geometrically interpret (1.68) and think of how to construct the difference $\partial_i \theta$. The construction is divided into several steps, compare Fig. 1.5:
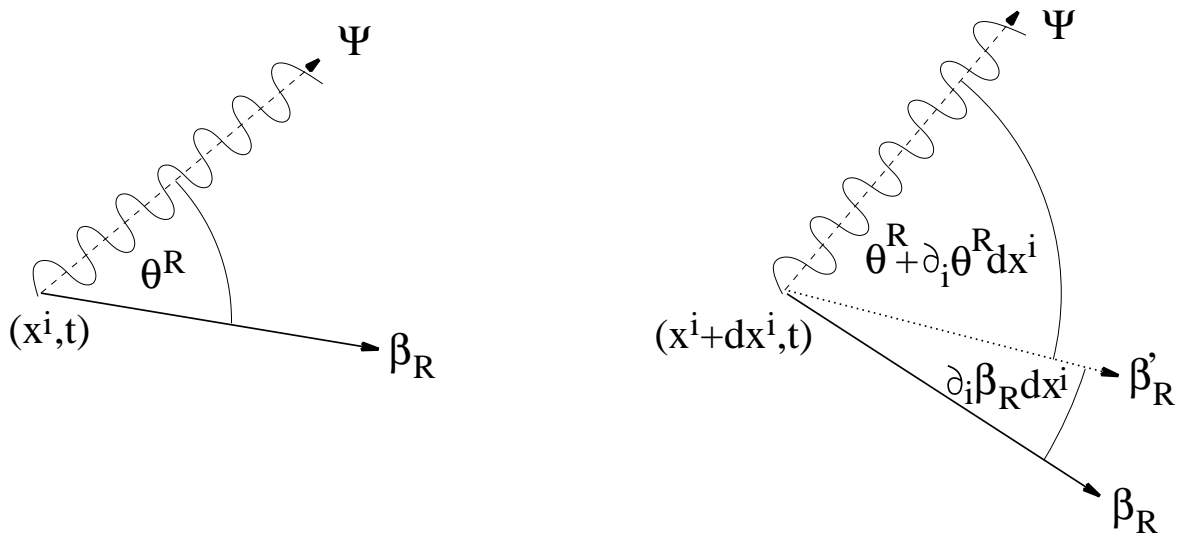
Figure 1.5: Determination of parallel reference systems at two spatially separated point (x^i, t) and $(x^i + dx^i, t)$. In the left part of the figure, at the point (x^i, t), the phase is given by $\theta = \theta^R + \beta_R$. In the right part of the figure, at the point $(x^i + dx^i, t)$, we have the phase $\theta = \theta^R + \partial_i \theta^R \, dx^i + \beta_R + \partial_i \beta_R \, dx^i$. This is explained in more detail in the text.

1. According to the previous subsection 1.2.3 we first determine the phase $\theta(x^i, t)$ at the point (x^i, t) by a reference system β_R according to $\theta(x^i, t) = \theta^R(x^i, t) + \beta_R(x^i, t)$.

2. At the point $(x^i + dx^i, t)$ we choose an arbitrary reference frame β_R'. In Fig. 1.5 this arbitrary reference frame is displayed by a dotted line. It can be used to read off the value

$$\theta'^R(x^i + dx^i, t) = \theta^R(x^i, t) + \partial_i \theta^R \, dx^i \,. \tag{1.69}$$

However, this value has no immediate physical relevance since its corresponding reference system has been arbitrarily chosen.

3. At the point $(x^i + dx^i, t)$ there exists a unique reference system β_R which is defined to be *unchanged* if compared to the reference system β_R at (x^i, t). In mathematical terms, such an unchanged reference system is named a *parallel* reference system. It is obtained from the arbitrary reference system β'_R by a rotation about an angle $\partial_i \beta_R$,

$$\beta_R(x^i + dx^i, t) = \beta'_R(x^i + dx^i, t) - \partial_i \beta_R \, dx^i \tag{1.70}$$

and has the *same* phase value as the reference system at the point (x^i, t),

$$\beta_R(x^i + dx^i, t) = \beta_R(x^i, t) \,. \tag{1.71}$$

With respect to this parallel reference system the component of the phase is given by

$$\theta^R(x^i + dx^i, t) = \theta^R(x^i, t) + \partial_i \theta^R \, dx^i + \partial_i \beta_R \, dx^i \,. \tag{1.72}$$

In summary we obtain the relations

$$\theta(x^i, t) = \theta^R(x^i, t) + \beta_R(x^i, t) \,, \tag{1.73}$$

$$\begin{aligned}\theta(x^i + dx^i, t) &= \theta'^R(x^i + dx^i, t) + \beta'_R(x^i + dx^i, t) \\ &= \theta^R(x^i, t) + \partial_i \theta^R \, dx^i + \beta_R(x^i, t) + \partial_i \beta_R \, dx^i \,, \end{aligned} \tag{1.74}$$

which lead us back to (1.68),

$$\partial_i \theta = \frac{\theta(x^i + dx^i, t) - \theta(x^i, t)}{dx^i} \tag{1.75}$$

$$= \partial_i \theta^R + \partial_i \beta_R \,. \tag{1.76}$$

The relation (1.76) determines the difference $\partial_i \theta$. But the contributions $\partial_i \theta^R$ and $\partial_i \beta_R$ need to be known explicitly. And this leads to the conceptual problem that has been described at the end of subsection 1.2.1: We may determine the change $\partial_i \theta^R$ if we properly read off the corresponding phases but the difference $\partial_i \beta_R$ a priori is not determined. We simply do not know which reference systems at different points are in parallel!

At this point the gauge fields come into play. In accordance to (1.45), (1.46) the difference $\partial_i \beta_R$ of the reference frame is determined from a gauge field A_i by

$$\partial_i \beta_R := -\frac{q}{\hbar} A_i \,. \tag{1.77}$$

Similar to (1.60) the factor $q/\hbar$ has been chosen to arrive at results that comply with quantum mechanics. Therefore, we obtain

$$\partial_i \theta = \partial_i \theta^R - \frac{q}{\hbar} A_i \,. \tag{1.78}$$

In view of (1.53) we realize that the momentum p_i of a microscopic particle can only be defined by means of the gauge field A_i. It follows from (1.78) that the SI-unit of the gauge field A_i is $[A_i] = \text{Vs/m}$.

So far we have considered the difference $\partial_i \theta$ between two spatially separated points (x^i, t) and $(x^i + dx^i, t)$. In the same way we can consider the difference $\partial_t \theta$ between two temporally separated points (x^i, t) and $(x^i, t + dt)$. This leads to the relation

$$\partial_t \theta = \partial_t \theta^i + \partial_t \beta_i \, . \tag{1.79}$$

Since $\partial_t \beta_i$ is undetermined this requires the introduction of a gauge field ϕ,

$$\partial_t \beta_R := \frac{q}{\hbar} \phi \, , \tag{1.80}$$

and we obtain the result

$$\partial_t \theta = \partial_t \theta^R + \frac{q}{\hbar} \phi \, . \tag{1.81}$$

Similar to above we realize, in view of (1.54), that the energy E of a microscopic particle can only be defined by means of the gauge field ϕ. Also it follows from (1.81) that the SI-unit of the gauge field ϕ is $[\phi] = \text{V}$.

A comparison between (1.78) and (1.81) reveals that we have chosen different signs in front of the fields A_i and Φ. This is done in order to be able to merge A_i and ϕ in accordance to common conventions into a single relativistically covariant quantity with four components.

At the end of this subsection we want to remark that the gauge fields A_i and ϕ are not gauge invariant. In fact, from (1.62) we have

$$\delta_\epsilon(\partial_i \beta_R) = -\frac{q}{\hbar} \partial_i \epsilon \, , \tag{1.82}$$

$$\delta_\epsilon(\partial_t \beta_R) = -\frac{q}{\hbar} \partial_t \epsilon \qquad , \tag{1.83}$$

and with the definitions (1.77), (1.80) we obtain for the behavior of A_i and ϕ under gauge transformations

$$\delta_\epsilon A_i = \partial_i \epsilon \, , \tag{1.84}$$

$$\delta_\epsilon \phi = -\partial_t \epsilon \, . \tag{1.85}$$

If we form the combinations

$$E_i := -\partial_i \phi - \partial_t A_i \, , \tag{1.86}$$

$$B^i := \epsilon^{ijk} \partial_j A_k \tag{1.87}$$

we can easily verify that these are invariant under gauge transformations,

$$\delta_\epsilon E_i = -\partial_i \partial_t \epsilon + \partial_t \partial_i \epsilon = 0 \, . \tag{1.88}$$

$$\delta_\epsilon B^i = -\epsilon^{ijk} \partial_j \partial_k \epsilon = 0 \, . \tag{1.89}$$

It follows from (1.86), (1.87) that the fields E_i and B^i carry the same SI-units as the electric and magnetic field strength, respectively, that have been introduced in subsection 1.1.2 within the axiomatic approach, $[E_i] = \mathrm{V/m}$, $[B^i] = \mathrm{Vs/m^2}$.

More important information is obtained if we consider the integrability conditions

$$\epsilon^{ijk}\partial_j E_k = -\underbrace{\epsilon^{ijk}\partial_j\partial_k\phi}_{=0} - \partial_t\epsilon^{ijk}\partial_j A_k$$

$$= -\partial_t B^i \tag{1.90}$$

$$\partial_i B^i = \epsilon^{ijk}\partial_i\partial_j A_k$$

$$= 0 \tag{1.91}$$

These conditions resemble exactly the *homogeneous Maxwell equations* (1.35) and (1.30).

1.2.5   The electromagnetic field as a gauge field

In the last subsection we introduced the gauge fields A_i and ϕ in order to define parallel reference frames at different points in spacetime. The approach was general and we still do not know the values of A_i and ϕ. How do we obtain these values?

If we assume that A_i and ϕ are physical fields we may further assume that they are determined by equations of motion which can be constructed according to the guidelines of classical field theory. These guidelines imply that equations of motion can often (but not always) be concisely characterized by a Lagrangian density

$$\mathcal{L} = \mathcal{L}(\Psi, \partial_i\Psi, \partial_t\Psi) \tag{1.92}$$

which, in the standard case, is a function of the fields Ψ of the theory and their first derivatives. Integration of the Lagrangian density $\mathcal{L}$ over space yields the Lagrangian L,

$$L = \int \mathcal{L}(\Psi, \partial_i\Psi, \partial_t\Psi)\, dv\,, \tag{1.93}$$

and further integration over time yields the action S,

$$S = \int L\, dt\,. \tag{1.94}$$

There are guiding principles that tell us how to obtain an appropriate Lagrangian density for a given theory. Once we have an appropriate Lagrangian density, we can conveniently derive the properties of the fields Ψ. For example, the equations of motion which determine the dynamics of Ψ follow from extremization of the action S with respect to variations of Ψ,

$$\delta_\Psi S = 0 \qquad \Longrightarrow \qquad \text{equations of motion for } \Psi\,. \tag{1.95}$$

More explicitly, the equations of motion are given by the well-known *Euler-Lagrange-equations*

$$\partial_j \left(\frac{\partial \mathcal{L}}{\partial_j \Psi} \right) - \partial_t \left(\frac{\partial \mathcal{L}}{\partial_t \Psi} \right) - \frac{\partial \mathcal{L}}{\partial \Psi} = 0 \tag{1.96}$$

Now we turn to the Lagrangian density $\mathcal{L}_{\text{gauge}}$ of the gauge fields A_i and ϕ. It has to fulfill a number of requirements:

- It should have the geometric character of a scalar density in order to be a proper volume integrand, compare section A.3.

- It should be gauge invariant.

- It should be relativistically invariant.

- It should have the SI-unit $[\mathcal{L}] = \text{VAs/m}^3$.

- It should be no more than quadratic in the fields A_i and ϕ.

- It should contain no higher order derivatives than first order derivatives to yield second order equations of motion.

These requirements are quite stringent. In fact, with the fields A_i, ϕ and the gauge invariant quantities E_i, B^i alone we cannot build a proper Lagrangian density. To construct a Lagrangian density we have to introduce, as in section 1.1.4, a metric structure $g_{ij} = g_{ji}$ that characterizes the geometry of spacetime. To get the dimensions right we further have to introduce two constants ε_0 and μ_0 with SI-units $[\varepsilon_0] = \text{As/Vm}$ and $[\mu_0] = \text{Vs/Am}$, respectively . Then the only meaningful combination of A_i and ϕ can be written in terms of E_i and B^i. It is given by the expression[5]

$$\mathcal{L}_{\text{gauge}} = \frac{1}{2} \left(\varepsilon_0 \sqrt{g} g^{ij} E_i E_j - (\mu_0 \sqrt{g})^{-1} g_{ij} B^i B^j \right) . \tag{1.97}$$

So far we only have taken into account the gauge fields. The Lagrangian density (1.97) corresponds to a free gauge field theory with no coupling to electrically charged matter fields. The inclusion of matter fields requires to set up a corresponding Lagrangian density $\mathcal{L}_{\text{matter}}$. As indicated in (1.92) this will involve derivatives $\partial_i \Psi$, $\partial_t \Psi$ of the fields Ψ. In order to be gauge invariant, i.e., to be independent of a specific choice of reference frames, these derivatives are expected from (1.47), (1.48) to involve the gauge fields. Indeed, if we recall the relations (1.57), (1.59) and (1.77) we can write

$$\begin{aligned} \partial_i \Psi = (\partial_i \Psi)^R e_R &= (\partial_i \Psi^R) e_R + \Psi^R (\partial_i e_R) \\ &= (\partial_i \Psi^R) e_R - j \partial_i \beta_R \Psi^R e_R \\ &= \left(\partial_i \Psi^R + j \frac{q}{\hbar} A_i \Psi^R \right) e_R \\ &= D_i^A \Psi^R e_R \end{aligned} \tag{1.98}$$

---

[5]The factor 1/2 is introduced to yield the correct Hamilton function (energy function) which can be obtained from the Lagrangian density by means of a Legendre transformation.

with the gauge covariant derivative

$$D_i^A := \partial_i + j\frac{q}{\hbar}A_i\,. \tag{1.99}$$

Analogously, it follows with (1.80)

$$\begin{aligned}
\partial_t \Psi = (\partial_t \Psi)^R e_R &= (\partial_t \Psi^R)e_R + \Psi^R(\partial_t e_R)\\
&= (\partial_t \Psi^R)e_R - j\partial_t\beta_R\Psi^R e_R\\
&= \left(\partial_t \Psi^R - j\frac{q}{\hbar}\phi\Psi^R\right)e_R\\
&= D_t^\phi \Psi^R e_R
\end{aligned} \tag{1.100}$$

with the gauge covariant derivative

$$D_t^\phi := \partial_t - j\frac{q}{\hbar}\phi\,. \tag{1.101}$$

Therefore, in order to be gauge invariant, the matter Lagrangian density may only contain derivatives and gauge fields as combinations of gauge covariant derivatives. To explicitly obtain the matter Lagrangian density requires advanced knowledge of relativistic quantum mechanics. We will refer at this point to the literature [21, 185] and quote as a result that the Lagrangian density of a certain class of electrically charged matter fields, like electrons, is given by[6]

$$\mathcal{L}_{\text{matter}} = -j\hbar c\overline{\Psi}^R\gamma^i\left(D_i^A - \frac{mc}{\hbar}\right)\Psi^R + j\hbar\overline{\Psi}^R\gamma^0\left(D_t^\phi - \frac{mc^2}{\hbar}\right)\Psi^R\,. \tag{1.102}$$

We may write the terms that couple the gauge fields to the matter fields as

$$\mathcal{L}_{\text{coupling}} = -A_i J^i - \phi\rho \tag{1.103}$$

where we defined

$$J^i := qc\overline{\Psi}^R\gamma^i\Psi^R \tag{1.104}$$

$$\rho := q\overline{\Psi}^R\gamma^0\Psi^R \tag{1.105}$$

Then the dynamics of the gauge fields A_i and ϕ is determined from the combined Lagrangian density

$$\mathcal{L}_{\text{em}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{coupling}}\,. \tag{1.106}$$

The equations of motion that follow from this Lagrangian density are given by the Euler-Lagrange equations (1.96). In our specific case they acquire the form

$$\partial_j\left(\frac{\partial\mathcal{L}_{\text{em}}}{\partial_j A_i}\right) - \partial_t\left(\frac{\partial\mathcal{L}_{\text{em}}}{\partial_t A_i}\right) - \frac{\partial\mathcal{L}_{\text{em}}}{\partial A_i} = 0\,, \tag{1.107}$$

$$\partial_j\left(\frac{\partial\mathcal{L}_{\text{em}}}{\partial_j \phi}\right) - \partial_t\left(\frac{\partial\mathcal{L}_{\text{em}}}{\partial_t \phi}\right) - \frac{\partial\mathcal{L}_{\text{em}}}{\partial \phi} = 0\,. \tag{1.108}$$

---

[6]In this expression Ψ^R denotes a 4-component spinor, $\overline{\Psi}^R$ is the adjoint spinor of Ψ^R, and γ^0, γ^i are 4×4-matrices [185].

We insert the expressions (1.97), (1.103) into (1.106) and obtain

$$\epsilon^{ijk} \partial_j H_k - \partial_t D^i = J^i \,, \tag{1.109}$$

$$\partial_i D^i = \rho \,, \tag{1.110}$$

with the definitions

$$H_k := (\mu_0 \sqrt{g})^{-1} g_{kl} \, B^l \,, \tag{1.111}$$

$$D^i := \varepsilon_0 \sqrt{g} \, g^{ij} \, E_j \,. \tag{1.112}$$

The results (1.109) and (1.110) are recognized as the inhomogeneous Maxwell equations (1.16) and (1.5), respectively. It follows that the gauge fields A_i and ϕ, that have been introduced in the formal context of reference frames, do indeed constitute the electromagnetic potentials that are familiar from classical electrodynamics. In this framework the homogeneous Maxwell equations turn out to be mathematical integrability conditions while the inhomogeneous Maxwell equations represent the equations of motion of the gauge fields A_i and ϕ.

1.3 On the relation between the axiomatics and the gauge field approach

In the following we want to comment on the interrelation between the previously presented axiomatic approach and the gauge field approach. It is interesting to see how the axioms find their proper place within the gauge approach.

1.3.1 Noether theorem and electric charge conservation

In field theory there is a famous result which connects symmetries of laws of nature to conserved quantities. This is the Noether theorem [161] which has been proven to be of great importance in both classical and quantum contexts. It is, in particular, discussed in books on classical electrodynamics, see [182, 199], for example. The Noether theorem connects the symmetry of a Lagrangian density $\mathcal{L}(\Psi, \partial_i \Psi, \partial_t \Psi)$, compare (1.93), to conserved quantities. Suppose, for example, that $\mathcal{L}$ is invariant under time translations δ_t. From our daily experience this assumption seems plausible since we do not expect that the laws of nature change in time. Then the Noether theorem implies a local conservation law which expresses the conservation of energy. Similarly, invariance under translations δ_{x^i} in space implies conservation of momentum, while invariance under rotations $\delta_{\omega^{ij}}$ yields the conservation of angular momentum,

$$\delta_t \mathcal{L} = 0 \qquad \Longrightarrow \qquad \text{conservation of energy} \,, \tag{1.113}$$

$$\delta_{x^i} \mathcal{L} = 0 \qquad \Longrightarrow \qquad \text{conservation of momentum} \,, \tag{1.114}$$

$$\delta_{\omega^{ij}} \mathcal{L} = 0 \qquad \Longrightarrow \qquad \text{conservation of angular momentum} \,. \tag{1.115}$$

These symmetries of spacetime are called *external* symmetries. But the Noether theorem also works for other types of symmetries that are called *internal* ones. Gauge symmetries often are internal symmetries. In this case, gauge invariance of the Lagrangian implies a conserved current with an associated charge. That is, if we denote a gauge transformation by δ_ϵ we conclude

$$\delta_\epsilon \mathcal{L} = 0 \qquad \Longrightarrow \qquad \text{charge conservation} . \tag{1.116}$$

If we apply this conclusion to electrodynamics we have to specify the Lagrangian density to be the one of matter fields that represent electrically charged particles. Then invariance of this Lagrangian density under the gauge symmetry of electrodynamics yields the conservation of electric charge. Thus, if we accept the validity of the Lagrangian formalism, we can arrive at electric charge conservation from gauge invariance via the Noether theorem.

1.3.2 Minimal coupling and the Lorentz force

We already have mentioned that, according to (1.95), we can derive the equations of motion (1.96) of a physical theory from a Lagrangian density and its associated action. We can use this scheme to derive the equations of motion of electrically charged particles. In this case, the corresponding Lagrangian density (that of the electrically charged particles) has to be gauge invariant.

For the electromagnetic case we have demonstrated that the Lagrangian density will be gauge invariant if we pass from partial derivatives to gauge covariant derivatives (1.99), (1.101) according to

$$\partial_i \longrightarrow D_i^A := \partial_i + j\frac{q}{\hbar}A_i , \tag{1.117}$$

$$\partial_t \longrightarrow D_t^\phi := \partial_t - j\frac{q}{\hbar}\phi , \tag{1.118}$$

with q the electric charge of the particle under consideration. The substitutions (1.117), (1.118) constitute the simplest way to ensure gauge invariance of the Lagrangian density of electrically charged particles. They constitute what commonly is called *minimal coupling*. Due to minimal coupling, we relate electrically charged particles and the electromagnetic field in a natural way that is dictated by the requirement of gauge invariance.

If we assume, as in Section 1.2.2,

$$\Psi^R = \Psi_0 \, e^{-\frac{j}{\hbar}(p_i x^i - Et)} , \tag{1.119}$$

we find

$$j\hbar\partial_i\Psi^R = p_i \, \Psi^R \tag{1.120}$$

$$-j\hbar\partial_t\Psi^R = E \, \Psi^R \tag{1.121}$$

These are relations that indicate how to pass from quantum physics to classical physics, i.e., how to pass from the action of the differential operators $j\hbar\partial_i$ and $-j\hbar\partial_t$ on a wavefunction to the momentum and energy of a classical particle. It follows that the classical analogues of (1.117) and (1.118) are given by

$$p_i \longrightarrow p_i - qA_i\,, \tag{1.122}$$

$$E \longrightarrow E - q\phi\,. \tag{1.123}$$

This is, if electrically charged particles are represented by classical particles, rather than by wave functions, we have to replace within the corresponding classical Lagrangian function the energy E and the momentum p_i of each particle according to (1.122) and (1.123).

As a general example we consider a non-relativistic classical particle with mass m and charge q. In absence of an electromagnetic field[7] the energy and momentum of the particle are related by

$$E = \frac{p_i p^i}{2m}\,. \tag{1.124}$$

In the presence of an electromagnetic field we have the replacements (1.122) and (1.123) which lead to

$$E = \frac{(p_i - qA_i)(p^i - qA^i)}{2m} + q\phi\,. \tag{1.125}$$

It follows from this expression for the energy E that the Lagrange function L of the particle is given by [14, p. 167]

$$L = \frac{m}{2}(\partial_t x_i)(\partial_t x^i) + qA_i(\partial_t x^i) - q\phi\,. \tag{1.126}$$

The equation of motion is obtained from the Lagrange function via

$$\frac{d}{dt}\frac{\partial L}{\partial(\partial_t x^i)} - \frac{\partial L}{\partial x^i} = 0\,. \tag{1.127}$$

This yields with (1.126)

$$m(\partial_t^2 x_i) + q\big(\partial_t A_i + \partial_t x^j(\partial_j A_i)\big) - q\partial_t x^j(\partial_i A_j) + q\partial_i\phi = 0\,. \tag{1.128}$$

The first term represents the force $F_i = m(\partial_t^2 x_i)$ that acts on the particle. We rearrange the other terms and note, in particular, the identity

$$\partial_i A_j - \partial_j A_i = \epsilon_{ijk}\epsilon^{klm}\partial_l A_m\,. \tag{1.129}$$

It is then immediate to arrive at

$$F_i = q(-\partial_i\phi - \partial_t A_i) + q(\epsilon_{ijk}\partial_t x^j \epsilon^{klm}\partial_l A_m)\,. \tag{1.130}$$

---

[7]We also assume the absence of a gravitational field.

Finally, we replace by means of (1.86) and (1.87) the gauge fields A_i and ϕ by the field strengths E_i and B^i. The result is the Lorentz force law (1.17),

$$F_i = q(E_i + \epsilon_{ijk}\partial_t x^j B^k)\,. \tag{1.131}$$

Therefore, in the gauge field approach the Lorentz force is a consequence of the minimal coupling procedure which couples electrically charged particles to the electromagnetic potentials.

1.3.3   Bianchi identity and magnetic flux conservation

The electromagnetic gauge fields A_i and ϕ are often introduced as mathematical tools to facilitate the integration of the Maxwell equations. Indeed, the equations (1.90) and (1.91) have revealed that the homogeneous Maxwell equations reduce to mere integrability conditions that automatically are fulfilled if the electromagnetic field strengths are expressed in terms of the gauge potentials. This is an interesting observation since within the gauge approach the gauge potentials are fundamental physical quantities and are not only the outcome of a mathematical trick. Thus we can state that the mathematical structure of the gauge potentials already implies the homogeneous Maxwell equations and, in turn, magnetic flux conservation. In this light, magnetic flux conservation, within the gauge approach, appears as the consequence of a geometric identity. This is in contrast to electric charge conservation that can be viewed as the consequence of gauge invariance, i.e., via the Noether theorem as the consequence of a physical symmetry.

The integrability conditions that are reflected in the homogeneous Maxwell equations are special cases of *Bianchi identities*. Bianchi identities are the result of differentiating a potential twice. For example, in electrostatics the electric field strength E_i can be derived from a scalar potential ϕ according to

$$E_i = -\partial_i\phi\,. \tag{1.132}$$

Differentiation reveals that the curl of E_i vanishes,

$$\epsilon^{ijk}\partial_j E_k = \epsilon^{ijk}\partial_j\partial_k\phi = 0\,, \tag{1.133}$$

which is due to the antisymmetry of ϵ^{ijk}. Again, this equation is a mathematical identity, a simple example of a Bianchi identity.

1.3.4   Gauge approach and constitutive relations

The gauge approach towards electrodynamics deals with the properties of gauge fields, which represent the electromagnetic field, and with matter fields. It does not reflect properties of spacetime. In contrast to this, the constitutive relations do reflect properties of spacetime, as can be already seen from the constitutive relations of vacuum that
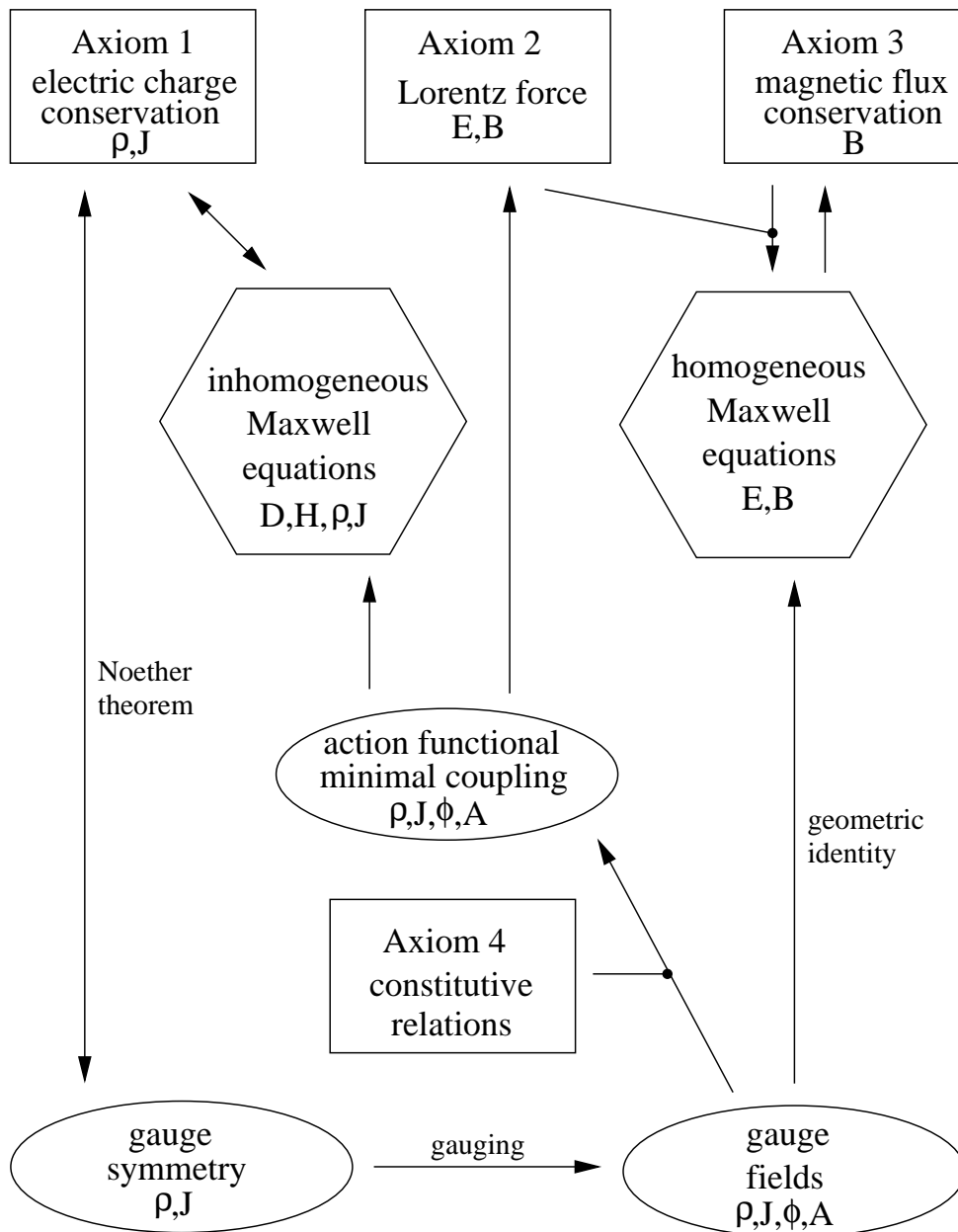
Figure 1.6: Interrelations between the axiomatic approach (rectangular frames) and the gauge field approach (elliptic frames) of classical electrodynamics. Both approaches yield the Maxwell equations. The gauge approach requires the knowledge of constitutive relations which represent the fourth axiom of the axiomatic approach. However, the first, second, and third axiom of the axiomatic approach can be obtained from the gauge approach. Electric charge conservation, the first axiom of the gauge approach, represents the gauge symmetry of electrodynamics by means of the Noether theorem.

involve the metric g_{ij}, compare (1.36) and (1.37). Thus, also in the gauge approach the constitutive relations have to be postulated as an axiom in some way. One should note that, according to (1.86), (1.87), the gauge potentials are directly related to the field strengths E_i and B^i. The excitations D^i and H_i are part of the inhomogeneous Maxwell equations which, within the gauge approach, are derived as equations of motion from an action principle, compare (1.109) and (1.110). Since the action itself involves the gauge potentials, one might wonder how it is possible to obtain equations of motion for the excitations rather than for the field strengths. The answer is that during the construction of the Lagrangian density (1.97) from the gauge potentials the constitutive relations are implicitly used.

Fig. 1.6 summarizes the interrelations between the axiomatic approach and the gauge approach.

1.4   More fundamental equations of electromagnetic field theory

From the axiomatic approach and from the gauge field approach to classical electrodynamics we obtained the Maxwell equations (1.5), (1.16), (1.30), and (1.35). Appropriate constitutive relations of the form (1.36), (1.37) or, more general, (1.38), (1.39) make Maxwell equations a set of determined partial differential equations. Within the limits of classical physics these equations model the interaction between electromagnetic sources ρ, J^i and the electromagnetic field, represented by (E_i, B^i) and (D^i, H_i). In order to explicitly formulate a specific problem we have to impose physically meaningful initial and boundary conditions that lead to a well-defined boundary value problem. The solution of such a boundary value problem, in turn, determines a unique solution of Maxwell equations.

For the solution of an electromagnetic boundary value problem it often is advantageous to first rewrite Maxwell equations as (second order) wave equations. This is straightforward as long as the constitutive relations are of a simple form[8]. Clearly, the solution of wave equations has been studied in many branches of physics and mathematics for a long time and a variety of corresponding solution procedures exist [146].

In the following we will no longer use the tensor notation of the previous sections but turn to the more conventional vector notation of ordinary vector analysis that has been adopted by many authors of standard textbooks, see for example [30, 43, 93, 183, 227]. The widespread use of this notation is the reason why we will adopt it as well. Hence, we replace the covariant vectors E_i, H_i by ordinary vectors in three-dimensional space,

---

[8]Constitutive relations usually are *not* of a simple form if they introduce material parameters that are space or time dependent, if they mix electric and magnetic fields, or if they introduce nonlinearities, for example. In such cases analytic solutions of Maxwell equations often cannot be found and it is required to directly apply numerical methods to the Maxwell equations.

$E_i \to \mathbf{E}$, $H_i \to \mathbf{H}$. Also the contravariant vector densities B^i, D^i, and J^i are replaced by ordinary vectors, $B^i \to \mathbf{B}$, $D^i \to \mathbf{D}$, and $J^i \to \mathbf{J}$, while the notation for the scalar density ρ remains unchanged. It is obvious that due to this transition we lose information on the geometric properties of the electromagnetic quantities. For example, integration along lines or over surfaces is no longer defined in a natural way since line integrals $\int_c \mathbf{E} \cdot d\mathbf{t}$ or surface integrals $\int_S \mathbf{B} \cdot d\mathbf{a}$ now require a metric structure of space which is represented by the scalar product. Formally, we may also form the expressions $\int_c \mathbf{D} \cdot d\mathbf{t}$ or $\int_S \mathbf{H} \cdot d\mathbf{a}$ but the physical interpretation of these integrals is not clear and certainly needs explanation since the electric excitation $\mathbf{D}$ or the magnetic excitation $\mathbf{H}$ are not natural integrands of line- or surface-integrals, respectively. This indicates that it is important to keep the limitations of the vector notation in mind in order to construct mathematical expressions of physical objects in a meaningful way.

In the following we will first consider electromagnetic quantities that are defined in the time domain. But quickly we will also pass to the frequency domain, see subsection 1.4.3, and to reciprocal space, see subsection 1.4.4. It is convenient to print the corresponding Fourier transforms using different fonts. Then it is not necessary to always keep the arguments $\mathbf{r}$, t, ω, or k in parentheses next to the symbols of the fields to indicate which domain or space they belong to. The letters of the different fonts are introduced in the table below. Greek letters remain unaltered, however, this should not lead to confusion since usually they do not appear isolated in an equation.

	Time domain	Frequency domain	Reciprocal space
electric excitation	$\mathbf{D}(\mathbf{r}, t)$	$\boldsymbol{D}(\mathbf{r}, \omega)$	$\mathsf{D}(k, t)$
magnetic excitation	$\mathbf{H}(\mathbf{r}, t)$	$\boldsymbol{H}(\mathbf{r}, \omega)$	$\mathsf{H}(k, t)$
electric field strength	$\mathbf{E}(\mathbf{r}, t)$	$\boldsymbol{E}(\mathbf{r}, \omega)$	$\mathsf{E}(k, t)$
magnetic field strength	$\mathbf{B}(\mathbf{r}, t)$	$\boldsymbol{B}(\mathbf{r}, \omega)$	$\mathsf{B}(k, t)$
charge density	$\rho(\mathbf{r}, t)$	$\rho(\mathbf{r}, \omega)$	$\rho(k, t)$
current density	$\mathbf{J}(\mathbf{r}, t)$	$\boldsymbol{J}(\mathbf{r}, \omega)$	$\mathsf{J}(k, t)$
vector potential	$\mathbf{A}(\mathbf{r}, t)$	$\boldsymbol{A}(\mathbf{r}, \omega)$	$\mathsf{A}(k, t)$
scalar potential	$\phi(\mathbf{r}, t)$	$\phi(\mathbf{r}, \omega)$	$\phi(k, t)$

Table 1.1: Different fonts that are used to distinguish between the basic electromagnetic quantities in time domain, frequency domain, and reciprocal space.

1.4.1   Decoupling of Maxwell equations and wave equations

We rewrite the Maxwell equations (1.5), (1.16), (1.30), and (1.35) in vector notation. This yields the familiar expressions

$$\boldsymbol{\nabla} \cdot \mathbf{D}(\boldsymbol{r}, t) = \rho(\boldsymbol{r}, t) \,, \tag{1.134}$$

$$\boldsymbol{\nabla} \times \mathbf{H}(\boldsymbol{r}, t) - \frac{\partial \mathbf{D}}{\partial t}(\boldsymbol{r}, t) = \mathbf{J}(\boldsymbol{r}, t) \,, \tag{1.135}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B}(\boldsymbol{r}, t) = 0 \,, \tag{1.136}$$

$$\boldsymbol{\nabla} \times \mathbf{E}(\boldsymbol{r}, t) + \frac{\partial \mathbf{B}}{\partial t}(\boldsymbol{r}, t) = \mathbf{0} \,. \tag{1.137}$$

To decouple the Maxwell equations we assume constitutive relations which characterize a homogeneous, isotropic medium. These are of the form

$$\mathbf{D}(\boldsymbol{r}, t) = \varepsilon \mathbf{E}(\boldsymbol{r}, t) \,, \tag{1.138}$$

$$\mathbf{B}(\boldsymbol{r}, t) = \mu \mathbf{H}(\boldsymbol{r}, t) \,, \tag{1.139}$$

with constant parameters ε and μ. Then we apply the curl operator $\boldsymbol{\nabla}\times$ to (1.135), (1.137) and combine the results. This yields

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}(\boldsymbol{r}, t) + \varepsilon\mu \frac{\partial^2}{\partial t^2} \mathbf{E}(\boldsymbol{r}, t) = -\mu \frac{\partial \mathbf{J}}{\partial t}(\boldsymbol{r}, t) \,, \tag{1.140}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{B}(\boldsymbol{r}, t) + \varepsilon\mu \frac{\partial^2}{\partial t^2} \mathbf{B}(\boldsymbol{r}, t) = \mu \boldsymbol{\nabla} \times \mathbf{J}(\boldsymbol{r}, t) \,. \tag{1.141}$$

These equations can be transformed into standard wave equations if the identity (B.6) is used, together with the constitutive relations (1.138), (1.139) and the Maxwell equations (1.134), (1.136). We obtain

$$\Delta \mathbf{E}(\boldsymbol{r}, t) - \varepsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2}(\boldsymbol{r}, t) = \frac{1}{\varepsilon} \boldsymbol{\nabla} \rho(\boldsymbol{r}, t) + \mu \frac{\partial \mathbf{J}}{\partial t}(\boldsymbol{r}, t) \,, \tag{1.142}$$

$$\Delta \mathbf{B}(\boldsymbol{r}, t) - \varepsilon\mu \frac{\partial^2 \mathbf{B}}{\partial t^2}(\boldsymbol{r}, t) = -\mu \boldsymbol{\nabla} \times \mathbf{J}(\boldsymbol{r}, t) \,. \tag{1.143}$$

Due to the constitutive relations (1.138), (1.139) two analogous equations are valid for $\mathbf{D}(\boldsymbol{r}, t)$ and $\mathbf{H}(\boldsymbol{r}, t)$ which furnish no additional information. Therefore we arrive at six scalar equations for six unknown field components. Equations (1.142), (1.143) constitute inhomogeneous wave equations[9] with phase velocity

$$c = \frac{1}{\sqrt{\varepsilon\mu}} \,. \tag{1.144}$$

---

[9]Since in equations (1.142), (1.143) the wave operator $\Delta - \varepsilon\mu\, \partial^2/\partial t^2$ acts on three-dimensional vectors it is clear that these equations are *vector wave equations*. However, also equations of the type (1.140), (1.141) often are called vector wave equations, even though the differential operator $-(\boldsymbol{\nabla} \times \boldsymbol{\nabla}\times) - \varepsilon\mu\, \partial^2/\partial t^2$ not always coincides with the wave operator. It does coincide with the wave operator if it acts on a divergence free vector field.

In vacuum ($\rho(\boldsymbol{r}, t) = 0, \mathbf{J}(\boldsymbol{r}, t) = \mathbf{0}$) the inhomogeneous terms on the right hand sides vanish and we obtain the homogeneous wave equations

$$\Delta\mathbf{E}(\boldsymbol{r}, t) - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}(\boldsymbol{r}, t) = \mathbf{0}\,, \tag{1.145}$$

$$\Delta\mathbf{B}(\boldsymbol{r}, t) - \frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2}(\boldsymbol{r}, t) = \mathbf{0}\,. \tag{1.146}$$

1.4.2   Equations of motion for the electromagnetic potentials

We already have seen that the field strength $\mathbf{E}(\boldsymbol{r}, t)$ and $\mathbf{B}(\boldsymbol{r}, t)$ can be derived from the scalar potential $\phi(\boldsymbol{r}, t)$ and the vector potential $\mathbf{A}(\boldsymbol{r}, t)$ via

$$\mathbf{E}(\boldsymbol{r}, t) = -\boldsymbol{\nabla}\phi(\boldsymbol{r}, t) - \frac{\partial \mathbf{A}}{\partial t}(\boldsymbol{r}, t)\,, \tag{1.147}$$

$$\mathbf{B}(\boldsymbol{r}, t) = \boldsymbol{\nabla} \times \mathbf{A}(\boldsymbol{r}, t)\,. \tag{1.148}$$

If the electromagnetic field is expressed by means of $\phi(\boldsymbol{r}, t)$ and $\mathbf{A}(\boldsymbol{r}, t)$ the homogeneous Maxwell equations (1.136), (1.137) are recognized as geometric identities which automatically are fulfilled. Then the remaining inhomogeneous Maxwell equations (1.134), (1.135) determine the electromagnetic field. We replace within the inhomogeneous Maxwell equations the excitations $\mathbf{D}(\boldsymbol{r}, t)$, $\mathbf{H}(\boldsymbol{r}, t)$ by means of the constitutive relations (1.138), (1.139) and the equations (1.147), (1.148) by $\phi(\boldsymbol{r}, t)$ and $\mathbf{A}(\boldsymbol{r}, t)$. This yields

$$\Delta\phi(\boldsymbol{r}, t) + \frac{\partial(\boldsymbol{\nabla} \cdot \mathbf{A}(\boldsymbol{r}, t))}{\partial t} = -\frac{\rho(\boldsymbol{r}, t)}{\varepsilon}\,, \tag{1.149}$$

$$\Delta\mathbf{A}(\boldsymbol{r}, t) - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}(\boldsymbol{r}, t)}{\partial t^2} - \boldsymbol{\nabla}\left(\boldsymbol{\nabla} \cdot \mathbf{A}(\boldsymbol{r}, t) + \frac{1}{c^2}\frac{\partial\phi(\boldsymbol{r}, t)}{\partial t}\right) = -\mu\mathbf{J}(\boldsymbol{r}, t)\,. \tag{1.150}$$

These are four scalar equations for the four unknown field components $\phi(\boldsymbol{r}, t)$ and $\mathbf{A}(\boldsymbol{r}, t)$. Since electrodynamics is invariant under the gauge transformations

$$\delta_\epsilon\phi(\boldsymbol{r}, t) = -\frac{\partial\epsilon(\boldsymbol{r}, t)}{\partial t}\,, \tag{1.151}$$

$$\delta_\epsilon\mathbf{A}(\boldsymbol{r}, t) = \boldsymbol{\nabla}\epsilon(\boldsymbol{r}, t)\,, \tag{1.152}$$

with an arbitrary function $\epsilon(\boldsymbol{r}, t)$ we may simplify (1.149) and (1.150) by the choice of a particular gauge. Common gauges are the Coulomb gauge

$$\boldsymbol{\nabla} \cdot \mathbf{A}(\boldsymbol{r}, t) = 0 \tag{1.153}$$

and the Lorenz gauge

$$\boldsymbol{\nabla} \cdot \mathbf{A}(\boldsymbol{r}, t) + \frac{1}{c^2}\frac{\partial\phi(\boldsymbol{r}, t)}{\partial t} = 0\,. \tag{1.154}$$

The Coulomb gauge leads to

$$\Delta\phi(\boldsymbol{r},t) = -\frac{\rho(\boldsymbol{r},t)}{\varepsilon}\,, \qquad \text{(Coulomb gauge)} \quad (1.155)$$

$$\Delta\mathbf{A}(\boldsymbol{r},t) - \frac{1}{c^2}\frac{\partial^2\mathbf{A}(\boldsymbol{r},t)}{\partial t^2} - \frac{1}{c^2}\frac{\partial(\boldsymbol{\nabla}\phi(\boldsymbol{r},t))}{\partial t} = -\mu\mathbf{J}(\boldsymbol{r},t)\,. \quad \text{(Coulomb gauge)} \quad (1.156)$$

while the Lorenz gauge yields a scalar and a vector wave equation,

$$\Delta\phi(\boldsymbol{r},t) - \frac{1}{c^2}\frac{\partial^2\phi(\boldsymbol{r},t)}{\partial t^2} = -\frac{\rho(\boldsymbol{r},t)}{\varepsilon}\,, \qquad \text{(Lorenz gauge)} \qquad (1.157)$$

$$\Delta\mathbf{A}(\boldsymbol{r},t) - \frac{1}{c^2}\frac{\partial^2\mathbf{A}(\boldsymbol{r},t)}{\partial t^2} = -\mu\mathbf{J}(\boldsymbol{r},t)\,. \qquad \text{(Lorenz gauge)} \qquad (1.158)$$

1.4.3 Maxwell equations in frequency domain and Helmholtz equations

Time harmonic fields with sinusoidal time dependency can be expressed as

$$\mathbf{F}_{\text{sinus}}(\boldsymbol{r},t) = \text{Re}\big[\boldsymbol{F}(\boldsymbol{r},\omega)e^{j\omega t}\big]\,. \tag{1.159}$$

This is a special case of the Fourier representation of a field with arbitrary time dependency,

$$\mathbf{F}(\boldsymbol{r},t) = \text{Re}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\boldsymbol{F}(\boldsymbol{r},\omega)e^{j\omega t}\,d\omega\right]\,. \tag{1.160}$$

In the time harmonic case we may pass to the frequency domain and write the Maxwell equations as

$$\boldsymbol{\nabla}\cdot\boldsymbol{D}(\boldsymbol{r},\omega) = \rho(\boldsymbol{r},\omega)\,, \tag{1.161}$$

$$\boldsymbol{\nabla}\times\boldsymbol{H}(\boldsymbol{r},\omega) - j\omega\boldsymbol{D}(\boldsymbol{r},\omega) = \boldsymbol{J}(\boldsymbol{r},\omega)\,, \tag{1.162}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B}(\boldsymbol{r},\omega) = 0\,, \tag{1.163}$$

$$\boldsymbol{\nabla}\times\boldsymbol{E}(\boldsymbol{r},\omega) + j\omega\boldsymbol{B}(\boldsymbol{r},\omega) = \boldsymbol{0}\,. \tag{1.164}$$

The vector wave equations (1.140) and (1.142) of the electric field **E**, for example, convert in the frequency domain to vector *Helmholtz equations*

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{E}(\boldsymbol{r},\omega) - k^2\boldsymbol{E}(\boldsymbol{r},\omega) = -j\omega\mu\boldsymbol{J}(\boldsymbol{r},\omega)\,, \tag{1.165}$$

$$\Delta\boldsymbol{E}(\boldsymbol{r},\omega) + k^2\boldsymbol{E}(\boldsymbol{r},\omega) = \frac{1}{\varepsilon}\boldsymbol{\nabla}\rho(\boldsymbol{r},\omega) + j\omega\mu\boldsymbol{J}(\boldsymbol{r},\omega)\,, \tag{1.166}$$

with $k = \omega/c$ the wave number. For the magnetic field **B** we have

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{B}(\boldsymbol{r},\omega) - k^2\boldsymbol{B}(\boldsymbol{r},\omega) = \mu\boldsymbol{\nabla}\times\boldsymbol{J}(\boldsymbol{r},\omega)\,, \tag{1.167}$$

$$\Delta\mathbf{B}(\boldsymbol{r},\omega) + k^2\boldsymbol{B}(\boldsymbol{r},\omega) = -\mu\boldsymbol{\nabla}\times\boldsymbol{J}(\boldsymbol{r},\omega)\,. \tag{1.168}$$

The Helmholtz equations for the scalar and vector potential in the Lorenz gauge assume the form

$$\Delta\phi(\boldsymbol{r},\omega) + k^2\phi(\boldsymbol{r},t) = -\frac{\rho(\boldsymbol{r},\omega)}{\varepsilon}\,, \qquad \text{(Lorenz gauge)} \qquad (1.169)$$

$$\Delta\boldsymbol{A}(\boldsymbol{r},\omega) + k^2\boldsymbol{A}(\boldsymbol{r},\omega) = -\mu\boldsymbol{J}(\boldsymbol{r},\omega)\,. \qquad \text{(Lorenz gauge)} \qquad (1.170)$$

If compared to (1.166), (1.168) these equations are more simple since now the source terms involve no derivatives.

1.4.4 Maxwell equations in reciprocal space

Fields $\mathbf{F}(\boldsymbol{r},t)$ that are defined in time domain are transformed to reciprocal space by a spatial Fourier transform according to

$$\mathsf{F}(\boldsymbol{k},t) = \frac{1}{(2\pi)^{3/2}}\int \mathbf{F}(\boldsymbol{r},t)e^{j\boldsymbol{k}\cdot\boldsymbol{r}}\,d^3r\,. \qquad (1.171)$$

The inverse transform is given by

$$\mathbf{F}(\boldsymbol{r},t) = \frac{1}{(2\pi)^{3/2}}\int \mathsf{F}(\boldsymbol{k},t)e^{-j\boldsymbol{k}\cdot\boldsymbol{r}}\,d^3k\,. \qquad (1.172)$$

The operator $\boldsymbol{\nabla}$ transforms to multiplication by $-j\boldsymbol{k}$ in reciprocal space. Therefore, the Maxwell equations in reciprocal space become

$$-j\boldsymbol{k}\cdot\mathsf{D}(\boldsymbol{k},t) = \rho(\boldsymbol{k},t)\,, \qquad (1.173)$$

$$-j\boldsymbol{k}\times\mathsf{H}(\boldsymbol{k},t) - \frac{\partial\mathsf{D}(\boldsymbol{k},t)}{\partial t} = \mathsf{J}(\boldsymbol{k},t)\,, \qquad (1.174)$$

$$-j\boldsymbol{k}\cdot\mathsf{B}(\boldsymbol{k},t) = 0\,, \qquad (1.175)$$

$$-j\boldsymbol{k}\times\mathsf{E}(\boldsymbol{k},t) + \frac{\partial\mathsf{B}(\boldsymbol{k},t)}{\partial t} = \mathsf{0}\,. \qquad (1.176)$$

This representation of Maxwell equations has the advantage that the fields and time derivatives all depend on the *same* point $\boldsymbol{k}$ in reciprocal space. Hence, the partial differential equations of real space become strictly local equations in reciprocal space.

1.4.5 Boundary conditions at interfaces

At the transition between two media with parameters $(\varepsilon_1, \mu_1, \sigma_1)$ and $(\varepsilon_2, \mu_2, \sigma_2)$ the boundary conditions can be derived from Maxwell equations by means of integration and application of Stokes' theorem. This is a standard procedure which is described in

many textbooks, see, for example, [61, § 7.3.6.]. We have

$$\boldsymbol{\nabla} \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0} \qquad \Longrightarrow \qquad (\mathbf{E}_1 - \mathbf{E}_2) \times \boldsymbol{e}_n = \mathbf{0}\,, \tag{1.177}$$

$$\boldsymbol{\nabla} \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} \qquad \Longrightarrow \qquad (\mathbf{H}_1 - \mathbf{H}_2) \times \boldsymbol{e}_n = \boldsymbol{J}_s\,, \tag{1.178}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \qquad \Longrightarrow \qquad (\mathbf{B}_1 - \mathbf{B}_2) \cdot \boldsymbol{e}_n = 0\,, \tag{1.179}$$

$$\boldsymbol{\nabla} \cdot \mathbf{D} = \rho \qquad \Longrightarrow \qquad (\mathbf{D}_1 - \mathbf{D}_2) \cdot \boldsymbol{e}_n = \rho_S\,. \tag{1.180}$$

These boundary conditions are valid both in time and frequency domain. The vector $\boldsymbol{e}_n$ denotes a normal unit vector that points on the interface between the different media from medium 1 to medium 2 and $\boldsymbol{J}_s$ denotes a surface current that may flow on the interface between the media. Accordingly, ρ_s denotes a surface charge density.

We also have a boundary condition for the magnetic vector potential $\mathbf{A}$,

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} \qquad \Longrightarrow \qquad (\mathbf{A}_1 - \mathbf{A}_2) \times \boldsymbol{e}_n = \mathbf{0}\,. \tag{1.181}$$

This boundary condition is gauge invariant since it is based on the rotational part of the magnetic vector potential.

1.5   Basic electromagnetic field properties

From the previous basic equations of the electromagnetic field it is already possible to derive some basic electromagnetic field properties without to explicitly solve electromagnetic boundary problems. An important issue is the dynamical content of the electromagnetic field which will be considered next.

1.5.1   Dynamical and nondynamical components of the electromagnetic field

(a) Helmholtz's vector theorem, longitudinal and transverse fields

In the study of vector fields $\mathbf{F}$ the *Helmholtz's vector theorem* is a useful tool [146]. It states that any vector field $\mathbf{F}$, which is finite, uniform, continuous and square integrable, may be split into a *longitudinal* or *irrotational* part $\mathbf{F}_\parallel$ and a *transverse* or *rotational* part $\mathbf{F}_\perp$,

$$\mathbf{F} = \mathbf{F}_\parallel + \mathbf{F}_\perp\,, \tag{1.182}$$

where $\mathbf{F}_\parallel$ and $\mathbf{F}_\perp$ are implicitly defined by

$$\boldsymbol{\nabla} \times \mathbf{F}_\parallel = \mathbf{0}\,, \tag{1.183}$$

$$\boldsymbol{\nabla} \cdot \mathbf{F}_\perp = 0\,. \tag{1.184}$$

This split is unique. A good discussion of the Helmholtz's vector theorem is contained in [206] where it is stressed that the theorem critically depends on the boundary conditions of the field $\mathbf{F}$. In the following we will assume that the boundary conditions are such that the Helmholtz's vector theorem can be applied[10]. The names *longitudinal* and *transverse* acquire a clear geometric interpretation in reciprocal space where (1.183) and (1.184) become

$$-j\boldsymbol{k} \times \mathsf{F}_{\parallel} = \mathbf{0}\,, \tag{1.185}$$

$$-j\boldsymbol{k} \cdot \mathsf{F}_{\perp} = 0\,, \tag{1.186}$$

that is, $\mathsf{F}_{\parallel}$ is parallel to $\boldsymbol{k}$ and $\mathsf{F}_{\perp}$ is perpendicular to $\boldsymbol{k}$.

By means of the vector identity (B.6) and the relation

$$\Delta\left(\frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|}\right) = -4\pi\delta(\boldsymbol{r} - \boldsymbol{r}') \tag{1.187}$$

it can be shown that in real space the longitudinal and transverse part of a vector field $\mathbf{F}(\boldsymbol{r}, t)$ are given by

$$\mathbf{F}_{\parallel}(\boldsymbol{r}, t) = -\frac{1}{4\pi}\boldsymbol{\nabla}\int \frac{\boldsymbol{\nabla}' \cdot \mathbf{F}(\boldsymbol{r}', t)}{|\boldsymbol{r} - \boldsymbol{r}'|}\, d^3r'\,, \tag{1.188}$$

$$\mathbf{F}_{\perp}(\boldsymbol{r}, t) = \frac{1}{4\pi}\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \int \frac{\mathbf{F}(\boldsymbol{r}', t)}{|\boldsymbol{r} - \boldsymbol{r}'|}\, d^3r'\,, \tag{1.189}$$

respectively. These explicit formulas show that the split $\mathbf{F} = \mathbf{F}_{\parallel} + \mathbf{F}_{\perp}$ introduces non-local effects: Both $\mathbf{F}_{\parallel}(\boldsymbol{r}, t)$ and $\mathbf{F}_{\perp}(\boldsymbol{r}, t)$, considered at a fixed time t and at a specific point $\boldsymbol{r}$, depend on the values of $\mathbf{F}(\boldsymbol{r}', t)$ at the *same* time and at *all* points $\boldsymbol{r}'$ in space. Conversely, even if $\mathbf{F}(\boldsymbol{r}, t)$ is localized in space, i.e., if it vanishes outside a compact region, the parts $\mathbf{F}_{\parallel}(\boldsymbol{r}, t)$ and $\mathbf{F}_{\perp}(\boldsymbol{r}, t)$ generally will extend over the whole space.

Example: We consider a point charge q at a position $\boldsymbol{R}(t)$ which moves with velocity $\boldsymbol{v}(t)$ and is observed from a position $\boldsymbol{r}$, compare Fig. 1.7

The charge density ρ and current density $\mathbf{J}$ of this point charge are given by

$$\rho(\boldsymbol{r}, t) = q\delta(\boldsymbol{r} - \boldsymbol{R}(t))\,, \tag{1.190}$$

$$\mathbf{J}(\boldsymbol{r}, t) = q\boldsymbol{v}(t)\delta(\boldsymbol{r} - \boldsymbol{R}(t))\,, \tag{1.191}$$

---

[10]From a microscopic but still classical point of view macroscopic boundary conditions are the result of the interaction between the electromagnetic field and electrically charged particles, like electrons or protons. Therefore, if we consider fundamental properties of the electromagnetic field, we may focus on these microscopic interactions and do not separately need to consider macroscopic boundary conditions. For example, we may replace the boundary conditions that are imposed by a perfect conductor by the interaction between the electromagnetic field and the electrons and protons that represent the electrically charged particles of the perfect conductor.
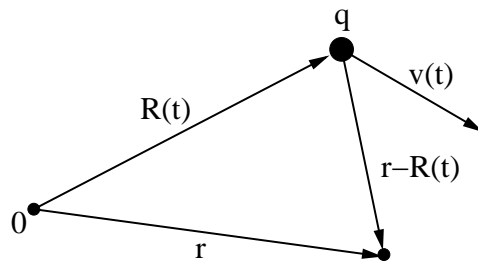
Figure 1.7: Coordinates of a moving point charge q. Position and velocity of the point charge are related by $\boldsymbol{v}(t) = \partial \boldsymbol{R}(t)/\partial t$.

respectively. To calculate the corresponding longitudinal current $\mathbf{J}_\parallel(\boldsymbol{r}, t)$ we use (1.188) and apply the continuity equation

$$\boldsymbol{\nabla}' \cdot \mathbf{J}(\boldsymbol{r}', t) = -\frac{\partial \rho}{\partial t}(\boldsymbol{r}', t) \,. \tag{1.192}$$

This yields

$$\mathbf{J}_\parallel(\boldsymbol{r}, t) = \frac{q}{4\pi} \frac{\partial}{\partial t} \boldsymbol{\nabla} \int \frac{\delta(\boldsymbol{r}' - \boldsymbol{R}(t))}{|\boldsymbol{r} - \boldsymbol{r}'|} \, d^3 r' \tag{1.193}$$

$$= \frac{q}{4\pi} \frac{\partial}{\partial t} \boldsymbol{\nabla} \left(\frac{1}{|\boldsymbol{r} - \boldsymbol{R}(t)|} \right) \tag{1.194}$$

$$= -\frac{q}{4\pi} \frac{\partial}{\partial t} \left(\frac{\boldsymbol{r} - \boldsymbol{R}(t)}{|\boldsymbol{r} - \boldsymbol{R}(t)|^3} \right) \tag{1.195}$$

$$= \frac{q}{4\pi} \left[\frac{\boldsymbol{v}(t)}{|\boldsymbol{r} - \boldsymbol{R}(t)|^3} - \frac{3(\boldsymbol{r} - \boldsymbol{R}(t))[(\boldsymbol{r} - \boldsymbol{R}(t)) \cdot \boldsymbol{v}(t)]}{|\boldsymbol{r} - \boldsymbol{R}(t)|^5} \right] \,. \tag{1.196}$$

Accordingly, due to $\mathbf{J}_\perp(\boldsymbol{r}, t) = \mathbf{J}(\boldsymbol{r}, t) - \mathbf{J}_\parallel(\boldsymbol{r}, t)$, we also have

$$\mathbf{J}_\perp(\boldsymbol{r}, t) = \frac{q}{4\pi} \left[4\pi \boldsymbol{v}(t) \delta(\boldsymbol{r} - \boldsymbol{R}(t)) - \frac{\boldsymbol{v}(t)}{|\boldsymbol{r} - \boldsymbol{R}(t)|^3} + \frac{3(\boldsymbol{r} - \boldsymbol{R}(t))[(\boldsymbol{r} - \boldsymbol{R}(t)) \cdot \boldsymbol{v}(t)]}{|\boldsymbol{r} - \boldsymbol{R}(t)|^5} \right] \tag{1.197}$$

and it is clearly seen that both $\mathbf{J}_\parallel(\boldsymbol{r}, t)$ and $\mathbf{J}_\perp(\boldsymbol{r}, t)$ extend over the whole space.

(b) Nondynamical Maxwell equations as boundary conditions in time

We turn to the complete set of Maxwell equations (1.134) – (1.137) and first note that (1.134) and (1.136) are no dynamical equations but rather so-called *boundary conditions* that determine appropriate initial conditions of the fields. By virtue of the remaining dynamical Maxwell equations they are fulfilled at all times if they are fulfilled at one time. To illustrate this circumstance for the boundary condition (1.136) we assume that at some initial time t_0 we have

$$\boldsymbol{\nabla} \cdot \mathbf{B}|_{t_0} = 0 \,. \tag{1.198}$$

It then needs to be shown that at an infinitesimally later time $t_0 + dt$ we have

$$\boldsymbol{\nabla} \cdot \mathbf{B}|_{t_0+dt} = 0 \tag{1.199}$$

i.e., that

$$\frac{\partial(\boldsymbol{\nabla} \cdot \mathbf{B})}{\partial t}\bigg|_{t_0} = 0 \,. \tag{1.200}$$

However, this condition immediately follows if we take the divergence of the dynamical Maxwell equation (1.137). Similarly, we find from (1.135)

$$\frac{\partial}{\partial t}\left(\boldsymbol{\nabla} \cdot \mathbf{D} - \rho\right)\bigg|_{t_0} = -\left(\boldsymbol{\nabla} \cdot \mathbf{J} + \frac{\partial \rho}{\partial t}\right)\bigg|_{t_0} \tag{1.201}$$

$$= 0 \,, \tag{1.202}$$

where in the second step the continuity equation (1.10) has been employed. Therefore, it is sufficient to calculate the solutions of (1.134) and (1.136) at an initial time t_0 and then solve with these solutions as boundary conditions the dynamical Maxwell equations (1.135) and (1.137) to obtain the time evolution of the electromagnetic field.

(c) Longitudinal part of the Maxwell equations

The Maxwell equation (1.134) can be written as

$$\boldsymbol{\nabla} \cdot \mathbf{D}_{\|}(\boldsymbol{r}, t) = \rho(\boldsymbol{r}, t) \tag{1.203}$$

and relates the longitudinal electric excitation $\mathbf{D}_{\|}$ to the charge density ρ. In reciprocal space this relation becomes

$$-j\boldsymbol{k} \cdot \mathsf{D}_{\|}(\boldsymbol{k}, t) = \rho(\boldsymbol{k}, t) \tag{1.204}$$

and can easily be solved for $\mathsf{D}_{\|}(\boldsymbol{k}, t)$ to yield

$$\mathsf{D}_{\|}(\boldsymbol{k}, t) = j\rho(\boldsymbol{k}, t)\frac{\boldsymbol{k}}{k^2} \,. \tag{1.205}$$

An inverse Fourier transform to real space gives the result

$$\mathbf{D}_{\|}(\boldsymbol{r}, t) = \frac{1}{4\pi} \int \rho(\boldsymbol{r}', t)\frac{\boldsymbol{r} - \boldsymbol{r}'}{|\boldsymbol{r} - \boldsymbol{r}'|^3}\, d^3r' \,. \tag{1.206}$$

This result is quite remarkable since it turns out that the longitudinal electric displacement is completely determined from the *instantaneous* Coulomb field of the charge distribution. With the constitutive relation (1.138) the same is true for the longitudinal electric field strength,

$$\mathbf{E}_{\|}(\boldsymbol{r}, t) = \frac{1}{4\pi\varepsilon} \int \rho(\boldsymbol{r}', t)\frac{\boldsymbol{r} - \boldsymbol{r}'}{|\boldsymbol{r} - \boldsymbol{r}'|^3}\, d^3r' \,. \tag{1.207}$$

The fact that $\mathbf{D}_\parallel(\boldsymbol{r}, t)$ and $\mathbf{E}_\parallel(\boldsymbol{r}, t)$ instantly respond to a change of the charge density does, at this point, not necessarily imply that causality is violated since we require the *complete* fields $\mathbf{D}(\boldsymbol{r}, t)$, $\mathbf{E}(\boldsymbol{r}, t)$ to be causal.

The longitudinal part of the second inhomogeneous Maxwell equation (1.135) is given by

$$-\frac{\partial \mathbf{D}_\parallel(\boldsymbol{r}, t)}{\partial t} = \mathbf{J}_\parallel(\boldsymbol{r}, t) \tag{1.208}$$

We take the divergence of this equation and reveal that it reduces to the continuity equation

$$\frac{\partial \rho(\boldsymbol{r}, t)}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{J}_\parallel(\boldsymbol{r}, t) = 0 \tag{1.209}$$

Therefore, (1.208) conveys no additional information.

We summarize that the longitudinal components of the electromagnetic field are determined from the instantaneous Coulomb field of the electric charge density. It follows that the longitudinal components do not have their own degrees of freedom, they are tied to the degrees of freedom of the electric charge density.

(d) Transverse part of the Maxwell equations

What is left to investigate are the transverse parts of the Maxwell equations (1.135) and (1.137),

$$\boldsymbol{\nabla} \times \mathbf{H}_\perp(\boldsymbol{r}, t) - \frac{\partial \mathbf{D}_\perp}{\partial t}(\boldsymbol{r}, t) = \mathbf{J}_\perp(\boldsymbol{r}, t) \,, \tag{1.210}$$

$$\boldsymbol{\nabla} \times \mathbf{E}_\perp(\boldsymbol{r}, t) + \frac{\partial \mathbf{B}_\perp}{\partial t}(\boldsymbol{r}, t) = \mathbf{0} \,. \tag{1.211}$$

With simple constitutive relations of the form (1.138), (1.139) these equations are easily decoupled and we arrive at the transverse part of the wave equations (1.142), (1.143),

$$\Delta \mathbf{E}_\perp(\boldsymbol{r}, t) - \varepsilon\mu \frac{\partial^2 \mathbf{E}_\perp}{\partial t^2}(\boldsymbol{r}, t) = \mu \frac{\partial \mathbf{J}_\perp}{\partial t}(\boldsymbol{r}, t) \,, \tag{1.212}$$

$$\Delta \mathbf{B}_\perp(\boldsymbol{r}, t) - \varepsilon\mu \frac{\partial^2 \mathbf{B}_\perp}{\partial t^2}(\boldsymbol{r}, t) = -\mu \boldsymbol{\nabla} \times \mathbf{J}_\perp(\boldsymbol{r}, t) \,. \tag{1.213}$$

Since

$$\mathbf{B}_\perp = \mathbf{B} \tag{1.214}$$

we will drop in the following the transverse index $\perp$ of the magnetic field strength.

From the wave equations (1.212) and (1.213) it appears that $\mathbf{E}_\perp$ and $\mathbf{B}$ are the dynamical quantities of the electromagnetic field with two independent components each. However, one needs to note that $\mathbf{E}_\perp$ and $\mathbf{B}$ are not independent of each other. To explicitly show how both quantities are interrelated we rewrite the dynamical Maxwell equations (1.210), (1.211) in reciprocal space. With the constitutive relations (1.138),

(1.139), the relations $c^2 = 1/(\epsilon\mu)$, $\omega = ck$, and the notation $\hat{\boldsymbol{k}} = \boldsymbol{k}/k$ we find, similar to (1.174), and (1.176) the equations

$$\frac{\partial \mathsf{E}_\perp(\boldsymbol{k},t)}{\partial t} = -j\omega c\hat{\boldsymbol{k}} \times \mathsf{B}(\boldsymbol{k},t) - \frac{\mathsf{J}_\perp(\boldsymbol{k},t)}{\varepsilon_0} \tag{1.215}$$

$$c\hat{\boldsymbol{k}} \times \frac{\partial \mathsf{B}(\boldsymbol{k},t)}{\partial t} = -j\omega \mathsf{E}_\perp(\boldsymbol{k},t) \tag{1.216}$$

In the sourceless case with $\mathsf{J}_\perp = 0$ one recognizes from these equations by addition and subtraction that eigenfunctions of this system are determined from

$$\frac{\partial}{\partial t}\left(\mathsf{E}_\perp(\boldsymbol{k},t) - c\hat{\boldsymbol{k}} \times \mathsf{B}(\boldsymbol{k},t)\right) = j\omega\left(\mathsf{E}_\perp(\boldsymbol{k},t) - c\hat{\boldsymbol{k}} \times \mathsf{B}(\boldsymbol{k},t)\right), \tag{1.217}$$

$$\frac{\partial}{\partial t}\left(\mathsf{E}_\perp(\boldsymbol{k},t) + c\hat{\boldsymbol{k}} \times \mathsf{B}(\boldsymbol{k},t)\right) = -j\omega\left(\mathsf{E}_\perp(\boldsymbol{k},t) + c\hat{\boldsymbol{k}} \times \mathsf{B}(\boldsymbol{k},t)\right). \tag{1.218}$$

To label these eigenfunctions we introduce variables $\mathsf{a}(\boldsymbol{k},t)$ and $\mathsf{b}(\boldsymbol{k},t)$ by

$$\mathsf{a}(\boldsymbol{k},t) := \frac{j}{2\mathsf{N}(k)}\left[\mathsf{E}_\perp(\boldsymbol{k},t) - c\hat{\boldsymbol{k}} \times \mathsf{B}(\boldsymbol{k},t)\right], \tag{1.219}$$

$$\mathsf{b}(\boldsymbol{k},t) := \frac{j}{2\mathsf{N}(k)}\left[\mathsf{E}_\perp(\boldsymbol{k},t) + c\hat{\boldsymbol{k}} \times \mathsf{B}(\boldsymbol{k},t)\right]. \tag{1.220}$$

The factor $j/2\mathsf{N}(k)$ denotes a normalization coefficient and is in accordance to a common notation that is used in the context of the quantization of the electromagnetic field [28]. In this context the function $\mathsf{N}(k)$ is related to the energy of a quantum state of the electromagnetic field. For our purposes the explicit form of $\mathsf{N}(k)$ is not important. Within expressions of the electromagnetic field in real space the function $\mathsf{N}(k)$ will cancel and drop out. It could also be absorbed in the definition of $\mathsf{a}(\boldsymbol{k},t)$ and $\mathsf{b}(\boldsymbol{k},t)$.

It is immediate to solve (1.219), (1.220) for $\mathsf{E}_\perp(\boldsymbol{k},t)$ and $\mathsf{B}(\boldsymbol{k},t)$. Since both quantities have to be real it turns out that we have to require

$$\mathsf{b}(\boldsymbol{k},t) = -\mathsf{a}^*(-\boldsymbol{k},t), \tag{1.221}$$

where the asterisk * denotes complex conjugation. Then we find

$$\mathsf{E}_\perp(\boldsymbol{k},t) = -j\mathsf{N}(k)\left[\mathsf{a}(\boldsymbol{k},t) - \mathsf{a}^*(-\boldsymbol{k},t)\right], \tag{1.222}$$

$$\mathsf{B}(\boldsymbol{k},t) = -\frac{j\mathsf{N}(k)}{c}\left[\hat{\boldsymbol{k}} \times \mathsf{a}(\boldsymbol{k},t) + \hat{\boldsymbol{k}} \times \mathsf{a}^*(-\boldsymbol{k},t)\right]. \tag{1.223}$$

Therefore, the transverse electromagnetic field is completely specified by the function $\mathsf{a}(\boldsymbol{k},t)$. Since $\mathsf{E}_\perp(\boldsymbol{k},t)$ and $\mathsf{B}(\boldsymbol{k},t)$ are transverse functions it follows that $\mathsf{a}(\boldsymbol{k},t)$ is a transverse function, too. Hence, we conclude that $\mathsf{a}(\boldsymbol{k},t)$ exhibits two degrees of freedom which are the two dynamical components of the electromagnetic field. The function $\mathsf{a}(\boldsymbol{k},t)$ is said to represent the *normal modes* of the electromagnetic field. This term

indicates that $\mathsf{a}(\boldsymbol{k}, t)$ represents a whole class of electromagnetic excitations which is parameterized by a discrete or continuous set of values for the wavenumber $\boldsymbol{k}$.

We may insert (1.222) and (1.223) in the Maxwell equations and obtain for the time evolution of $\mathsf{a}(\boldsymbol{k}, t)$ the equation

$$\frac{\partial \mathsf{a}(\boldsymbol{k}, t)}{\partial t} - j\omega \mathsf{a}(\boldsymbol{k}, t) = -\frac{j}{2\varepsilon \mathsf{N}(k)} \mathsf{J}_\perp(\boldsymbol{k}, t) \,. \tag{1.224}$$

This equation of motion for the normal modes represents in fact the motion of harmonic oscillation: If we implicitly introduce a new variable $\mathsf{c}(\boldsymbol{k}, t)$ via

$$\mathsf{a}(\boldsymbol{k}, t) = \mathsf{c}(\boldsymbol{k}, t) - \frac{j}{\omega} \frac{\partial \mathsf{c}(\boldsymbol{k}, t)}{\partial t} \tag{1.225}$$

we find from (1.224) the familiar equation of motion of a harmonic oscillator

$$\frac{\partial^2 \mathsf{c}(\boldsymbol{k}, t)}{\partial t^2} + \omega^2 \mathsf{c}(\boldsymbol{k}, t) = \frac{\omega}{2\varepsilon \mathsf{N}(k)} \mathsf{J}_\perp(\boldsymbol{k}, t) \,. \tag{1.226}$$

From (1.222) and (1.223) we also obtain for the fields $\mathbf{E}_\perp(\boldsymbol{r}, t)$ and $\mathbf{B}(\boldsymbol{t}, t)$ from a Fourier transformation the expansions

$$\mathbf{E}_\perp(\boldsymbol{r}, t) = -\frac{j}{(2\pi)^{3/2}} \int \mathsf{N}(k) \Big[\mathsf{a}(\boldsymbol{k}, t) e^{-j\boldsymbol{k}\cdot\boldsymbol{r}} - \mathsf{a}^*(\boldsymbol{k}, t) e^{j\boldsymbol{k}\cdot\boldsymbol{r}} \Big] d^3k \,, \tag{1.227}$$

$$\mathbf{B}(\boldsymbol{r}, t) = -\frac{j}{(2\pi)^{3/2}} \int \frac{\mathsf{N}(k)}{c} \Big[\hat{\boldsymbol{k}} \times \mathsf{a}(\boldsymbol{k}, t) e^{-j\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{\boldsymbol{k}} \times \mathsf{a}^*(\boldsymbol{k}, t) e^{j\boldsymbol{k}\cdot\boldsymbol{r}} \Big] d^3k \,. \tag{1.228}$$

We remind ourselves that $\mathsf{a}(\boldsymbol{k}, t)$ and $\mathsf{a}^*(\boldsymbol{k}, t)$ are purely complex quantities such that the fields $\mathbf{E}_\perp(\boldsymbol{r}, t)$ and $\mathbf{B}(\boldsymbol{r}, t)$ are real.

In the absence of sources we have $\mathsf{J}_\perp(\boldsymbol{k}, t) = \mathsf{0}$. Then the equation of motion (1.224) yields the solution

$$\mathsf{a}(\boldsymbol{k}, t) = \mathsf{a}(\boldsymbol{k}) e^{j\omega t} \tag{1.229}$$

and the expansions (1.227), (1.228) turn into expansions in traveling plane waves,

$$\mathbf{E}_{\perp\text{free}}(\boldsymbol{r}, t) = -\frac{j}{(2\pi)^{3/2}} \int \mathsf{N}(k) \Big[\mathsf{a}(\boldsymbol{k}) e^{j(\omega t - \boldsymbol{k}\cdot\boldsymbol{r})} - \mathsf{a}^*(\boldsymbol{k}) e^{-j(\omega t - \boldsymbol{k}\cdot\boldsymbol{r})} \Big] d^3k \,, \tag{1.230}$$

$$\mathbf{B}_{\text{free}}(\boldsymbol{r}, t) = -\frac{j}{(2\pi)^{3/2}} \int \frac{\mathsf{N}(k)}{c} \Big[\hat{\boldsymbol{k}} \times \mathsf{a}(\boldsymbol{k}) e^{j(\omega t - \boldsymbol{k}\cdot\boldsymbol{r})} - \hat{\boldsymbol{k}} \times \mathsf{a}^*(\boldsymbol{k}) e^{-j(\omega t - \boldsymbol{k}\cdot\boldsymbol{r})} \Big] d^3k \,.$$
$$\tag{1.231}$$

In these expansions of the free fields the functions $\mathsf{a}(\boldsymbol{k}, t) = \mathsf{a}(\boldsymbol{k}) e^{j\omega t}$ corresponding to different $\boldsymbol{k}$ are completely decoupled. This holds also true if the electromagnetic sources, represented by $\mathsf{J}_\perp(\boldsymbol{k}, t)$, are independent of $\mathsf{a}(\boldsymbol{k}, t)$, i.e., independent of the electromagnetic field. However, if the electromagnetic sources do interact with the electromagnetic field the time evolution of $\mathsf{J}_\perp(\boldsymbol{k}, t)$ will depend on $\mathsf{a}(\boldsymbol{k}, t)$ and, in general, lead to a coupling between $\mathsf{a}(\boldsymbol{k}, t)$ with different $\boldsymbol{k}$.

1.5.2 Electromagnetic energy and the singularities of the electromagnetic field

The energy density $w(\boldsymbol{r}, t)$ of the electromagnetic field is given by

$$w_{\mathrm{em}}(\boldsymbol{r}, t) = \frac{1}{2} \Big(\mathbf{E}(\boldsymbol{r}, t) \cdot \mathbf{D}(\boldsymbol{r}, t) + \mathbf{B}(\boldsymbol{r}, t) \cdot \mathbf{H}(\boldsymbol{r}, t) \Big) \tag{1.232}$$

and the corresponding energy $W(t)$ is

$$W_{\mathrm{em}}(t) = \int w_{\mathrm{em}}(\boldsymbol{r}, t) \, d^3 r \tag{1.233}$$

$$= \frac{1}{2} \int \big(\mathbf{E}(\boldsymbol{r}, t) \cdot \mathbf{D}(\boldsymbol{r}, t) + \mathbf{B}(\boldsymbol{r}, t) \cdot \mathbf{H}(\boldsymbol{r}, t) \big) \, d^3 r \,. \tag{1.234}$$

This energy, in general, is not a constant of motion if the electromagnetic field interacts with electromagnetic sources.

We want to split the electromagnetic field energy into a contribution of the longitudinal fields $\mathbf{E}_\parallel$, $\mathbf{D}_\parallel$ and a contribution of the transverse fields $\mathbf{E}_\perp$, $\mathbf{D}_\perp$, $\mathbf{B}$, and $\mathbf{H}$. Clearly, the magnetic part of $W_{\mathrm{em}}(t)$ only involves the transverse fields $\mathbf{B}$, $\mathbf{H}$. The electric part can be written as

$$\frac{1}{2} \int \mathbf{E}(\boldsymbol{r}, t) \cdot \mathbf{D}(\boldsymbol{r}, t) \, d^3 r$$

$$= \frac{1}{2} \int \mathsf{E}^*(\boldsymbol{k}, t) \cdot \mathsf{D}(\boldsymbol{k}, t) \, d^3 k \tag{1.235}$$

$$= \frac{1}{2} \int \big(\mathsf{E}_\parallel{}^*(\boldsymbol{k}, t) + \mathsf{E}_\perp{}^*(\boldsymbol{k}, t) \big) \cdot \big(\mathsf{D}_\parallel(\boldsymbol{k}, t) + \mathsf{D}_\perp(\boldsymbol{k}, t) \big) \, d^3 k \tag{1.236}$$

$$= \frac{1}{2} \int \mathsf{E}_\parallel{}^*(\boldsymbol{k}, t) \cdot \mathsf{D}_\parallel(\boldsymbol{k}, t) \, d^3 k + \frac{1}{2} \int \mathsf{E}_\perp{}^*(\boldsymbol{k}, t) \cdot \mathsf{D}_\perp(\boldsymbol{k}, t) \, d^3 k \tag{1.237}$$

$$= \frac{1}{2} \int \mathbf{E}_\parallel(\boldsymbol{r}, t) \cdot \mathbf{D}_\parallel(\boldsymbol{r}, t) \, d^3 r + \frac{1}{2} \int \mathbf{E}_\perp(\boldsymbol{r}, t) \cdot \mathbf{D}_\perp(\boldsymbol{r}, t) \, d^3 r \,. \tag{1.238}$$

The result is the desired split into longitudinal and transverse contributions to $W_{\mathrm{em}}(t)$.

If we assume a constitutive relation of the form (1.138) and take into account the results (1.206), (1.207) it is easy to see that the longitudinal contribution $W_{\mathrm{em}\parallel}$ to the electromagnetic field energy W_{em} is just given by the electrostatic Coulomb energy,

$$W_{\mathrm{em}\parallel}(t) = \frac{1}{2} \int \mathbf{E}_\parallel(\boldsymbol{r}, t) \cdot \mathbf{D}_\parallel(\boldsymbol{r}, t) \, d^3 r$$

$$= \frac{1}{8\pi\varepsilon} \iint \frac{\rho(\boldsymbol{r}, t) \rho(\boldsymbol{r}', t)}{|\boldsymbol{r} - \boldsymbol{r}'|} \, d^3 r \, d^3 r' \,. \tag{1.239}$$

For the transverse contribution

$$W_{\mathrm{em}\perp}(t) = \frac{1}{2} \int \big(\mathbf{E}_\perp(\boldsymbol{r}, t) \cdot \mathbf{D}_\perp(\boldsymbol{r}, t) + \mathbf{B}(\boldsymbol{r}, t) \cdot \mathbf{H}(\boldsymbol{r}, t) \big) \, d^3 r \tag{1.240}$$

we cannot find in real space a result of similar simplicity. But if we assume the validity of the constitutive relations (1.138), (1.139) and shift to reciprocal space we can write

$$W_{\mathrm{em}\perp}(t) = \frac{\varepsilon}{2} \int \left(\mathsf{E}_\perp{}^*(\boldsymbol{k}, t) \cdot \mathsf{E}_\perp(\boldsymbol{k}, t) + \mathsf{B}^*(\boldsymbol{k}, t) \cdot \mathsf{B}(\boldsymbol{k}, t) \right) d^3 k \qquad (1.241)$$

and insert into this expression the relations (1.222) and (1.223). From simple vector algebra we then obtain

$$W_{\mathrm{em}\perp}(t) = \varepsilon \int \mathsf{N}^2(k) \left[\mathsf{a}^*(\boldsymbol{k}, t) \cdot \mathsf{a}(\boldsymbol{k}, t) + \mathsf{a}(-\boldsymbol{k}, t) \cdot \mathsf{a}^*(-\boldsymbol{k}, t) \right] d^3 k \,. \qquad (1.242)$$

Therefore, the transverse contribution to the electromagnetic energy is completely determined from the normal modes $\mathsf{a}(\boldsymbol{k}, t)$.

Next we inspect under which conditions the electromagnetic energy of a system becomes divergent: From the expression (1.239) for the electrostatic Coulomb energy of the longitudinal fields it is seen that this energy depends on the relative position of the electric sources within a system. It diverges in the limit $|\boldsymbol{r} \to \boldsymbol{r}'|$ in which case two sources become arbitrarily close. This behavior characterizes the *Coulomb singularity* of classical electrodynamics and is related to the infinite amount of Coulomb energy that is carried by any electric source. It is independent of the kinematical or dynamical state of a source, i.e., it is independent of velocity or acceleration. In particular, in the time-harmonic case it is independent of frequency. The expression (1.242) for the energy of the transverse fields exhibits no spatial singularity. We have seen from (1.224) and (1.226) that the normal modes $\mathsf{a}(\boldsymbol{k}, t)$ represent oscillations of the electromagnetic field that are driven by the transverse part $\mathsf{J}_\perp(\boldsymbol{k}, t)$ of the electric current. Dominant contributions to the energy (1.242) occur if the excitation is such that it operates at an eigenfrequency ω of the system. In this case of an *electromagnetic resonance* it is known from the elementary solution of the equation of motion of a forced harmonic oscillator that the amplitude of the resulting oscillation will tend to infinity if no loss mechanism is present [14]. Then the energy (1.242) will tend to infinity as well. It follows that in the lossless case the oscillations of the electromagnetic field become singular at resonance. This is the second type of electromagnetic singularity which leads to diverging electromagnetic energy. Since forced harmonic motion is the only solution of the equation of motion (1.226) this type of singularity is the only one that is contained in (1.242).

As a result we can state that the electromagnetic field exhibits two types of singularities that lead to diverging electromagnetic energy densities. These are

- Coulomb singularities that are related to the mutual position of electric sources

- electromagnetic resonances that are related to forced oscillations of the electromagnetic field

1.5.3 Coulomb fields and radiation fields

The split of the electromagnetic fields $\mathbf{D}$, $\mathbf{H}$, $\mathbf{E}$, $\mathbf{B}$ and the electric current $\mathbf{J}$ in longitudinal and transverse parts already has been proven to be useful in our study of basic electromagnetic field properties. Unfortunately, it also has some disadvantages. Two major disadvantages, which, in fact, are interrelated, are the following:

- The split in longitudinal and transverse parts is not relativistically invariant. A field which is purely longitudinal if observed from a first reference system will contain transverse contributions if observed from a second reference system which is in relative motion to the first one.

- The decomposition in longitudinal and transverse parts introduces non-local effects. This has been noted in subsection 1.5.1 (a) where the explicit form of the decomposition is given by (1.188) and (1.189).

The physical relevance of these disadvantages becomes apparent if we return to the analysis of the longitudinal and transverse parts of the Maxwell equations as outlined in subsections 1.5.1 (c) and (d): It was found that the longitudinal electric field $\mathbf{E}_{\parallel}$ represents the instantaneous Coulomb field of the electric charge density ρ. Since the complete electric field $\mathbf{E}$, according to (1.142), fulfills a proper wave equation which leads to causal solutions it follows that the transverse electric part $\mathbf{E}_{\perp}$ must contain an instantaneous contribution which exactly cancels the instantaneous contribution of $\mathbf{E}_{\parallel}$. Indeed, it can be shown that this is the case [28]. At first sight this is surprising since $\mathbf{E}_{\perp}$ fulfills the wave equation (1.212). But the solution of this wave equation contains instantaneous contributions since the source term $\mathbf{J}_{\perp}$ is non-local and contains instantaneous contributions as well, compare the short discussion for the general transverse field $\mathbf{F}_{\perp}$ after (1.189). Therefore, the split in longitudinal and transverse electromagnetic fields does not separate physically independent electromagnetic field contributions since longitudinal and transverse electromagnetic fields are inseparably connected to each other.

To illustrate this circumstance we consider the electromagnetic field of a moving point charge where we assume the validity of the constitutive relations (1.36) and (1.37). At a fixed time the point charge is located at a position $\boldsymbol{r}'$ where it moves with velocity $\boldsymbol{v}$, and an observer is positioned at a position $\boldsymbol{r}$. The unit vector $\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}} := (\boldsymbol{r}-\boldsymbol{r}')/|\boldsymbol{r}-\boldsymbol{r}'|$ points from the charge to the observer, compare Fig. 1.8, and we also introduce as an abbreviation $\boldsymbol{\beta} := \boldsymbol{v}/c$. Then the electromagnetic field, expressed in terms of the field strengths $\mathbf{E}$, $\mathbf{B}$, that is noticed by the observer is given by [93]

$$\mathbf{E}(\boldsymbol{r},t) = \underbrace{\frac{q}{4\pi\varepsilon_0}\left[\frac{(\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}}-\boldsymbol{\beta})(1-\beta^2)}{(1-\boldsymbol{\beta}\cdot\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}})^3\,|\boldsymbol{r}-\boldsymbol{r}'|^2}\right]_{\text{ret}}}_{\text{velocity field (Coulomb field)}} + \underbrace{\frac{q}{4\pi\varepsilon_0}\left[\frac{\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}}\times\left((\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}}-\boldsymbol{\beta})\times\frac{\partial\boldsymbol{\beta}}{\partial t}\right)}{c(1-\boldsymbol{\beta}\cdot\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}})^3\,|\boldsymbol{r}-\boldsymbol{r}'|}\right]_{\text{ret}}}_{\text{acceleration field (radiation field)}},$$

$$(1.243)$$

$$\mathbf{B}(\boldsymbol{r}, t) = \frac{1}{c}\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}} \times \mathbf{E}(\boldsymbol{r}, t) \tag{1.244}$$

$$= \underbrace{\frac{q}{4\pi\varepsilon_0}\left[\frac{(\boldsymbol{\beta} \times \boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}})(1-\beta^2)}{c(1-\boldsymbol{\beta}\cdot\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}})^3\,|\boldsymbol{r}-\boldsymbol{r}'|^2}\right]_{\text{ret}}}_{\text{velocity field (Coulomb field)}} + \underbrace{\frac{q}{4\pi\varepsilon_0}\left[\frac{\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}} \times \left(\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}} \times \left((\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}} - \boldsymbol{\beta}) \times \frac{\partial\boldsymbol{\beta}}{\partial t}\right)\right)}{c^2(1-\boldsymbol{\beta}\cdot\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}})^3\,|\boldsymbol{r}-\boldsymbol{r}'|}\right]_{\text{ret}}}_{\text{radiation field (acceleration field)}}.$$

The brackets $[\ ]_{\text{ret}}$ indicate that the enclosed quantities have to be taken at the retarded time t_{ret} that is introduced in the caption of Fig. 1.8. Each of the fields (1.243) and (1.244) nicely splits into a first part which depends on the velocity of the charge and a second part which depends on both the velocity and the acceleration of the charge.
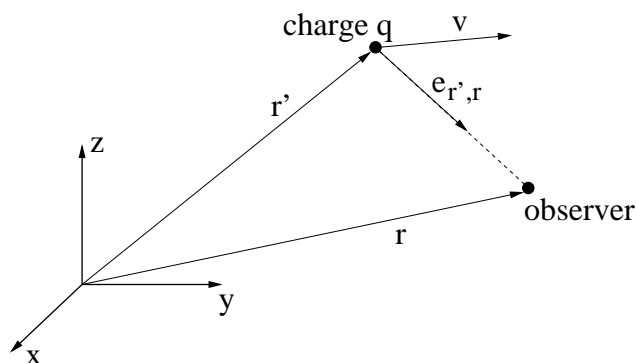

Figure 1.8: A charge q moves with velocity $\boldsymbol{v}$ in the presence of an observer. We assume that the observer does not move with respect to the inertial system xyz. The electromagnetic field that is generated by the electric charge requires the time $|\boldsymbol{r}-\boldsymbol{r}'|/c$ to reach the observer. Therefore, the electromagnetic field that is noticed by the observer at a time t has been generated by the electric charge at the earlier, *retarded time* $t_{\text{ret}} = t - |\boldsymbol{r}-\boldsymbol{r}'|/c$

The velocity fields can also be obtained by a Lorentz transformation of the static fields

$$\mathbf{E}_{\text{static}}(\boldsymbol{r}) = \frac{q}{4\pi\varepsilon}\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\,, \tag{1.245}$$

$$\mathbf{B}_{\text{static}}(\boldsymbol{r}) = \mathbf{0}\,. \tag{1.246}$$

This suggests the name *Coulomb fields* for the electromagnetic velocity fields. They constitute the static Coulomb field of a point charge as noticed by an observer which is in relative motion to the charge. The remaining acceleration fields are commonly called *radiation fields*.

We can also split the fields (1.243) and (1.244) into longitudinal and transverse parts. Since the magnetic field already is purely transverse we only consider the electric field.

The longitudinal component $\mathbf{E}_{\parallel}$ is given by the instantaneous Coulomb field, see (1.207), and it follows

$$\mathbf{E}_{\parallel}(\boldsymbol{r},t) = \underbrace{\frac{q}{4\pi\varepsilon_0}\frac{\boldsymbol{r}-\boldsymbol{r}'(t)}{|\boldsymbol{r}-\boldsymbol{r}'(t)|^3}}_{\text{instantaneous Coulomb field}}, \tag{1.247}$$

$$\mathbf{E}_{\perp}(\boldsymbol{r},t) = \frac{q}{4\pi\varepsilon_0}\underbrace{\left(\left[\frac{(\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}}-\boldsymbol{\beta})(1-\beta^2)}{(1-\boldsymbol{\beta}\cdot\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}})^3\,|\boldsymbol{r}-\boldsymbol{r}'|^2}\right]_{\text{ret}}-\frac{\boldsymbol{r}-\boldsymbol{r}'(t)}{|\boldsymbol{r}-\boldsymbol{r}'(t)|^3}\right)}_{\text{remaining part of the Coulomb field}}$$

$$+\frac{q}{4\pi\varepsilon_0}\underbrace{\left[\frac{\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}}\times\left((\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}}-\boldsymbol{\beta})\times\frac{\partial\beta}{\partial t}\right)}{c(1-\boldsymbol{\beta}\cdot\boldsymbol{e}_{\boldsymbol{r}',\boldsymbol{r}})^3\,|\boldsymbol{r}-\boldsymbol{r}'|}\right]_{\text{ret}}}_{\text{radiation field}}. \tag{1.248}$$

Therefore, the split of the electric field in longitudinal and transverse parts also splits the Coulomb field in two parts and assigns the afore-mentioned instantaneous field contributions to both the longitudinal and the transverse component of the electric field.

In many electrical engineering applications of classical electrodynamics we do not have available the microscopic picture of single electric charges that generate an electromagnetic field. Then we express electromagnetic sources by the charge and current density ρ and $\mathbf{J}$, respectively. In this case we can no longer characterize a resulting electromagnetic field by the velocities and accelerations of microscopic charges and, accordingly, do no longer have the notion of velocity fields and acceleration fields. This implies that, in general, we can no longer split an electromagnetic field into its Coulomb part and its radiation part. We still have the split into longitudinal and transverse parts, but the transverse part contains contributions of the Coulomb part and the complete radiation part.

In principle, we could isolate the radiation part of the electromagnetic field if, at a particular time, we were able to switch off the coupling between electric charges and the electromagnetic field, i.e., if we could switch off electric charges and their accompanying Coulomb fields. Then the remaining radiation field would be the solution of the sourceless Maxwell equations with $\rho = 0$, $\boldsymbol{J} = \boldsymbol{0}$, and nontrivial initial conditions. A solution of the sourceless Maxwell equations commonly is called *free electromagnetic field*. It is a pure radiation field and fulfills the homogeneous wave equations (1.145) and (1.146). A free electromagnetic field has no longitudinal components since its longitudinal electric field component (1.207) vanishes by virtue of $\rho = 0$. The solutions for the transverse components are characterized by oscillatory motion, as is recognized from the solutions (1.230) and (1.231) which, in turn, reflect the solutions of the harmonic oscillator equations (1.224) and (1.226) for a vanishing transverse current. The solutions (1.230) and (1.231) also exhibit that for a radiation field the wave vector $\boldsymbol{k}$, the electric field $\mathbf{E} = \mathbf{E}_{\perp}$, and the magnetic field $\mathbf{B}$ always are mutually orthogonal to each other. It is in this way that a radiation field propagates electromagnetic energy with phase velocity $c = 1/\sqrt{\varepsilon\mu}$ through space.

However, in practice we cannot simply switch off electric charges to neglect the coupling between electric charges and the electromagnetic field. Instead, we have to consider the non-local transverse electric current $\mathbf{J}_\perp$ which drives the transverse electromagnetic field components and, in general, extends through the whole of space. Then the concept of a free electromagnetic field turns to an ideal which, nevertheless, often is a useful one. An example is given by the far field of an antenna in free space. In free space the transverse electric current falls off faster in intensity than the electromagnetic field does. Then the electromagnetic field becomes asymptotically free at large distances where it constitutes the common radiation field.

The fact that we generally cannot split a given electromagnetic field into a Coulomb part and a radiation part indicates that there are some conceptual difficulties in classical electrodynamics which cannot be resolved. These do not necessarily have to be a matter of concern. If we solve an electromagnetic boundary value problem we will usually solve for the complete fields and it might be of no practical interest to know which part of the solution constitutes a Coulomb field and which part represents a radiation field. However, the fact that there are two different categories of electric fields with different, and often complementary, properties will be the reason for many difficulties that we encounter during the solution of practical problems in electrical engineering. As two examples we mention the problematic numerical evaluation of electromagnetic fields that are generated by antennas within cavities (see Chapter 3) and the derivation of conventional transmission line theory from the complete Maxwell theory which requires some inelegant approximations (see Chapter 5).

Chapter 2

Linear Operator Theory, Green's Function Method, and Numerical Methods

In the previous chapter we introduced the main physical quantities of the Maxwell theory. These are the electromagnetic sources $\rho, \mathbf{J}$ and the electromagnetic field which is represented by the electromagnetic excitations $\mathbf{D}, \mathbf{H}$ and the electromagnetic field strengths $\mathbf{E}, \mathbf{B}$. The mathematical setting has been a geometric one: Physical quantities are geometric objects that are defined in space and time, the relevant Maxwell equations can be derived from symmetry principles, and also the fundamental constitutive relations are closely connected to the structure of space and time. In such a geometric field-theoretical framework we can always expand a physical field $\mathbf{F}$ into a basis $\boldsymbol{e}_i$ and corresponding components F^i,

$$\mathbf{F} = \sum_{i=1}^{D} F^i \boldsymbol{e}_i \,. \tag{2.1}$$

A formal expansion of this kind already has been employed in Section 1.2.1, compare (1.42). The basis $\boldsymbol{e}_i$ characterizes the geometric properties of the field $\mathbf{F}$. If $\mathbf{F}$ constitutes a vector in three-dimensional space, for example, the basis $\boldsymbol{e}_i$ constitutes a three-dimensional vector frame, the components F^i are real or complex valued functions, and the dimension of the vector space where $\mathbf{F}$ resides is that of space itself, $D = 3$. An expansion of the form (2.1) explicitly shows, by means of the basis $\boldsymbol{e}_i$, in which way a physical field $\mathbf{F}$ is connected to space or time.

With this understanding of the Maxwell theory we may formulate proper electromagnetic boundary value problems. These will involve the Maxwell equations, together with physically meaningful initial and boundary conditions. Then the solution of an electromagnetic boundary value problems requires to determine the components F^i for a specific basis $\boldsymbol{e}_i$. Corresponding solution procedures are contained in the classic realm of mathematical methods in theoretical physics [32, 146]. In the development of these

solution procedures it has been one of the main mathematical achievements of the twentieth century to merge the three principal branches of mathematics, namely algebra, geometry, and analysis, in order to arrive at the unified framework of *functional analysis*. Functional analysis provides a natural setting for the formulation and solution of electromagnetic boundary value problems.

In functional analysis functions are viewed as elements of a vector space, also called a *linear space* or a *function space*, that, in most cases of practical interest, is infinite dimensional. In such a space a function F^i is expanded in terms of an infinite dimensional set of given basis function ϕ_n according to

$$F^i = \sum_{n=1}^{\infty} F_n^i \phi_n \tag{2.2}$$

with real or complex numbers F_n^i. Then a complete expansion of a physical field $\mathbf{F}$ is written as

$$\mathbf{F} = \sum_{i=1}^{D} \sum_{n=1}^{\infty} F_n^i \phi_n \boldsymbol{e}_i \,. \tag{2.3}$$

In Fig. 2.1 it is schematically shown how the physical space and the function space merge to yield the complete space where $\mathbf{F}$ resides.

The main benefit of an expansion of the form (2.3) is that the determination of a physical field $\mathbf{F}$, that is, the determination of a finite set of functions $\{F^i\}$, is reduced to the determination of an infinite set of real or complex numbers $\{F_n^i\}$. The fact that the set $\{F_n^i\}$ is infinite is, in practice, not too much of a disadvantage. In some cases analytic solution methods may yield the complete infinite set of coefficients F_n^i in analytic form. If no such solution can be found the function F^i can often be approximated to a sufficiently accurate degree by a *finite* expansion

$$F^i \approx \sum_{n=1}^{N} F_n^i \phi_n \,. \tag{2.4}$$

Finite expansions of this form are the basis of *numerical solution methods*. These characterize the solution of a boundary value problem by a finite number of numerically calculated coefficients F_n^i.

2.1   Elements of functional analysis

In order to make the introductory remarks and ideas of this chapter more precise we first need to define function spaces that are appropriate to accommodate functions that represent solutions of (electromagnetic) boundary value problems. These spaces are characterized by a number of algebraic, geometric, and analytic properties. In most cases the function spaces of physical interest are given by so-called *Hilbert spaces* [32].
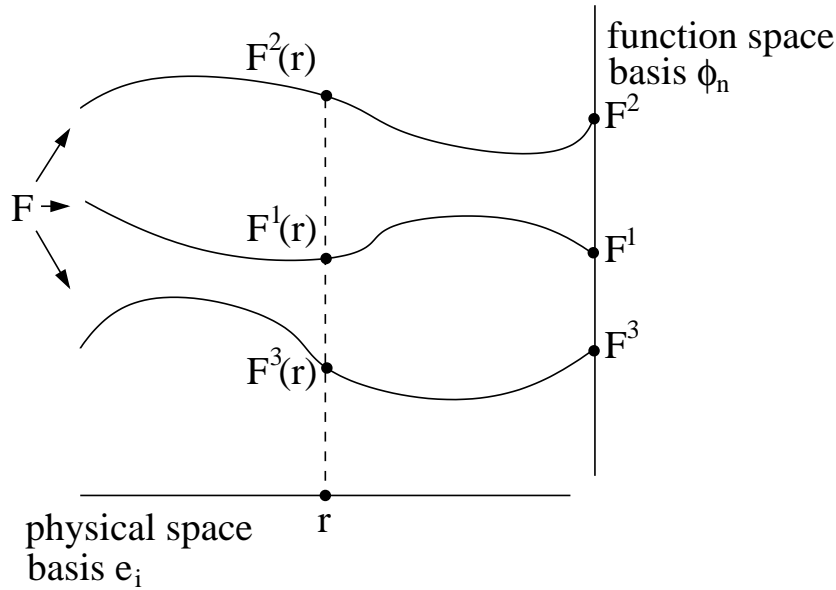
Figure 2.1: Expansion of a vector field $\mathbf{F}$ in its three components F^i via $\mathbf{F} = \sum_{i=1}^{3} F^i \boldsymbol{e}_i$. The number of components matches the dimension of the physical space. The components itself are viewed as elements of a function space. This function space usually is infinite dimensional. In the figure each function F^i is represented by a specific curve which is obtained from an expansion $F^i = \sum_{n=1}^{\infty} F_n^i \phi_n$. The representation of F^i by smooth curves indicates the infinite number of degrees of freedom of each component F^i.

We will outline in the following the requirements for a function space to be a Hilbert space. Then the next step will be to define the notion of a *linear operator* that acts on the elements of a Hilbert space. This serves to reformulate a linear boundary value problem in terms of a linear operator equation. A linear operator equation will be viewed as a linear mapping between two Hilbert spaces. Then the solution of a linear boundary value problem reduces to the construction of the corresponding inverse mapping. This formulation will lead to a proper understanding of the Green's function method and numerical solution procedures.

There is a wealth of literature on the subject of functional analysis. Many treatments extend on a mathematically solid and abstract level. Thorough introductions include [41, 2]. Historically, functional analytic methods have been developed for the solution of partial differential equations that are important in mathematical physics and many books focus on these applications, see for instance [32, 141, 208, 17]. Functional analysis has also been of fundamental importance for the formulation of quantum mechanics [134, 140, 37]. The mathematical framework of quantum mechanics necessarily is a functional analytic one and any serious book on quantum mechanics requires to introduce the corresponding mathematical concepts to some degree. Electromagnetic theory can

be formulated without prior knowledge of functional analysis. But functional analysis unifies the various solution methods for electromagnetic boundary value problem into a single framework and makes it possible to find solutions in a systematic way. This is the main motivation to study functional analysis which, at first sight, seems to be a rather abstract topic and unrelated to electrical engineering problems. Monographs that explain the advantages of functional analysis in an electrical engineering context are given by [40, 96, 81, 243], among others.

> *This Section 2.1 collects a number of mathematical definitions and their use for the solution of electromagnetic boundary value problems is not obvious from the beginning. It might be motivating to know that already in Section 2.2 we will apply these mathematical concepts to the Maxwell theory.*

2.1.1   Function spaces

We consider *sets*, equivalently termed *spaces*, with elements f, g, and h. For the following mathematical definitions these elements are not required to represent physical quantities (even though we have in mind that they do).

Example: A function space that is important in both mathematics and physics is denoted by $\boldsymbol{L}^p(\Omega)^m$ [51, 81]. This is the set of all functions $\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{r}) = (f_1(\boldsymbol{r}), f_2(\boldsymbol{r}), \ldots, f_m(\boldsymbol{r})) \in \mathbb{C}^m$ that are Lebesgue integrable[1] to the pth power on $\boldsymbol{r} = (x_1, x_2, \ldots x_n) \in \Omega \subseteq \mathbb{R}^n$ such that

$$\int_\Omega |\boldsymbol{f}(\boldsymbol{r})|^p \, d\Omega < \infty \,, \tag{2.5}$$

with $1 \le p \le \infty$. Here the absolute value $|\boldsymbol{f}(\boldsymbol{r})| \in \mathbb{R}$ is defined by

$$|\boldsymbol{f}(\boldsymbol{r})| := \left(\boldsymbol{f}(\boldsymbol{r}) \cdot \boldsymbol{f}^*(\boldsymbol{r})\right)^{1/2} = \left(\sum_{i=1}^m f_i(\boldsymbol{r}) f_i^*(\boldsymbol{r})\right)^{1/2}. \tag{2.6}$$

Again, the asterisk * denotes complex conjugation. Of particular interest in electromagnetic applications is the function space $\boldsymbol{L}^2(\Omega)^3$. It accommodates electromagnetic field configurations that are represented by both a three-component electric field and a three component magnetic field and, additionally, are square integrable. Square integrability implies that within the integration domain the electromagnetic energy is finite, compare (1.234). This is a necessary physical requirement [228].

We now add structures to general spaces in order to be able to define Hilbert spaces. This, in turn, requires the definitions of

---

[1]The theory of Lebesgue integration is necessary if we want to integrate functions that are *discontinuous* or *unbounded* [205]. It extends the theory of Riemann integration that is limited to *continuous* functions. In practice, if we wish to explicitly calculate the value of an integral, the differences between Lebesgue integration and Riemann integration play no major role and it is enough to keep in mind the few facts on Lebesgue integration that are listed in [208].

(a) metric spaces

(b) linear spaces or, equivalently, vector spaces

(c) normed spaces

(d) inner product spaces

These different categories of spaces are, in fact, interrelated. This is illustrated in Fig. 2.2. In the following we will first define metric spaces and then proceed towards the definition of Hilbert spaces.
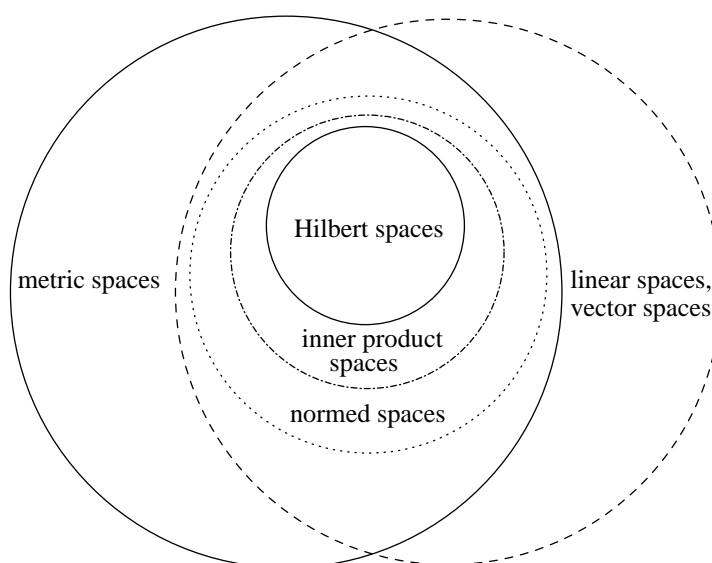

Figure 2.2: Function spaces that are important in functional analysis. A Hilbert space turns out to be a special case of an inner product space which, in turn, is a special case of a normed space. A normed space is both a linear space and a metric space. But a linear space is not necessarily a metric space and vice versa.

(a) Metric spaces

A space S is called a *metric space* if there exists a mapping $d : S \times S \to \mathbb{R}$, $d(f, g) \in \mathbb{R}$ with properties

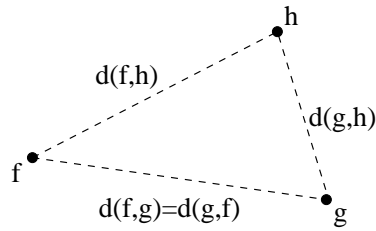$$d(f, g) \geq 0 \,, \tag{2.7}$$

$$d(f, g) = 0 \quad \text{if and only if} \quad f = g \,, \tag{2.8}$$

$$d(f, g) = d(g, f) \,, \tag{2.9}$$

$$d(f, g) \leq d(f, h) + d(g, h) \qquad (\text{triangle inequality}) \,. \tag{2.10}$$

A metric space introduces the concept of *distance*, see Fig. 2.3. The mapping d is called a *metric*.

Figure 2.3: In a metric space a distance is associated to any two elements. In particular, the metric d respects the triangle inequality (2.10).

Example: The function space $\boldsymbol{L}^p(\Omega)^m$ is a metric space. A class of metrics for $\boldsymbol{L}^p(\Omega)^m$ is given by the p-metric d_p,

$$d_p(\boldsymbol{f}, \boldsymbol{g}) := \left(\int_\Omega |\boldsymbol{f}(\boldsymbol{r}) - \boldsymbol{g}(\boldsymbol{r})|^p \, d\Omega \right)^{1/p} \tag{2.11}$$

with $1 \le p < \infty$. In this definition the absolute value (2.6) has been used. The p-metric is also defined for $p = \infty$ via

$$d_\infty(\boldsymbol{f}, \boldsymbol{g}) := \sup_{\boldsymbol{r} \in \Omega} |\boldsymbol{f}(\boldsymbol{r}) - \boldsymbol{g}(\boldsymbol{r})| \,. \tag{2.12}$$

In a metric space the topological notions of 1. continuity, 2. convergence, 3. Cauchy convergence, 4. completeness, 5. denseness, 6. closure, 7. boundedness, and 8. compactness can be defined:

1. If $L : S_1 \to S_2$ is a mapping between two metric spaces S_1, S_2 with metrics d_{S_1}, d_{S_2}, respectively, then L is *continuous* at $f \in S_1$ if for every number $\varepsilon > 0$ there exists a number $\delta > 0$ such that $d_{S_2}(L(f), L(g)) < \varepsilon$ whenever $d_{S_1}(f, g) < \delta$ with $g \in S_1$.

2. A sequence of elements $\{f_n\} = f_1, f_2, \ldots$ in a metric space S with metric d is *convergent* if there is an element f such that for every number $\varepsilon > 0$ there is an integer N with $d(f_n, f) < \varepsilon$ whenever $n > N$.

3. A sequence of elements $\{f_n\} = f_1, f_2, \ldots$ in a metric space S with metric d is *Cauchy convergent* or a *Cauchy series* if for every number $\varepsilon > 0$ there is an integer N with $d(f_n, f_m) < \varepsilon$ whenever $n, m > N$.

4. A metric space S with metric d is *complete* if each Cauchy series in S is a convergent sequence in S.

5. If we consider two subspaces S_1 and S_2 of a metric space S with $S_1 \subset S_2$ then the set S_1 is said to be *dense* in S_2 if for each $g \in S_2$ and each $\epsilon > 0$ there exists an element $f \in S_1$ such that $d(f, g) < \epsilon$. This states that every element of S_2 can be approximated arbitrarily close by elements of the set S_1.

6. The *closure* $\overline{S}$ of a set S consists of the limits of all sequences that can be constructed from S. Then a set S is called *closed* if $\tilde{S} = S$.

7. A set S is *bounded* if there is a real number k such that $d(f, g) < k$ for all $f, g \in S$.

8. A set S is called *compact* if each sequence of elements in S, which is not necessarily convergent, has a subsequence that converges to an element of S. Compact sets are closed and bounded.

For simple illustrations and examples of these notions we refer to [37].

(b) Linear Spaces, Vector Spaces

In a *linear space S*, also called *vector space*, the operations *addition* and *scalar multiplication* are defined, compare Fig. 2.4. Moreover, with real or complex scalars α, β, the following properties are valid,

$$f + g = g + f \,, \tag{2.13}$$

$$(f + g) + h = f + (g + h) \,, \tag{2.14}$$

$$\text{There is a zero element } 0 \text{ such that} \quad f + 0 = f \,, \tag{2.15}$$

$$\text{For every } f \text{ there is an element } \tilde{f} = -f \text{ such that } f + \tilde{f} = 0 \,, \tag{2.16}$$

$$\alpha(f + g) = \alpha f + \alpha g \,, \tag{2.17}$$

$$(\alpha + \beta)f = \alpha f + \beta f \,, \tag{2.18}$$

$$1 \cdot f = f \,, \tag{2.19}$$

$$\alpha(\beta f) = (\alpha\beta)f \,. \tag{2.20}$$

Example: The function space $\boldsymbol{L}^p(\Omega)^m$ is a linear space by the common method of adding functions and multiplying functions with scalars. These additions and scalar multiplications are induced by the usual rules for addition and scalar multiplication in $\mathbb{C}^n$. Therefore, the functions of $\boldsymbol{L}^p(\Omega)^m$, or the functions of some other linear function space, can be referred to as vectors.

Also the spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are linear spaces. In particular, from the common use of vector calculus we are very familiar to consider $\mathbb{R}^3$ as a vector space and its elements as vectors.
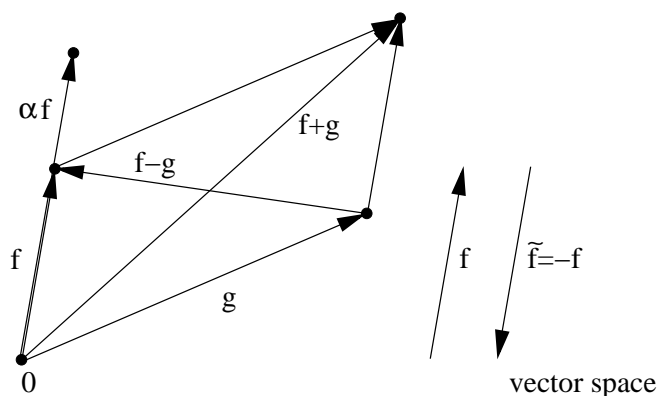
Figure 2.4: In a linear space the operations *addition* and *scalar multiplication* can be performed. Also the notions of a *zero element* and the *negative of an element* are defined.

(c) Normed spaces

A linear space S is *normed* if there is a real valued function $\|f\|$, the norm of f, with properties

$$\|f\| \geq 0 \quad \text{and} \quad \|f\| = 0 \text{ if and only if } f = 0 \,, \tag{2.21}$$

$$\|\alpha f\| = |\alpha| \, \|f\| \,, \tag{2.22}$$

$$\|f + g\| \leq \|f\| + \|g\| \quad \text{for all} \quad f, g \in S \,. \tag{2.23}$$

A normed space introduces the concept of *length*, as illustrated in Fig. 2.5. We note that any normed space is a metric space with $d(f, g) := \|f - g\|$.
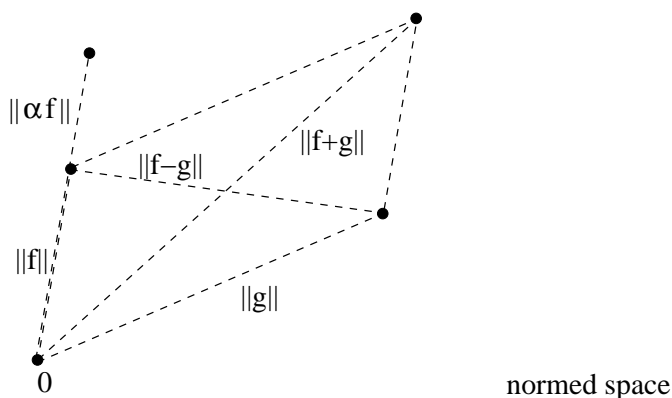

Figure 2.5: In a normed space a *length* $\|f\|$ is associated to each vector f. A norm induces via $d(f, g) := \|f - g\|$ a metric in a natural way. Therefore, any normed space is also a metric space.

Example: The function space $\boldsymbol{L}^p(\Omega)^m$ is a normed space. With $\boldsymbol{f} \in \boldsymbol{L}^p(\Omega)^m$ a class of norms for $\boldsymbol{L}^p(\Omega)^m$ is defined by the p-norm according to

$$\|\boldsymbol{f}\|_p := d_p(\boldsymbol{f}, \boldsymbol{0}) = \left(\int_\Omega |\boldsymbol{f}(\boldsymbol{r})|^p \, d\Omega \right)^{1/p} \tag{2.24}$$

with $1 \le p < \infty$. For $p = \infty$ we define

$$\|\boldsymbol{f}\|_\infty := d_\infty(\boldsymbol{f}, \boldsymbol{0}) = \sup_{\boldsymbol{r} \in \Omega} |\boldsymbol{f}(\boldsymbol{r})| \,. \tag{2.25}$$

(d) Inner product spaces and pseudo inner product spaces

An *inner product space* is a linear space S with an inner product. An inner product is a mapping $S \times S \to \mathbb{C}$ which associates to each ordered pair $f, g \in S$ a complex scalar $\langle f, g \rangle$ with

$$\langle f, g \rangle = \langle g, f \rangle^* \,, \tag{2.26}$$

$$\langle f, f \rangle \ge 0 \quad \text{and} \quad \langle f, f \rangle = 0 \text{ if and only if } f = 0 \,, \tag{2.27}$$

$$\langle \alpha f, g \rangle = \alpha \langle f, g \rangle \,, \tag{2.28}$$

$$\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle \,. \tag{2.29}$$

Within an inner product space the notion of *orthogonality* is defined. Two elements f, g are defined to be orthogonal if $\langle f, g \rangle = 0$. An inner product space also is a normed space since an inner product induces a norm by means of

$$\|f\| := \langle f, f \rangle^{1/2} \,. \tag{2.30}$$

Example: The function space $\boldsymbol{L}^2(\Omega)^m$ is an inner product space. For $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{L}^2(\Omega)^m$ an inner product is given by

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle := \int_\Omega \boldsymbol{f}(\boldsymbol{r}) \cdot \boldsymbol{g}^*(\boldsymbol{r}) \, d\Omega \,. \tag{2.31}$$

The norm induced by this inner product is the two norm $\|\boldsymbol{f}\|_2$,

$$\langle \boldsymbol{f}, \boldsymbol{f} \rangle^{1/2} = \left(\int_\Omega |\boldsymbol{f}(\boldsymbol{r})|^2 \, d\Omega \right)^{1/2} = \|\boldsymbol{f}\|_2 \,. \tag{2.32}$$

One should note that the definition (2.31) presupposes square integrability and, therefore, we require $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{L}^2(\Omega)^m$ rather than $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{L}^p(\Omega)^m$ for arbitrary p. However, if $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{L}^p(\Omega)^m$ for $p \ne 2$ we still can have $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{L}^2(\Omega)^m$, but this is not true in general.

We can attribute a further geometric interpretation to the inner product if we consider the expressions

$$f^{\|_g} := \frac{\langle f, g \rangle g}{\|g\|^2} \,, \tag{2.33}$$

$$f^{\perp_g} := f - \frac{\langle f, g \rangle g}{\|g\|^2} = f - f^{\|_g} \,. \tag{2.34}$$

Here we assume that $\|g\| \neq 0$ and that the norm $\|\ \|$ is defined according to (2.30). From the properties of the inner product we then obtain

$$\langle f^{\|g}, f^{\|g} \rangle + \langle f^{\perp g}, f^{\perp g} \rangle = \langle f, f \rangle \,, \tag{2.35}$$

or

$$\|f^{\|g}\|^2 + \|f^{\perp g}\|^2 = \|f\|^2 \,. \tag{2.36}$$

It is also easy to check that

$$\langle f^{\perp g}, g \rangle = 0 \,. \tag{2.37}$$

These relations are explained in terms of the Pythagorean theorem in Fig. 2.6 where the expression $f^{\|g}$ geometrically is explained as the *projection* of f onto g. In the special case of $\|g\| = 1$ the absolute value $|\langle f, g \rangle|$ of the inner product $\langle f, g \rangle$ coincides with the norm of this projection,

$$\|f^{\|g}\| = \left\| \frac{\langle f, g \rangle g}{\|g\|^2} \right\| \tag{2.38}$$

$$= |\langle f, g \rangle| \frac{\|g\|}{\|g\|^2} \tag{2.39}$$

$$= |\langle f, g \rangle| \,, \tag{2.40}$$

where in the last line $\|g\| = 1$ was used.
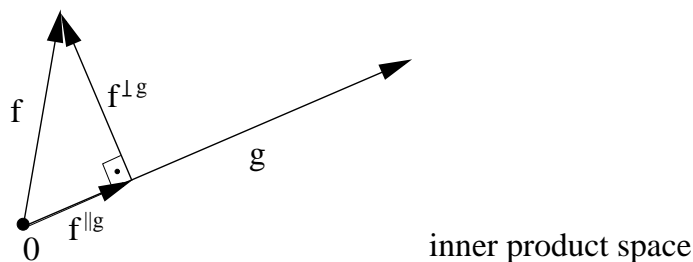
Figure 2.6: Within an inner product space the inner product $\langle \ , \ \rangle$ allows to define the *projection* $f^{\|g}$ of one element f onto another element g. The complement $f^{\perp g}$ is orthogonal to $f^{\|g}$ and we have $f = f^{\|g} + f^{\perp g}$.

In electromagnetic theory also the notion of a *pseudo inner product space* is of importance. A pseudo inner product space is a linear space S which is equipped with a *pseudo inner product*. A pseudo inner product, in turn, is a mapping $S \times S \to \mathbb{C}$ which

associates to each pair f, $g \in S$ a scalar $\langle f, g \rangle_p$ with

$$\langle f, g \rangle_p = \langle g, f \rangle_p \,, \tag{2.41}$$

$$\langle \alpha f, g \rangle_p = \alpha \langle f, g \rangle_p \,, \tag{2.42}$$

$$\langle f + g, h \rangle_p = \langle f, h \rangle_p + \langle g, h \rangle_p \,. \tag{2.43}$$

These properties imply that $\langle f, f \rangle_p$ is not necessarily positive or real-valued. Therefore, a pseudo inner product does not always generate a norm.

Example: The function space $\boldsymbol{L}^2(\Omega)^m$ is a pseudo inner product space. For $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{L}^2(\Omega)^m$ a pseudo inner product is defined via

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle_p := \int_\Omega \boldsymbol{f}(\boldsymbol{r}) \cdot \boldsymbol{g}(\boldsymbol{r}) \, d\Omega \,. \tag{2.44}$$

(e) Hilbert spaces

It was indicated in (2.2) that we want to expand a physical quantity in terms of an infinite series of known basis functions. From a physical point of view it is important to choose basis function with physically meaningful properties. A physical meaningful property can be "continuity", "differentiability", or "square-integrability", for example. Once we have chosen appropriate basis functions we want to be sure that an infinite linear combination of these functions shares the same properties. This requires that the space of the basis functions is complete[2].

We consider as a simple illustration the set $\mathbb{Q}$ of rational numbers. This set is not complete. To show this we consider the sequence $\{\sum_{m=1}^n \frac{1}{m!}\}_{n=1}^\infty$. With the metric $d(f, g) = |f - g|$ this sequence is easily recognized as a Cauchy series. But due to the relation

$$\sum_{n=1}^\infty \underbrace{\frac{1}{n!}}_{\text{rational}} = \underbrace{e}_{\text{irrational}} = 2.71828\ldots \,, \tag{2.45}$$

it does not converge in $\mathbb{Q}$ and it follows that $\mathbb{Q}$ is not complete. Therefore, the result of an infinite linear combination of rational numbers is not necessarily a rational number.

We now turn to the definition of a Hilbert space: An inner product space is called a *Hilbert space* H if it is complete in the induced norm $\|f\| = \langle f, f \rangle^{1/2}$. This means that every Cauchy sequence in H converges to an element of H, that is, for every sequence $\{f_n\} \subset H$ with $\|f_n - f_m\| \to 0$ there exists an $f \in H$ such that $\|f_n - f\| \to 0$. This is illustrated in Fig. 2.7.

Example: The function space $\boldsymbol{L}^2(\Omega)^m$ with the inner product (2.31) is a Hilbert space. This implies that $\boldsymbol{L}^2(\Omega)^m$ needs to be complete. The completeness of $\boldsymbol{L}^2(\Omega)^m$

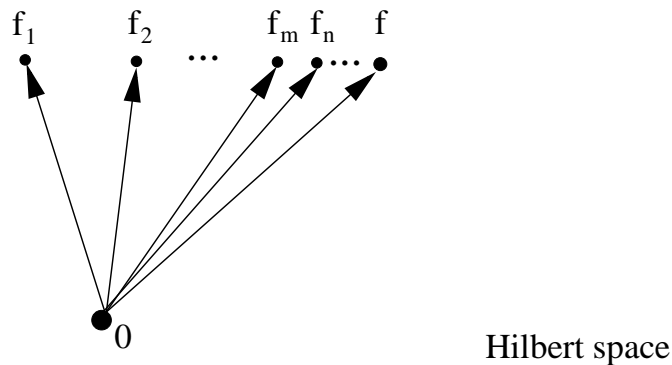[2]Completeness has been defined in Sec. 2.1.1 (a).

Hilbert space

Figure 2.7: An inner product space is a Hilbert space H if any Cauchy sequence $\{f_n\} \subset H$ converges to an element $f \in H$.

follows from the completeness of $\boldsymbol{L}^p(\Omega)^m$ as a metric space. The proof that $\boldsymbol{L}^p(\Omega)^m$ as a metric space is complete is known as the *Riesz-Fischer theorem* and can be found in [51, pp. 99–100], for example.

(f) Finite expansions and best approximation

We now consider the approximation of a given element f of a Hilbert space H by a *finite* set of mutually orthonormal elements $g_m \in H$, $m = 1, \ldots, N$. To this end we form the linear combination $\tilde{f} := \sum_{m=1}^N \alpha_m g_m$ with a sequence of coefficients α_m. The difference $e := f - \tilde{f}$ is defined as an *error* with norm $\|e\| = \|f - \tilde{f}\|$.

Obviously, it is of interest to know which choice of coefficients α_m minimizes the error. We have

$$\|f - \tilde{f}\| = \langle f - \tilde{f}, f - \tilde{f} \rangle \tag{2.46}$$

$$= \langle f, f \rangle + \langle \tilde{f}, \tilde{f} \rangle - \langle \tilde{f}, f \rangle - \langle f, \tilde{f} \rangle \tag{2.47}$$

$$= \|f\|^2 + \sum_{m=1}^N |\alpha_m|^2 - \sum_{m=1}^N \alpha_m \langle f, g_m \rangle^* - \sum_{m=1}^N \alpha_m^* \langle f, g_m \rangle \tag{2.48}$$

$$= \|f\|^2 + \sum_{m=1}^N |\alpha_m - \langle f, g_m \rangle|^2 - \sum_{m=1}^N |\langle f, g_m \rangle|^2, \tag{2.49}$$

and this leads to the conclusion that the choice

$$\alpha_m = \langle f, g_m \rangle \tag{2.50}$$

minimizes the norm of the error. These coefficients are known as *generalized Fourier*

coefficients and the linear combination

$$\tilde{f} = \sum_{m=1}^{N} \langle f, g_m \rangle g_m \tag{2.51}$$

is the *expansion* of f with respect to the elements g_m. It is the best possible expansion of f in terms of the set g_m. A finite-dimensional example of this circumstance is provided by Fig. 2.8.
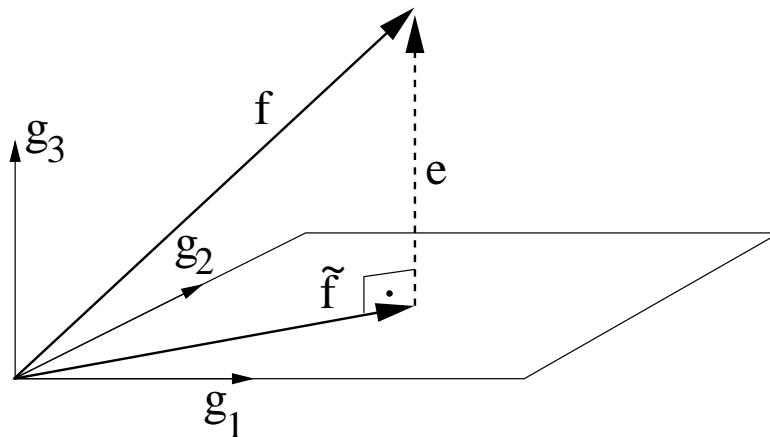
Figure 2.8: Illustration of the "best approximation" of an element $f = \sum_{i=1}^{3} f^i g_i$ by an element $\tilde{f} = \sum_{i=1}^{2} \langle f, g_i \rangle g_i$. The error $e = f - \tilde{f}$ is orthogonal to the approximation $\tilde{f}$.

The error e is orthogonal to the finite expansion, $\langle e, \tilde{f} \rangle = 0$. Since the set of all elements that can be obtained from linear combinations of the set g_m forms a closed linear subspace M of H we arrive at an illustration of the Projection theorem.

(g) Projection theorem:

- If M is a closed linear subspace of a Hilbert space H any element $f \in H$ can uniquely be written as the sum $f = \tilde{f} + e$ of an element $\tilde{f} \in M$ and an element $e \in M^\perp$.

Here the *orthogonal complement* $M^\perp$ is the set of all $f \in H$ such that $\langle f, g \rangle = 0$ for all $g \in M$. A proof of the projection theorem can be found in [37, p. 123].

(h) Basis of a Hilbert space:

A basis of a Hilbert space H is defined as a set of orthonormal elements f_n such that any $f \in H$ can uniquely be written as

$$f = \sum_{n=1}^{\infty} \langle f, f_n \rangle f_n . \tag{2.52}$$

Here we have assumed that the Hilbert space H is infinite-dimensional and, thus, requires an infinite number of basis elements. However, there also exist finite dimensional Hilbert spaces. Examples for finite dimensional Hilbert spaces are $\mathbb{R}^n$ and $\mathbb{C}^n$.

2.1.2   Linear operators

With the definition of a Hilbert space we have provided a function space that will be suitable to accommodate the solutions of (electromagnetic) boundary value problems. However, the structures of a Hilbert space are not sufficient to model the equations of electromagnetic field theory. Therefore, it is required to introduce operators that relate elements of a Hilbert space in a general way.

In what follows we will concentrate on linear operators. This is sufficient as long as the equations which we want to model and, eventually, to solve are linear. Electromagnetic field theory is a linear theory if the constitutive relations are linear. In this case we may apply the methods that are provided by linear operator theory.

(a) Definition of a linear operator, domain and range of an operator

A *linear operator* $\mathcal{L}$ is defined as a linear mapping $\mathcal{L} : S_1 \to S_2$ between linear spaces S_1, S_2. With $f, g \in S_1$, $\mathcal{L}f$, $\mathcal{L}g \in S_2$, and $\alpha, \beta \in \mathbb{C}$ linearity implies

$$\mathcal{L}(\alpha f + \beta g) = \alpha \, \mathcal{L}f + \beta \, \mathcal{L}g . \tag{2.53}$$

The *domain* $D_{\mathcal{L}}$ of an operator $\mathcal{L}$ is the set of all elements $f \in S_1$ for which the operator is defined, while the *range* $R_{\mathcal{L}}$ of the operator $\mathcal{L}$ is the set of elements of S_2 that result from the mapping of the domain.

We will consider in the following linear operators that act between Hilbert spaces H_1, H_2.

(b) Bounded operators and the norm of an operator

We define a linear operator $\mathcal{L} : H_1 \to H_2$ as *bounded* if for all elements $f \in H_1$ there is a real number k such that

$$\|\mathcal{L}f\| \leq k \, \|f\| . \tag{2.54}$$

Operators that are not bounded are called *unbounded.* Closely related to the definition of bounded operators is the definition of the *norm* of an operator: The norm $\|\mathcal{L}\|$ of a linear operator $\mathcal{L} : H_1 \to H_2$ is the smallest number k that satisfies $\|\mathcal{L}f\| \le k \|f\|$ for all $f \in H_1$,

$$\|\mathcal{L}\| := \sup_{\|f\| \neq 0} \frac{\|\mathcal{L}f\|}{\|f\|} . \tag{2.55}$$

From this definition the relation $\|\mathcal{L}f\| \le \|\mathcal{L}\| \|f\|$ follows immediately.

(c) Continuous operators

The *continuity* of an operator that acts between two Hilbert spaces is defined in analogy to the continuity of a mapping between two metric spaces: A linear operator $\mathcal{L} : H_1 \to H_2$ is continuous at an element $f_0 \in H_1$ if for every ε there is a δ such that $\|\mathcal{L}f - \mathcal{L}f_0\| < \varepsilon$ if $\|f - f_0\| < \delta$.

It can be shown that a linear operator $\mathcal{L} : H_1 \to H_2$ is continuous if and only if it is bounded [208, p. 318]. Moreover, if a linear operator $\mathcal{L} : H_1 \to H_2$ is defined on a finite dimensional Hilbert space it is continuous and, thus, also bounded.

(d) Linear functionals

Linear functionals are special cases of linear operators. They map elements of a linear space into the set $\mathbb{C}$ of complex numbers. If we focus on mappings between Hilbert spaces we may define a functional $\mathcal{I}$ as a mapping $\mathcal{I} : H \to \mathbb{C}$ with the property

$$\mathcal{I}(\alpha f + \beta g) = \alpha \mathcal{I}(f) + \beta \mathcal{I}(g) \tag{2.56}$$

for $f, g \in H$ and $\alpha, \beta \in \mathbb{C}$. Clearly, the notions of boundedness, norm, and continuity are defined for linear functionals in the same way as for linear operators. For each element g of a Hilbert space there is a natural bounded linear functional which is defined via the inner product and given by

$$\mathcal{I}_g(f) = \langle f, g \rangle . \tag{2.57}$$

That the converse is also true is the content of the Riesz representation theorem.

(e) Riesz representation theorem:

- For a bounded linear functional $\mathcal{I}$ on a Hilbert space H there is a unique element $g \in H$ such that $\mathcal{I}_g(f) = \langle f, g \rangle$ for all $f \in H$. In this case it also follows that $\|\mathcal{I}\| = \|g\|$.

The Riesz representation theorem is proven in [37, p. 126], for example.

(f) Adjoint and pseudo adjoint operators

To approach the definition of an adjoint operator we first consider a bounded linear operator $\mathcal{L} : H_1 \to H_2$. Then $\mathcal{I}_g(f) = \langle \mathcal{L}f, g \rangle$ is a bounded linear functional $\mathcal{I}_g : H_1 \to \mathbb{C}$ for all $f \in H_1$ and it follows from the Riesz representation theorem that there is a unique $g^\star \in H_1$ such that for all $f \in H_1$

$$\underbrace{\langle \mathcal{L}f, g \rangle}_{\in H_2} = \underbrace{\langle f, g^\star \rangle}_{\in H_1} . \tag{2.58}$$

Since $g^\star$ depends on g we introduce a new operator $\mathcal{L}^\star : H_2 \to H_1$, the so-called *adjoint operator* of $\mathcal{L}$, which is defined by

$$\mathcal{L}^\star g := g^\star . \tag{2.59}$$

This implies

$$\langle \mathcal{L}f, g \rangle = \langle f, \mathcal{L}^\star g \rangle \tag{2.60}$$

and it follows that for any bounded linear operator $\mathcal{L}$ there is a unique adjoint $\mathcal{L}^\star$.

For an unbounded linear operator the Riesz representation theorem does not necessarily hold. However, even if $\mathcal{L}$ is unbounded we still may relate an element $g \in H_2$ to an element $g^\star \in H_1$ such that for $f \in H_1$ the property (2.60) is valid.

Pseudo adjoint operators $\mathcal{L}^{\star p} : S_1 \to S_2$ that act between pseudo inner product spaces S_1, S_2 can also be considered. By means of a pseudo inner product they are introduced by the relation

$$\langle \mathcal{L}f, g \rangle_p = \langle f, \mathcal{L}^{\star p} g \rangle_p \tag{2.61}$$

(g) Compact operators

A bounded linear operator $\mathcal{L} : H_1 \to H_2$ is *compact* if it maps any bounded set of H_1 into a compact set of H_2. From the definition of a compact set, as given in section 2.1.1 (a), it follows that for a compact operator for each bounded sequence $\{f_n\} \subset H_1$ there is a subsequence $\{f_{n_i}\} \subset H_1$ such that $\mathcal{L}\{f_{n_i}\}$ converges in H_2. Compact operators that act on infinite dimensional spaces are comparatively well understood and have advantageous properties if compared to other operators that act on infinite dimensional spaces and are not compact.

(h) Invertible operators, resolvent operator

We often want to solve operator equations of the form $\mathcal{L}f = g$. If $\mathcal{L}$ possesses a continuous inverse $\mathcal{L}^{-1}$ we find the unique solution $f = \mathcal{L}^{-1}g$. This formal solution procedure naturally leads to the notion of invertible operators: An operator $\mathcal{L} : H_1 \to H_2$ is *invertible* if there exists an operator $\mathcal{L}^{-1} : H_2 \to H_1$ such that

$$\mathcal{L}^{-1}\mathcal{L}f = f \tag{2.62}$$

for all $f \in H_1$ and

$$\mathcal{L}\mathcal{L}^{-1}g = g \tag{2.63}$$

for all $g \in H_2$. The operator $\mathcal{L}^{-1}$ is the *inverse* of $\mathcal{L}$. It is easy to see that if $\mathcal{L} : H_1 \to H_2$ is an invertible linear operator then $\mathcal{L}^{-1}$ is a linear operator.

A class of operators that often occurs in the formulation of boundary value problems in terms of integral equations has the form $\mathcal{L} - \lambda I$, where $\mathcal{L}$ is a compact operator, I is the identity operator, and $\lambda \in \mathbb{C}$, $\lambda \neq 0$. Operators of this form are invertible under certain conditions. In particular, we have the result [208, p. 401] that an operator $\mathcal{L} - \lambda I$, with $\mathcal{L} : H \to H$ bounded and $|\lambda| > \|\mathcal{L}\|$, is invertible with a bounded inverse

$$(\mathcal{L} - \lambda I)^{-1} = -\sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n \tag{2.64}$$

and, furthermore,

$$\|(\mathcal{L} - \lambda I)^{-1}\| \leq (|\lambda| - \|\mathcal{L}\|)^{-1} . \tag{2.65}$$

The series expansion (2.64) is often called *Neumann series* and the operator $(\mathcal{L} - \lambda I)^{-1}$ is known as the *resolvent operator.*

(i) Self-adjoint, normal, and unitary operators

A linear operator $\mathcal{L} : H \to H$ is called *self-adjoint* or *Hermitian* if

$$\mathcal{L} = \mathcal{L}^{\star} \tag{2.66}$$

A linear operator $\mathcal{L} : H \to H$ is called *normal* if it is bounded and

$$\mathcal{L}\mathcal{L}^{\star} = \mathcal{L}^{\star}\mathcal{L} . \tag{2.67}$$

Moreover, a linear operator $\mathcal{L} : H \to H$ is *unitary* if and only if the adjoint operator $\mathcal{L}^{\star}$ is equal to the inverse $\mathcal{L}^{-1}$,

$$\mathcal{L}\mathcal{L}^{\star} = \mathcal{L}^{\star}\mathcal{L} = I . \tag{2.68}$$

Unitary operators preserve sizes, distances, and angles since

$$\langle \mathcal{L}f, \mathcal{L}g \rangle = \langle f, \mathcal{L}^{\star}\mathcal{L}g \rangle = \langle f, g \rangle . \tag{2.69}$$

2.1.3 Spectrum of a linear operator

(a) Standard eigenvalue problem, spectrum, and resolvent set

We consider a linear operator $\mathcal{L} : H \to H$ that maps a Hilbert space onto itself. The *standard eigenvalue problem* involves to find nontrivial solutions of the equation

$$\mathcal{L}f = \lambda f \tag{2.70}$$

with $\lambda \in \mathbb{C}$ an *eigenvalue*, $f \in D_{\mathcal{L}}$ and $f \neq 0$ an *eigenfunction* or *eigenvector*. Trivially, the standard eigenvalue problem can also be written as

$$\mathcal{L}_\lambda f := (\mathcal{L} - \lambda I) f = 0 \qquad (2.71)$$

with I the identity operator on H.

For finite dimensional Hilbert spaces H the properties of eigenvalues and eigenvectors are well-known from elementary linear algebra. Hence, for the definition of the spectrum of a linear operator it is instructive to first consider the finite dimensional case and then move on to the infinite dimensional case.

Finite dimensional case: Suppose the resolvent operator $\mathcal{L}_\lambda^{-1}$ exists for a particular $\lambda \in \mathbb{C}$. Then λ cannot be an eigenvalue since

$$\mathcal{L}_\lambda^{-1} \mathcal{L}_\lambda f = 0 \qquad (2.72)$$

implies $f = 0$, i.e., (2.70) or (2.71) would only have trivial solutions. Conversely, if $\mathcal{L}_\lambda^{-1}$ does not exist we have a nontrivial solution for $\mathcal{L}_\lambda f = 0$ and λ is an eigenvalue. The set of all eigenvalues of $\mathcal{L}$ makes up the *spectrum* of $\mathcal{L}$ and, in the finite dimensional case considered here, is denoted by $\sigma_{\text{finite}}(\mathcal{L})$,

$$\sigma_{\text{finite}}(\mathcal{L}) := \left\{ \lambda \in \mathbb{C} : \mathcal{L}_\lambda^{-1} = (\mathcal{L} - \lambda I)^{-1} \text{ does not exist} \right\}. \qquad (2.73)$$

For all other values of λ the operator $\mathcal{L}_\lambda$ is invertible and (2.70) or (2.71) do only admit trivial solutions. This complement of the spectrum is called the *resolvent set* and denoted by $\rho(\mathcal{L})$,

$$\rho(\mathcal{L}) := \mathbb{C} - \sigma_{\text{finite}}(\mathcal{L}). \qquad (2.74)$$

It follows that for finite dimensional H the resolvent set is the set of all $\lambda \in \mathbb{C}$ that make $\mathcal{L}_\lambda^{-1}$ exist. In this case $\mathcal{L}_\lambda^{-1}$ is bounded since all linear operators that act between finite dimensional Hilbert spaces are bounded. Additionally, the range of $\mathcal{L}_\lambda$ constitutes the complete space H, i.e. the dimension of $R_{\mathcal{L}_\lambda}$ is equal to the dimension of H since the kernel of the linear mapping $\mathcal{L}_\lambda$ only contains the zero element $\{0\}$. In mathematical terms the latter statement is written as $\overline{R_{\mathcal{L}_\lambda}} = H$ where in this context the bar indicates the closure of a set[3]. As a result, the inverse mapping $\mathcal{L}_\lambda^{-1}$ is defined on the complete space H and not just on a subset of H.

Infinite dimensional case: We have just observed that in the finite dimensional case the resolvent set is the set of $\lambda \in \mathbb{C}$ for which $\mathcal{L}_\lambda^{-1}$ exists and it followed that $\mathcal{L}_\lambda^{-1}$ is bounded and is defined on the whole space H, i.e., $\overline{R_{\mathcal{L}_\lambda}} = H$. In the infinite dimensional case it will no longer be true that the existence of $\mathcal{L}_\lambda^{-1}$ implies boundedness and $\overline{R_{\mathcal{L}_\lambda}} = H$. If the existence of $\mathcal{L}_\lambda^{-1}$ implies boundedness and

---

[3]The closure of a set has been defined in Sec. 2.1.1 (a).

$\overline{R_{\mathcal{L}_\lambda}} = H$ we call λ a *regular value* of $\mathcal{L}$. Then the resolvent set is defined as the set of all regular values of $\mathcal{L}_\lambda$,

$$\rho(\mathcal{L}) := \left\{ \lambda \in \mathbb{C} : \lambda \text{ a regular value of } \mathcal{L}_\lambda \right\}. \tag{2.75}$$

The set of all $\lambda \in \mathbb{C}$ which are not regular values of $\mathcal{L}$ is defined as the spectrum $\sigma(\mathcal{L})$ of $\mathcal{L}$,

$$\sigma(\mathcal{L}) := \mathbb{C} - \rho(\mathcal{L}). \tag{2.76}$$

These definitions are generalizations of the definitions of a resolvent set and a spectrum in the finite dimensional case.

(b) Classification of spectra by operator properties

In the following we will outline a specific classification of operators and their corresponding spectra which has been introduced and discussed by a number of authors, see, for example, [50, p. 125], [51, p. 194], [150, p. 412] and [81, p. 223]. Other types of classification do exist [208, p. 347]. An advantage of the classification which is described below is the division of the corresponding spectra into disjoint sets. It is based on the consideration of $\mathcal{L}_\lambda$, that is, we first fix both $\mathcal{L}$ and λ, put $\mathcal{L}_\lambda$ into one of the following categories, and then identify λ as part of a spectrum or part of the resolvent set.

1. *The operator $\mathcal{L}_\lambda$ is not invertible.*
   Then the standard eigenvalue problem (2.70) admits a nontrivial solution f. For a fixed $\mathcal{L}$ the set of all eigenvalues λ which makes $\mathcal{L}_\lambda$ not invertible form the *point spectrum* of $\mathcal{L}$ which we denote by $\sigma_p(\mathcal{L})$. That is, the point spectrum is exactly the set of all eigenvalues. For a given eigenvalue λ the corresponding nontrivial solution f is an eigenvector corresponding to that eigenvalue. Eigenvalues λ of an operator $\mathcal{L}$ are also often called *poles* of the resolvent operator $\mathcal{L}_\lambda^{-1}$.

2. *The operator $\mathcal{L}_\lambda$ is invertible.*
   Then the standard eigenvalue problem (2.70) admits only the trivial solution $f = 0$ and we distinguish the following subcases:

   (a) *The range of $\mathcal{L}_\lambda$ is dense in H, $\overline{R_{\mathcal{L}_\lambda}} = H$.*
   In this subcase we further distinguish between the possibilities that the resolvent $\mathcal{L}_\lambda^{-1}$ is bounded or unbounded.

       i. *The resolvent $\mathcal{L}_\lambda^{-1}$ is bounded.*
       Then the value λ is called a *regular value* of $\mathcal{L}$. The set of all regular values forms the resolvent set $\rho(\mathcal{L})$.

       ii. *The resolvent $\mathcal{L}_\lambda^{-1}$ is unbounded.*
       Then the value λ is part of the so-called *continuous spectrum* which we denote by $\sigma_c(\mathcal{L})$.

(b) *The range of $\mathcal{L}_\lambda$ is not dense in H, $\overline{R_{\mathcal{L}_\lambda}} \neq H$*

Then the range of $\mathcal{L}_\lambda$ is a proper subset of H, $R_{\mathcal{L}_\lambda} \subset H$, $R_{\mathcal{L}_\lambda} \neq H$, and the corresponding set of values of λ forms the *residual spectrum* of $\mathcal{L}$ which we denote by $\sigma_r(\mathcal{L})$.
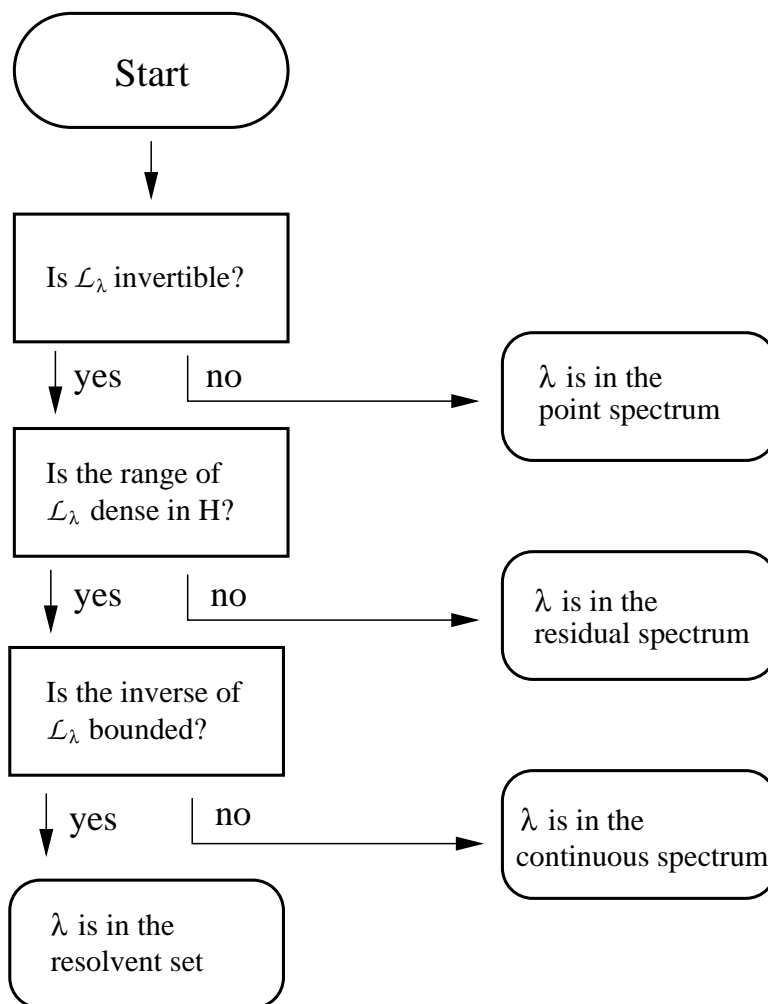
Figure 2.9: Illustration of the definition of the point spectrum, residual spectrum, continuous spectrum, and the resolvent set, as adapted from [150].

The diagram of Fig. 2.9 illustrates these definitions. The *total spectrum* is the union of the disjoint sets that form the point spectrum, continuous spectrum, and residual spectrum,

$$\sigma(\mathcal{L}) = \sigma_p(\mathcal{L}) \cup \sigma_c(\mathcal{L}) \cup \sigma_r(\mathcal{L}), \tag{2.77}$$
$$= \mathbb{C} - \rho(\mathcal{L}). \tag{2.78}$$

This mathematical classification is important because it defines and categorizes the spectrum of a linear operator. However, it does not tell us how to actually calculate the spectrum.

For illustrative examples of the point, continuous, and residual spectrum we refer to [150, § 6.6]. In view of applications to electromagnetics we are interested in the calculation of the spectrum of linear operators that occur in the formulation of electromagnetic boundary value problems. An important class of such operators is given by second order differential operators that are known as *Sturm-Liouville operators*. The properties of these operators have been investigated in the context of electromagnetic applications and this includes the calculation of their spectra for specified boundary value problems. It then turns out that the residual spectrum usually does not occur, but the point and continuous spectrum are of great importance in electromagnetic field analysis. Relevant examples and calculations are not detailed here but can be found in the literature, see [40] and, in particular, [81, § 5]. Our main motivation to introduce the spectrum of a linear operator is to arrive at the method of eigenfunction expansion.

2.1.4 Spectral expansions and representations

(a) Linear independence of eigenfunctions

We have in mind to construct expansions of elements of a Hilbert space in terms of eigenfunctions of a specific operator. In this context the following statements are important:

- For a linear operator $\mathcal{L} : H \to H$ the eigenvectors $f_1, f_2, \ldots, f_n$ that correspond to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ form a linearly independent set in H. A simple proof of this statement is given in [81, p. 228].

- If the operator $\mathcal{L} : H \to H$ is not only linear but additionally self-adjoint and has eigenvalues then these eigenvalues are real and the eigenvectors corresponding to distinct eigenvalues are orthogonal. This is proven in [37, p. 182], for example.

- Moreover, if the linear operator $\mathcal{L} : H \to H$, with H infinite-dimensional, is both compact and self-adjoint the *Hilbert-Schmidt theorem* states that there exists an orthonormal system of eigenvectors $\{e_n\}$ corresponding to nonzero eigenvalues $\{\lambda_n\}$ such that any $f \in H$ can uniquely be represented in the form

$$f = f_0 + \sum_{n=1}^{\infty} \langle f, e_n \rangle e_n \,, \tag{2.79}$$

  where the element f_0 satisfies $\mathcal{L} f_0 = 0$.
  A proof of the Hilbert-Schmidt theorem can be found in [37, p. 188]. It should be noted that the eigenvectors e_n not necessarily form a basis of H. In this case the element f_0 is the projection of f on the space which is orthogonal to the closed linear subspace of H which is spanned by the eigenvectors $\{e_n\}$.

The Hilbert-Schmidt theorem can be generalized if the eigenvectors corresponding to zero eigenvalues are included. This leads to the spectral theorem for compact and self-adjoint operators.

(b) Spectral theorem for compact and self-adjoint operators:

- If $\mathcal{L} : H \to H$ is a compact and self-adjoint linear operator acting on an infinite-dimensional Hilbert space H then there exists an orthonormal basis H of eigenvectors $\{f_n\}$ with corresponding eigenvalues $\{\lambda_n\}$. It follows for every $f \in H$

$$f = \sum_{n=1}^{\infty} \langle f, f_n \rangle f_n \tag{2.80}$$

and

$$\mathcal{L}f = \sum_{n=1}^{\infty} \lambda_n \langle f, f_n \rangle f_n. \tag{2.81}$$

This is the spectral theorem for compact and self-adjoint operators which is a standard piece of linear operator theory. A proof is given in [37, p. 190] and [60, p. 243], for example. The *eigenfunction expansions* (2.80), (2.81) are particularly useful because they allow to reduce the action of a linear operator to an algebraic mapping. The corresponding matrix with respect to the eigenvectors $\{f_n\}$ is given by (2.81). It is diagonal and, in the usual cases, infinite dimensional.

(c) Remarks on the relation between differential and integral operators

The differential operators that are involved in electromagnetic boundary value problems can be self-adjoint, compare Sec. 2.2.2 below, but they are usually not compact. This is unfortunate since in these situations we cannot apply the spectral theorem for compact and self-adjoint operators. Consequently, we do not know if the eigenfunctions of a differential operator, if they exist, form a basis within the Hilbert space considered.

At this point we find a loophole if we consider the following statement:

- If $\mathcal{L} : H \to H$ is an invertible linear operator with associated eigenvalues λ_n and eigenfunctions f_n then $\mathcal{L}^{-1} : H \to H$ has eigenvalues $1/\lambda_n$ corresponding to the eigenfunctions f_n.

A proof of this statement can be found in [81, p. 233]. Therefore, we can show that the eigenfunctions of a differential operator form a basis if we are able to construct its inverse operator and demonstrate that this inverse operator is compact and self-adjoint.

It might be clear from intuition that the inverse operators of many differential operators are given by integral operators.

- If the inverse integral operator of a self-adjoint differential operator can be constructed it has a symmetric kernel and it follows as a consequence that it is self-adjoint and compact.

This fundamental result is obtained in [32, Chap. V, § 14]. The mathematical details of this circumstance, like the precise assumptions that are needed to derive this results, are elaborated in the relevant mathematical literature, see [2, 141, 60, 50], among others.

We conclude that the standard way to justify the eigenfunction expansion method, if applied to a linear differential operator, is to construct the corresponding inverse integral operator and to show that it is self-adjoint and compact. We also note that this general strategy to convert a differential boundary value problem to an integral boundary value problem is contained in the Green's function method which will be explained in the following Sec. 2.2.

In electromagnetics the negative Laplace operator $\mathcal{L}_D = -\Delta$ and the double-curl operator $\mathcal{L}_D = \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times$ are of particular interest, as is evident from Sec. 1.4.3. Solutions of the homogeneous Helmholtz equations

$$\Delta \boldsymbol{F} + k^2 \boldsymbol{F} = 0 \,, \tag{2.82}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{F} - k^2 \boldsymbol{F} = 0 \,, \tag{2.83}$$

are determined from solutions of the spectral problems

$$-\Delta \boldsymbol{F} = k^2 \boldsymbol{F} \,, \tag{2.84}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{F} = k^2 \boldsymbol{F} \,, \tag{2.85}$$

respectively. Therefore, it is important to study the spectral properties of these operators. There are no general or simple results that characterize the spectra of the negative Laplace operator or the double-curl operator because the definition of an operator includes the definition of its domain, which, in turn, is determined from the specific boundary value problem. Self-adjointness of the negative Laplace operator and the double-curl operator will depend on their domains and we first have to specify these in order to be able to arrive at explicit results.

(d) A comment on Sobolev spaces

It seems to be inconvenient that the spectral theorem for compact and self-adjoint operators cannot directly be applied to differential operators which, usually, are not compact. This reflects the fundamental difficulty that differential operators often produce singular effects if acting on functions that are not differentiable in their whole domain. As a rule, integral operators are more well-behaved, they tend to "smooth out" singular behavior and therefore it is plausible that we require integral operators in order to show that the eigenfunctions of a differential operator are complete and form a basis.

However, it is possible to avoid the reduction from differential to integral equations if we pass to *distributional Hilbert spaces* which contain generalized functions that are defined in a distributional sense. Such spaces are called *Sobolev spaces*. The Sobolev-space approach can be applied to the direct solution of partial differential equations [141, 60]. We will not pursue this approach since we will focus in the following on the conventional Green's function approach which is based on integral equations.

2.2   The Green's function method

The construction of solutions to linear differential equations with specified sources and given boundary conditions belongs to the fundamental problems of the Maxwell theory and other physical field theories. The Green's function method provides a technique to find these solutions.

Again, we formally express a linear differential equation as an operator equation of the form $\mathcal{L}_D f = g$ with a linear differential operator $\mathcal{L}_D$, a source function g which is assumed to be known, and an unknown function f. The Green's function method consists in finding a Green's function G such that the unknown function f is expressed as an integral over the source function g, weighted with the Green's function. From a physical point of view the Green's function method is a representation of the superposition principle: The Green's function is the solution of the given linear differential equation with respect to a unit source which is placed at a specific position. Then the solution with respect to a general source is obtained from the superposition of known solutions of individual unit sources at various positions.

2.2.1   Basic ideas

To introduce the Green's function method we consider a real, self-adjoint differential operator $\mathcal{L}_D$ with

$$(\mathcal{L}_D f)(\boldsymbol{r}) = g(\boldsymbol{r}). \tag{2.86}$$

Here, we explicitly indicate that the functions depend on a variable $\boldsymbol{r}$ which usually represents a position in space. It is also common to have the time t as an additional parameter.

A *Green's function* is implicitly defined by

$$\mathcal{L}_D G(\boldsymbol{r}, \boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') \tag{2.87}$$

with

$$\delta_{\boldsymbol{r}} := \delta(\boldsymbol{r} - \boldsymbol{r}') \tag{2.88}$$

the *Dirac delta function* which is a generalized function that is defined in the distributional sense [55, 198]. In a Hilbert space H with an inner product $\langle \ , \ \rangle$ it can be

introduced via the relationship

$$\langle f, \delta_{\boldsymbol{r}} \rangle = f(\boldsymbol{r}) \,. \tag{2.89}$$

Let us suppose that we have constructed a Green's function that fulfills (2.87). Then we may consider the expression

$$\langle \mathcal{L}_D f, G^* \rangle = \langle f, \mathcal{L}_D G^* \rangle \tag{2.90}$$
$$= \langle f, (\mathcal{L}_D G)^* \rangle \,. \tag{2.91}$$

Here we used that $\mathcal{L}_D$ is self-adjoint and real. We apply (2.87) and notice that the delta function is a real function. This yields

$$\langle \mathcal{L}_D f, G^* \rangle = \langle f, \delta \rangle \tag{2.92}$$
$$= f \,. \tag{2.93}$$

With (2.86) we find the solution of the original problem as

$$f = \langle g, G^* \rangle \,. \tag{2.94}$$

This establishes the Green's function method for solving differential equations that are represented by a linear, self-adjoint, and real differential operator $\mathcal{L}_D$.

Example: We consider as an example the Hilbert space $\boldsymbol{L}^2(\Omega)^m$ with inner product (2.31) and $\boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{L}^2(\Omega)^m$. In analogy to (2.86) and (2.87) we assume a linear differential equation

$$(\mathcal{L}_D \boldsymbol{f})(\boldsymbol{r}) = \boldsymbol{g}(\boldsymbol{r}) \,. \tag{2.95}$$

For the corresponding *Green's function* we make the ansatz

$$\mathcal{L}_D \overline{\boldsymbol{G}}(\boldsymbol{r}, \boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') \overline{\boldsymbol{I}} \tag{2.96}$$

where now the Green's function $\overline{\boldsymbol{G}}(\boldsymbol{r}, \boldsymbol{r}')$ is represented as a dyadic and $\overline{\boldsymbol{I}}$ denotes the unit dyad [213]. With the inner product (2.31) the delta function acts according to

$$\langle \boldsymbol{f}, \delta_{\boldsymbol{r}} \rangle = \int_\Omega \delta(\boldsymbol{r} - \boldsymbol{r}') \boldsymbol{f}(\boldsymbol{r}') \, d\Omega' \tag{2.97}$$
$$= \boldsymbol{f}(\boldsymbol{r}) \,. \tag{2.98}$$

We repeat the steps that led from (2.90) to (2.94) and obtain the solution of (2.95) in the form

$$\boldsymbol{f}(\boldsymbol{r}) = \int_\Omega \boldsymbol{g}(\boldsymbol{r}') \cdot \overline{\boldsymbol{G}}(\boldsymbol{r}, \boldsymbol{r}') \, d\Omega' \,. \tag{2.99}$$

2.2.2   Self-adjointness of differential operators and boundary conditions

The general solutions (2.94), (2.99) that we obtained by means of the Green's function method look deceptively simple because they only involve to weight the source function g with the Green's function G. However, from the theory of differential equations we know that boundary conditions play a fundamental role for the determination of a unique solution. Therefore, the information on boundary conditions must have been incorporated in the derivation of (2.94), (2.99). Indeed, we presupposed self-adjointness of the real differential operator $\mathcal{L}_D$,

$$\langle \mathcal{L}_D f, g \rangle = \langle f, \mathcal{L}_D g \rangle . \tag{2.100}$$

Since in most function spaces of physical interest the inner product is represented by means of integration it follows that self-adjointness is closely connected to "generalized partial integration", i.e., to the generalized Green's identity

$$\int_\Omega (\mathcal{L}_D f) g^* \, d\Omega = \int_\Omega f (\mathcal{L}_D g)^* \, d\Omega + \int_{\Gamma=\partial\Omega} J(f,g) \, d\Gamma \tag{2.101}$$

where Ω is the integration volume and $\partial\Omega = \Gamma$ its boundary. We can write this identity in terms of the inner product as

$$\langle \mathcal{L}_D f, g \rangle = \langle f, \mathcal{L}_D g \rangle + \int_{\Gamma=\partial\Omega} J(f,g) \, d\Gamma . \tag{2.102}$$

It follows that $\mathcal{L}_D$ is self-adjoint if and only if the integral on the right side of this equation vanishes. This requirement will pose restrictions on the a priori unknown function f and the Green's function G.

We demonstrate this circumstance by the following general application of the Green's function method: Let us consider again a differential equation of the form

$$\mathcal{L}_D f = g \tag{2.103}$$

and a corresponding Green's function G which satisfies

$$\mathcal{L}_D G = \delta . \tag{2.104}$$

At this point we do not require $\mathcal{L}_D$ to be self-adjoint. We take the inner product of (2.103) with G^* and the inner product of the complex conjugate of (2.104) with f in order to obtain

$$\langle \mathcal{L}_D f, G^* \rangle = \langle g, G^* \rangle , \tag{2.105}$$

$$\langle f, (\mathcal{L}_D G)^* \rangle = \langle f, \delta^* \rangle = f . \tag{2.106}$$

We form the differences of both equations and find

$$\langle \mathcal{L}_D f, G^* \rangle - \langle f, (\mathcal{L}_D G)^* \rangle = \langle g, G^* \rangle - f \tag{2.107}$$

In the notation of (2.102) we are thus led to

$$\langle g, G^* \rangle - f = \int_{\Gamma = \partial\Omega} J(f, G) \, d\Gamma \,, \tag{2.108}$$

i.e., we will obtain the simple solution (2.94) if and only if the boundary integral, which involves the so-called *conjunct* $J(f, G)$, vanishes. This explicitly shows the relation between self-adjointness of the differential operator $\mathcal{L}_D$ and boundary conditions of f and G.

Example 1: We consider the Helmholtz equation (1.169) for the magnetic vector potential $\boldsymbol{A}$ in the Lorentz gauge,

$$\Delta \boldsymbol{A}(\boldsymbol{r}) + k^2 \boldsymbol{A}(\boldsymbol{r}) = -\mu \boldsymbol{J}(\boldsymbol{r}) \,. \quad \text{(Lorenz gauge)} \tag{2.109}$$

Up to a factor μ the dyadic Green's function of this equation has to fulfill

$$\Delta \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') + k^2 \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') = -\overline{\boldsymbol{I}} \delta(\boldsymbol{r} - \boldsymbol{r}') \,. \tag{2.110}$$

With the inner product (2.31) of $\boldsymbol{L}^2(\Omega)^3$ the general equation (2.107) yields

$$\int_\Omega \left[(\Delta \boldsymbol{A}(\boldsymbol{r}')) \cdot \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') - \boldsymbol{A}(\boldsymbol{r}') \cdot \Delta \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') \right] d^3 r' = \tag{2.111}$$

$$-\mu \int_\Omega \boldsymbol{J}(\boldsymbol{r}') \cdot \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') \, d^3 r' + \boldsymbol{A}(\boldsymbol{r}) \,.$$

By means of the second vector-dyadic Green's second theorem (B.21) the integral on the left side can be transformed to a boundary integral. We obtain

$$\boldsymbol{A}(\boldsymbol{r}) = \mu \int_\Omega \boldsymbol{J}(\boldsymbol{r}') \cdot \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') \, d^3 r' +$$

$$\oint_\Gamma \left[(\boldsymbol{e}_n \times \boldsymbol{A}(\boldsymbol{r}')) \cdot (\boldsymbol{\nabla}' \times \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')) - (\boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r}')) \cdot (\boldsymbol{e}_n \times \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')) + \right.$$

$$\left. \boldsymbol{e}_n \cdot \boldsymbol{A}(\boldsymbol{r}')(\boldsymbol{\nabla}' \cdot \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')) - \boldsymbol{e}_n \cdot \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')(\boldsymbol{\nabla}' \cdot \boldsymbol{A}(\boldsymbol{r}')) \right] d^2 r' \,. \tag{2.112}$$

Since we want to make the boundary integral vanish we have to think about appropriate boundary conditions. If we suppose that the boundary is perfectly conducting it follows

$$\boldsymbol{e}_n \times \boldsymbol{A}(\boldsymbol{r})|_{\boldsymbol{r} \in \Gamma} = \boldsymbol{0} \,, \tag{2.113}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r})|_{\boldsymbol{r} \in \Gamma} = 0 \,, \tag{2.114}$$

such that the first and last term within the surface integral will vanish. We also have the corresponding boundary conditions

$$\boldsymbol{e}_n \times \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')\Big|_{\boldsymbol{r} \in \Gamma} = \boldsymbol{0} \,, \tag{2.115}$$

$$\boldsymbol{\nabla} \cdot \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')\Big|_{\boldsymbol{r} \in \Gamma} = 0 \tag{2.116}$$

of the dyadic Green's function such that the surface integral completely vanishes. Therefore, the magnetic vector potential can be calculated from the expression

$$\boldsymbol{A}(\boldsymbol{r}) = \mu \int \boldsymbol{J}(\boldsymbol{r}') \cdot \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') \, d^3 r' \,. \tag{2.117}$$

which involves no boundary terms.

    Example 2: We consider the vector Helmholtz equation (1.165) for the electric field $\boldsymbol{E}$,

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}) - k^2 \boldsymbol{E}(\boldsymbol{r}) = -j\omega\mu \boldsymbol{J}(\boldsymbol{r}) \,. \tag{2.118}$$

Up to a factor $-j\omega\mu$ the corresponding Green's function needs to fulfill

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') - k^2 \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') = \overline{\boldsymbol{I}} \delta(\boldsymbol{r}, \boldsymbol{r}') \,. \tag{2.119}$$

The general equation (2.107) yields

$$\int_\Omega \left[(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}')) \cdot \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') - \boldsymbol{E}(\boldsymbol{r}') \cdot (\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')) \right] d^3 r' =$$
$$-j\omega\mu \int_\Omega \boldsymbol{J}(\boldsymbol{r}') \cdot \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') \, d^3 r' - \boldsymbol{E}(\boldsymbol{r}) \,. \tag{2.120}$$

From application of the Green's theorem (B.20) and the identity (B.15) we find

$$\boldsymbol{E}(\boldsymbol{r}) = -j\omega\mu \int_\Omega \boldsymbol{J}(\boldsymbol{r}') \cdot \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') \, d^3 r' + \tag{2.121}$$
$$\oint_\Gamma \left[(\boldsymbol{e}_n \times \boldsymbol{E}(\boldsymbol{r}')) \cdot (\boldsymbol{\nabla} \times \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')) - (\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}')) \cdot (\boldsymbol{e}_n \times \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')) \right] d^2 r' \,.$$

If we suppose again that the interior of the boundary Γ is perfectly conducting we have the boundary condition

$$\boldsymbol{e}_n \times \boldsymbol{E}(\boldsymbol{r})|_{\boldsymbol{r} \in \Gamma} = \boldsymbol{0} \,, \tag{2.122}$$

The corresponding boundary condition of the Green's function is

$$\boldsymbol{e}_n \times \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')\Big|_{\boldsymbol{r} \in \Gamma} = \boldsymbol{0} \tag{2.123}$$

and, as a consequence, the surface integral vanishes such that the electric field can be calculated according to

$$\boldsymbol{E}(\boldsymbol{r}) = -j\omega\mu \int_\Omega \boldsymbol{J}(\boldsymbol{r}') \cdot \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') \, d^3 r' \,. \tag{2.124}$$

Example 1 and Example 2 show that the Helmholtz equations for the magnetic vector potential and the electric field strength, respectively, form self-adjoint boundary value problems if the fields are defined in a finite domain which is enclosed by a perfectly conducting boundary.

2.2.3 Spectral representation of the Green's function

We have seen by means of the Green's function method that it is straightforward to construct a formal solution of a linear differential equation. However, so far nothing has been said about the explicit construction of a Green's function. In this respect the spectral properties of the relevant differential operator are of great help.

To illustrate this circumstance we consider again an operator $\mathcal{L}_D : H \to H$ which determines a differential equation of the form

$$(\mathcal{L}_D f)(\boldsymbol{r}) = g(\boldsymbol{r}). \tag{2.125}$$

Let $\mathcal{L}_D$ be self-adjoint and let us further assume that, along the discussion of Sec. 2.1.4 (c), the eigenvectors $\{f_n\}$ of $\mathcal{L}_D$ form an orthonormal basis of the Hilbert space H,

$$(\mathcal{L}_D f_n)(\boldsymbol{r}) = \lambda_n f_n(\boldsymbol{r}). \tag{2.126}$$

Since $\mathcal{L}_D$ is self-adjoint the eigenvalues λ_n are real. It is then possible to expand the unknown function f and the known source g of (2.125) according to

$$f(\boldsymbol{r}) = \sum_{n=1}^{\infty} \alpha_n f_n(\boldsymbol{r}), \tag{2.127}$$

$$g(\boldsymbol{r}) = \sum_{n=1}^{\infty} \langle g, f_n \rangle f_n(\boldsymbol{r}), \tag{2.128}$$

with a priori undetermined expansion coefficients α_n. These expansions can be inserted into (2.125) to yield

$$\sum_{n=1}^{\infty} \alpha_n \lambda_n f_n(\boldsymbol{r}) = \sum_{n=1}^{\infty} \langle g, f_n \rangle f_n(\boldsymbol{r}). \tag{2.129}$$

It follows for the coefficients α_n

$$\alpha_n = \frac{\langle g, f_n \rangle}{\lambda_n} \tag{2.130}$$

where we assumed $\lambda_n \neq 0$. If $\lambda_n = 0$ the corresponding eigenfunction is in the null space of the operator $\mathcal{L}_D$ and an arbitrary multiple of it can be added to the solution of (2.125) to yield another solution. Therefore, our assumption $\lambda_n \neq 0$ is based on the assumption that (2.125) has a unique solution. This is physically meaningful since, according to the uniqueness theorem [82], well-posed electromagnetic boundary value problems have a unique solution.

With (2.130) we have obtained a formal solution for the function f,

$$f(\boldsymbol{r}) = \sum_{n=1}^{\infty} \frac{\langle g, f_n \rangle}{\lambda_n} f_n(\boldsymbol{r}) \tag{2.131}$$

$$= \left\langle g, \sum_{n=1}^{\infty} \frac{f_n^*(\boldsymbol{r}) f_n}{\lambda_n} \right\rangle . \tag{2.132}$$

In the last step we have used the fact that the eigenvalues λ_n are real. If we compare this solution to the general solution (2.94) which is written in terms of a Green's function we find the identification

$$G(\boldsymbol{r}, \boldsymbol{r}') = \sum_{n=1}^{\infty} \frac{f_n(\boldsymbol{r}) f_n^*(\boldsymbol{r}')}{\lambda_n} \tag{2.133}$$

This is a very useful and important result: If we know the eigenfunctions and eigenvalues of a self-adjoint differential operator with eigenfunctions that form a basis we automatically know its corresponding Green's function. A Green's function of the form (2.133) is referred to as *spectral representation* since it is based on the spectral properties of the corresponding differential operator. It is also referred to as *mode representation* if, in physical applications, the eigenfunctions $f_n(\boldsymbol{r})$ are identified as eigenmodes or eigenoscillations of a physical system.

2.2.4   General solutions of Maxwell equations

We recall that we have in mind to solve electromagnetic boundary value problems. In Sections 1.4.1 and 1.4.3 we have seen that for homogeneous and isotropic media it is immediate to decouple the Maxwell equations and rewrite them in the form of wave equations or, in the time harmonic case, as Helmholtz equations. Then linear operator theory suggests to study the spectral properties of the relevant wave operator or Helmholtz operator. For a given boundary value problem we need to explictly determine the eigenvalues and eigenfunctions of the generalized eigenvalue problem that has been defined in Sec. 2.1.3. Once we know the eigenvalues and eigenfunctions we may construct the Green's function that yields the formal solution of the boundary value problem. In summary, we arrive at a fairly general method to obtain analytical expressions for the solution of Maxwell equations.

However, depending on the type of boundary conditions, the construction of the Green's function can be arbitrarily difficult. And even if we can find a Green's function it still can be a nontrivial numerical task to explicitly compute the electromagnetic quantities that we are interested in. This will become evident in Chapter 3 where antenna boundary value problems will explicitly be solved.

In the absence of boundaries, i.e., in free space, and within a homogeneous medium the general solution of the Maxwell equations is given in terms of the solution of the

scalar Helmholtz equation. This follows from (1.169) and (1.170) which constitute in free space four independent scalar Helmholtz equations. These have the general structure

$$(\Delta + k^2) f(\boldsymbol{r}, \omega) = -g(\boldsymbol{r}, \omega) \tag{2.134}$$

and the appropriate Green's function $G_0(\boldsymbol{r}, \boldsymbol{r}')$ needs to satisfy

$$(\Delta + k^2) G_0(\boldsymbol{r}, \boldsymbol{r}') = -\delta(\boldsymbol{r} - \boldsymbol{r}'). \tag{2.135}$$

The solution for $G_0(\boldsymbol{r}, \boldsymbol{r}')$ is most easily obtained in spherical coordinates, taking advantage of the symmetries of free space. This yields the retarded solution [93, p. 243]

$$G_0(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{4\pi} \frac{e^{-jk|\boldsymbol{r} - \boldsymbol{r}'|}}{|\boldsymbol{r} - \boldsymbol{r}'|}. \tag{2.136}$$

Therefore, in free space the general solution of the Maxwell equations is represented by the equations

$$\begin{aligned} \phi(\boldsymbol{r}) &= \int G_0(\boldsymbol{r}, \boldsymbol{r}') \rho(\boldsymbol{r}') \, d^3 r' \\ &= \frac{1}{4\pi\varepsilon} \int \frac{e^{-jk|\boldsymbol{r} - \boldsymbol{r}'|}}{|\boldsymbol{r} - \boldsymbol{r}'|} \rho(\boldsymbol{r}') \, d^3 r', \end{aligned} \tag{2.137}$$

$$\begin{aligned} \boldsymbol{A}(\boldsymbol{r}) &= \int G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}(\boldsymbol{r}') \, d^3 r' \\ &= \frac{\mu}{4\pi} \int \frac{e^{-jk|\boldsymbol{r} - \boldsymbol{r}'|}}{|\boldsymbol{r} - \boldsymbol{r}'|} \boldsymbol{J}(\boldsymbol{r}') \, d^3 r', \end{aligned} \tag{2.138}$$

which relate the electromagnetic sources ρ, $\boldsymbol{J}$ to the electromagnetic field, expressed by ϕ and $\boldsymbol{A}$ in the Lorenz gauge.

2.3 Green's functions of electromagnetic cavities

To solve the Maxwell equations within a electromagnetic cavity we recall again from Sec. 1.4.3 that in a linear, isotropic, and homogeneous medium the Maxwell equations in frequency domain can be reduced to Helmholtz equations. For the vector potential $\boldsymbol{A}(\boldsymbol{r})$ in the Lorenz gauge, the electric field $\boldsymbol{E}(\boldsymbol{r})$, and the magnetic field $\boldsymbol{B}(\boldsymbol{r})$, respectively, we found the (vector) Helmholtz equations

$$\Delta \boldsymbol{A}(\boldsymbol{r}, \omega) + k^2 \boldsymbol{A}(\boldsymbol{r}, \omega) = -\mu \boldsymbol{J}(\boldsymbol{r}, \omega), \tag{2.139}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}, \omega) - k^2 \boldsymbol{E}(\boldsymbol{r}, \omega) = -j\omega\mu \boldsymbol{J}(\boldsymbol{r}, \omega), \tag{2.140}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{B}(\boldsymbol{r}, \omega) - k^2 \boldsymbol{B}(\boldsymbol{r}, \omega) = \mu \boldsymbol{\nabla} \times \boldsymbol{J}(\boldsymbol{r}, \omega). \tag{2.141}$$

The corresponding dyadic Green's functions obey the differential equations[4]

$$\Delta \overline{G}^A(r, r') + k^2 \overline{G}^A(r, r') = -\overline{I}\delta(r - r') \,, \tag{2.142}$$

$$\nabla \times \nabla \times \overline{G}^E(r, r') - k^2 \overline{G}^E(r, r') = \overline{I}\delta(r - r') \,, \tag{2.143}$$

$$\nabla \times \nabla \times \overline{G}^B - k^2 \overline{G}^B(r, r') = \nabla\delta(r - r') \times \overline{I} \,. \tag{2.144}$$

Equations (2.142) and (2.143) already are familiar from the examples of Sec. 2.2.2.

Suppose we are able to construct $\overline{G}^A(r, r')$. Then we obtain $\overline{G}^E(r, r')$ and $\overline{G}^B(r, r')$ via

$$\overline{G}^E(r, r') = \left(\overline{I} + \frac{1}{k^2} \nabla\nabla \right) \overline{G}^A(r, r') \,, \tag{2.145}$$

$$\overline{G}^B(r, r') = \nabla \times \overline{G}^A(r, r') \,. \tag{2.146}$$

Clearly, this is an immediate consequence of the relations

$$E(r) = -j\omega \left(1 + \frac{1}{k^2} \nabla\nabla\cdot \right) A(r) \,, \tag{2.147}$$

$$B(r) = \nabla \times A(r) \,. \tag{2.148}$$

If we construct $\overline{G}^E(r, r')$ rather than $\overline{G}^A(r, r')$ we obtain $\overline{G}^B(r, r')$ from

$$\overline{G}^B(r, r') = \nabla \times \overline{G}^E(r, r') \tag{2.149}$$

since

$$B(r) = -\frac{1}{j\omega} \nabla \times E(r) \,. \tag{2.150}$$

In the following we will outline how to construct the Green's functions $\overline{G}^A(r, r')$ and $\overline{G}^E(r, r')$ inside a closed and perfectly conducting cavity by means of the eigenfunction expansion method. These are important standard procedures which are contained in a number of textbooks. We mention in particular Morse & Feshbach [146, § 13], Van Bladel [227, § 10], and Collin [30, § 2]. However, these standard procedures presuppose that the eigenfunctions of the relevant Helmholtz operators form a basis of the Hilbert space $L^2(\Omega)^3$ with inner product (2.31). To assure this property requires to extent the spectral theorem for compact and self-adjoint operators to non-compact self-adjoint differential operators. Such extensions exist in the mathematical literature, see [160, 2, 35, 215, 111], and have partly been reformulated to be easier accessible to scientists and engineers [150, 208, 81]. They critically depend on the domain of the differential operator considered and there is no single theorem which covers all cases of physical interest. Fortunately, the solution of the Helmholtz equations (2.139), (2.140) within a

---

[4]It is a convention to skip in the definitions of the Green's functions the factors μ and $-j\omega\mu$ that appear on the right hand sides of the corresponding Helmholtz equations.

finite three-dimensional volume which is enclosed by a perfectly conducting boundary belongs to a class of problems which is known as three-dimensional, regular Sturm-Liouville problems [40]. These problems are rather well understood and characterized by eigenfunctions which form a basis in the Hilbert space $\boldsymbol{L}^2(\Omega)^3$ [81, p. 262]. This is, in fact, the mathematical foundation of the spectral representations which we will consider now.

2.3.1 Spectral representations of perfectly conducting cavities' Green's functions

(a) General considerations and a canonical example

We consider a closed cavity that surrounds a finite three-dimensional volume Ω by a perfectly conducting cavity wall Γ. We further assume that the interior of the cavity is given by a linear, homogeneous, and isotropic medium. Electromagnetic fields within the cavity have to fulfill the Helmholtz equations (2.139) – (2.141) together with Dirichlet boundary conditions for the magnetic vector potential,

$$\boldsymbol{e}_n \times \boldsymbol{A}(\boldsymbol{r})|_{\boldsymbol{r}\in\Gamma} = \boldsymbol{0}\,, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r})|_{\boldsymbol{r}\in\Gamma} = 0\,, \tag{2.151}$$

or a Dirichlet boundary condition for the electric field,

$$\boldsymbol{e}_n \times \boldsymbol{E}(\boldsymbol{r})|_{\boldsymbol{r}\in\Gamma} = \boldsymbol{0}\,. \tag{2.152}$$

For the corresponding Green's functions this implies

$$\boldsymbol{e}_n \times \overline{\boldsymbol{G}}^A(\boldsymbol{r},\boldsymbol{r}')\Big|_{\boldsymbol{r}\in\Gamma} = \boldsymbol{0}\,, \qquad \boldsymbol{\nabla} \cdot \overline{\boldsymbol{G}}^A(\boldsymbol{r},\boldsymbol{r}')\Big|_{\boldsymbol{r}\in\Gamma} = 0\,, \tag{2.153}$$

or

$$\boldsymbol{e}_n \times \overline{\boldsymbol{G}}^E(\boldsymbol{r},\boldsymbol{r}')\Big|_{\boldsymbol{r}\in\Gamma} = \boldsymbol{0}\,. \tag{2.154}$$

With these boundary conditions it follows from Example 1 and Example 2 of Sec. 2.2.2 that for a real wavenumber k the Helmholtz operators $\Delta + k^2$ and $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times + k^2$ are self-adjoint. As mentioned above, the eigenfunctions of these operators form bases and we have to explicitly calculate these eigenfunctions in order to construct the spectral representation of the corresponding Green's functions along the lines of Sec. 2.2.3. This is a general way to solve the Maxwell equations inside a perfectly conducting cavity.

Example 1: The Helmholtz equation for the magnetic vector potential $\boldsymbol{A}(\boldsymbol{r})$

$$\Delta\boldsymbol{A} + k^2\boldsymbol{A} = -\mu\boldsymbol{J} \tag{2.155}$$

has, according to (2.126), the related eigenvalue problem

$$\Delta\boldsymbol{f}_n + k^2\boldsymbol{f}_n = -\tilde{k}_n^2\boldsymbol{f}_n \tag{2.156}$$

with eigenfunctions $\boldsymbol{f}_n$ and eigenvalues $\tilde{k}_n^2$. We note that, trivially, this can be rewritten as

$$-\Delta \boldsymbol{f}_n = k_n^2 \boldsymbol{f}_n, \qquad k_n^2 := \tilde{k}_n^2 + k^2. \qquad (2.157)$$

It follows that the eigenfunctions of the Helmholtz operator can be determined from the eigenfunctions of the negative Laplace operator. The eigenvalues are different, though.

To proceed we have to specify boundary conditions, that is, the geometry of the cavity. As a canonical example we consider a rectangular cavity of dimensions l_x, l_y and l_z. In Cartesian coordinates its enclosed volume is characterized by $0 \leq x \leq l_x$, $0 \leq y \leq l_y$, and $0 \leq z \leq l_z$. The eigenvalue problem (2.157) reduces to three single equations,

$$-\Delta f_n^x \boldsymbol{e}_x = k_n^2 f_n^x \boldsymbol{e}_x, \qquad (2.158)$$

$$-\Delta f_n^y \boldsymbol{e}_y = k_n^2 f_n^y \boldsymbol{e}_y, \qquad (2.159)$$

$$-\Delta f_n^z \boldsymbol{e}_z = k_n^2 f_n^z \boldsymbol{e}_z, \qquad (2.160)$$

where the unit vectors $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, and $\boldsymbol{e}_z$ could also be dropped. The boundary condition $\boldsymbol{e}_n \times \boldsymbol{A}(\boldsymbol{r})|_{\boldsymbol{r} \in \Gamma} = \boldsymbol{0}$ yields

$$f_n^y(x = 0) = f_n^z(x = 0) = 0, \qquad (2.161)$$

$$f_n^z(y = 0) = f_n^x(y = 0) = 0, \qquad (2.162)$$

$$f_n^x(z = 0) = f_n^y(z = 0) = 0. \qquad (2.163)$$

The solutions for the eigenfunctions are easily found, e.g., by separation of variables, and are given by

$$f_{mnp}^x(\boldsymbol{r}) = \cos\left(\frac{m\pi}{l_x}x\right) \sin\left(\frac{n\pi}{l_y}y\right) \sin\left(\frac{p\pi}{l_z}z\right), \qquad (2.164)$$

$$f_{mnp}^y(\boldsymbol{r}) = \sin\left(\frac{m\pi}{l_x}x\right) \cos\left(\frac{n\pi}{l_y}y\right) \sin\left(\frac{p\pi}{l_z}z\right), \qquad (2.165)$$

$$f_{mnp}^z(\boldsymbol{r}) = \sin\left(\frac{m\pi}{l_x}x\right) \sin\left(\frac{n\pi}{l_y}y\right) \cos\left(\frac{p\pi}{l_z}z\right). \qquad (2.166)$$

We observe that the former collective index n has turned to a triple index mnp where m, n, and p assume integer values that run from 0 to infinity. It is customary to introduce the abbreviations

$$k_x := \frac{m\pi}{l_x}, \qquad k_y := \frac{n\pi}{l_y}, \qquad k_z := \frac{p\pi}{l_z}. \qquad (2.167)$$

Then the eigenvalues k_{mnp}^2 are written as

$$k_{mnp}^2 = k_x^2 + k_y^2 + k_z^2. \qquad (2.168)$$

We note from (2.164) – (2.166) that $\boldsymbol{\nabla} \cdot \boldsymbol{f}(\boldsymbol{r})|_{\boldsymbol{r} \in \Gamma} = 0$ such that our solution respects the boundary condition $\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r})|_{\boldsymbol{r} \in \Gamma} = 0$. The eigenfunctions $f_{mnp}^{x,y,z}(\boldsymbol{r})$ are square integrable

in the domain Ω, i.e., within the rectangular cavity, but not normalized. In fact, we have

$$\int_0^{l_x} \int_0^{l_y} \int_0^{l_z} (f_{mnp}^{x,y,z})^2(\boldsymbol{r}) \, d^3r = \frac{l_x l_y l_z}{\epsilon_{0m} \epsilon_{0n} \epsilon_{0p}} \tag{2.169}$$

with

$$\epsilon_{0N} = \begin{cases} 1 & \text{for } N = 0, \\ 2 & \text{for } N > 0. \end{cases} \tag{2.170}$$

Thus, the normalization factor of the functions $f_{mnp}^{x,y,z}(\boldsymbol{r})$ is given by $\sqrt{\frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z}}$. The orthonormal eigenfunctions $\sqrt{\frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z}} f_{mnp}^{x,y,z}(\boldsymbol{r})$ determine the spectral representation of the dyadic Green's function $\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')$ by means of the central result (2.133) with $\lambda_n = \tilde{k}_n^2 = k_{mnp}^2 - k^2$. It follows

$$\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') = \tag{2.171}$$

$$\sum_{m,n,p=0}^{\infty} \frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z} \frac{f_{mnp}^x(\boldsymbol{r}) f_{mnp}^x(\boldsymbol{r}') \boldsymbol{e}_x \boldsymbol{e}_x + f_{mnp}^y(\boldsymbol{r}) f_{mnp}^y(\boldsymbol{r}') \boldsymbol{e}_y \boldsymbol{e}_y + f_{mnp}^z(\boldsymbol{r}) f_{mnp}^z(\boldsymbol{r}') \boldsymbol{e}_z \boldsymbol{e}_z}{k_{mnp}^2 - k^2}$$

We have mentioned that it is possible to immediately obtain from $\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')$ the dyadic Green's functions $\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')$ and $\overline{\boldsymbol{G}}^B(\boldsymbol{r}, \boldsymbol{r}')$ via (2.145) and (2.146). For example, the spectral representation of the electric dyadic Green's function that follows from (2.171) is given by

$$\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') = \sum_{m,n,p=0}^{\infty} \frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z} \times$$

$$\left[\frac{f_{mnp}^x(\boldsymbol{r}) f_{mnp}^x(\boldsymbol{r}') \boldsymbol{e}_x \boldsymbol{e}_x + f_{mnp}^y(\boldsymbol{r}) f_{mnp}^y(\boldsymbol{r}') \boldsymbol{e}_y \boldsymbol{e}_y + f_{mnp}^z(\boldsymbol{r}) f_{mnp}^z(\boldsymbol{r}') \boldsymbol{e}_z \boldsymbol{e}_z}{k_{mnp}^2 - k^2} \right.$$

$$\left. - \frac{1}{k^2} \frac{\boldsymbol{\nabla} f_{mnp}(\boldsymbol{r}) \boldsymbol{\nabla} f_{mnp}(\boldsymbol{r}')}{k_{mnp}^2 - k^2} \right] \tag{2.172}$$

where we defined for notational convenience

$$f_{mnp}(\boldsymbol{r}) := \sin(k_x x) \sin(k_y y) \sin(k_z z). \tag{2.173}$$

In summary, we arrive at spectral representations for both $\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')$ and $\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')$ within a perfectly conducting, rectangular cavity.

(b) Spectral representations in terms of longitudinal and transverse eigenfunctions

In the previous Example 1 we have been fortunate that the vector Helmholtz equation, together with the associated boundary conditions, decoupled to three independent scalar

Helmholtz equations which could straightforwardly be solved. In general, vector wave equations and their associated boundary conditions will couple the different components of the unknown vector field. This will lead, in turn, to complicated partial differential equations that are not solved easily. This problem has been discussed in particular by Morse & Feshbach [146, § 13] where it is suggested to reduce the solution of a vector wave equation to the solution of scalar wave equations. As a general solution strategy, and based on previous work of Hansen [80] and Stratton [210], Morse & Feshbach show how to set up, in a first step, an ansatz for the solution which consists of the sum of a longitudinal part and a transverse part. That such a split is unique is guaranteed by the Helmholtz theorem, see Sec. 1.5.1, where the decomposition of a vector field into its longitudinal and transverse part already has been proven to be useful for the investigation of the dynamical properties of the electromagnetic field.

The longitudinal part of the solution of a vector wave equation is the gradient of a scalar potential. Many techniques to obtain solutions for scalar potentials are available. From a physical point of view we know that the longitudinal part of a solution of the Maxwell equations for the electromagnetic field is nondynamical and determined from the position of electric charges.

The transverse part of the solution of a vector wave equation is the curl of a vector potential. The constraint that the divergence of a curl vanishes indicates that the transverse part of the solution, and thus the vector potential, may always be derived from two scalar fields. In electrodynamics these two scalar fields represent the two dynamical degrees of freedom of the electromagnetic field that were exhibited in Sec. 1.5.1 in terms of the normal variables $\mathsf{a}(\boldsymbol{k}, t)$.

We illustrate this solution strategy by a specific example, compare [30, § 5].

Example 2: The vector Helmholtz equation for the electric field strength $\boldsymbol{E}(\boldsymbol{r})$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}, \omega) - k^2 \boldsymbol{E}(\boldsymbol{r}, \omega) = -j\omega\mu \boldsymbol{J}(\boldsymbol{r}, \omega) \tag{2.174}$$

has the associated eigenvalue problem

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{f}_n = \lambda_n \boldsymbol{f}_n \tag{2.175}$$

We have in mind to find the eigenfunctions of this equation in order to construct a spectral representation of the electric dyadic Green's function $\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')$, if viewed as a solution of (2.143). As a prerequisite we recall that a split of the corresponding vector wave equation (2.140) into its longitudinal and its transverse part yields

$$-k^2 \boldsymbol{E}_\| = -j\omega\mu \boldsymbol{J}_\| , \tag{2.176}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{E}_\perp - k^2 \boldsymbol{E}_\perp = -j\omega\mu \boldsymbol{J}_\perp . \tag{2.177}$$

If we introduce a scalar potential φ in order to express $\boldsymbol{E}_\|$ according to

$$\boldsymbol{E}_\| = -\boldsymbol{\nabla}\varphi \tag{2.178}$$

we identify (2.176) as a Poisson equation,

$$-\boldsymbol{\nabla} \cdot \boldsymbol{E}_{\|} = \Delta\varphi = -\frac{j\omega\mu}{k^2}\boldsymbol{\nabla} \cdot \boldsymbol{J}_{\|} \tag{2.179}$$

$$= -\mu\frac{\omega^2}{k^2}\rho \tag{2.180}$$

$$= -\frac{\rho}{\varepsilon} \tag{2.181}$$

where we used the continuity equation (1.209) and the relation $\omega^2/k^2 = c^2 = 1/(\varepsilon\mu)$.

Now the eigenfunctions $\boldsymbol{f}_n$ of the eigenvalue problem are divided into longitudinal eigenfunctions $\boldsymbol{L}_n$,

$$\boldsymbol{\nabla} \times \boldsymbol{L}_n = \boldsymbol{0}\,, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{L} \neq \boldsymbol{0}\,, \tag{2.182}$$

and transverse eigenfunctions $\boldsymbol{F}_n$,

$$\boldsymbol{\nabla} \times \boldsymbol{F}_n \neq \boldsymbol{0}\,, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{F}_n = \boldsymbol{0}\,. \tag{2.183}$$

Clearly, the longitudinal eigenfunctions $\boldsymbol{L}_n$ are in the null space of $\boldsymbol{\nabla} \times \boldsymbol{\nabla}\times$. However, they are not in the null space of $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times\, -k^2$ and need to be obtained via the solution of the Poisson equation (2.181) in terms of scalar functions φ_n. As in the previous Example 1 this involves the eigenfunctions of the (negative) Laplace operator.

The transverse eigenfunctions $\boldsymbol{F}_n$ are further divided into functions $\boldsymbol{M}_n$ and $\boldsymbol{N}_n$. The functions $\boldsymbol{M}_n$ are obtained via the ansatz

$$\boldsymbol{M}_n = \boldsymbol{\nabla} \times (\psi_n\boldsymbol{c}) \tag{2.184}$$

with a scalar potential ψ_n and a constant vector $\boldsymbol{c}$ which, in this context, is often referred to as *pilot vector* or *piloting vector*. If this ansatz is inserted into the vector wave equation one finds

$$\boldsymbol{\nabla} \times \left(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times (\psi_n\boldsymbol{c}) - k^2(\psi_n\boldsymbol{c})\right) = -\boldsymbol{\nabla} \times [\boldsymbol{c}(\Delta\psi_n + k^2\psi_n)] \tag{2.185}$$

$$= \boldsymbol{0} \tag{2.186}$$

and it follows

$$\Delta\psi_n + k^2\psi_n = 0\,. \tag{2.187}$$

Hence, $\boldsymbol{M}_n$ is obtained from the solution of a scalar Helmholtz equation. Similarly, we obtain the functions $\boldsymbol{N}_n$ from the ansatz

$$\boldsymbol{N}_n = \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times (\chi_n\boldsymbol{c}) \tag{2.188}$$

with another scalar potential χ_n which fulfills

$$\Delta\chi_n + k^2\chi_n = 0\,. \tag{2.189}$$

For a rectangular cavity we may choose the piloting vector $\boldsymbol{c}$ as $\boldsymbol{e}_z$. This yields the transverse functions $\boldsymbol{M}_n$ and $\boldsymbol{N}_n$ as TE- and TM-modes, respectively, with reference to the $z-$ coordinate. It follows for the the scalar functions φ_n, ψ_n, and χ_n that [146, p. 1774]

$$\varphi_{mnp}(\boldsymbol{r}) = \sin\left(\frac{m\pi}{l_x}x\right)\sin\left(\frac{n\pi}{l_y}y\right)\sin\left(\frac{p\pi}{l_z}z\right) , \tag{2.190}$$

$$\psi_{mnp}(\boldsymbol{r}) = \cos\left(\frac{m\pi}{l_x}x\right)\cos\left(\frac{n\pi}{l_y}y\right)\sin\left(\frac{p\pi}{l_z}z\right) , \tag{2.191}$$

$$\chi_{mnp}(\boldsymbol{r}) = \sin\left(\frac{m\pi}{l_x}x\right)\sin\left(\frac{n\pi}{l_y}y\right)\cos\left(\frac{p\pi}{l_z}z\right) . \tag{2.192}$$

This yields the eigenfunctions

$$\boldsymbol{L}_{mnp}(\boldsymbol{r}) = \boldsymbol{\nabla}\left(\sin(k_x x)\sin(k_y y)\sin(k_z z)\right) , \tag{2.193}$$

$$\boldsymbol{M}_{mnp}(\boldsymbol{r}) = \boldsymbol{\nabla}\times\left(\cos(k_x x)\cos(k_y y)\sin(k_z z)\boldsymbol{e}_z\right) , \tag{2.194}$$

$$\boldsymbol{N}_{mnp}(\boldsymbol{r}) = \boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\left(\sin(k_x x)\sin(k_y y)\cos(k_z z)\boldsymbol{e}_z\right) , \tag{2.195}$$

$$\tag{2.196}$$

which form a complete set. After normalization we obtain the corresponding spectral representation of the Green's function $\overline{\boldsymbol{G}}^E(\boldsymbol{r},\boldsymbol{r}')$ in the form

$$\overline{\boldsymbol{G}}^E(\boldsymbol{r},\boldsymbol{r}') = \sum_{m,n,p=0}^{\infty} \frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z}\left[\frac{\boldsymbol{M}_{mnp}(\boldsymbol{r})\boldsymbol{M}_{mnp}(\boldsymbol{r}')}{(k_{mnp}^2-k^2)k_{mn}^2} + \right.$$
$$\left.\frac{\boldsymbol{N}_{mnp}(\boldsymbol{r})\boldsymbol{N}_{mnp}(\boldsymbol{r}')}{(k_{mnp}^2-k^2)k_{mnp}^2 k_{mn}^2} - \frac{\boldsymbol{L}_{mnp}(\boldsymbol{r})\boldsymbol{L}_{mnp}(\boldsymbol{r}')}{k^2 k_{mnp}^2}\right] \tag{2.197}$$

where k_{mnp}^2 is defined as in (2.168) and, furthermore,

$$k_{mn}^2 = k_x^2 + k_y^2 . \tag{2.198}$$

We may compare this result to the spectral representation (2.172). Both representations are equivalent but derived from different eigenfunctions. The first representation (2.172) utilizes the eigenfunctions of the negative Laplace operator $-\Delta$ while the second representation (2.197) utilizes the eigenfunctions of the double-curl operator $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times$. These two sets of eigenfunctions constitute equivalent bases of the Hilbert space $\boldsymbol{L}^2(\Omega)^3$ with inner product (2.44). More equivalent representations of dyadic Green's functions for a rectangular cavity are given in [212].

2.3.2   Spectral representations of lossy cavities' Green's functions and the quality factor of a cavity

So far we considered lossless cavities. The fundamental difference between solutions of electromagnetic wave equations in lossless cavities and corresponding solutions in

lossy cavities is the circumstance that in the lossy case the wave operators $\Delta + k^2$ and $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \; -k^2$ are no longer self-adjoint. To see this explicitly, we recall that self-adjointness of a differential operator requires

$$\langle \mathcal{L}_D f, g \rangle = \int_\Omega (\mathcal{L}_D f) g^* \, d\Omega = \int_\Omega f (\mathcal{L}_D g)^* \, d\Omega = \langle f, \mathcal{L}_D g \rangle \tag{2.199}$$

If we take $\mathcal{L}_D$ as $\Delta + k^2$ or $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \; -k^2$ we notice that the shift of $\mathcal{L}_D$ from f to g within the inner product requires to have a real wavenumber k. Also we have to apply partial integration to the differential operators Δ and $\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times$. This will produce boundary terms which we require to vanish in order to arrive at a self-adjoint boundary value problem. We will discuss in the following that, in general, these two requirements are not met if losses are present.

(a) Dissipative media and complex wavenumber

In electrodynamics, dissipative media exhibit electric and magnetic losses which are characterized by a complex permittivity ε and complex permeability μ, respectively [92]. Electric losses are more common than magnetic losses and the complex permittivity is often written in the form

$$\varepsilon = \varepsilon' - j \left(\varepsilon'' + \frac{\sigma}{\omega} \right) \tag{2.200}$$

with σ the conductivity of the medium considered. Permittivity ε and permeability μ yield the *intrinsic impedance*[5] $\mathcal{Z}_{\mathrm{int}}$ and the relation between wavenumber k and angular frequency ω,

$$\mathcal{Z}_{\mathrm{int}} = \sqrt{\frac{\mu}{\varepsilon}}\,. \tag{2.202}$$

$$k = \omega \sqrt{\varepsilon \mu}\,, \tag{2.203}$$

It follows that in a lossy medium $\mathcal{Z}_{\mathrm{int}}$ and k may become complex quantities. In principle, the wavenumber k can be kept real since, trivially, we can always write (2.203) as $\omega = k/\sqrt{\varepsilon\mu}$, define a complex angular frequency, and keep the wavenumber real. In

---

[5]Here we mention the intrinsic impedance only in passing. This quantity represents a common notation that is applied if freely propagating electromagnetic waves, i.e., radiation fields, are considered. If the amplitudes of the corresponding electric field strength and magnetic excitation are denoted by $\mathcal{E}$ and $\mathcal{H}$, respectively, then it follows from the sourceless Maxwell equations that [93, §7.1]

$$\sqrt{\frac{\mu}{\varepsilon}} = \frac{\mathcal{E}}{\mathcal{H}}\,. \tag{2.201}$$

The expressions on both sides of this equation exhibit the dimension Ω and, thus, are defined as intrinsic impedance. The intrinsic impedance is an intrinsic property of the space or medium where the electromagnetic field is propagating.

fact, both a complex wavenumber or a complex frequency are physically meaningful. Since a plane wave that propagates, say, in the positive x-direction is of the form

$$\boldsymbol{E}(x,t) = \boldsymbol{E}_0 e^{-j(kx-\omega t)} \tag{2.204}$$

the imaginary part of a complex wavenumber will describe an attenuation in space while the imaginary part of a complex angular frequency will describe a complementary attenuation in time. The relation between both imaginary parts is given by (2.203), of course.

If we work in the frequency domain and assume a time dependency $e^{j\omega t}$ we need to keep ω real and have a complex wavenumber k. Then the wave operators $\Delta + k^2$ and $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times -k^2$ are no longer self-adjoint and there is no more guarantee that the associated eigenfunctions form a complete set that can be used to obtain a spectral representation of the corresponding Green's function. If we consider the special case that a lossy medium is enclosed by a *perfectly conducting* cavity it is clear that the boundary conditions are the same as in the lossless case. Then the solution of the Helmholtz equations

$$\Delta\boldsymbol{A} + k^2\boldsymbol{A} = -\mu\boldsymbol{J} \, , \tag{2.205}$$

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{E} - k^2\boldsymbol{E} = -j\omega\mu\boldsymbol{J} \tag{2.206}$$

in terms of the solution of an eigenvalue problem, as described by Example 1 and Example 2 of Sec. 2.3.1, is the same for both real and complex wavenumbers k. It follows that the eigenfunctions are of the same form in both cases and can be used to obtain the spectral representations of the associated Green's functions [95]. Thus, to generalize within a perfectly conducting cavity the solution of Helmholtz equations for lossless media to the solution of Helmholtz equations for lossy media is simple. We only have to employ a complex wavenumber in the relevant expressions. This is, in particular, true for the spectral representations of the Green's functions.

(b) Lossy cavity walls

It was pointed out in connection with (2.199) that self-adjointness of a differential operator requires the vanishing of boundary terms that are produced by partial integration. In Example 1 and Example 2 of Sec. 2.2.2 the boundary terms associated to the operators $\Delta + k^2$ and $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times -k^2$ vanished since we assumed Dirichlet boundary conditions that correspond to a perfect conductor and, thus, imply no losses at the cavity walls. However, the boundary terms will not generally vanish if losses at the cavity walls are present. These losses are commonly modeled by approximate boundary condition that often are given in terms of impedance boundary conditions [201, 92, 89]. Already the relatively simple first order boundary condition of Leontovich [124],

$$\boldsymbol{e}_n \times \boldsymbol{E} = -\mathscr{Z}_{\mathrm{s}}\,\boldsymbol{e}_n \times (\boldsymbol{e}_n \times \boldsymbol{H})\,, \tag{2.207}$$

with $\mathcal{Z}_{\mathrm{s}}$ the *surface impedance*[6], leads to a coupling between different field components [201]. This considerably complicates the solution of a Helmholtz equation within a lossy cavity. As a rule, there are no longer exact solutions available and one needs to rely on approximate solution techniques.

An important approximate technique to find the electromagnetic field inside a cavity with lossy walls is to expand it in terms of the orthogonal eigenfunctions of a lossless cavity [92, §4]. This is reasonable as long as the losses are rather small and do not perturb the lossless modes too much. To illustrate this technique we consider within a lossy cavity the source-free Maxwell equations

$$\boldsymbol{\nabla} \times \boldsymbol{E} + j\omega\mu\boldsymbol{H} = \boldsymbol{0} \,, \tag{2.208}$$

$$\boldsymbol{\nabla} \times \boldsymbol{H} - j\omega\varepsilon\boldsymbol{E} = \boldsymbol{0} \,, \tag{2.209}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are subject to the impedance boundary condition (2.207). Here we assumed a harmonic time dependency $e^{j\omega t}$ that characterizes the electromagnetic modes *after* external sources have been switched off. Due to the losses at the cavity walls the electromagnetic field will be attenuated in time such that the angular frequency needs to be a complex quantity, $\omega = \omega' + j\omega''$. It is of primary interest to find the possible values of ω since these values characterize the electromagnetic properties of the lossy cavity.

In order to proceed we introduce eigenfunctions of the lossless problem that yield expansions of the electric and magnetic field within a perfectly conducting cavity. We denote electric and magnetic eigenfunctions by $\boldsymbol{E}_n = \{\boldsymbol{L}_n^E, \boldsymbol{F}_n^E\}$ and $\boldsymbol{H}_n = \{\boldsymbol{L}_n^H, \boldsymbol{F}_n^H\}$, respectively. In view of (2.208), (2.209) these eigenfunctions fulfill[7], compare [227, §10],

$$\boldsymbol{\nabla} \times \boldsymbol{L}_n^E = \boldsymbol{0} \,, \tag{2.210}$$

$$\boldsymbol{\nabla} \times \boldsymbol{L}_n^H = \boldsymbol{0} \,, \tag{2.211}$$

$$\boldsymbol{\nabla} \times \boldsymbol{F}_n^E + j\omega_{\mathrm{p}}\mu\boldsymbol{F}_n^H = \boldsymbol{0} \,, \tag{2.212}$$

$$\boldsymbol{\nabla} \times \boldsymbol{F}_n^H - j\omega_{\mathrm{p}}\varepsilon\boldsymbol{F}_n^E = \boldsymbol{0} \,. \tag{2.213}$$

As eigenvectors of self-adjoint operators these eigenvectors are mutually orthogonal and

---

[6]The surface impedance $\mathcal{Z}_{\mathrm{s}}$ is defined as the coefficient of proportionality that links on the surface of a conductor the tangential component $\boldsymbol{e}_n \times \boldsymbol{E}$ of the electric field strength to the (effective) surface current $\boldsymbol{J}_s = \boldsymbol{e}_n \times (\boldsymbol{e}_n \times \boldsymbol{H})$, compare (1.177) and (1.178). The concept of surface impedance is an approximate one. For a good conductor it follows from Maxwell equations that $\mathcal{Z}_{\mathrm{s}} \approx (1 + j)/\sigma\delta$, with σ the conductivity of the conductor and δ the relevant skin depth [93, §8.1]. In (2.207) the sign in front of $\mathcal{Z}_{\mathrm{s}}$ depends on the orientation of the normal unit vector $\boldsymbol{e}_n$. Here, the orientation is chosen such that $\boldsymbol{e}_n$ points to the interior of the cavity.

[7]The eigenfunctions $\boldsymbol{E}_n$ and $\boldsymbol{H}_n$ can also be viewed as eigenfunctions of the double-curl operator $\boldsymbol{\nabla} \times \boldsymbol{\nabla}\times$ and constructed along the lines of Example 2 of Sec. 2.3.1. This is easily verified if, for the case of position-independent material parameters ε and μ, the equations (2.212) and (2.213) are decoupled by the application of the curl operator $\boldsymbol{\nabla}\times$. For the specific Example 2 of Sec. 2.3.1 we have the correspondences $\boldsymbol{L}_n^E \longrightarrow \boldsymbol{L}_{mnp}$ and $\boldsymbol{F}_n^E \longrightarrow \boldsymbol{M}_{mnp}, \boldsymbol{N}_{mnp}$.

can be taken as normalized as well[8],

$$\int_\Omega \boldsymbol{E}_m(\boldsymbol{r}) \cdot \boldsymbol{E}_n^*(\boldsymbol{r})\, d^3r = \int_\Omega \boldsymbol{H}_m(\boldsymbol{r}) \cdot \boldsymbol{H}_n^*(\boldsymbol{r})\, d^3r = \delta_{mn}\,, \qquad (2.214)$$

with δ_{mn} the Kronecker symbol which assumes the value 1 if $m = n$ and 0 otherwise. In (2.212), (2.213) the index p ("perfect") indicates that the eigenfrequencies ω_{p} of the perfectly conducting cavity will be different from the eigenfrequencies ω of the lossy cavity.

We expand the solutions of (2.208), (2.209) in terms of the eigenfunctions $\boldsymbol{E}_n$, $\boldsymbol{H}_n$ and unknown coefficients α_n, β_n,

$$\boldsymbol{E}(\boldsymbol{r}) = \sum_n \alpha_n \boldsymbol{E}_n(\boldsymbol{r})\,, \qquad \boldsymbol{H}(\boldsymbol{r}) = \sum_n \beta_n \boldsymbol{H}_n(\boldsymbol{r})\,, \qquad (2.215)$$

and consider the expressions

$$\int_\Omega \big(\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}) + j\omega\mu \boldsymbol{H}(\boldsymbol{r})\big) \cdot \boldsymbol{H}_n(\boldsymbol{r})\, d^3r = 0\,, \qquad (2.216)$$

$$\int_\Omega \big(\boldsymbol{\nabla} \times \boldsymbol{H}(\boldsymbol{r}) - j\omega\varepsilon \boldsymbol{E}(\boldsymbol{r})\big) \cdot \boldsymbol{E}_n(\boldsymbol{r})\, d^3r = 0\,, \qquad (2.217)$$

for $n = 1, 2, \ldots$. By means of the vector identity (B.10) and after application of Stokes theorem we find

$$\int_\Omega \big[\boldsymbol{E}(\boldsymbol{r}) \cdot \big(\boldsymbol{\nabla} \times \boldsymbol{H}_n(\boldsymbol{r})\big) + j\omega\mu \boldsymbol{H}(\boldsymbol{r}) \cdot \boldsymbol{H}_n(\boldsymbol{r})\big]\, d^3r + \int_\Gamma \big(\boldsymbol{E}(\boldsymbol{r}) \times \boldsymbol{H}_n(\boldsymbol{r})\big) \cdot \boldsymbol{e}_n\, d^2r = 0\,, \tag{2.218}$$

$$\int_\Omega \big[\boldsymbol{H}(\boldsymbol{r}) \cdot \big(\boldsymbol{\nabla} \times \boldsymbol{E}_n(\boldsymbol{r})\big) - j\omega\varepsilon \boldsymbol{E}(\boldsymbol{r}) \cdot \boldsymbol{E}_n(\boldsymbol{r})\big]\, d^3r + \int_\Gamma \big(\boldsymbol{H}(\boldsymbol{r}) \times \boldsymbol{E}_n(\boldsymbol{r})\big) \cdot \boldsymbol{e}_n\, d^2r = 0\,. \tag{2.219}$$

The surface integral of (2.218) can be rewritten as

$$\int_\Gamma \big(\boldsymbol{E}(\boldsymbol{r}) \times \boldsymbol{H}_n(\boldsymbol{r})\big) \cdot \boldsymbol{e}_n\, d^2r = -\mathscr{Z}_{\mathrm{s}} \int \big(\boldsymbol{e}_n \times \big(\boldsymbol{e}_n \times \boldsymbol{H}(\boldsymbol{r})\big)\big) \cdot \boldsymbol{H}_n(\boldsymbol{r})\, d^2r \qquad (2.220)$$

$$= \mathscr{Z}_{\mathrm{s}} \sum_n \beta_n \underbrace{\int_\Gamma \big(\boldsymbol{e}_n \times \boldsymbol{H}_m(\boldsymbol{r})\big) \cdot \big(\boldsymbol{e}_n \times \boldsymbol{H}_n(\boldsymbol{r})\big)\, d^2r}_{=:\Lambda_{mn}} \quad (2.221)$$

$$= \mathscr{Z}_{\mathrm{s}} \sum_n \beta_n \Lambda_{mn} \qquad (2.222)$$

where we used the impedance boundary condition (2.207) and the expansion (2.215) for the magnetic field. Due to the boundary condition $\boldsymbol{e}_n \times \boldsymbol{E}_n = \boldsymbol{0}$ the surface integral

---

[8]This statement could also be written in terms of the inner product, $\langle \boldsymbol{E}_m, \boldsymbol{E}_n \rangle = \langle \boldsymbol{H}_m, \boldsymbol{H}_n \rangle = \delta_{mn}$

of (2.219) vanishes. It remains to eliminate in the volume integrals by means of (2.212)–(2.211) the curl-operators, to expand the fields $\boldsymbol{E}$ and $\boldsymbol{H}$ by means of (2.215), and to use the orthonormal property (2.214). This yields

$$j(\omega_{\mathrm{p}}\varepsilon\alpha_n + \omega\mu\beta_n) + \mathcal{Z}_{\mathrm{s}}\sum_n \beta_n\Lambda_{mn} = 0\,, \tag{2.223}$$

$$\omega\varepsilon\alpha_n + \omega_{\mathrm{p}}\mu\beta_n = 0\,. \tag{2.224}$$

In this system of equations the variables α_n, β_n and ω are unknown. We are particularly interested in the time-dependent case $\omega \neq 0$. In this case (2.224) can be used to eliminate α_n in (2.224). This leads us to

$$\sum_m \beta_m \left(j\mu\frac{\omega^2 - \omega_{\mathrm{p}}^2}{\omega}\delta_{mn} + \mathcal{Z}_{\mathrm{s}}\Lambda_{mn} \right) = 0\,. \tag{2.225}$$

Nontrivial solutions for the unknowns β_n are obtained if and only if

$$\mathrm{Det}\left[j\mu\frac{\omega^2 - \omega_{\mathrm{p}}^2}{\omega}\delta_{mn} + \mathcal{Z}_{\mathrm{s}}\Lambda_{mn} \right] = 0 \tag{2.226}$$

and this condition yields the possible values for ω.

The condition (2.226) is readily evaluated if simplifying assumptions are made: Let us suppose that the modes ω_{p} are not degenerate and sufficiently separated such that ω can be considered to be close to the mth mode $\omega_{\mathrm{p}}^{(m)}$. If, furthermore, the surface impedance $\mathcal{Z}_{\mathrm{s}}$ is small the coupling between different modes is small as well and it is sufficient to consider the single equation

$$j\mu\frac{\omega^2 - (\omega_{\mathrm{p}}^{(m)})^2}{\omega} + \mathcal{Z}_{\mathrm{s}}\Lambda_{mm} = 0 \tag{2.227}$$

which has the solution

$$\omega = \sqrt{\left(\omega_{\mathrm{p}}^{(m)}\right)^2 - \left(\frac{\mathcal{Z}_{\mathrm{s}}\Lambda_{mm}}{2\mu}\right)^2} + \frac{j}{2}\frac{\mathcal{Z}_{\mathrm{s}}\Lambda_{mm}}{\mu}\,. \tag{2.228}$$

This formula exemplifies that a finite surface impedance $\mathcal{Z}_{\mathrm{s}}$ modifies the possible resonance frequencies of a cavity if compared to the lossless case:

- Resonance frequencies are shifted towards smaller values.

- The angular frequency ω acquires a positive imaginary part which leads, due to the time dependency $e^{j\omega t}$, to an attenuation in time.

These qualitative results could have been expected from the well-known features of a damped harmonic oscillator [14, §18]. However, the difficult part of the analysis is to explicitly calculate the values of the complex resonance frequencies. In general, this is only possible by means of perturbative and approximate methods.

(c) Quality factor of a cavity

In a perfectly conducting cavity the discrete eigenvalues of the Helmholtz equations for the electromagnetic field determine the discrete resonance frequencies. In principle, only those time-harmonic excitations which exactly match a resonance frequency may excite an electromagnetic field and then the resonance curve is modeled by a delta-function peak. If losses occur these will smear out the sharp resonance curves and there will be band of frequencies around each resonance frequency that can lead to an excitation. The half-width of the corresponding realistic resonance curves depends on the amount of power loss in the cavity. It is common to characterize this circumstance by a dimensionless quantity which is called the *quality factor Q* of a cavity. It is defined by [93]

$$Q = \omega_{\mathrm{p}} \frac{\text{stored energy}}{\text{power loss}} . \tag{2.229}$$

If we denote the energy that is stored in the cavity by W_{cav} it follows from energy conservation[9]

$$\frac{dW_{\mathrm{cav}}}{dt} = -\frac{\omega_{\mathrm{p}}}{Q} W_{\mathrm{cav}}(t) \tag{2.230}$$

Therefore, the time dependency of W_{cav} is given by

$$W_{\mathrm{cav}}(t) = W_{0\,\mathrm{cav}} e^{-\omega_{\mathrm{p}} t/Q} . \tag{2.231}$$

The electromagnetic energy is proportional to the square of the electric and magnetic fields in the cavity. It follows that the electromagnetic fields have a time dependency of the form

$$\boldsymbol{E}(t) = \boldsymbol{E}_0 e^{-\omega_{\mathrm{p}} t/(2Q)} e^{j(\omega_{\mathrm{p}} + \Delta\omega)t} \tag{2.232}$$

$$= \boldsymbol{E}_0 e^{j(\omega_{\mathrm{p}} + \Delta\omega + j\omega_{\mathrm{p}}/(2Q))t} \tag{2.233}$$

where we incorporated the shift $\Delta\omega$ of the resonance frequency which is due to the losses. Since the complex eigenfrequencies of a lossy cavity often are written as $\omega = \omega' + j\omega''$ we have from (2.233)

$$\omega'' = \frac{\omega' - \Delta\omega}{2Q} \approx \frac{\omega'}{2Q} , \tag{2.234}$$

and the quality factor can be expressed according to

$$Q \approx \frac{\omega'}{2\omega''} . \tag{2.235}$$

We also have from (2.234) the relation

$$\omega \approx \omega' \left(1 + \frac{j}{2Q}\right) . \tag{2.236}$$

---

[9] As in the previous subsection 2.3.2 (b) we still exclude external sources.

(d) Time harmonic sources and lossy cavities

If time harmonic sources are present inside a lossy cavities we will have a stationary process where the power loss is compensated by the time harmonic sources. In this case the electromagnetic fields will no longer be attenuated in time. But we can think of the electromagnetic field inside the cavity as the result of propagating fields that scatter many times at the cavity walls and eventually superimpose to form the eigenmodes that we observe in the time harmonic case. This interpretation is justified by the scattering expansion of the next Sec. 2.3.3. Then a propagating field inside a lossy cavity will decay in space while it propagates through a lossy medium or gets scattered at a lossy cavity walls. In Sec. 2.3.2 we have already mentioned that due to the relation (2.203) the losses in a dissipative medium can be accounted for by either a complex wavenumber $k = k' - jk''$ or by a complex angular frequency $\omega = \omega' + j\omega''$. That is, if we consider the phase $\phi = \boldsymbol{k} \cdot \boldsymbol{r} - \omega t$ of a propagating field it is, from a mathematical point of view, a matter of taste to accommodate a negative imaginary part of the phase in the wavenumber or in the angular frequency. If we have lossy cavity walls the situation is, from a physical point of view, more complicated since we require perturbation theory in order to infer from a given surface impedance $\mathcal{Z}_\mathrm{s}$ that the angular frequency becomes a complex quantity. However, a rather simple result such as (2.228) or, in terms of the quality factor, (2.236) shows that for small losses the effect of lossy cavity walls is equivalent to that of a lossy medium. Indeed, for small losses a propagating wave will scatter many times before there is a noticeable attenuation and, in the mean, this is equivalent to a continuous attenuation due to a lossy medium.

It follows that both for dissipative media and lossy cavities we can describe small losses by a complex wavenumber. To convert between complex angular frequencies and complex wavenumbers we use the relation (2.203) and, for complex wavenumbers, write

$$k' - jk'' = \omega'\sqrt{\varepsilon\mu} \tag{2.237}$$

while, for complex angular frequencies, we write

$$\omega' + j\omega'' = \frac{k'}{\sqrt{\varepsilon\mu}} \tag{2.238}$$

For small losses, $\mathrm{Re}(\sqrt{\varepsilon\mu}) \gg \mathrm{Im}(\sqrt{\varepsilon\mu})$, it is easy to find that

$$\frac{k'}{k''} = \frac{\omega'}{\omega''} \tag{2.239}$$

It follows, for example, that the relation (2.236), which is valid if we work with complex frequencies, is equivalent to

$$k \approx k'\left(1 - \frac{j}{2Q}\right) \tag{2.240}$$

if we work with complex wavenumbers.

The bottom line is that in the time harmonic case we split off the time dependency $e^{j\omega't}$ and incorporate losses in the complex wavenumber $k = k' - jk''$. From the discussion above it follows that small losses at the cavity walls can approximatively be treated as losses in a dissipative medium. In particular, we then can use the spectral representations for the Green's functions of the Helmholtz equations to write down the corresponding solution of the Maxwell equations. In the lossless case the wavenumber in the solution is real, in the lossy case it becomes complex.

2.3.3 Ray representations of cavities' Green's functions from scattering expansions

So far we have considered spectral representations of Green's functions. This type of representations emerged as a natural result of the eigenfunction expansion method where we employed the spectral properties of the relevant wave operators. In this section we want to point out another type of representation which is based on the Green's function $G_0(\boldsymbol{r}, \boldsymbol{r}')$ of free space, as defined by (2.136), and allows to construct cavities' Green's functions in terms of a so-called *ray representation*. We can think of a *ray* as an electromagnetic field which propagates freely, i.e., without scattering, between a source point $\boldsymbol{r}'$ and an observation point $\boldsymbol{r}$. Therefore, a ray is represented by the Green's function $G_0(\boldsymbol{r}, \boldsymbol{r}')$ which, in this context, is also referred to as a *propagator*.

If we put an electromagnetic source within a cavity at a position $\boldsymbol{r}'$ rays will emanate from this source. One of these rays will directly lead to an observation point $\boldsymbol{r}$. Other rays will scatter at the cavity wall and might eventually reach the observation $\boldsymbol{r}'$. These rays do not necessarily correspond to the one of geometric optics, all possible paths that lead via scattering processes from the source to the observation point need to be considered in order to obtain the complete effect of the source on the observation point. It follows that the Green's function of a cavity can be obtained if we sum up all rays that reach an observation point in the presence of this cavity. The corresponding ray representation of the Green's function is the result of a scattering expansion which has been introduced by Balian & Duplantier [6]. In their paper Balian & Duplantier constructed scattering expansions for the dyadic magnetic and electric Green's function. To find an analogue representation for the dyadic vector potential Green's function is rather immediate and will be considered in the following.

Suppose that electric currents $\boldsymbol{J}(\boldsymbol{r}')$ are enclosed by a cavity within a volume Ω. Then the vector potential $\boldsymbol{A}(\boldsymbol{r})$ at an observation point $\boldsymbol{r}$ is given by

$$\boldsymbol{A}(\boldsymbol{r}) = \mu \int_\Omega G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}(\boldsymbol{r}') \, d^3r' + \mu \int_{\Gamma=\partial\Omega} G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}_s(\boldsymbol{r}') \, d^2r' \tag{2.241}$$

$$=: \boldsymbol{A}^{\mathrm{inc}}(\boldsymbol{r}) + \mu \int_{\Gamma=\partial\Omega} G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}_s(\boldsymbol{r}') \, d^2r' \,. \tag{2.242}$$

Here $\boldsymbol{J}_s(\boldsymbol{r}')$ denotes the surface current that is induced by the currents $\boldsymbol{J}(\boldsymbol{r}')$ on the

inner surface of the cavity. The incident field $\boldsymbol{A}^{\mathrm{inc}}(\boldsymbol{r})$ is the magnetic vector potential that results from the direct ray contributions of the sources $\boldsymbol{J}(\boldsymbol{r}')$. If we are able to express $\boldsymbol{J}_s(\boldsymbol{r}')$ in terms of $\boldsymbol{J}(\boldsymbol{r}')$ we can find from (2.241) an expression for the cavities' Green's function.

The surface currents can be determined from a magnetic field integral equation which is based on the boundary condition (1.178), compare [227, §11.6],

$$\boldsymbol{J}_s(\boldsymbol{r}) = 2\boldsymbol{e}_n \times \boldsymbol{H}^{\mathrm{inc}}(\boldsymbol{r}) + 2\int_\Gamma \boldsymbol{e}_n \times \boldsymbol{\nabla} \times \big(G_0(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{J}_s(\boldsymbol{r}')\big)\, d^2 r'\,. \tag{2.243}$$

In this equation we defined

$$\boldsymbol{H}^{\mathrm{inc}}(\boldsymbol{r}) = \frac{1}{\mu}\boldsymbol{\nabla} \times \boldsymbol{A}^{\mathrm{inc}}(\boldsymbol{r}) \tag{2.244}$$

$$= \int_\Omega \boldsymbol{\nabla} \times (G_0(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{J}(\boldsymbol{r}'))\, d^3 r'\,. \tag{2.245}$$

The integral kernel of the last expression can be rewritten if we use the identity (B.8) and take advantage of the symmetry property $G_0(\boldsymbol{r}, \boldsymbol{r}') = G_0(|\boldsymbol{r} - \boldsymbol{r}'|)$ of the Green's function of free space. Then we have

$$\boldsymbol{H}^{\mathrm{inc}}(\boldsymbol{r}) = \int_\Omega G'_0(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{e}_{\boldsymbol{r}, \boldsymbol{r}'} \times \boldsymbol{J}(\boldsymbol{r}'))\, d^3 r'\,, \tag{2.246}$$

where we introduced the abbreviation

$$G'_0(\boldsymbol{r}, \boldsymbol{r}') := \frac{\partial G_0(|\boldsymbol{r} - \boldsymbol{r}'|)}{\partial |\boldsymbol{r} - \boldsymbol{r}'|} \tag{2.247}$$

and $\boldsymbol{e}_{\boldsymbol{r}, \boldsymbol{r}'}$ denotes a unit vector that points from $\boldsymbol{r}$ to $\boldsymbol{r}'$. Let us now define the first order current $\boldsymbol{J}_{1s}$ as

$$\boldsymbol{J}_{1s}(\boldsymbol{r}) := 2\boldsymbol{e}_n \times \boldsymbol{H}^{\mathrm{inc}}(\boldsymbol{r}) = \frac{2}{\mu}\,\boldsymbol{e}_n \times \big(\boldsymbol{\nabla} \times \boldsymbol{A}^{\mathrm{inc}}(\boldsymbol{r})\big)\,. \tag{2.248}$$

It can be used to solve the magnetic field integral equation (2.243) by iteration. This yields a solution for $\boldsymbol{J}_s(\boldsymbol{r})$ in terms of a Neumann series (2.64)[10],

$$\boldsymbol{J}_s(\boldsymbol{r}) = \boldsymbol{J}_{1s}(\boldsymbol{r}) + \underbrace{2\int_\Gamma \boldsymbol{e}_{n_r} \times (G'_0(\boldsymbol{r}, \boldsymbol{r}_1)\boldsymbol{e}_{\boldsymbol{r}, \boldsymbol{r}_1}) \times \boldsymbol{J}_{1s}(\boldsymbol{r}_1)\, d^2 r_1}_{=:\boldsymbol{J}_{2s}}$$

$$+ \underbrace{4\iint_\Gamma \boldsymbol{e}_{n_r} \times (G'_0(\boldsymbol{r}, \boldsymbol{r}_1)\boldsymbol{e}_{\boldsymbol{r}, \boldsymbol{r}_1}) \times \boldsymbol{e}_{n_{r_1}} \times (G'_0(\boldsymbol{r}_1, \boldsymbol{r}_2)\boldsymbol{e}_{\boldsymbol{r}_1, \boldsymbol{r}_2}) \times \boldsymbol{J}_{1s}(\boldsymbol{r}_2)\, d^2 r_1\, d^2 r_2}_{=:\boldsymbol{J}_{3s}} + \dots$$

$$= \boldsymbol{J}_{1s} + \boldsymbol{J}_{2s} + \boldsymbol{J}_{3s} + \dots$$

$$= \sum_{n=1}^\infty \boldsymbol{J}_{ns}\,. \tag{2.249}$$

---

[10]To avoid too many parenthesisses we assume in the following formulas that multiple vector products are evaluated from the right to the left, i.e., $\boldsymbol{a} \times \boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$

In this equation we explicitly indicated the position of a normal vector by an index, i.e., $\boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}}}$ is the normal vector at position $\boldsymbol{r}$ at the cavity wall. The surface current $\boldsymbol{J}_{ns}$ is the result of electromagnetic fields that have been scattered n times at the cavity wall. Therefore, (2.249) constitutes a scattering expansion of the surface current $\boldsymbol{J}_s$.

We insert (2.249) into (2.241) and obtain a scattering expansion for the magnetic vector potential,

$$
\begin{aligned}
\boldsymbol{A}(\boldsymbol{r}) = {}& \mu \int_{\Omega} G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}(\boldsymbol{r}')\, d^3 r' + \mu \int_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}_{1s}(\boldsymbol{r}')\, d^2 r' \\
&+ \mu \int_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}_{2s}(\boldsymbol{r}')\, d^2 r' + \mu \int_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}_{3s}(\boldsymbol{r}')\, d^2 r' \\
&+ \ldots \\
= {}& \mu \int_{\Omega} G_0(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}(\boldsymbol{r}')\, d^3 r' + 2 \int_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}_1) \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_1}} \times \boldsymbol{\nabla} \times \boldsymbol{A}^{\mathrm{inc}}(\boldsymbol{r}_1)\, d^2 r_1 \\
&+ 4 \iint_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}_1) \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_1}} \times (G_0'(\boldsymbol{r}_1, \boldsymbol{r}_2) \boldsymbol{e}_{\boldsymbol{r}_1, \boldsymbol{r}_2}) \times \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_2}} \times \boldsymbol{\nabla} \times \boldsymbol{A}^{\mathrm{inc}}(\boldsymbol{r}_2)\, d^2 r_1\, d^2 r_2 \\
&+ 8 \iiint_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}_1) \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_1}} \times (G_0'(\boldsymbol{r}_1, \boldsymbol{r}_2) \boldsymbol{e}_{\boldsymbol{r}_1, \boldsymbol{r}_2}) \times \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_2}} \times (G_0'(\boldsymbol{r}_2, \boldsymbol{r}_3) \boldsymbol{e}_{\boldsymbol{r}_2, \boldsymbol{r}_3}) \\
&\hspace{4cm} \times \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_3}} \times \boldsymbol{\nabla} \times \boldsymbol{A}^{\mathrm{inc}}(\boldsymbol{r}_3)\, d^2 r_1\, d^2 r_2\, d^2 r_3 \\
&+ \ldots
\end{aligned}
\tag{2.250}
$$

This result yields a ray representation in terms of a scattering expansion for the dyadic cavities' Green's function $\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')$,

$$
\begin{aligned}
\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') = {}& \overline{\boldsymbol{G}}_0(\boldsymbol{r}, \boldsymbol{r}') + 2 \int_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}_1) \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}}} \times \boldsymbol{\nabla} \times \overline{\boldsymbol{G}}_0(\boldsymbol{r}_1, \boldsymbol{r}')\, d^2 r_1 \\
&+ \frac{4}{\mu} \iint_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}_2) \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_2}} \times (G_0'(\boldsymbol{r}_2, \boldsymbol{r}_1) \boldsymbol{e}_{\boldsymbol{r}_2, \boldsymbol{r}_1}) \times \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_1}} \times \boldsymbol{\nabla} \times \overline{\boldsymbol{G}}_0(\boldsymbol{r}_1, \boldsymbol{r}')\, d^2 r_1\, d^2 r_2 \\
&+ \frac{8}{\mu} \iiint_{\Gamma} G_0(\boldsymbol{r}, \boldsymbol{r}_3) \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_3}} \times (G_0'(\boldsymbol{r}_3, \boldsymbol{r}_2) \boldsymbol{e}_{\boldsymbol{r}_3, \boldsymbol{r}_2}) \times \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_2}} \times (G_0'(\boldsymbol{r}_2, \boldsymbol{r}_1) \boldsymbol{e}_{\boldsymbol{r}_2, \boldsymbol{r}_1}) \\
&\hspace{4cm} \times \boldsymbol{e}_{\boldsymbol{n}_{\boldsymbol{r}_1}} \times \boldsymbol{\nabla} \times \overline{\boldsymbol{G}}_0(\boldsymbol{r}_1, \boldsymbol{r}')\, d^2 r_1\, d^2 r_2\, d^2 r_3 \\
&+ \ldots
\end{aligned}
\tag{2.251}
$$

with

$$
\overline{\boldsymbol{G}}_0(\boldsymbol{r}, \boldsymbol{r}') = G_0(\boldsymbol{r}, \boldsymbol{r}') \overline{\boldsymbol{I}}
\tag{2.252}
$$

the dyadic Green's function of free space.

Ray representations and scattering expansions are interesting in their own rights since they make precise the intuitive idea to construct a cavities' Green's function by the superposition of elementary scattering processes. Thus they are reminiscent of the Huygens' principle. It should be admitted, however, that an exact evaluation of the integrals that occur in (2.251) will be impossible in most cases. In physical contexts, the

scattering expansion has approximately been evaluated in high-frequency limits, yielding corrections to ray optics [6]. Clearly, these corrections will involve geometric factors that characterize the geometry of the scattering surface. This connection of electromagnetic properties and geometric properties has been used in an engineering context to show that for the description of current propagation along linear antenna and transmission line structures radiation effects become important if the wavelength is of the order or shorter than the radii of curvature of these structures [63, 65].

A considerable simplification occurs if the geometry of the cavity is such that the mirror principle can be applied. Then it is possible to express the scattering contributions in terms of image sources.

Example 1: We consider a rectangular cavity of dimensions l_x, l_y, and l_z and want to find a ray representation for its Green's function $\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')$. From the mode representation (2.171) we already know that among the nine components of $\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')$ only three components are non-vanishing,

$$\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') = G_{xx}^A(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{e}_x\boldsymbol{e}_x + G_{yy}^A(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{e}_y\boldsymbol{e}_y + G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{e}_z\boldsymbol{e}_z \,. \tag{2.253}$$

original currents

mirror plane
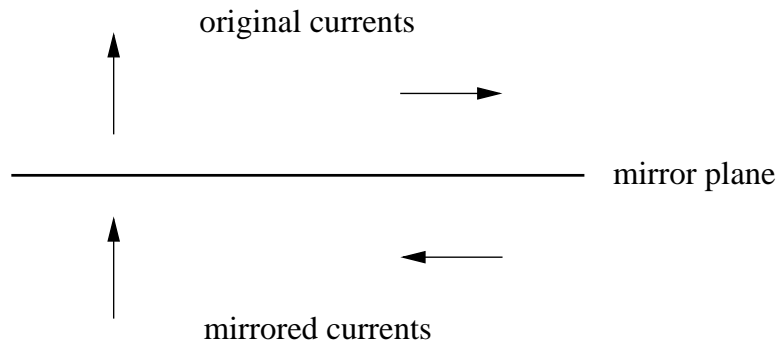
mirrored currents

Figure 2.10: Illustration of the mirror principle if applied to electric current sources. Electric currents that are perpendicular to the mirror plane do not reverse direction if mirrored, currents that are in parallel to the mirror plane do reverse direction if mirrored. The mirrored sources are imaginary sources that allow to replace the actual system (original sources plus mirror plane) by an equivalent system (original sources plus mirrored currents) where the mirror plane is removed. For the equivalent system the mirrored sources are chosen such that the electromagnetic field at the location of the (removed) mirror plane is the same as in the original system, i.e., the equivalent system respects the boundary conditions of the actual system, as is required by the uniqueness theorem. Textbooks discuss the mirror principle in the context of image theory, see, for example, [4, §7.4].
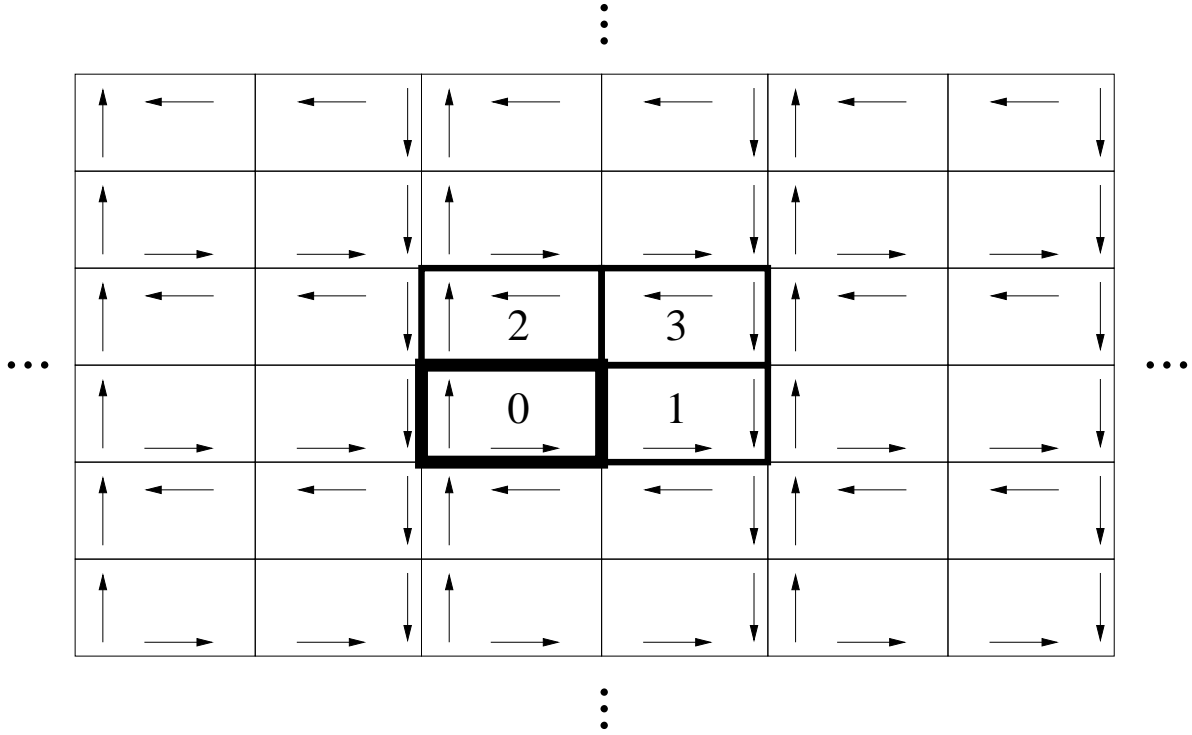
Figure 2.11: Illustration of the mirror principle if applied to a two-dimensional periodic structure. The original sources are located in the cell which is marked by "0" and printed in bold. They are mirrored three times to yield the mirrored sources in the cells "1", "2", and "3". The four numbered cells constitute a basic block. Any other mirrored source can be obtained from one of the sources of this basic block by an *even* number of mirror reflections, such that no more directional changes need to be taken into account. It follows that all image sources can be obtained from copies of the basic block. In the figure eight copies of the basic block are drawn. For a three-dimensional periodic structure one uses a basic block which consists of eight cells that form a rectangular parallelepiped and are numbered from "0" to "7". In the formulas (2.254) - (2.256) the summation with index i yields this basic block, where the coefficients A_i^{xx}, A_i^{yy}, and A_i^{zz} take into account the directional changes that are due to mirror reflections. Finally, the triple sum with indices m, n, and p generates the copies of the basic block.

As explained in Figures 2.10 and 2.11, application of the mirror principle yields

$$G_{xx}^A(\boldsymbol{r}, \boldsymbol{r}') = \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{xx} G_0\big(R_{i,mnp}(\boldsymbol{r}, \boldsymbol{r}')\big), \tag{2.254}$$

$$G_{yy}^A(\boldsymbol{r}, \boldsymbol{r}') = \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{yy} G_0\big(R_{i,mnp}(\boldsymbol{r}, \boldsymbol{r}')\big), \tag{2.255}$$

$$G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}') = \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} G_0\big(R_{i,mnp}(\boldsymbol{r}, \boldsymbol{r}')\big), \tag{2.256}$$

with

$$G_0\big(R_{i,mnp}(\boldsymbol{r},\boldsymbol{r}')\big) = \frac{e^{-jkR_{i,mnp}(\boldsymbol{r},\boldsymbol{r}')}}{4\pi R_{i,mnp}(\boldsymbol{r},\boldsymbol{r}')}\,. \tag{2.257}$$

The length $R_{i,mnp}(\boldsymbol{r},\boldsymbol{r}')$ represents the distances of the source and its mirror sources to the observation point $\boldsymbol{r} = (x,y,z)$. It is given by

$$R_{i,mnp}(\boldsymbol{r},\boldsymbol{r}') := \sqrt{(X_i + 2ml_x)^2 + (Y_i + 2nl_y)^2 + (Z_i + 2pl_z)^2}\,, \tag{2.258}$$

with

$$X_i := \begin{cases} x - x'\,, & i = 0,1,2,3 \\ x + x'\,, & i = 4,5,6,7 \end{cases}, \tag{2.259}$$

$$Y_i := \begin{cases} y - y'\,, & i = 0,1,4,5 \\ y + y'\,, & i = 2,3,6,7 \end{cases}, \tag{2.260}$$

$$Z_i := \begin{cases} z - z'\,, & i = 0,2,4,6 \\ z + z'\,, & i = 1,3,5,7 \end{cases}. \tag{2.261}$$

The coefficients A_i^{xx}, A_i^{yy}, and A_i^{zz} are defined by

$$A_i^{xx} := \begin{cases} +1\,, & i = 0,3,4,7 \\ -1\,, & i = 1,2,5,6 \end{cases}, \tag{2.262}$$

$$A_i^{yy} := \begin{cases} +1\,, & i = 0,2,5,7 \\ -1\,, & i = 1,3,4,6 \end{cases}, \tag{2.263}$$

$$A_i^{zz} := \begin{cases} +1\,, & i = 0,1,6,7 \\ -1\,, & i = 2,3,4,5 \end{cases}. \tag{2.264}$$

In the summation series (2.254) – (2.256) the term obtained for $m = n = p = i = 0$, respectively, corresponds to the Green's function of free space. The numerical properties of this ray representation will be studied and compared to those of the spectral representation (2.171) in Sec. 3.3.

2.4 Numerical methods and the method of moments

With the Green's function approach we have been able to formally solve the Maxwell equations, if rewritten as Helmholtz equations, in free space and within cavities. For the special case of rectangular cavities we have given explicit expressions for the vector potential and electric Green's function. The study of these canonical solutions is important in order to understand electromagnetic phenomena on a fundamental level. In practice, however, we have to go beyond the scope of canonical problems. Then the Green's function approach becomes intricate for two major reasons:

- In general, we will not be able to find an explicit expression for the Green's function of a given boundary value problem.

- Even if we explicitly know the Green's function we have not necessarily completed the solution of the given boundary value problem. The Green's function tells us the field that is generated by a *given* source, and this is important information. But we not always know the source since, in general, it is coupled to the field and interacts with the field. In this case we still have to solve some operator equation, typically an integral equation, which contains the Green's function. And the corresponding solution often cannot be found in closed form.

As a result, we face some operator equation which we cannot solve in closed form or which is too complicated to be solved by approximate analytical methods. It is at this point where numerical methods become a useful option.

A numerical method approximates the solution of a boundary value problem by a finite set of numerically calculated real or complex numbers[11]. These numbers are the coefficients F_n^i of an expansion with respect to some basis functions ϕ_n, as shown in (2.4). The choice of basis functions ϕ_n and the way of calculating the coefficients F_n^i largely depends on the specific numerical method that is chosen.

A very general type of a numerical method is provided by the method of moments. Due to its generality, its abbreviation "MoM" is often interpreted as "*M*other *o*f all *M*ethods". We will see in the following how the method of moments results from the approximate solution of a linear operator equation.

2.4.1 Derivation of the method of moments

We consider a linear operator $\mathcal{L} : H_1 \rightarrow H_2$ that acts between two Hilbert spaces H_1, H_2 and examine the linear operator equation

$$\mathcal{L}f = g \tag{2.265}$$

with an unknown function f and a known function g. First, we choose bases $\{\psi_n\}_{n=1}^\infty \in H_1$ and $\{w_n\}_{n=1}^\infty \in H_2$. Then we form the approximations

$$\tilde{f} = \sum_{k=1}^N \alpha_k \psi_k \,, \tag{2.266}$$

$$\tilde{g} = \sum_{j=1}^N \langle g, w_j \rangle w_j \,. \tag{2.267}$$

Since f is unknown the coefficients α_k are unknown as well. In (2.267) $\tilde{g}$ approaches g in the limit $N \longrightarrow \infty$. The functions ψ_k are commonly called *basis functions* or *expansion functions*, while the functions w_j are known as *weighting functions*.

---

[11]We might be lucky and find that such a finite set of numbers yields the *exact* solution, but this case is not common.

In order to determine the unknowns α_k we consider the approximate problem

$$\mathcal{L}\tilde{f} = \tilde{g}\,.$$

This equation can further be approximated by N algebraic equations for N unknowns α_k. To this end we first consider the approximation

$$\mathcal{L}\tilde{f} = \sum_{k=1}^{N} \alpha_k (\mathcal{L}\psi_k) \tag{2.268}$$

$$= \sum_{k=1}^{N} \sum_{j=1}^{\infty} \alpha_k \langle \mathcal{L}\psi_k, w_j \rangle w_j\,, \tag{2.269}$$

$$\approx \sum_{k=1}^{N} \sum_{j=1}^{N} \alpha_k \langle \mathcal{L}\psi_k, w_j \rangle w_j\,. \tag{2.270}$$

We set the last expression equal to the approximation (2.267). It follows

$$\sum_{k=1}^{N} \alpha_k \langle \mathcal{L}\psi_k, w_j \rangle = \langle g, w_j \rangle \tag{2.271}$$

for $j = 1 \ldots N$. This already is the desired linear system of equations for the unknowns α_k. We introduce the abbreviations

$$A_{jk} := \langle \mathcal{L}\psi_k, w_j \rangle\,, \tag{2.272}$$

$$\beta_j := \langle g, w_j \rangle\,, \tag{2.273}$$

and write (2.271) in the succinct form

$$[A][\alpha] = [\beta]\,. \tag{2.274}$$

If the inverse matrix $[A]^{-1}$ exists we find the solution

$$[\alpha] = [A]^{-1}[\beta]$$

which determines the approximate solution (2.266) of the original problem (2.265).

In the application of the method of moments the main difficulty is to choose the basis functions ψ_k and weighting functions w_j such that the approximations (2.266), (2.267) are physically meaningful and, additionally, make an evaluation of the matrix elements (2.272) feasible. It is not obvious how to find "good" choices of basis and weighting functions. In practice, the calculation of the matrix $[A]$ will be the most time consuming part of the method of moment algorithm.

2.4.2 General remarks on numerical methods for linear problems in electromagnetic field theory

A variety of numerical methods for solving electromagnetic boundary value problems has been developed during the last decades. Since at the beginning of the 1970s the use of digital computers became common the ratio between computer cost and computer power has constantly decreased. This makes it more and more attractive to use numerical methods.

In the development of numerical solution procedures we begin with a mathematical model which normally is expressed in terms of differential equations, integral equations, or variational expressions. This mathematical model should reflect the underlying physical model as well as possible. Then, in a second step, the mathematical model is discretized by an appropriate approximation of the solution domain such that the desired accuracy of the solution can be achieved. Finally, the solution algorithm needs to be implemented as a computer program which should use the computer resources in an efficient way. Usually, at the end of the solution algorithm there will be a matrix equation, as exemplified by (2.274), which reflects a discretized linear operator equation and needs to be inverted to yield the approximate, numerical solution.

There is a vast amount on literature on the subject of numerical methods in electromagnetic field analysis among we mention in particular [243, 11, 187, 209]. The various numerical methods can roughly be divided into *domain methods* and *boundary methods*. Among the domain methods are the *Finite Difference Method* and the *Finite Element Method* which are based on differential equations or variational expressions. Domain methods involve the discretization of the domain where the electromagnetic fields are defined and yield, as a result, approximate expressions for the electromagnetic fields. Boundary methods are based on integral equations or variational expressions and involve the discretization of the boundary of the electromagnetic field problem. Usually they yield approximate expressions for the electromagnetic sources. Examples are the *Boundary Element Method* and the *Method of Moments*, if understood as a method for solving integral equations.

Formally, the different numerical solution methods have a unified description in the framework of functional analysis. They all yield approximate solutions in the form (2.266) with a finite number of numerically calculated coefficients α_k that refer to a finite number of basis functions ψ_k. Different numerical solution methods consists of different choices of basis functions and different ways to compute the corresponding coefficients. As is evident from Sec. 2.1.1 (f), it is advantageous to choose basis functions which are orthonormal. For orthonormal basis functions it has been shown that the generalized Fourier coefficients (2.50) yield the best possible approximation to the exact solution which minimizes the error. It also follows from the projection theorem of Sec. 2.1.1 (g) that an approximate solution will be a projection of the exact solution onto a finite dimensional linear subspace of the Hilbert space considered. For that reason cer-

tain numerical solution procedures, as the Finite Element Method or integral equation methods, are often referred to as *projection methods* [145, 209]. The mathematical analysis of such methods exemplifies that the functional analytic framework is essential in order to understand and optimize the numerical methods of electromagnetic field theory.

Chapter 3

Antenna Theory in Resonating Systems

The previous two chapters dealt with the electromagnetic interaction on a general level. From the formulation of the Maxwell equations in terms of three-dimensional wave or Helmholtz equations together with the Green's function method it has been possible to represent the electromagnetic interaction by Green's functions.

Now we apply these theoretical concepts to formulate antenna theory in resonating systems. It has been mentioned in the introduction before Chapter 1 that a main motivation to introduce this subject is to consider in the framework of Electromagnetic Compatibility the coupling between EMI-sources and -victims in environments where resonances can be excited. In this context, an EMI-source represents a transmitting antenna and an EMI-victim represents a receiving antenna. The class of problems in Electromagnetic Compatibility that involve resonating environments are often referred to as *interior problems*.

To arrive at an antenna theory in resonating systems it is necessary to merge concepts of conventional antenna theory with those of microwave theory. This is illustrated in Fig. 3.1. In antenna theory the main task is to calculate antenna currents. *The usual approach to do this is to solve for a given electromagnetic excitation an integral equation for the unknown current* and we will shortly review this approach in Sec. 3.2. Once the antenna current is determined it is straightforward to obtain the related electromagnetic field by means of integration and the Green's function of free space, compare (2.138). On the other hand, in microwave theory the main task is to determine the electromagnetic field within a waveguide or cavity. *Here, the usual approach is to construct an appropriate Green's function.* Once the Green's function is constructed from the solution of a boundary value problem, which reflects the geometry of the waveguide or cavity, the principal structure of the electromagnetic field is known. It is given in terms of the eigenmodes of the resonating system. Which eigenmodes actually are excited depends on the electromagnetic excitation which is often provided by a fixed electric current that
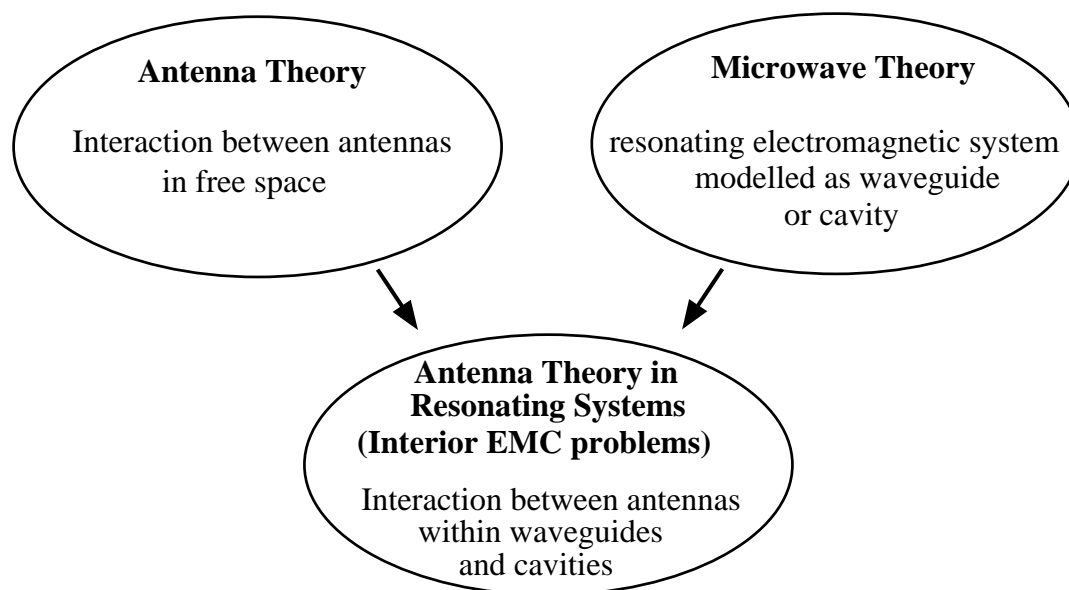
Figure 3.1: Conventional antenna theory and microwave theory merge into an antenna theory of resonating systems which is suitable to model interior problems of Electromagnetic Compatibility.

is assumed to be independent of the electromagnetic field that it generates. Therefore, antenna theory puts focus on the determination of *electric currents* while microwave theory puts focus on the determination of *electromagnetic fields*[1]

From these considerations it becomes clear what an antenna theory in resonating systems consists of: First, it consists of the determination of the Green's function of a resonating system, i.e., it consists of the characterization of the electromagnetic properties of the resonating system. Second, it consists of the determination of an antenna current within the resonating system. This involves the solution of an integral equation with the Green's function of the resonating system as kernel. Since the integral equation will be defined on the antenna surface it is at this point where the properties of the antenna enter.

It follows that the main task of antenna theory in resonating systems is to solve integral equations with Green's functions as kernels which incorporate the properties of electromagnetic resonances. The main difficulties that are encountered in the solution of this class of integral equations are of a numerical nature. The physical reason behind these numerical difficulties is implied by the discussion of Sec. 1.5.2 and given by the fact that the Green's function of a resonating system exhibits the two complementary singu-

---

[1]Of course, also electromagnetic fields are of great importance and interest in antenna theory. But it is straightforward to calculate from a known antenna current the associated electromagnetic field by means of an appropriate Green's function. In antenna theory the difficult part is to calculate the, a priori, unknown antenna current, and this is why focus needs to be put on this aspect.

larities of the electromagnetic field, namely the Coulomb singularity and electromagnetic resonances.

Antenna theory in free space involves a continuous electromagnetic spectrum and no discrete resonances occur. In this situation the Coulomb singularity is the only electromagnetic field singularities that occurs. In a resonating system the presence of resonances adds singular field effects and these additional effects will also be reflected by antenna characteristics. It is of primary interest to calculate these additional effects. Single resonances can have a dominating influence and drastically change the behavior of antenna configurations inside resonating systems if compared to free space.

"Antenna theory in resonating systems" is a notion which is not established in the current textbook literature. Some of the ideas and solution strategies that we will introduce in the following have been applied to rather specific electrical engineering problems. These include the characterization of cavity backed antennas [125, 59, 202], antenna arrays [5, 118], printed circuit boards [224, 78, 123], reverberation chambers [240, 88, 232, 233, 18, 242], and multilayered media [22, 129, 241]. There are also recent investigations of wire antennas [179, 180, 72] and transmission lines [207, 222, 221] within cavities. These works differ in the type of (simplifying) assumptions that are made to arrive at specific results. The results obtained seem to become increasingly important to the field of Electromagnetic Compatibility since they add insights to the mostly numerical and experimental studies of electromagnetic coupling through cavity apertures and inside cavities [19, 20, 107, 166, 49, 151, 203, 164, 126, 127].

3.1   Basic concepts of antenna theory

The definition and calculation of antenna characteristics require a few basic concepts of antenna theory that are summarized in the following subsections. We will introduce important measures of electromagnetic coupling, review the notion of reciprocity, give mathematical expressions that are suitable to calculate the self and mutual impedance of antennas, and finally relate antenna impedances to the spectral properties of the electromagnetic field within a cavity.

3.1.1   Measures of electromagnetic coupling

A transmitting antenna represents an electric current source $\boldsymbol{J}(\boldsymbol{r}')$ which generates an electromagnetic field $\boldsymbol{E}(\boldsymbol{r})$. This "current to field" coupling is expressed by the electric Green's function $\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')$ that has been introduced in (2.124),

$$\boldsymbol{E}(\boldsymbol{r}) = -j\omega\mu \int \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}(\boldsymbol{r}') \, d^3 r' \, . \tag{3.1}$$

The electric Green's function is a fundamental quantity since it represents the solution of the Maxwell equations with respect to prescribed boundary conditions. It is a local

function and does not reflect any specific antenna properties.

In practice, the field $\boldsymbol{E}(\boldsymbol{r})$ will interact with an antenna not locally at one point but along the extent of the antenna where it will induce a current $\boldsymbol{J}(\boldsymbol{r})$. A measure of this "field to current" coupling is the *reaction* which is defined by [82][§ 3]

$$\langle a, b \rangle_p := \int \boldsymbol{E}^a(\boldsymbol{r}) \cdot \boldsymbol{J}^b(\boldsymbol{r}) \, d^3r \tag{3.2}$$

$$= \langle \boldsymbol{E}^a, \boldsymbol{J}^b \rangle_p . \tag{3.3}$$

Here we do not consider artificial magnetic sources which can also be incorporated in the definition of reciprocity. In (3.2) the first entry "a" represents the electric field and the second entry "b" represents the electric current that reacts on the electric field. The index "p" indicates that the reaction is defined in terms of a *pseudo* inner product, compare Sec. 2.1.1 (d). Since the field $\boldsymbol{E}^a(\boldsymbol{r})$ will be generated by some source $\boldsymbol{J}^a(\boldsymbol{r}')$ it follows that the reaction is closely related to the Green's function $\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')$,

$$\langle a, b \rangle_p = -j\omega\mu \iint \left(\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}^a(\boldsymbol{r}') \right) \cdot \boldsymbol{J}^b(\boldsymbol{r}) \, d^3r' d^3r . \tag{3.4}$$

The reaction is a more practical measure of electromagnetic coupling than the mere electric Green's function. It can be thought of as the coupling between a measured field and the antenna or probe that is used for the measurement.

The reaction itself is not sufficient to characterize the mutual coupling between two antennas since it accounts for the properties of only one antenna. To characterize the mutual coupling between two antennas 1 and 2 it is customary to employ the notion of mutual impedance Z_{12}. The definition of mutual impedance involves two antennas that are excited by some electromagnetic source. Then the voltages V_1, V_2 and currents I_1, I_2 at the antenna input terminals are related to each other by a 2×2 impedance matrix [42]

$$Z = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} \tag{3.5}$$

according to

$$V_1 = Z_{11}I_1 + Z_{12}I_2 , \tag{3.6}$$

$$V_2 = Z_{21}I_1 + Z_{22}I_2 . \tag{3.7}$$

This network representation is illustrated in Fig. 3.2.

The impedance matrix concisely characterizes the antenna configuration which is determined from the antenna geometries, the antenna positions and orientations, and their electromagnetic environment. The impedances Z_{11} and Z_{22} are the self impedances of antenna 1 and antenna 2 in the presence of antenna 2 and antenna 1, respectively. The
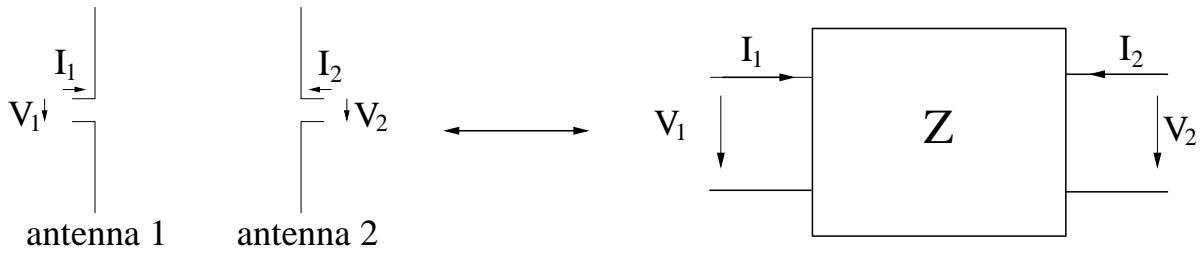
Figure 3.2: Two coupled antennas and their network equivalent

mutual impedances Z_{12} and Z_{21} specify the electromagnetic coupling between both antennas. Even though impedance is a very basic concept one should keep in mind that it is only defined for time harmonic fields.

As will be recalled in Sec. 3.1.3, the mutual impedance Z_{12} can be calculated in terms of the reaction according to [42, 149]

$$Z_{12} = -\frac{\langle b, a \rangle_p}{I_2^a I_1^b} \tag{3.8}$$

$$= -\frac{\int_{\text{antenna 2}} \boldsymbol{E}^b(\boldsymbol{r}) \cdot \boldsymbol{J}^a(\boldsymbol{r})\, d^3r}{I_2^a I_1^b}, \tag{3.9}$$

where the indices a and b refer to two different situations. Similar to the electric Green's function, the mutual impedance incorporates the electromagnetic properties of the antenna environment and, additionally, takes into account the properties of the transmitting and receiving antenna. To actually calculate Z_{12} from (3.9) is not trivial since the required electric field and electric currents have to be obtained as the solution of complete boundary value problems.

The notions "electric Green's function", "reaction", and "mutual impedance" are collected and concisely characterized in Fig. 3.3.
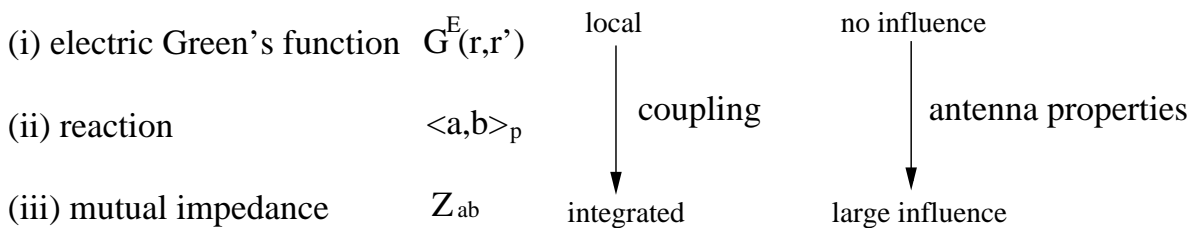
Figure 3.3: The electric Green's function is a mathematical ideal and not influenced by antenna properties. Reaction and mutual coupling are more realistic measures for electromagnetic coupling that take into account the properties of transmitting and/or receiving antennas.

3.1.2   Reciprocity

Reciprocity is a symmetry principle which represents the invariance of an interaction under the exchange of source and observation point. In electrical engineering contexts the concept of reaction is frequently used to exploit the associated invariance under the exchange of a source ("transmitting antenna") and a probe ("receiving antenna"). To formulate this definition in mathematical terms we consider a source $\boldsymbol{J}^a(\boldsymbol{r}')$ at a point $\boldsymbol{r}'$ which produces a field $\boldsymbol{E}^a(\boldsymbol{r})$ at a point $\boldsymbol{r}$ and a source $\boldsymbol{J}^b(\boldsymbol{r})$ at the same point $\boldsymbol{r}$ which produces a field $\boldsymbol{E}^b(\boldsymbol{r}')$ at the point $\boldsymbol{r}'$ where the source $\boldsymbol{J}^a(\boldsymbol{r}')$ is located. Then reciprocity is fulfilled if the reaction between $\boldsymbol{E}^a(\boldsymbol{r})$ and $\boldsymbol{J}^b(\boldsymbol{r})$ is the same as the reaction between $\boldsymbol{E}^b(\boldsymbol{r}')$ and $\boldsymbol{J}^a(\boldsymbol{r}')$,

$$\boldsymbol{E}^a(\boldsymbol{r}) \cdot \boldsymbol{J}^b(\boldsymbol{r}) = \boldsymbol{E}^b(\boldsymbol{r}') \cdot \boldsymbol{J}^a(\boldsymbol{r}') \,. \tag{3.10}$$

This statement of reciprocity is particularly useful to show the well-known equivalence of antenna pattern in transmission and reception [82][§ 3]. It should be noted that reciprocity also implies a number of additional fundamental relationships between receiving and transmitting antenna properties that are of practical interest [8].

Whether or not reciprocity is valid depends on the differential operator that determines the equation of motion of the fields and it depends on the boundary conditions that are imposed on the fields [146, §7]. To find in the electromagnetic case a criterion for reciprocity we follow a standard approach and examine two sets of electric currents $\boldsymbol{J}^a$ and $\boldsymbol{J}^b$ within a region Ω which produce electric and magnetic fields $\boldsymbol{E}^a$, $\boldsymbol{H}^a$ and $\boldsymbol{E}^b$, $\boldsymbol{H}^b$, respectively. Then, in the frequency domain, Maxwell equations

$$\boldsymbol{\nabla} \times \boldsymbol{E}^{a,b}(\boldsymbol{r}) + j\omega\mu \boldsymbol{H}^{a,b}(\boldsymbol{r}) = \boldsymbol{0} \,, \tag{3.11}$$

$$\boldsymbol{\nabla} \times \boldsymbol{H}^{a,b}(\boldsymbol{r}) - j\omega\varepsilon \boldsymbol{E}^{a,b}(\boldsymbol{r}) = \boldsymbol{J}^{a,b}(\boldsymbol{r}) \tag{3.12}$$

are valid. Within Faraday's induction law (3.11) the magnetic field strength $\boldsymbol{B}$ has already been replaced by the magnetic excitation $\boldsymbol{H}$ with the help of a linear constitutive relation of the form $\boldsymbol{B} = \mu\boldsymbol{H}$. Similarly, a linear constitutive relation of the form $\boldsymbol{D} = \varepsilon\boldsymbol{E}$ has been employed to write down the Ampére-Maxwell law (3.12).

Dot multiplication of (3.11) with $\boldsymbol{H}^{b,a}$ and of (3.12) with $\boldsymbol{E}^{b,a}$ yields, after some elementary algebraic manipulations and the application of (B.10), the single equation

$$\boldsymbol{\nabla} \cdot (\boldsymbol{E}^b(\boldsymbol{r}) \times \boldsymbol{H}^a(\boldsymbol{r}) - \boldsymbol{E}^a(\boldsymbol{r}) \times \boldsymbol{H}^b(\boldsymbol{r})) = \boldsymbol{E}^a(\boldsymbol{r}) \cdot \boldsymbol{J}^b(\boldsymbol{r}) - \boldsymbol{E}^b(\boldsymbol{r}) \cdot \boldsymbol{J}^a(\boldsymbol{r}) \,. \tag{3.13}$$

Integration of this local relation over a simply connected volume Ω with surface Γ that encloses all sources $\boldsymbol{J}^a$, $\boldsymbol{J}^b$ and subsequent application of Gauss's law lead to

$$\int_{\Gamma=\partial\Omega} \left(\boldsymbol{E}^b(\boldsymbol{r}) \times \boldsymbol{H}^a(\boldsymbol{r}) - \boldsymbol{E}^a(\boldsymbol{r}) \times \boldsymbol{H}^b(\boldsymbol{r})\right) \cdot d^2\boldsymbol{r} = \int_{\Omega} \left(\boldsymbol{E}^a(\boldsymbol{r}) \cdot \boldsymbol{J}^b(\boldsymbol{r}) - \boldsymbol{E}^b(\boldsymbol{r}) \cdot \boldsymbol{J}^a(\boldsymbol{r})\right) d^3r \,,$$

$$= \langle a, b\rangle_p - \langle b, a\rangle_p \,. \tag{3.14}$$

This relation is often called *reciprocity theorem* [42]. It is analogous to the relation (2.102) which yields conditions for the self-adjointness of a linear differential operator $\mathcal{L}_D$. The condition for reciprocity resembles (2.100) and is given by

$$\langle a, b \rangle_p = \langle b, a \rangle_p. \tag{3.15}$$

Obviously, there are two possibilities to verify if reciprocity is fulfilled. We can check if either the surface integral or the volume integral of (3.14) identically vanishes. In both cases we need to know the dynamical equations (3.11), (3.12) and the boundary conditions that are imposed on the fields [71].

We first consider the surface integral

$$I_\Gamma := \int_\Gamma \left(\boldsymbol{E}^b(\boldsymbol{r}) \times \boldsymbol{H}^a(\boldsymbol{r}) - \boldsymbol{E}^a(\boldsymbol{r}) \times \boldsymbol{H}^b(\boldsymbol{r}) \right) \cdot d^2\boldsymbol{r} \tag{3.16}$$

of (3.14). It vanishes both in free space and inside a cavity with lossy walls:

- The case of free space is the standard case. If the sources of the fields are located in free space the surface Γ can be chosen as the surface of a sphere with radius r that tends to infinity, $r \to \infty$. Then the appropriate surface boundary condition on the fields is given by the radiation condition. It ensures that the integrand $(\boldsymbol{E}^b \times \boldsymbol{H}^a - \boldsymbol{E}^a \times \boldsymbol{H}^b)$ falls off faster than $1/r^2$ [227]. It follows that the surface integral I_Γ vanishes and reciprocity is fulfilled.

- Inside a cavity with lossy walls we choose Γ as the interior of the cavity wall and assume that the Leontovich boundary condition (2.207) is valid,

$$\boldsymbol{e}_n \times \boldsymbol{E}(\boldsymbol{r}) = -\mathcal{Z}_s \boldsymbol{e}_n \times \left(\boldsymbol{e}_n \times \boldsymbol{H}(\boldsymbol{r}) \right). \tag{3.17}$$

  This yields

$$\left(\boldsymbol{E}^b(\boldsymbol{r}) \times \boldsymbol{H}^a(\boldsymbol{r}) \right) \cdot \boldsymbol{e}_n = \left(\boldsymbol{e}_n \times \boldsymbol{E}^b(\boldsymbol{r}) \right) \cdot \boldsymbol{H}^a(\boldsymbol{r}) \tag{3.18}$$

$$= -\mathcal{Z}_s \left[\left(\boldsymbol{e}_n \times (\boldsymbol{e}_n \times \boldsymbol{H}^b(\boldsymbol{r})) \right) \right] \cdot \boldsymbol{H}^a(\boldsymbol{r}) \tag{3.19}$$

$$= -\mathcal{Z}_s \left[\left(\boldsymbol{e}_n \cdot \boldsymbol{H}^a(\boldsymbol{r}) \right) \left(\boldsymbol{e}_n \cdot \boldsymbol{H}^b(\boldsymbol{r}) \right) + \boldsymbol{H}^a(\boldsymbol{r}) \cdot \boldsymbol{H}^b(\boldsymbol{r}) \right]. \tag{3.20}$$

  Clearly, the last expression is invariant under exchange of a and b. Due to $d^2\boldsymbol{r} = \boldsymbol{e}_n \, d^2 r$ it follows that the surface integral vanishes. Therefore, the Leontovich boundary condition leads to reciprocity as well. This includes the special case of perfectly conducting cavity walls.

Alternatively, we consider the volume integral

$$I_\Omega := \int_\Omega \left(\boldsymbol{E}^a(\boldsymbol{r}) \cdot \boldsymbol{J}^b(\boldsymbol{r}) - \boldsymbol{E}^b(\boldsymbol{r}) \cdot \boldsymbol{J}^a(\boldsymbol{r}) \right) d^3 r \tag{3.21}$$

$$= \langle a, b \rangle_p - \langle b, a \rangle_p \tag{3.22}$$

and employ the electric Green's function $\overline{\boldsymbol{G}}(\boldsymbol{r}, \boldsymbol{r}')$ to relate the electric fields to the electric currents,

$$\boldsymbol{E}^{a,b}(\boldsymbol{r}) = \int_\Omega \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}^{a,b}(\boldsymbol{r}') \, d^3r' \,. \tag{3.23}$$

It follows

$$I_\Omega = \iint_\Omega \left(\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}^a(\boldsymbol{r}') \right) \cdot \boldsymbol{J}^b(\boldsymbol{r}) - \left(\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}^b(\boldsymbol{r}') \right) \cdot \boldsymbol{J}^a(\boldsymbol{r}) \, d^3r \, d^3r' \,. \tag{3.24}$$

Therefore, the volume integral vanishes and reciprocity, $\langle a, b \rangle_p = \langle b, a \rangle_p$, is fulfilled if the condition

$$\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}') = \left(\overline{\boldsymbol{G}}^E \right)^T (\boldsymbol{r}', \boldsymbol{r}) \tag{3.25}$$

holds. Here, $\left(\overline{\boldsymbol{G}}^E \right)^T$ denotes the transpose of the dyad $\overline{\boldsymbol{G}}^E$. It is easy to see that the Green's functions (2.171), (2.172), and (2.197) fulfill the condition (3.25) since they are both symmetric and invariant under exchange of $\boldsymbol{r}$ and $\boldsymbol{r}'$.

3.1.3 Integral expressions for self and mutual impedances of antenna elements

We return to the impedance matrix $[Z]$ that has been introduced in (3.5) to characterize the coupling between two antennas by a two-port network. It is evident from (3.6) and (3.7) that the matrix elements Z_{ij} can be calculated as ratios between open-circuit voltages and input currents. Explicitly, we have

$$Z_{11} = \left. \frac{V_1}{I_1} \right|_{I_2=0}, \qquad\qquad Z_{22} = \left. \frac{V_2}{I_2} \right|_{I_1=0}, \tag{3.26}$$

and

$$Z_{12} = \left. \frac{V_1}{I_2} \right|_{I_1=0}, \qquad\qquad Z_{21} = \left. \frac{V_2}{I_1} \right|_{I_2=0}. \tag{3.27}$$

It is a standard proof of network theory to show that reciprocity implies $Z_{12} = Z_{21}$ [25, §13].

The matrix components Z_{11} and Z_{22} are the self impedances of antenna 1 and 2 in the presence of an open-circuited antenna 2 and 1, respectively. Hence, to calculate Z_{11} or Z_{22} it is necessary to take into account the coupling between an active and an open-circuited antenna. In free space this coupling can often be neglected. However, it can be strong if the antennas are located within a resonating environment. To explicitly calculate Z_{11}, for example, one excites antenna 1 by an input voltage V_1 and then calculates the current I_1 from coupled integral equations.

To calculate the mutual impedance $Z_{12} = Z_{21}$ we may focus on the left equation of (3.27). Since in this expression it is assumed that antenna 1 is open-circuited it is required to excite antenna 2 by a given source and to calculate with this excitation the

current I_2 and the open-circuit voltage V_1. Again, the current I_2 can be obtained as the solution of an integral equation. But the calculation of V_1 is problematic since this open-circuit voltage is given by the displacement current of an infinitesimally small gap which represents the antenna input terminal. An adequate modeling of this gap and a calculation of the corresponding displacement current is, in practice, too complicated. In order to circumvent this difficulty it is possible to employ reciprocity and express V_1 in terms of quantities that are easier to calculate.

To see how reciprocity is applied to this problem we consider Fig. 3.4 and denote by "a-situation" and "b-situation" two different antenna configurations. The a-situation corresponds to the original configuration that is used to calculate Z_{12}, i.e., the configuration consists of a transmitting antenna 2 which induces an electric field $\boldsymbol{E}^a$ at an open circuited antenna 1. The b-situation consists of a *single* transmitting antenna 1 which is identical to antenna 1 of the a-situation. Antenna 1 generates an electric field $\boldsymbol{E}^b$ *at the position* of antenna 2 that is present in the a-situation. If we invoke reciprocity of the a- and b-situation we have $\langle a, b \rangle_p = \langle b, a \rangle_p$, i.e.,

$$\int_{\text{antenna 1}} \boldsymbol{E}^a(\boldsymbol{r}) \cdot \boldsymbol{J}^b(\boldsymbol{r}) \, d^3r = \int_{\text{antenna 2}} \boldsymbol{E}^b(\boldsymbol{r}) \cdot \boldsymbol{J}^a(\boldsymbol{r}) \, d^3r \,. \tag{3.28}$$

If we assume that antenna 1 is perfectly conducting the tangential electric field $\boldsymbol{E}^a$ at the antenna surface will vanish except at the gap that constitutes the antenna input terminal. Then we find

$$\int_{\text{antenna 1}} \boldsymbol{E}^a(\boldsymbol{r}) \cdot \boldsymbol{J}^b(\boldsymbol{r}) \, d^3r = -V_1^a I_1^b \tag{3.29}$$

with V_1^a the open-circuit voltage of antenna 1 in the a-situation and I_1^b the input current of antenna 1 in the b-situation. The minus sign in (3.29) results from the convention that at a network port the voltage and the current point in different directions, compare Fig. 3.2. It follows that the open-circuited voltage V_1^a can be calculated from the expression

$$V_1^a = -\frac{\int_{\text{antenna 2}} \boldsymbol{E}^b(\boldsymbol{r}) \cdot \boldsymbol{J}^a(\boldsymbol{r}) \, d^3r}{I_1^b} \,. \tag{3.30}$$

With this result we obtain from (3.27) the expression for the mutual impedance Z_{12} that already has been quoted in (3.9),

$$Z_{12} = -\frac{\int_{\text{antenna 2}} \boldsymbol{E}^b(\boldsymbol{r}) \cdot \boldsymbol{J}^a(\boldsymbol{r}) \, d^3r}{I_2^a I_1^b}$$

$$= -\frac{\langle \boldsymbol{E}^b, \boldsymbol{J}^a \rangle_p}{I_2^a I_1^b} \,. \tag{3.31}$$

We conclude that in order to calculate the mutual impedance between two antennas we have to determine, for a given excitation, the current distributions $\boldsymbol{J}^a$ and $\boldsymbol{J}^b$. Then
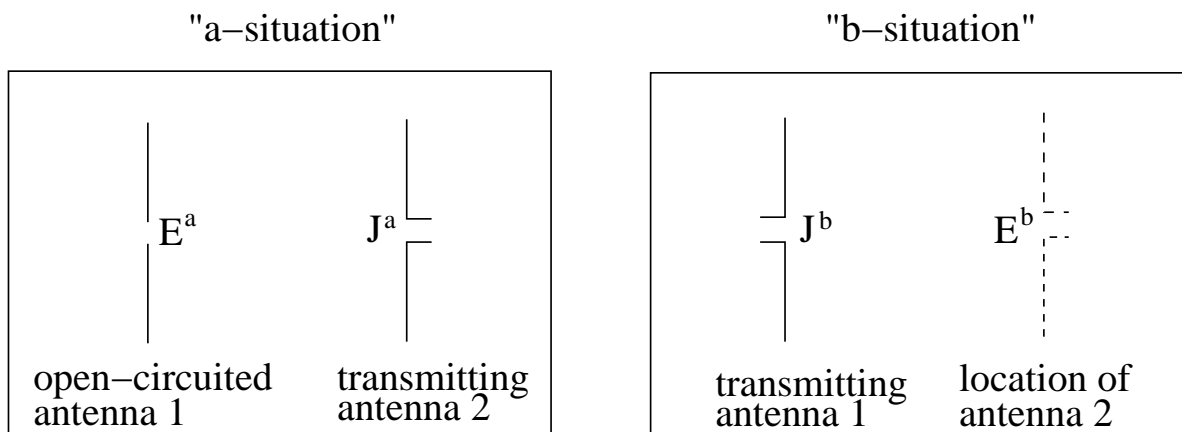
Figure 3.4: If reciprocity is applied to the a- and b-situation the calculation of the reaction at the input terminal of antenna 1 can be expressed in terms of the reaction along antenna 2. This facilitates to calculate the open-circuit voltage in the a-situation.

the electric field $\boldsymbol{E}^b$ that enters (3.31) is obtained from the current distribution $\boldsymbol{J}^b$ via the electric dyadic Green's function $\overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')$ and the relation (3.1).

A formula which is analogous to (3.31) can be derived for the self impedance Z_{self} of a *single* antenna. There are at least two approaches to arrive at such a formula. The first is to apply (3.31) to the mutual impedance between two identical antennas which are brought together to form a single antenna. The second approach is to apply reciprocity to an "a-situation" which consists of a single transmitting antenna and to a "b-situation" which consists of an auxiliary current distribution. In both cases one obtains the common result [97, § 14]

$$Z_{\text{self}} = -\frac{\int_{\text{antenna}} \boldsymbol{E}(\boldsymbol{r}) \cdot \boldsymbol{J}(\boldsymbol{r}) \, d^3r}{I^2} \tag{3.32}$$

$$= -\frac{\langle \boldsymbol{E}, \boldsymbol{J} \rangle_p}{I^2} \, . \tag{3.33}$$

Here, the current distribution $\boldsymbol{J}$, electric field $\boldsymbol{E}$, and input current I refer to a fixed excitation which typically is taken as a unit voltage source.

3.1.4   Antenna impedances and spectral properties

The impedance formulas (3.31) and (3.33) connect the fundamental electromagnetic quantities $\boldsymbol{E}$ and $\boldsymbol{J}$ to the lumped network quantities Z_{12} and Z_{self}, respectively. A specific electromagnetic boundary value problem which involves antenna elements will be characterized by a certain electromagnetic spectrum which, in turn, will be reflected by the impedances of the antenna elements. From the material that has been collected in Chapter 2 it is straightforward to relate antenna impedances to the spectral properties

of the electromagnetic field.

(a) Expansion of antenna impedances with respect to the eigenfunctions of the Helmholtz equation for the magnetic vector potential

We consider, along the lines of Sections 2.2.3 and 2.3.1 (a), the Helmholtz equation of the magnetic vector potential $\boldsymbol{A}$ in the Lorenz gauge,

$$\Delta \boldsymbol{A} + k^2 \boldsymbol{A} = -\mu \boldsymbol{J} \,. \tag{3.34}$$

The magnetic vector potential has an eigenfunction expansion of the form[2]

$$\boldsymbol{A} = \sum_n \alpha_n^A \boldsymbol{f}_n^A \tag{3.35}$$

where the eigenfunctions obey

$$-\Delta \boldsymbol{f}_n^A = k_n^2 \boldsymbol{f}_n^A \,. \tag{3.36}$$

With an expansion for the electric current,

$$\boldsymbol{J} = \sum_n \langle \boldsymbol{J}, \boldsymbol{f}_n^A \rangle \boldsymbol{f}_n^A \,, \tag{3.37}$$

we write the Helmholtz equation (3.34) as

$$\sum_n \alpha_n^A (-k_n^2 + k^2) \boldsymbol{f}_n^A = -\mu \sum_n \langle \boldsymbol{J}, \boldsymbol{f}_n^A \rangle \boldsymbol{f}_n^A \tag{3.38}$$

and find for $k_n^2 \neq k^2$ the coefficients

$$\alpha_n^A = \mu \frac{\langle \boldsymbol{J}, \boldsymbol{f}_n^A \rangle}{k^2 - k_n^2} \,. \tag{3.39}$$

Therefore, the magnetic vector potential can be written in terms of the eigenfunction expansion

$$\boldsymbol{A} = \mu \sum_n \frac{\langle \boldsymbol{J}, \boldsymbol{f}_n^A \rangle}{k^2 - k_n^2} \boldsymbol{f}_n^A \,. \tag{3.40}$$

This also yields an eigenfunction expansion for the electric field: Generally, we have

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - j\omega \boldsymbol{A} \tag{3.41}$$

---

[2]It should be noted that the summation over the index n represents a condensed notation. In Example 1 of Sec. 2.3.1 the eigenfunction expansion of the magnetic vector potential within a rectangular cavity was discussed. It became clear that the index n turns to a triple index mnp and that an eigenfunction $\boldsymbol{f}_n$ comprises three independent eigenfunctions $f_n^x \boldsymbol{e}_x$, $f_n^y \boldsymbol{e}_y$, and $f_n^z \boldsymbol{e}_z$. Hence, in (3.35) each summand is of the form $\alpha_{mnp}^{x_i} f_{mnp}^{x_i} \boldsymbol{e}_{x_i}$ with $x_i \in \{x, y, z\}$.

and in the Lorenz gauge, $\boldsymbol{\nabla} \cdot \boldsymbol{A} + j\omega\phi/c^2 = 0$, this relation turns to

$$\boldsymbol{E} = -\frac{j\omega}{k^2}\left(\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{A}) + k^2\boldsymbol{A}\right). \tag{3.42}$$

From (3.40) we can also obtain an eigenfunction expansion of the electric field,

$$\boldsymbol{E} = -\frac{j\omega\mu}{k^2}\sum_n \frac{\langle \boldsymbol{J}, \boldsymbol{f}_n^A\rangle}{k^2 - k_n^2}\left(\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{f}_n^A) + k^2\boldsymbol{f}_n^A\right). \tag{3.43}$$

The electric field enters the impedance formulas (3.31) and (3.33). If we insert (3.43) in (3.31), for example, we find

$$Z_{12} = \frac{j\omega\mu}{I_2^a I_1^b k^2}\sum_n \frac{\langle \boldsymbol{J}^b, \boldsymbol{f}_n^A\rangle}{k^2 - k_n^2}\langle \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{f}_n^A) + k^2\boldsymbol{f}_n^A, \boldsymbol{J}^a\rangle_p \tag{3.44}$$

$$= \frac{j\omega\mu}{I_2^a I_1^b k^2}\sum_n \frac{\langle \boldsymbol{J}^b, \boldsymbol{f}_n^A\rangle\langle \boldsymbol{J}^a, \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{f}_n^A) + k^2\boldsymbol{f}_n^A\rangle}{k^2 - k_n^2}. \tag{3.45}$$

In the last line we used the property (2.41) of the pseudo inner product and the fact that the eigenfunctions $\boldsymbol{f}_n^A$, as eigenfunctions of a linear, self-adjoint operator, are real, compare Sec. 2.1.4. For the self impedance we find a similar expression,

$$Z_{\text{self}} = \frac{j\omega\mu}{I^2 k^2}\sum_n \frac{\langle \boldsymbol{J}, \boldsymbol{f}_n^A\rangle\langle \boldsymbol{J}, \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{f}_n^A) + k^2\boldsymbol{f}_n^A\rangle}{k^2 - k_n^2}. \tag{3.46}$$

From (3.45) and (3.46) it appears that the mutual and self impedance encounter a pole whenever the wavenumber k approaches an eigenvalue k_n. We will discuss this feature after the next subsection.

(b) Expansion of antenna impedances with respect to the eigenfunctions of the Helmholtz equation for the electric field

We now consider the Helmholtz equation of the electric field $\boldsymbol{E}$,

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{E} - k^2\boldsymbol{E} = -j\omega\mu\boldsymbol{J}. \tag{3.47}$$

From Sec. 2.3.1 (b) it follows that the electric field has an eigenfunction expansion of the form

$$\boldsymbol{E} = \boldsymbol{E}_\parallel + \boldsymbol{E}_\perp \tag{3.48}$$

$$= \sum_n \alpha_n^{E\parallel}\boldsymbol{L}_n + \sum_n \alpha_n^{E\perp}\boldsymbol{F}_n \tag{3.49}$$

where the longitudinal eigenfunctions $\boldsymbol{L}_n$ and the transverse eigenfunctions $\boldsymbol{F}_n$ fulfill

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{L}_n = \boldsymbol{0}, \tag{3.50}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{F}_n = k_n^2\boldsymbol{F}_n. \tag{3.51}$$

With an expansion for the electric current we write the Helmholtz equation (3.47) as

$$-\sum_n \alpha_n^{E\parallel} k^2 \boldsymbol{L}_n + \sum_n \alpha_n^{E\perp}(k_n^2 - k^2)\boldsymbol{F}_n = -j\omega\mu \sum_n \langle \boldsymbol{J}, \boldsymbol{L}_n\rangle \boldsymbol{L}_n - j\omega\mu \sum_n \langle \boldsymbol{J}, \boldsymbol{F}_n\rangle \boldsymbol{F}_n \,.$$
(3.52)

For $k_n^2 \neq k^2$ this yields the coefficients

$$\alpha_n^{E\parallel} = \frac{j\omega\mu}{k^2}\langle \boldsymbol{J}, \boldsymbol{L}_n\rangle \,,$$
(3.53)

$$\alpha_n^{E\perp} = j\omega\mu \frac{\langle \boldsymbol{J}, \boldsymbol{F}_n\rangle}{k^2 - k_n^2} \,.$$
(3.54)

Therefore, the electric field can be written in terms of the eigenfunction expansion

$$\boldsymbol{E} = \frac{j\omega\mu}{k^2}\sum_n \langle \boldsymbol{J}, \boldsymbol{L}_n\rangle \boldsymbol{L}_n + j\omega\mu \sum_n \frac{\langle \boldsymbol{J}, \boldsymbol{F}_n\rangle}{k^2 - k_n^2}\boldsymbol{F}_n \,.$$
(3.55)

If we insert (3.55) in (3.31) we find for the mutual impedance the expansion

$$Z_{12} = -\frac{j\omega\mu}{k^2 I_2^a I_1^b}\sum_n \langle \boldsymbol{J}^b, \boldsymbol{L}_n\rangle\langle \boldsymbol{J}^a, \boldsymbol{L}_n\rangle - \frac{j\omega\mu}{I_2^a I_1^b}\sum_n \frac{\langle \boldsymbol{J}^b, \boldsymbol{F}_n\rangle\langle \boldsymbol{J}^a, \boldsymbol{F}_n\rangle}{k^2 - k_n^2} \,,$$
(3.56)

$$=: Z_{12\parallel} + Z_{12\perp} \,.$$
(3.57)

Here we denote by $Z_{12\parallel}$ and $Z_{12\perp}$, respectively, the longitudinal and transverse contributions to the mutual impedance. Similarly, we find for the self impedance (3.33)

$$Z_{\text{self}} = -\frac{j\omega\mu}{k^2 I^2}\sum_n \langle \boldsymbol{J}, \boldsymbol{L}_n\rangle^2 - \frac{j\omega\mu}{I^2}\sum_n \frac{\langle \boldsymbol{J}, \boldsymbol{F}_n\rangle^2}{k^2 - k_n^2}$$
(3.58)

$$=: Z_{\text{self}\parallel} + Z_{\text{self}\perp} \,.$$
(3.59)

It is recognized again that the mutual and self impedance encounter a pole whenever the wavenumber k approaches an eigenvalue k_n.

(c) Network representation of electromagnetic cavities

The expressions (3.46) and (3.58) for the self impedance of an antenna element have been obtained by different eigenfunction expansions but have a similar structure. This structure finds a simple interpretation in terms of equivalent circuit elements. To exhibit this relation we consider a parallel RLC-circuit, as shown in Fig. 3.5. Its impedance Z_{circuit} is given by

$$\frac{1}{Z_{\text{circuit}}} = \frac{1}{R_n} + \frac{1}{j\omega L_n} + j\omega C_n$$
(3.60)

or

$$Z_{\text{circuit}} = -\frac{j\omega\frac{1}{C_n}}{\omega^2 - \frac{1}{L_n C_n}(1 + j\frac{\omega L_n}{R_n})} \,.$$
(3.61)
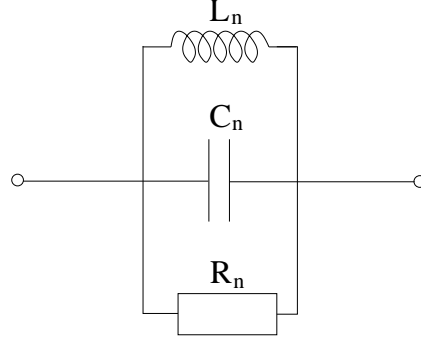
Figure 3.5: A parallel *RLC*-circuit. Its impedance resembles the contribution of an electromagnetic resonance to the self impedance of an antenna element.

We compare this expression to the dynamical, transverse part $Z_{\text{self}\perp}$ of the self impedance of an antenna element, multiply both the nominator and denominator of (3.58) by $1/\varepsilon\mu$, and allow for ohmic losses. This yields (3.58) in the form

$$Z_{\text{self}\perp} = \sum_n -\frac{j\omega\frac{\langle \boldsymbol{J},\boldsymbol{F}_n\rangle^2}{\varepsilon I^2}}{\omega^2 - \omega_n^2(1 + j\frac{1}{2Q})^2} \tag{3.62}$$

$$\approx \sum_n -\frac{j\omega\frac{\langle \boldsymbol{J},\boldsymbol{F}_n\rangle^2}{\varepsilon I^2}}{\omega^2 - \omega_n^2(1 + j\frac{1}{Q})} \tag{3.63}$$

In view of (3.61) we formally have the identifications

$$C_n = \frac{\varepsilon I^2}{\langle \boldsymbol{J},\boldsymbol{F}_n\rangle^2}, \tag{3.64}$$

$$L_n = \frac{\langle \boldsymbol{J},\boldsymbol{F}_n\rangle^2}{\varepsilon(I\omega_n)^2}, \tag{3.65}$$

$$R_n = \frac{Q\langle \boldsymbol{J},\boldsymbol{F}_n\rangle^2}{\varepsilon I^2\omega_n}. \tag{3.66}$$

In these equations the current density $\boldsymbol{J}$ is, a priori, not known but needs to be obtained by the complete solution of the actual antenna problem.

Suppose we specialize an antenna element to a small probe where, up to an unknown amplitude, the shape of the current distribution is approximated by a known function. A practical choice for such a function is given by a sinusoidal current distribution. Then in the expansions (3.58) and (3.63) the unknown quadratic current amplitudes of the nominator and denominator will cancel and the self impedance turns to a quantity which is characterized by cavity properties alone. It then follows that the expression (3.63) constitutes the equivalent circuit for a probe-fed cavity, compare [181, §11] and [82, §8], which is given by an infinite sum of parallel *RLC*-circuit impedances. Therefore, an empty cavity has a representation by an equivalent circuit which consists of an infinite

number of RLC-circuits in series, with each RLC-circuit representing one cavity mode[3]. If an extended antenna is brought into the cavity the single RLC-circuits will not be independent of each other since the lumped elements (3.64)–(3.66) depend on the antenna current $\boldsymbol{J}$ and this current involves *all* modes of the cavity[4].

The formal identification of an antenna impedance of the form (3.63) with circuit impedances such as (3.61) does not facilitate the actual calculation of antenna impedances. However, since RLC-circuits are well-understood this identification is helpful to interpret the qualitative features of calculated antenna impedances. For example, close to a resonance $\omega \approx \omega_m$ the expression for the self impedance (3.58) can be written in the form

$$Z_{\mathrm{self}} \approx Z_{\mathrm{self}\|} - \frac{j\omega \frac{\langle \boldsymbol{J}, \boldsymbol{f}_m^E \rangle^2}{\varepsilon I^2}}{\omega^2 - \omega_m^2 \left(1 + j\frac{1}{Q}\right)} + \sum_{n \neq m} -\frac{j\omega \frac{\langle \boldsymbol{J}, \boldsymbol{f}_n^E \rangle^2}{\varepsilon I^2}}{\omega^2 - \omega_n^2 \left(1 + j\frac{1}{Q}\right)} \tag{3.67}$$

where the second term will dominate the variation of Z_{self} with respect to ω. Therefore, close to a resonance the characteristics of Z_{self} will be those of a standard resonance curve, as shown in Fig. 3.6, which is well-known from usual circuit theory [25]. The same will be true for the mutual impedance Z_{12}, as is evident from the expansions (3.56) and (3.58) which have an identical structure.

3.2 Integral equations for the electric current on linear antennas

The problem to calculate the electric current on an antenna is a special case of a scattering problem. An antenna current is generated by primary sources that produce incident electromagnetic fields $\boldsymbol{E}^{\mathrm{inc}}$, $\boldsymbol{H}^{\mathrm{inc}}$. Then the total fields $\boldsymbol{E}$, $\boldsymbol{H}$ in the presence of the antenna, i.e., in the presence of the scatterer, are a superposition of the incident fields and scattered fields $\boldsymbol{E}^{\mathrm{sca}}$, $\boldsymbol{H}^{\mathrm{sca}}$,

$$\boldsymbol{E} = \boldsymbol{E}^{\mathrm{inc}} + \boldsymbol{E}^{\mathrm{sca}}, \tag{3.68}$$

$$\boldsymbol{H} = \boldsymbol{H}^{\mathrm{inc}} + \boldsymbol{H}^{\mathrm{sca}}. \tag{3.69}$$

The general strategy to find from this decomposition an equation for the unknown antenna current consists of three steps which require the surface equivalence principle, the source-field relationships that follow from the solution of the Maxwell equations, and the boundary conditions for the total electromagnetic fields [174, §1]:

---

[3]This statement does not take into account the longitudinal eigenfunctions $\boldsymbol{L}_n$ which can be supported if the cavity is not simply connected

[4]This is reminiscent of the fundamental discussion of the transverse part of the Maxwell equations of Sec. 1.5.1 (d), compare in particular the remarks after (1.231).
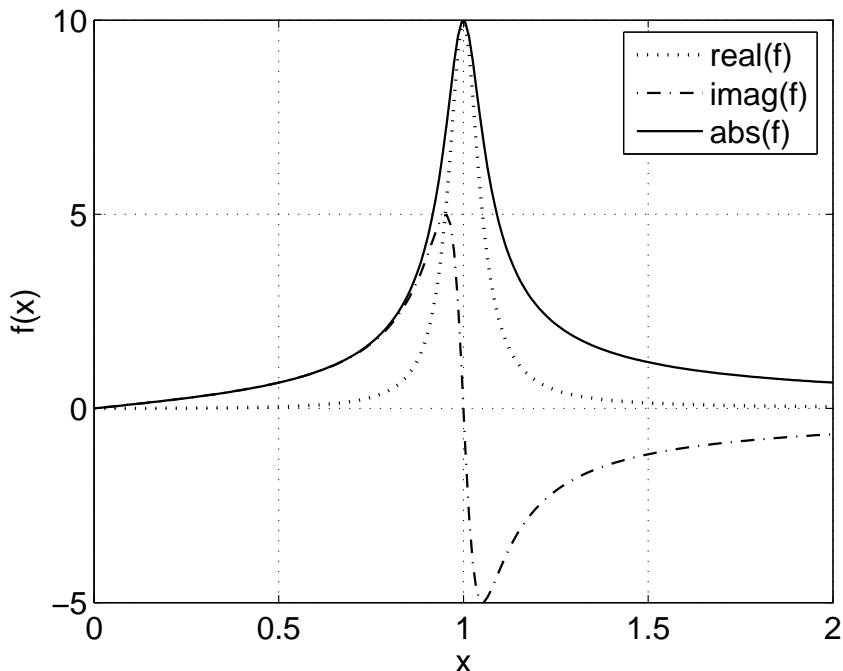
Figure 3.6: Plot of the function $f(x) = -\frac{jx}{x^2-(1+jx/10)}$ which resembles the circuit impedance (3.61). Shown are the absolute value, the real part, and the imaginary part of $f(x)$. At resonance, both the absolute value and the real part attain a maximum which is finite as long as losses are present. The imaginary part changes sign at resonance where the dominant behavior of the resonant circuit shifts between inductive and capacitive.

1. By virtue of the surface equivalence principle the antenna is replaced by equivalent electromagnetic sources which, a priori, are unknown. If the antenna is assumed to be perfectly conducting the equivalent electromagnetic sources are represented by an electric antenna current. In general, the equivalent electromagnetic sources will be determined from the incident electromagnetic fields.

2. The equivalent electromagnetic sources generate a scattered electromagnetic field according to the source-field relationships that express a field by the integral over a source, weighted with the appropriate Green's function. This allows to replace the unknown scattered electromagnetic field, which usually is defined within an entire volume, by the unknown equivalent sources, which usually are defined on a boundary surface.

3. On the boundary surface, where the unknown equivalent sources are defined, the boundary conditions (3.68), (3.69) for the total fields must be enforced. This relates on the boundary surface the known incident electromagnetic field to integrals over the unknown equivalent sources

These three steps comprise, in short, how to construct field integral equations for unknown electromagnetic sources that are induced by primary electromagnetic fields.

In the following we will focus on linear antennas. These constitute the classical antenna prototype and generalize the concept of an idealized, mathematical electric dipole to an actual engineering device. In the development of antenna theory linear antennas have played a dominant role because many of their properties can be modeled by analytic methods. However, even in the simplest realistic cases simplifying approximations have to be made in order to arrive at analytic results [104, 192, 29, 106].

3.2.1   Pocklington's equation

It is remarkable that already ten years after the discovery of electromagnetic radiation by Hertz [87] in 1887 an integral equation for the current distribution along cylindrical wire dipole antennas was published by Pocklington [175] in 1897. Pocklington's equation constitutes an electric field integral equation that is adapted to cylindrical, thin-wire antennas. To formulate this integral equation we follow the three-step procedure of the last paragraph and first introduce a surface current $\boldsymbol{J}_s$ which is related to a scattered electric field $\boldsymbol{E}^{\mathrm{sca}}$ via the electric Green's function $\overline{\boldsymbol{G}}^E$,

$$\boldsymbol{E}^{\mathrm{sca}}(\boldsymbol{r}) = -j\omega\mu \int_\Gamma \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')\, \boldsymbol{J}_s(\boldsymbol{r}')\, d^2 r' \,. \tag{3.70}$$

Here the antenna surface is denoted by Γ. The boundary condition for the total electric field $\boldsymbol{E}$ on a perfectly conducting surface is $\boldsymbol{e}_n \times \boldsymbol{E} = \boldsymbol{0}$ or, alternatively, $E_t = 0$, with $E_t = \boldsymbol{E} \cdot \boldsymbol{e}_t$ the projection of $\boldsymbol{E}$ on a given tangential vector $\boldsymbol{e}_t$. With this boundary condition and (3.68) we obtain an electric field integral equation,

$$j\omega\mu \left[\int_\Gamma \overline{\boldsymbol{G}}^E(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{J}_s(\boldsymbol{r}')\, d^2 r' \right] \cdot \boldsymbol{e}_t(\boldsymbol{r}) = E_t^{\mathrm{inc}}(\boldsymbol{r}) \,. \tag{3.71}$$

This equation simplifies if the antenna geometry is that of a thin cylindrical wire. Then a thin-wire approximation can be performed where azimuthal currents are neglected and the surface current $\boldsymbol{J}_s$ turns to a filamentary current I that flows along the cylinder axis [16, 235]. If furthermore the wire is assumed to be straight and, in Cartesian coordinates, directed along the z-axis we obtain from (3.71)

$$j\omega\mu \int_{-L/2}^{L/2} G_{zz}^E(z, z')I(z')\, dz' = E_z^{\mathrm{inc}}(z) \tag{3.72}$$

with L the length of the antenna. We finally use (2.145) to replace the zz-component of the electric dyadic Green's function by the zz-component of the dyadic Green's function for the magnetic vector potential in the Lorenz gauge. This yields Pocklington's equation in the form

$$-\frac{1}{j\omega\varepsilon} \int_{-L/2}^{L/2} \left(\frac{\partial^2}{\partial z^2} + k^2 \right) G_{zz}^A(z, z')I(z')\, dz' = E_z^{\mathrm{inc}}(z) \,. \tag{3.73}$$

3.2.2 Hallén's equation

Not until forty years after the publication of Pocklington's integral equation a different integral equation for the calculation of antenna currents was proposed by Hallén [77]. We derive this integral equation in close analogy to Pocklington's equation and first introduce an equivalent surface current $\boldsymbol{J}_s$ on the boundary of the antenna surface. This current is related to the scattered magnetic vector potential via

$$\boldsymbol{A}^{\mathrm{sca}}(\boldsymbol{r}) = \mu \int_{\Gamma} \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') \, \boldsymbol{J}_s(\boldsymbol{r}') \, d^2 r' \,. \tag{3.74}$$

The boundary condition for the total magnetic vector potential $\boldsymbol{A}$ on a perfectly conducting surface is $\boldsymbol{e}_n \times \boldsymbol{A} = \boldsymbol{0}$ or, equivalently, $A_t = 0$. With this boundary condition and the relation

$$\boldsymbol{A} = \boldsymbol{A}^{\mathrm{inc}} + \boldsymbol{A}^{\mathrm{sca}} \tag{3.75}$$

we obtain

$$\mu \left[\int_{\Gamma} \overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}_s(\boldsymbol{r}') \, d^2 r' \right] \cdot \boldsymbol{e}_t(\boldsymbol{r}) = -A_t^{\mathrm{inc}}(\boldsymbol{r}) \,. \tag{3.76}$$

Again, we may consider the special case of a thin, straight cylindrical wire which is directed along the z-axis and find, similar to (3.72),

$$\mu \int_{-L/2}^{L/2} G_{zz}^A(z, z') I(z') \, dz' = -A_z^{\mathrm{inc}}(z) \,. \tag{3.77}$$

This result looks rather simple but it must be noted that, in practice, the incident electromagnetic field usually will be given in terms of the electric field strength $\boldsymbol{E}^{\mathrm{inc}}$ rather than in terms of the magnetic vector potential $\boldsymbol{A}^{\mathrm{inc}}$. In the Lorenz gauge we have

$$\boldsymbol{E}^{\mathrm{inc}}(\boldsymbol{r}) = -\frac{j\omega}{k^2} \left(\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{A}^{\mathrm{inc}}(z)) + k^2 \boldsymbol{A}^{\mathrm{inc}}(z) \right) \tag{3.78}$$

and this second order partial differential equation needs to be solved in order to obtain $\boldsymbol{A}^{\mathrm{inc}}$ from $\boldsymbol{E}^{\mathrm{inc}}$. In case of a z-directed thin-wire antenna (3.78) simplifies to

$$E_z^{\mathrm{inc}}(z) = -\frac{j\omega}{k^2} \left(\frac{\partial^2}{\partial z^2} + k^2 \right) A_z^{\mathrm{inc}}(z) \,. \tag{3.79}$$

This ordinary differential equation has well-known solutions that are given by sum of a general solution of the homogeneous problem and a special solution of the inhomogeneous problem. It follows

$$-A_z^{\mathrm{inc}}(z) = C_1 e^{jkz} + C_2 e^{-jkz} + \frac{k}{2j\omega} \int_{-L/2}^{L/2} \sin(k|z - z'|) E_z^{\mathrm{inc}}(z') \, dz' \,, \tag{3.80}$$

where C_1, C_2 denote two integration constants that need to be determined from the boundary condition that the antenna current vanishes at the antenna ends. The function

$G(z) = \sin(k|z|)$ that appears in the special solution of the inhomogeneous problem is, up to a constant factor, a Green's function for the differential equation (3.79). It fulfills $(\partial^2/\partial z^2 + k^2)G(z) = 2k\delta(z)$. With (3.80) we obtain from (3.77) Hallén's equation in the form

$$\mu \int_{-L/2}^{L/2} G_{zz}^A(z, z') I(z')\, dz' = C_1 e^{jkz} + C_2 e^{-jkz} + \frac{k}{2j\omega} \int_{-L/2}^{L/2} \sin(k|z - z'|) E_z^{\text{inc}}(z')\, dz' . \tag{3.81}$$

Compared to Pocklington's equation (3.73) the integral kernel of Hallén's equation is less singular and thus preferable for numerical evaluations. The integral kernel of Pocklington's equation exhibits a spatial singularity that is proportional to $1/|\boldsymbol{r} - \boldsymbol{r}'|^3$ while in the case of Hallén's equation the spatial singularity is proportional to $1/|\boldsymbol{r} - \boldsymbol{r}'|$. However, to determine in Hallén's integral equation the integration constants C_1, C_2 can be cumbersome and in such cases Pocklington's equation might be the more practical choice.

3.2.3 Mixed-potential integral equation

The mixed-potential integral equation represents another version of an electric-field integral equation. It is often used in numerical calculations since its integral kernel is proportional to $1/|\boldsymbol{r} - \boldsymbol{r}'|^2$, that is, the singularity of the integral kernel is weaker than in the case of Pocklington's equation. Additionally, the mixed-potential integral equation does not require to determine integration constants as in the case of Hallén's equation.

To derive the mixed-potential integral equation we consider the relation (1.147) for the scattered electromagnetic field in the frequency domain,

$$\boldsymbol{E}^{\text{sca}}(\boldsymbol{r}) = -\boldsymbol{\nabla}\phi^{\text{sca}} - j\omega \boldsymbol{A}^{\text{sca}} . \tag{3.82}$$

If we replace by means of the Lorenz gauge the scalar potential ϕ^{sca} by the vector potential $\boldsymbol{A}^{\text{sca}}$ and employ (3.74) we will be led back to (3.70) and obtain nothing new. Alternatively, we consider, besides (3.74), the source-field relation

$$\phi^{\text{sca}}(\boldsymbol{r}) = \mu \int_\Gamma G^\phi(\boldsymbol{r}, \boldsymbol{r}')\, \rho_s(\boldsymbol{r}')\, d^2 r' \tag{3.83}$$

with $G^\phi(\boldsymbol{r}, \boldsymbol{r}')$ the scalar Green's function of the scalar Helmholtz equation, compare (1.169), and ρ_s a surface charge density. This surface charge density is related to a surface current $\boldsymbol{J}_s$ by a continuity equation which, in integral form, reads

$$j\omega \int_\Gamma \rho_s\, dA + \int_{\partial\Gamma} \boldsymbol{J}_s \cdot d\boldsymbol{A} = 0 . \tag{3.84}$$

If the surface Γ is simply connected Stokes theorem (A.25) can be applied to yield the local continuity equation

$$j\omega\rho_s(\boldsymbol{r}) + (\boldsymbol{\nabla} \times \boldsymbol{J}_s(\boldsymbol{r})) \cdot \boldsymbol{e}_n = 0 \tag{3.85}$$

with $\boldsymbol{e}_n$ a normal vector on Γ. This is a special case of the fundamental continuity equation (1.10). It follows that (3.82) can be rewritten in the form

$$\boldsymbol{E}^{\text{sca}}(\boldsymbol{r}) = \frac{1}{j\omega\varepsilon} \int_\Gamma \left(\left(\boldsymbol{\nabla} G^\phi(\boldsymbol{r},\boldsymbol{r}') \right) \left(\boldsymbol{\nabla}' \times \boldsymbol{J}_s(\boldsymbol{r}') \right) \cdot \boldsymbol{e}_n + k^2 \overline{\boldsymbol{G}}^A(\boldsymbol{r},\boldsymbol{r}') \boldsymbol{J}_s(\boldsymbol{r}') \right) d^2 r' . \quad (3.86)$$

This yields the electric field integral equation

$$\frac{1}{j\omega\varepsilon} \left[\int_\Gamma \left(\left(\boldsymbol{\nabla} G^\phi(\boldsymbol{r},\boldsymbol{r}') \right) \left(\boldsymbol{\nabla}' \times \boldsymbol{J}_s(\boldsymbol{r}') \right) \cdot \boldsymbol{e}_n + k^2 \overline{\boldsymbol{G}}^A(\boldsymbol{r},\boldsymbol{r}') \boldsymbol{J}_s(\boldsymbol{r}') \right) d^2 r' \right] \cdot \boldsymbol{e}_t(\boldsymbol{r}) = -E_t^{\text{inc}}(\boldsymbol{r}) . \quad (3.87)$$

Again, we may consider the special case of a z-directed, cylindrical and straight antenna. In this case (3.85) is not valid since the surface of a cylinder is not simply connected. But after a thin-wire approximation we can replace (3.85) by

$$j\omega q' + \frac{dI}{dz} = 0 , \quad (3.88)$$

with q' the electric charge per unit length and I the total electric current on the antenna, and find in analogy to (3.87) the mixed-potential integral equation

$$\frac{1}{j\omega\varepsilon} \int_{-L/2}^{L/2} \left[\frac{\partial G^\phi(z,z')}{\partial z} \frac{\partial I(z')}{\partial z'} + k^2 G_{zz}^A(z,z') I(z') \right] dz' = -E_z^{\text{inc}}(z) . \quad (3.89)$$

In free space $G^\phi(z,z')$ and $G^A(z,z')$ are the same functions and the mixed-potential integral equation further simplifies.

3.2.4   Schelkunoff's equation

For completeness we mention another electric field integral equation that is known as Schelkunoff's equation. It requires the condition

$$\frac{\partial G^\phi(z,z')}{\partial z} = -\frac{\partial G^\phi(z,z')}{\partial z'} \quad (3.90)$$

which implies translational invariance of the Green's function, $G^\phi(z,z') = G^\phi(|z-z'|)$. This condition is fulfilled for the Green's function of free space, compare (2.137), but it will not be valid in general.

If (3.90) holds we may integrate by parts the first term in the integral of (3.89),

$$\int_{-L/2}^{L/2} \frac{\partial G^\phi(z,z')}{\partial z'} \frac{\partial I(z')}{\partial z'} dz' = G^\phi(z,z') \frac{\partial I(z')}{\partial z'} \bigg|_{z'=-L/2}^{z'=L/2} - \int_{-L/2}^{L/2} G^\phi(z,z') \frac{\partial^2 I(z')}{\partial z'^2} . \quad (3.91)$$

Then the mixed-potential integral equation can be rewritten to yield Schelkunoff's equation,

$$\frac{1}{j\omega\epsilon} \int_{-L/2}^{L/2} \left[G^\phi(z,z') \frac{\partial^2 I(z')}{\partial z'^2} + k^2 G_{zz}^A(z,z') I(z') \right] dz'$$
$$-\frac{1}{j\omega\varepsilon} G^\phi(z,z') \frac{\partial I(z')}{\partial z'} \bigg|_{z'=-L/2}^{z'=L/2} = -E_z^{\text{inc}}(z) . \quad (3.92)$$

As in the case of Hallén's equation the integral kernel of Schelkunoff's equation is proportional to $1/|\boldsymbol{r} - \boldsymbol{r}'|$ and, thus, advantageous for numerical evaluation.

3.2.5 General remarks on solution methods

The integral equations (3.73), (3.81), (3.89), and (3.92) are considered to be the four standard equations for the current determination on cylindrical thin-wire antennas [176, 177, 189]. To obtain an unknown antenna current it is necessary to solve for a given excitation one of these equations. Corresponding solution methods can be divided into analytical and numerical methods.

(a) Analytical methods

After its publication Pocklington's equation was mainly used to justify that the antenna current on a linear dipole antenna can be approximated by a sinusoidal current distribution. This rather crude approximation yields satisfying results if the current distribution is required to calculate electromagnetic far-field patterns, but it fails for the calculation of self impedances. Hallén provided refined results for the antenna current in terms of series solutions [77]. His approach stimulated further research activities on wire antennas. Especially King added to this subject and provided detailed results in his unparalleled monograph [104]. In particular, King's three-term approximation to Hallén's equation, which involves the three sinusoidal terms $\sin(k|z|)$, $\cos(kz)$, and $\cos(kz/2)$, made it possible to obtain satisfying, approximate results for the calculation of antenna impedances [105]. It should be noted in this connection that the theory of linear antennas is accompanied by many subtleties that are rooted in the difficulties to properly model details of antenna geometries or antenna excitations. To get an impression of the related analytic approaches we refer, besides to the classic book of King [104], to the monographs by Schelkunoff & Friis [192], Collin & Zucker [29], and King & Harrison [106]. It is telling that these thorough works appeared just before Harrington introduced the method of moments to the electromagnetic research community and the use of digital computers became common [83]. From that time there was a significant shift from analytical to numerical methods in antenna theory.

(b) Numerical methods

A combination of the integral equations for wire antennas with the method of moments, as described in Sec. 2.4.1, turns out to be an effective way to calculate antenna currents. The integral equations are defined on antenna surfaces and, thus, for their numerical solution it is sufficient to discretize these boundaries rather than the complete volume where the electromagnetic field is defined. Once we have calculated a numerical solution for the antenna current the corresponding electromagnetic field is found from a Green's function that relates the electric current to the electromagnetic field.

In the method of moments an unknown antenna current is approximated by a finite linear combination of basis functions ψ_k with unknown coefficients α_k, compare (2.266),

$$I \approx \sum_{k=1}^{N} \alpha_k \psi_k \,. \tag{3.93}$$

If this ansatz is inserted in one of the linear integral equations of the form

$$\mathcal{L}I = -E \tag{3.94}$$

one finds, after the introduction of weighting functions w_j, the linear algebraic system of equations (2.271),

$$\sum_{k=1}^{N} \alpha_k \langle \mathcal{L}\psi_k, w_j \rangle = -\langle E, w_j \rangle \,. \tag{3.95}$$

In practice, the difficult numerical part is to evaluate the matrix components $\langle \mathcal{L}\psi_k, w_j \rangle$. Once we know these components it is immediate to solve the system (3.95) by s simple matrix inversion for the unknown expansion coefficients α_k. In this way we arrive at the following simple solution scheme:

1. Choose an integral equation for the unknown antenna current. This integral equation will contain the physical information of the antenna problem.

2. Choose appropriate basis functions ψ_k and weighting functions w_j. The choice should be such that the subsequent numerical evaluation of the components $\langle \mathcal{L}\psi_k, w_j \rangle$ can effectively be done. This step mainly requires mathematical skills and experience.

3. Evaluate the elements $\langle \mathcal{L}\psi_k, w_j \rangle$ and $-\langle E, w_j \rangle$.

4. Solve the linear system of equations and obtain an approximate solution for the unknown antenna current.

Early and important results that have been obtained by this method are discussed and summarized in [142, 143, 16, 237]. More recent results and references are provided by [177, 189]. Nowadays, the numerical analysis of wire antennas in free space can be considered to be a standard task and a variety of free and commercial software packages that are based on the method of moments are available. However, the analysis of wire antennas can still be difficult if specific geometric details, like wire ends, wire junctions, or connections to metallic surfaces, need to be properly taken into account or if the antenna environment is characterized by complex media or complicated boundaries.

(c) Solution methods for antenna theory in cavities

In Chapter 2 we have considered Green's function of self-adjoint differential operators that relate a primary source g and a field f by a simple relation of the form

$$f(r) = \int_\Omega g(r') \cdot \overline{G}(r, r')\, d\Omega\,, \tag{3.96}$$

see (2.117) and (2.124), for example. In particular, spectral and ray representations of cavities' Green's function have been constructed in Sec. 2.3. These Green's functions naturally incorporate the boundary conditions at the cavity walls. As a consequence, it is not necessary to explicitly take into account the interaction of primary sources $g(r')$ with the cavity walls.

Source-field relations of the form (3.96) are the basis of the integral equations of antenna theory, compare (3.70), (3.74), and (3.83). Therefore, any configuration of boundaries with a known Green's function and associated source-field relations of the form (3.96) can be considered. This generality is an obvious benefit of the Green's function approach.

It already has been mentioned various times that antenna theory normally is formulated in free space. That is, most results of antenna theory are obtained on the basis of integral equations that employ the Green's function of free space. It seems logical to build upon these results and look for the *modifications* that occur if we pass to the inside of a resonating environment. It has been shown by means of the general ray representation (2.251) that a cavities' Green's function contains the Green's function of free space as a zeroth order approximation. This suggests the possibility to split an antenna problem within a resonating environment into an antenna problem in free space and an additional problem that is determined from the characteristics of the resonances. This intuitive idea is realized by a mathematical solution procedure, the so-called method of analytical regularization.

3.2.6   Method of analytical regularization

Halléns equation, Pocklington's equation, and the mixed-potential integral equation are integral equations of the first kind and of the general form

$$0 = E(z) + \int k(z, z') I(z')\, dz'\,, \tag{3.97}$$

with $k(z, z')$ an integral kernel that contains the relevant Green's function. In operator notation we write (3.97) as

$$\mathcal{L} I = -E\,. \tag{3.98}$$

The linear operator $\mathcal{L}$ is an integral or integro-differential operator and contains the Coulomb singularity. As a consequence, the first order integral equations of antenna theory turn out to be ill-conditioned to various degrees.

At this point the method of analytical regularization can be beneficial [162, 229, 230]. By this method an ill-conditioned first integral equation is converted to a well-conditioned second order integral equation of the general form

$$I(z) = I_0(z) + \int \tilde{k}(z, z')I(z')\, dz'. \tag{3.99}$$

The method of analytical regularization requires that the linear operator $\mathcal{L}$ can be split into two parts,

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1\,, \tag{3.100}$$

where $\mathcal{L}_0$ contains the singularity that needs to be regularized. It is also required that the inverse operator $\mathcal{L}_0^{-1}$ can be constructed. If this operator is applied to the equation

$$\mathcal{L}_0 I + \mathcal{L}_1 I = -E \tag{3.101}$$

we obtain an integral equation of the second kind,

$$I = -\mathcal{L}_0^{-1}E - \mathcal{L}_0^{-1}\mathcal{L}_1 I \tag{3.102}$$

$$= I_0 - \mathcal{L}_0^{-1}\mathcal{L}_1 I\,, \tag{3.103}$$

where we defined $I_0 := -\mathcal{L}_0^{-1}E$. This second order integral equation needs to be solved for the unknown current I. It usually is well-conditioned and accompanied by stable solutions.

In the application of the method of analytical regularization to antenna theory in cavities, or possibly some other resonating environment, it is reasonable to choose $\mathcal{L}_0$ as the "free space" operator of the problem, that is, $\mathcal{L}_0$ becomes the integral or integro-differential operator which results if the Green's function of free space is employed in the kernel of the relevant first order integral equation. Then $\mathcal{L}_0$ contains the Coulomb singularity and the methods and results of conventional antenna theory are applied to obtain $\mathcal{L}_0^{-1}$. The subsequent solution of the resulting second order integral equation (3.103) takes into account the cavity properties. The kernel of this second order integral equation will involve the difference between the cavities' Green's function and the Green's function of free space and will contain no spatial singularity since the Coulomb singularity gets subtracted. The method of analytical regularization can straightforwardly be combined with the method of moments where the operators $\mathcal{L}_0$ and $\mathcal{L}_1$ turn to finite-dimensional matrices.

3.3 Accelerating the convergence rate of series Green's functions

The application of Green's functions to actual physical problems usually requires to evaluate these functions many times for various arguments. Unfortunately, most Green's

functions of physical interest are given in terms of mathematical series. These series often slowly converge and make it impossible to calculate accurate values within a reasonable time frame. It is then necessary to either find an alternative to the Green's function approach or transform a slowly converging series into another series with better convergence properties. In this section we discuss the second possibility.

3.3.1   Mode, ray, and hybrid representations

There is no general mathematical recipe for the transformation of a slowly converging series into a rapidly converging series. A series will have good convergence properties if the characteristics of the corresponding Green's functions are reflected by the summands of the series. In the electromagnetic case dominant characteristics are the two types of electromagnetic field singularities that have been identified in Sec. 1.5.2. Close to the source region, where the distance between observation point r and source point r' is electrically small, $k|r - r'| \ll 1$, the Coulomb singularity will be strong. In this case the series Green's function should contain a summand which incorporates the Coulomb singularity. This requirement is fulfilled by a ray representation, as discussed in Sec. 2.3.3. Conversely, close to a resonance frequency, where the wavenumber k approaches an eigenvalue k_n, the Green's function will be dominated by a resonance. Then the series Green's function should contain a summand which incorporates this resonance such that in this case a spectral mode representation of the Green's function will be suitable, compare Sec. 2.3.1.

Modes	Rays
oscillations yield *global* information of a system	scattering processes yield *local* information of a system
characterize *late* response in time-domain	characterize *early* response in time-domain
advantageous for *low-frequency* regime where the mode-density is low and a small number of modes characterizes the field	advantageous for *high- frequency regime* where the mode-density is high and rays of geometrical optics characterize the field
advantageous to model resonances	advantageous to model Coulomb singularities

Table 3.1: Complementary properties of modes and rays. A more detailed discussion is provided by Felsen [47].

These simple arguments indicate that representations of electromagnetic series Green's functions which incorporate both spectral and ray properties will have advantageous convergence properties both at resonance and close to the source region. Such representations are known as *hybrid representations.* The first systematic investigation of hybrid representations is mainly due to Felsen [46, 47]. He pointed out that modes and rays are related to each other by infinite Poisson transformations and this explains their complementary properties which, from a physical point of view, reflect the complementarity of resonances and the Coulomb singularity. Mode properties are compared to ray properties in Tab. 3.1.

In one dimension an infinite Poisson transformation is expressed by the infinite Poisson sum formula [32, Chap. 2, §5]

$$\sum_{n=-\infty}^{\infty} f(2n\pi + t) = \frac{1}{2\pi} \sum_{\nu=-\infty}^{\infty} e^{j\nu t} \int_{-\infty}^{\infty} f(\tau)e^{-j\nu\tau}\,d\tau\,. \tag{3.104}$$

For $t = 0$ this formula reduces to

$$\sum_{n=-\infty}^{\infty} f(2n\pi) = \frac{1}{2\pi} \sum_{\nu=-\infty}^{\infty} \int_{-\infty}^{\infty} f(\tau)e^{-j\nu\tau}\,d\tau \tag{3.105}$$

and a three-dimensional version of this result is

$$\sum_{m,n,p=-\infty}^{\infty} f(2m\pi, 2n\pi, 2p\pi) = \tag{3.106}$$

$$\frac{1}{(2\pi)^3} \sum_{m,n,p=-\infty}^{\infty} \iiint_{-\infty}^{\infty} f(\tau_1, \tau_2, \tau_3)e^{-j(m\tau_1+n\tau_2+p\tau_3)}\,d\tau_1\,d\tau_2\,d\tau_3\,.$$

The following example shows the transformation between modes and rays within a rectangular cavity.

Example: We consider the dyadic Green's function $\overline{\boldsymbol{G}}^A$ of a rectangular cavity, compare (2.171). This Green's function is diagonal and the three components G_{xx}^A, G_{yy}^A, and G_{zz}^A are related to each other by cyclic exchange of x, y, and z. The explicit form of the G_{zz}^A component, for example, is given by

$$G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}') = \tag{3.107}$$

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{p=0}^{\infty} \frac{\epsilon_{0p}}{l_x l_y l_z} \frac{\sin(\frac{m\pi x}{l_x})\sin(\frac{m\pi x'}{l_x})\sin(\frac{n\pi y}{l_y})\sin(\frac{n\pi y'}{l_y})\cos(\frac{p\pi z}{l_z})\cos(\frac{p\pi z'}{l_z})}{(\frac{m\pi}{l_x})^2 + (\frac{n\pi}{l_y})^2 + (\frac{p\pi}{l_z})^2 - k^2}$$

$$= \frac{1}{8} \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\sum_{p=-\infty}^{\infty} \frac{\epsilon_{0p}}{l_x l_y l_z} \frac{\sin(\frac{m\pi x}{l_x})\sin(\frac{m\pi x'}{l_x})\sin(\frac{n\pi y}{l_y})\sin(\frac{n\pi y'}{l_y})\cos(\frac{p\pi z}{l_z})\cos(\frac{p\pi z'}{l_z})}{(\frac{m\pi}{l_x})^2 + (\frac{n\pi}{l_y})^2 + (\frac{p\pi}{l_z})^2 - k^2}\,.$$

The trigonometric functions of each summand of the last expression can be recombined to yield eight exponential terms. In this exponential form the mode representation

(3.107) of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')$ is transformed by means of the three-dimensional infinite Poisson summation formula (3.106). The resulting integrals on the right hand side of (3.106) have closed form solutions and yield the transformed result [239, 240]

$$G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}') = \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} G_0\big(R_{i,mnp}(\boldsymbol{r}, \boldsymbol{r}')\big), \qquad (3.108)$$

with $R_{i,mnp}(\boldsymbol{r}, \boldsymbol{r}')$ and A_i^{zz} as defined in (2.258) and (2.264), respectively. Therefore, we obtain the ray representation (2.256) as the infinite Poisson transform of the mode representation (3.107). Additionally, it is recognized that single rays are represented by integrals of the form $\iiint_{-\infty}^{\infty} f(\tau_1, \tau_2, \tau_3) e^{-j(m\tau_1 + n\tau_2 + p\tau_3)} \, d\tau_1 \, d\tau_2 \, d\tau_3$ and it follows that rays and modes are essentially each others Fourier transform.

The infinite Poisson transformation mediates between mode and ray representation but does not yield a hybrid representation which contains, by definition, both a mode and a ray part. To obtain a hybrid representation for the Green's function of a rectangular cavity it is possible to consider a finite Poisson transformation which is a generalization of (3.106) and given by [33, 240]

$$\sum_{m=m_0}^{M} \sum_{n=n_0}^{N} \sum_{p=p_0}^{P} f(2m\pi, 2n\pi, 2p\pi) = \qquad (3.109)$$

$$\sum_{m,n,p=-\infty}^{\infty} \int_{m_0-\frac{1}{2}}^{M+\frac{1}{2}} \int_{n_0-\frac{1}{2}}^{N+\frac{1}{2}} \int_{p_0-\frac{1}{2}}^{P+\frac{1}{2}} f(\tau_1, \tau_2, \tau_3) e^{-2\pi j(m\tau_1 + n\tau_2 + p\tau_3)} \, d\tau_1 \, d\tau_2 \, d\tau_3 \,.$$

This formula allows to transform only a subset of modes into a ray contribution. Wu & Chang calculated in this way a hybrid representation for the Green's function $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')$ of a rectangular cavity [239, 240]. They noticed during the calculation that the integrations of (3.109) cannot be done in closed form and obtained as a result a hybrid representation as an *approximate* sum of a mode part, a ray part, and a correction term. This mathematical feature is in accordance to the observation of Felsen [47] that pure mode and ray contributions are inseparably entangled and cannot be split into two disjunct sets. The convergence properties of the hybrid representation of Wu & Chang have been compared to those of the standard mode and ray representations and the expected advantageous convergence properties close to the source region and close at resonance could be confirmed [68]. However, we will later make use of a different hybrid representation which is based on the so-called Ewald transformation and will be introduced below.

Having mentioned the Ewald transformation it is already indicated that besides the infinite and finite Poisson transformation there are other transformations and methods that have been found useful to improve the convergence of series Green's functions. A variety of results for the series Green's functions of the *two-dimensional* Helmholtz equation in periodic domains includes the application of the so-called Kummer's transformation, Veysoglu's transformation, Shanks' transformation, or Melnikov's method

[130, 38, 138, 99, 204]. These results can be applied to the solution of two-dimensional problems which involve the analysis of antenna arrays [118], electromagnetic coupling on printed circuit boards [224], the description of cavity backed antennas [202], and electromagnetic fields in layered media [22], for example.

Obviously, series Green's functions of the *three-dimensional* Helmholtz equation are more involved than their two-dimensional counterparts. A direct evaluation of the triple sum in (3.107), for example, is not practical. But in this and similar cases it is possible to perform one of the required summations in analytic form. To simplify (3.107), for example, we first apply the trigonometric identity

$$\cos(\alpha)\cos(\beta) = \frac{1}{2}\big(\cos(\alpha-\beta) + \cos(\alpha+\beta)\big) \tag{3.110}$$

to each summand and then use the formula

$$\sum_{p=0}^{\infty} \frac{\cos(px)}{p^2 - a^2} = -\frac{\pi}{2a}\frac{\cos(x-\pi)a}{\sin(\pi a)} \tag{3.111}$$

which is valid for $0 \le x \le 2\pi$. After some algebra this reduces the triple sum representation (3.107) to the double sum representation

$$G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}') = \tag{3.112}$$

$$\frac{4}{l_x l_y} \sum_{m,n=1}^{\infty} \frac{\sin(\frac{m\pi}{l_x}x)\sin(\frac{m\pi}{l_x}x')\sin(\frac{n\pi}{l_y}y)\sin(\frac{n\pi}{l_y}y')}{\gamma \sinh(\gamma l_z)} \cdot \left\{ \begin{array}{ll} \cosh(\gamma(l_z - z))\cosh(\gamma z'), & z \ge z' \\ \cosh(\gamma(l_z - z'))\cosh(\gamma z), & z < z' \end{array} \right.,$$

where we defined $\gamma := \sqrt{(\frac{m\pi}{l_x})^2 + (\frac{n\pi}{l_y})^2 - k^2}$. In this expression the hyperbolic terms lead, in general, to satisfying convergence properties. However, since the double sum representation basically is a mode representation convergence properties deteriorate for $\boldsymbol{r} \to \boldsymbol{r}'$, i.e., if the Coulomb singularity is approached. Moreover, the representation (3.112) diverges for $z \to z'$ even if $\boldsymbol{r}$ and $\boldsymbol{r}'$ are not close to each other.

As an alternative, representations of the three-dimensional periodic Green's function that are based on the Ewald transformation have been proven to be advantageous. This transformation was introduced by Ewald [44, 45] to theoretical solid state physics in the beginning of the 20$^{\text{th}}$ century. Decades later, the benefit of the Ewald transformation for the numerical evaluation of the Green's function for the Helmholtz operator on periodic structures was pointed out by Jordan et al. [98], see also Cohen [27]. This led to an Ewald representation for the electromagnetic Green's function of a perfectly conducting cavity where the wavenumber k is assumed to be real [167, 137]. A generalization of this result to lossy cavities which are characterized by complex wavenumbers is given in [69]. As shown in Appendix C, the component G_{zz}^A of the dyadic cavity's Green's function $\overline{\boldsymbol{G}}^A$ can be split in two parts,

$$G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}') = \underbrace{G_{zz1}^A(\boldsymbol{r}, \boldsymbol{r}')}_{\text{"mode part"}} + \underbrace{G_{zz2}^A(\boldsymbol{r}, \boldsymbol{r}')}_{\text{"ray part"}}, \tag{3.113}$$

with

$$G_{zz1}^A(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{8l_x l_y l_z} \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} \frac{e^{-\frac{k_{mnp}^2 - k^2}{4E^2}}}{k_{mnp}^2 - k^2} e^{j(k_x X_i + k_y Y_i + k_z Z_i)} , \qquad (3.114)$$

and

$$G_{zz2}^A(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{8\pi} \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} \left[\frac{e^{jkR_{i,mnp}} \mathrm{erfc}(R_{i,mnp}E + jk/2E)}{R_{i,mnp}} \right. $$
$$\left. + \frac{e^{-jkR_{i,mnp}} \mathrm{erfc}(R_{i,mnp}E - jk/2E)}{R_{i,mnp}} \right] , \quad (3.115)$$

where A_i^{zz} is defined as in (2.264) and erfc(z) denotes the complex complementary error function. Analogous expressions are valid for the components G_{xx}^A and G_{yy}^A. The parameter E is adjustable and balances the contributions of both the mode part and the ray part. Increasing values of E make the mode part contribute more to (3.113) and decrease the influence of the ray part while decreasing values of E make the ray part contribute more to (3.113) and increase the influence of the mode part. From a numerical point of view it is best to choose E such that the decay of both series is balanced [98, 110]. Some care must be taken since the first terms of (3.114) and (3.115) can become much larger than the actual value of the Green's function. These large terms cancel each other but due to the limited numerical accuracy of a computer this cancellation is susceptible to rounding errors. It should also be noted that the calculation of the complex complementary error function erfc(z) is nontrivial. The various algorithms that lead to an efficient evaluation are still subject to research in numerical mathematics [170, 195, 234].

3.3.2 Numerical examples

For illustrational purposes we calculate in this section the electromagnetic Green's function of a lossy rectangular cavity by means of the mode representation (3.112), the ray representation (2.254)–(2.256), and the Ewald representation (3.113). We take as an example a cubic cavity with edges $l_x = l_y = l_z = L$ and place a unit source at $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, compare Fig. 3.7

As observation points we choose close to the source the point $\boldsymbol{r} = (0.26L, 0.26L, 0.26L)$ and, for comparison, distant from the source the point $\boldsymbol{r} = (0.7L, 0.7L, 0.7L)$. We also choose two wavenumbers. The first one, $k = 9.42/L$, is near to the resonances $k_{122} = k_{212} = k_{221} = 3\pi/L \approx 9.42477796/L$, the second one, $k = 8.50/L$, is inbetween the resonances above and its neighboring resonances $k_{112} = k_{121} = k_{211} = \sqrt{6}\pi/L$, i.e., this value is *not* near a resonance. Combining both the two observation points and the two wavenumbers yields four different combinations. For a fixed quality factor $Q = 1000$ we first calculate for these combinations the four complex values of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k)$ by means of the mode representation (3.112) and the
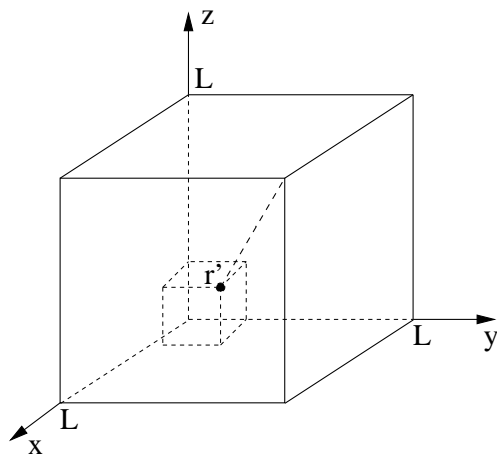
Figure 3.7: Sample cavity of dimensions $l_x = l_y = l_z = L$.

Ewald representation (3.113). It is known, and subsequently will be confirmed, that in case of high quality factors the convergence of the ray representation (2.254)–(2.256) is very poor .

We plot the relative error versus the number of terms that are included in the calculation. To number the terms in a systematic way all terms are first calculated and then ordered by their absolute value in descending order, that is, the term $N = 1$ corresponds to the term which has the largest absolute value. Reference values are calculated by the mode representation (3.112) and the Ewald representation (3.113), including a sufficient number of terms. Within the plots the calculated relative errors vary by about 15 orders and do not tend to zero due to numerical inaccuracies.

(a) Convergence distant to source region and off resonance

For the combination $\boldsymbol{r} = (0.7L, 0.7L, 0.7L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$ and $k = 8.50/L$ the resulting relative errors of the real and imaginary part of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k)$ are shown in Fig. 3.8. From a numerical and physical point of view this combination is not problematic. In this case the two-dimensional mode representation performs considerably better than the Ewald representation.

(b) Convergence distant to source region and close to resonance

Next we consider the combination $\boldsymbol{r} = (0.7L, 0.7L, 0.7L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, and $k = 9.42/L$. The relative errors of the real and imaginary part of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k)$ are shown in Fig. 3.9. Since both the two-dimensional mode representation and the Ewald representation are suitable to model resonances the convergence properties even slightly improve if compared to Fig. 3.8.
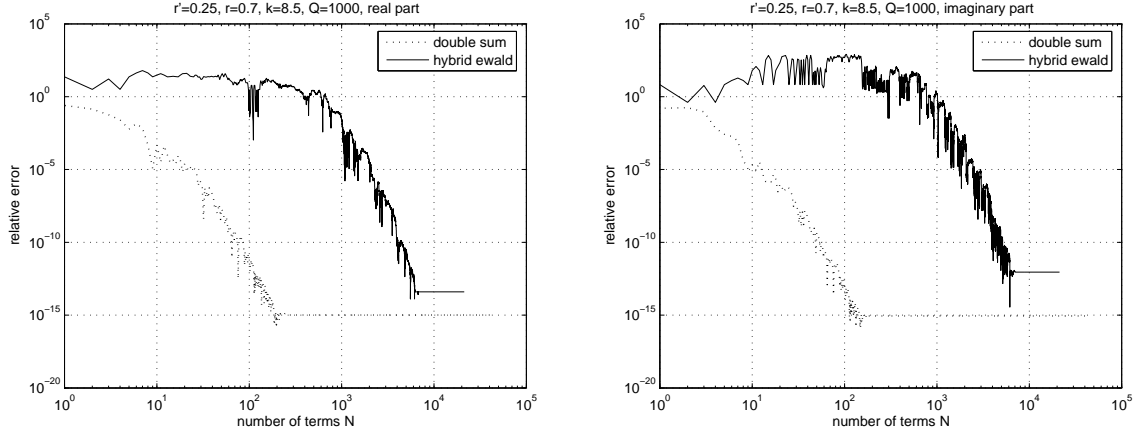
Figure 3.8: Convergence properties of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k)$ for $\boldsymbol{r} = (0.7L, 0.7L, 0.7L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, $k = 8.50/L$ and $Q = 1000$. The real part is plotted to the left, the imaginary part to the right.
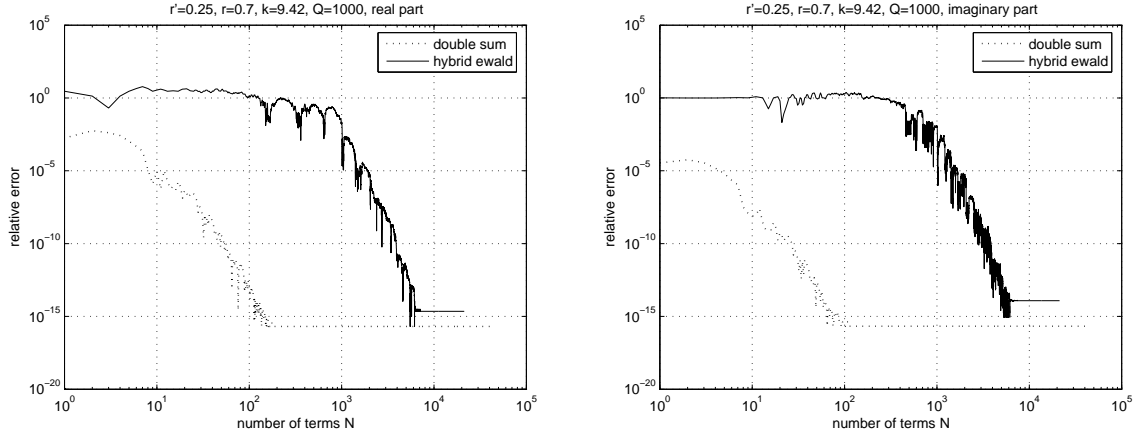
Figure 3.9: Convergence properties of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k)$ for $\boldsymbol{r} = (0.7L, 0.7L, 0.7L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, $k = 9.42/L$, and $Q = 1000$.

(c) Convergence close to source region and off resonance

We consider the combination $\boldsymbol{r} = (0.26L, 0.26L, 0.26L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, and $k = 9.42/L$. The resulting relative errors of the real and imaginary part of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k)$ are shown in Fig. 3.10. In this case the observation point is close to the source point and it is seen that the convergence properties of the two-dimensional mode representation become worse. The convergence properties of the Ewald representation are practically unaffected.
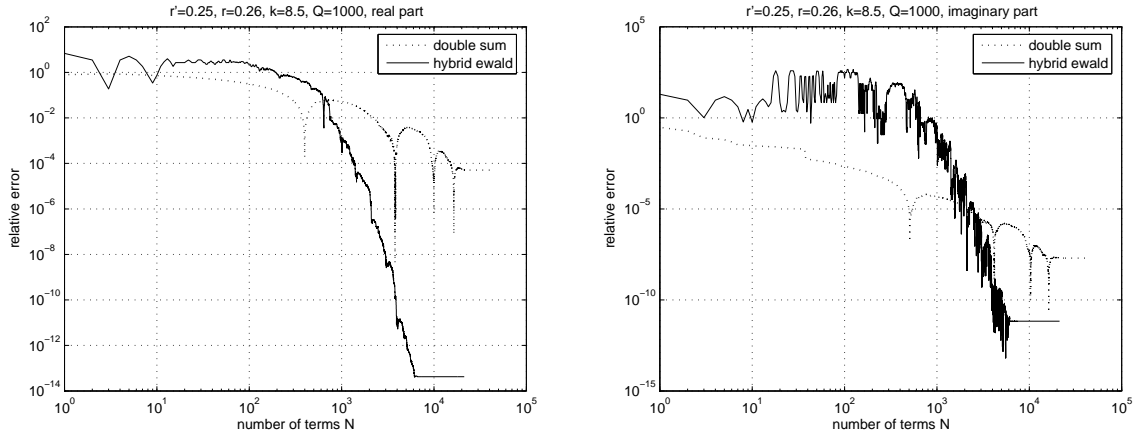
Figure 3.10: Convergence properties of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k)$ for $\boldsymbol{r} = (0.26L, 0.26L, 0.26L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, $k = 8.50/L$, and $Q = 1000$.
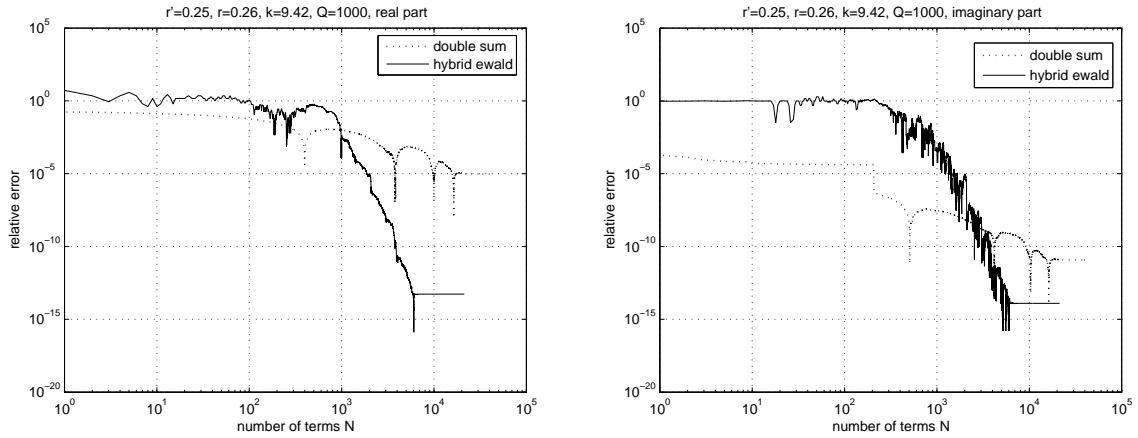
Figure 3.11: Convergence properties of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k_c)$ for $\boldsymbol{r} = (0.26L, 0.26L, 0.26L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, $k = 9.42/L$, and $Q = 1000$.

(d) Convergence close to source region and close resonance

We take the fourth combination $\boldsymbol{r} = (0.26L, 0.26L, 0.26L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, and $k = 9.42/L$. The resulting relative errors of the real and imaginary part of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k)$ are shown in Fig. 3.11. Similar to the previous case the convergence of the two-dimensional mode representation becomes worse while the Ewald representation still allows to calculate highly accurate values.

(e) Convergence if the source region is approached

To more clearly exhibit the influence of the source singularity we plot in Fig. 3.12 and Fig. 3.13 the convergence of the two-dimensional mode representation and the Ewald representation, respectively, for decreasing distances $|\boldsymbol{r} - \boldsymbol{r}'|$ where the position of $\boldsymbol{r}'$
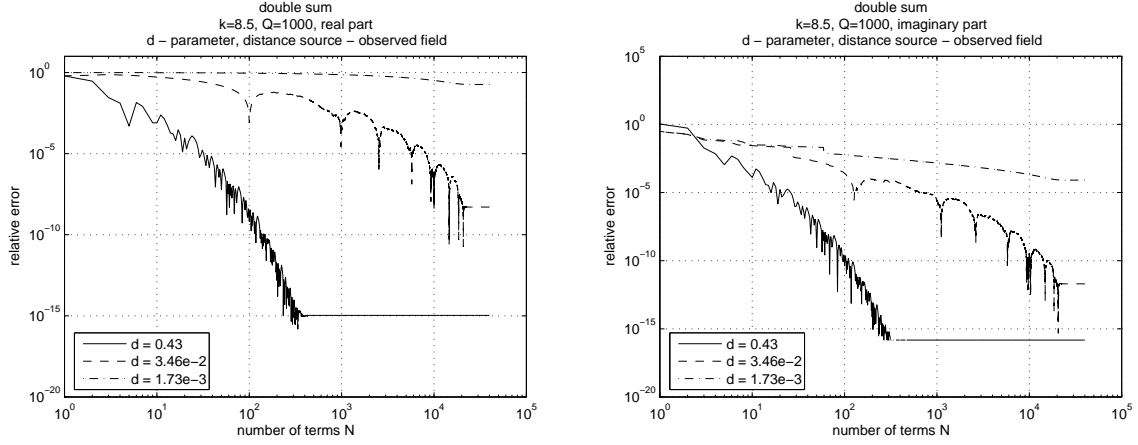
Figure 3.12: Convergence properties of the double sum representation for $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k_c)$ with $\boldsymbol{r} \to \boldsymbol{r}'$, $k = 8.5/L$, and $Q = 1000$. The parameter $d := |\boldsymbol{r} - \boldsymbol{r}'|$ assumes the values $0.43L$, $3.46 \cdot 10^{-2}L$, and $1.73 \cdot 10^{-3}L$.
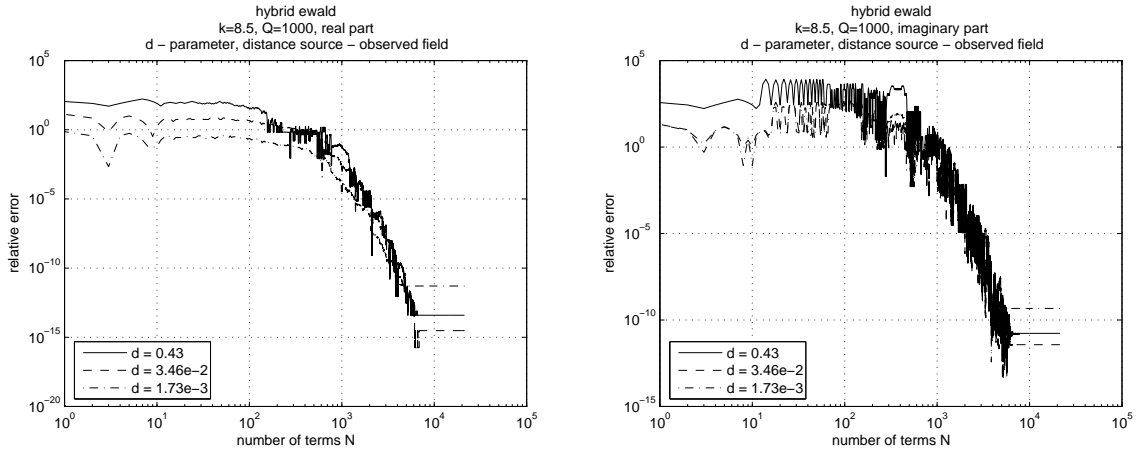
Figure 3.13: Convergence properties of the Ewald representation for $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k_c)$ with $\boldsymbol{r} \to \boldsymbol{r}'$, $k = 8.5/L$, and $Q = 1000$. The parameter $d := |\boldsymbol{r} - \boldsymbol{r}'|$ assumes the values $0.43L$, $3.46 \cdot 10^{-2}L$, and $1.73 \cdot 10^{-3}L$.

is kept and $\boldsymbol{r}$ approaches $\boldsymbol{r}'$ along the diagonal that is indicated in Fig. 3.7. For the mode representation the decreasing convergence for smaller distances between source and observation point is evident, while the Ewald representation appears to be insensitive towards variation of this distance.

(f) Convergence if the two-dimensional singularity of the two-dimensional mode representation is approached

It has been mentioned that the two-dimensional mode representation contains a singularity for $z \to z'$. That is, even if $\boldsymbol{r}$ and $\boldsymbol{r}'$ are not close to each other the two-dimensional

mode representation will no longer converge if z approaches z'. This is exemplified in Fig. 3.14. For the same situation the Ewald representation still has good convergence properties, as can be seen from Fig. 3.15.
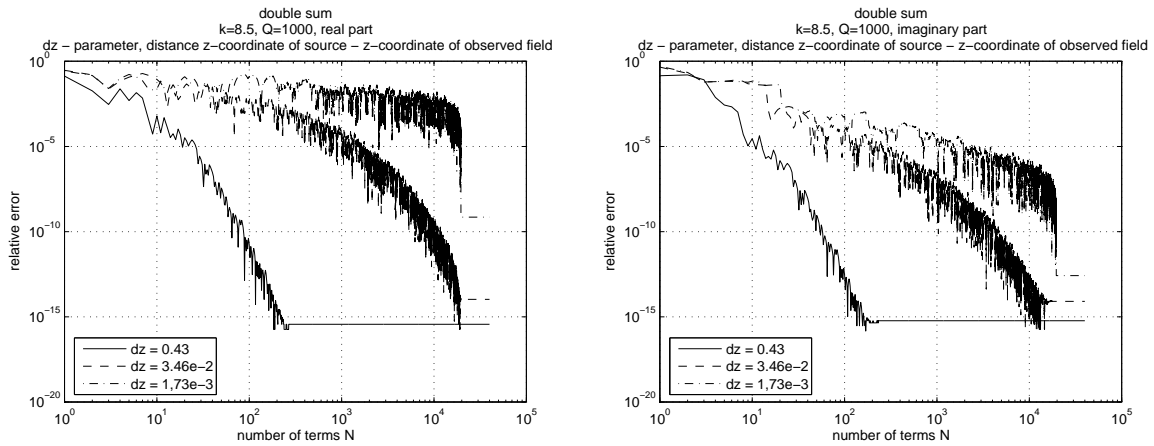
Figure 3.14: Convergence properties of the double sum representation for $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k_c)$ with $z \to z'$, $k = 8.5/L$, and $Q = 1000$. Fixed coordinates are $x' = y' = z' = 0.25L$ and $x = y = 0.7L$. The distance $dz := z - z'$ assumes the values $0.43L$, $3.46 \cdot 10^{-2}L$, and $1.73 \cdot 10^{-3}L$.
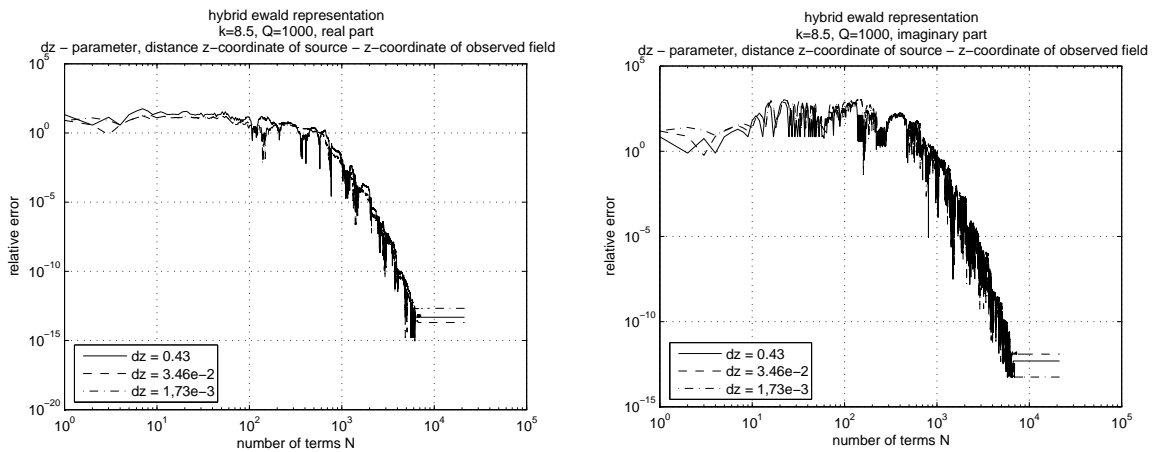
Figure 3.15: Convergence properties of the Ewald representation for $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k_c)$ with $z \to z'$, $k = 8.5/L$, and $Q = 1000$. Fixed coordinates are $x' = y' = z' = 0.25L$ and $x = y = 0.7L$. The distance $dz := z - z'$ assumes the values $0.43L$, $3.46 \cdot 10^{-2}L$, and $1.73 \cdot 10^{-3}L$.

(g) Dependency of the quality factor on the convergence

Finally, we illustrate in Fig. 3.16 and 3.17 the dependency of the quality factor on the accuracy of all three different representations. We take a fixed number of terms for each
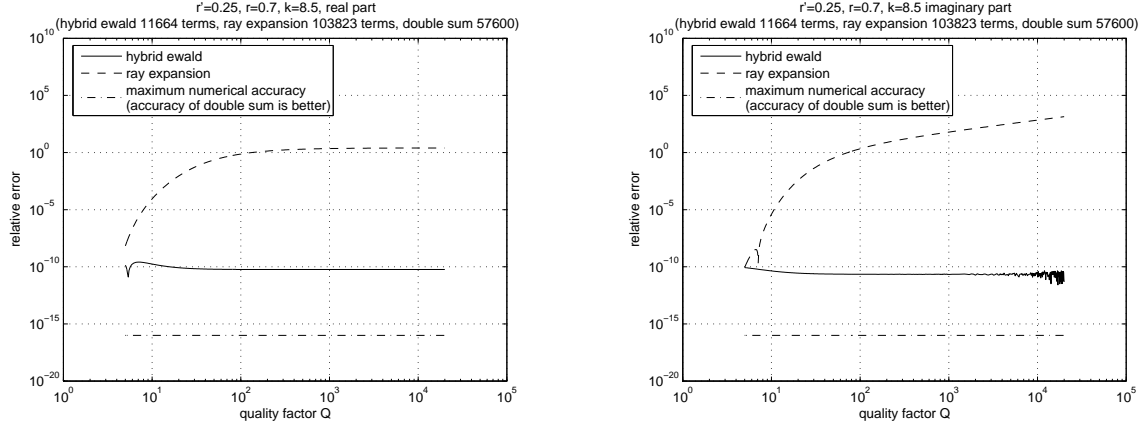
Figure 3.16: Convergence properties of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k_c)$ for $\boldsymbol{r} = (0.7L, 0.7L, 0.7L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, $k = 8.50/L$ for varying Q. For each representation the number of terms is fixed. Ewald representation: 11664 terms; ray representation: 103823 terms; two-dimensional mode representation: 57600 terms
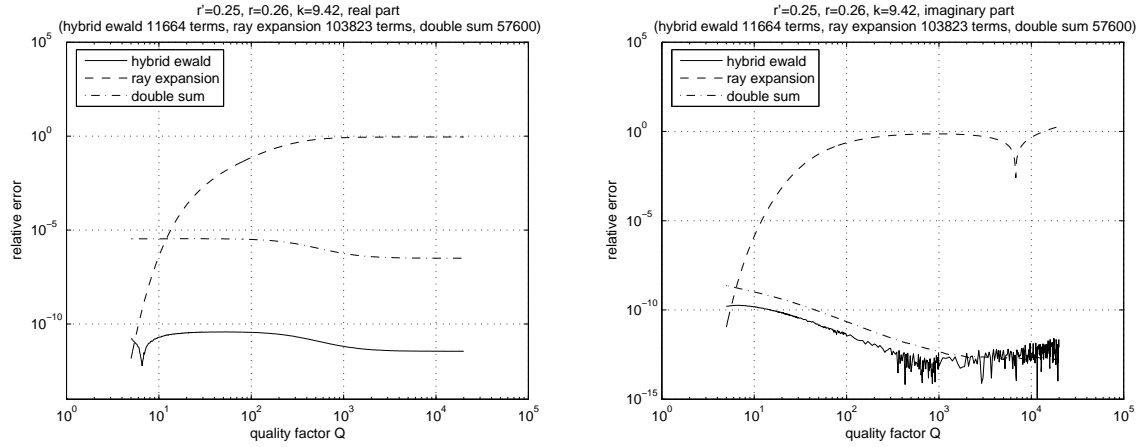
Figure 3.17: Convergence properties of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}', k_c)$ for $\boldsymbol{r} = (0.26L, 0.26L, 0.26L)$, $\boldsymbol{r}' = (0.25L, 0.25L, 025L)$, $k = 9.42/L$ for varying Q. For each representation the number of terms is fixed. Ewald representation: 11664 terms; ray representation: 103823 terms; two-dimensional mode representation: 57600 terms

representation which are chosen such that the calculation time for a specific value of G_{zz}^A is about the same order for all three representations. Since a single term of the ray sum is calculated faster than single terms of the other representations we include in the ray sum the highest number of terms. Accordingly, the lowest number of terms is included in the Ewald representation since the required calculation of the complementary error function for each single term is comparatively time consuming. As a result, it is recognized that the convergence of the mode representation and the Ewald representation does not significantly depend on the quality factor. The ray representation converges to a satisfactory degree only in case of low quality factors.

3.3.3   Approximation and interpolation of series Green's function

The use of hybrid representations for series Green's functions eliminates major convergence problems that commonly are encountered if Green's functions are evaluated close to the source region or close to a resonance. But in spite of this major advantage it still is necessary to sum up hundreds or thousands of series terms to calculate a single accurate value of a series Green's function. This is not very satisfying. To improve this situation it suggests itself to approximate or interpolate a series Green's function by polynomial functions.

The *approximation* of a function is its representation by an element of a finite dimensional function space that is expressed as a linear combination of suitable basis functions. We already have discussed such finite dimensional expansions in an appropriate functional analytic setting in Sec. 2.1.1 (f). Efficient approximations can be achieved by means of orthonormal polynomials [23]. Historically, the approximation of solutions of boundary value problems by means of orthonormal polynomials has been of great importance[5]. However, these approximation techniques have their limitations. If fine discretization are required we usually need a large number of polynomials to model a given function. This involves polynomials of high order which are characterized by strong oscillations and, in general, are not suitable for efficient approximations. The resulting convergence problems are analogous to those discussed in Sec. 3.3.1.

Similar to approximation schemes are *interpolation* schemes[6]. An interpolation scheme requires the calculation of a finite number of function values at discrete nodes that form a grid. In order to find function values between neighboring nodes the function is interpolated between these nodes by a piecewise polynomial of low order. Therefore, an interpolation scheme approximates a given function by a number of piecewise polynomials which are defined inbetween discrete nodes rather than over a whole grid. The interpolation by piecewise polynomials leads to the concept of *spline interpolation* [196, 79]. In practice, the most important splines are piecewise polynomials of zeroth order (step functions), first order (linear splines), or third order (cubic splines).

The implementation of an approximation or interpolation scheme requires to consider the following nontrivial issues:

- Usually, we consider a fixed frequency and need to evaluate a Green's function for various source and observation points r' and r, respectively. Each point depends on

---

[5] The possible polynomial solutions of the general second order Sturm-Liouville differential equation play a major role in mathematical physics and encompass a variety of canonical solutions in electromagnetic theory. They are given by the Jacobi, Gegenbauer, Tschebycheff, Legendre, Laguerre, and Hermite polynomials [17, § 5.10].

[6] Formally, the distinction between approximation and interpolation schemes is not clearly cut since both types constitute finite expansions of the form (2.4). In the applied mathematical literature it is nevertheless common to introduce both notions separately [79].

three spatial coordinates such that the Green's function depends on six variables. It follows that a polynomial approximation to the Green's function involves six-dimensional polynomials of the form $P(\boldsymbol{r}, \boldsymbol{r}') : \mathbb{R}^6 \to \mathbb{C}$. This is not a problem in principle, but it is desirable to reduce the number of variables, e.g., by symmetry considerations.

- An electromagnetic Green's function contains the Coulomb singularity and this singularity is not suitable to be modeled by polynomials. Hence, it is necessary to subtract terms that characterize the Coulomb singularity before an approximation or interpolation is done. The subtracted terms can later be added to the fitted polynomial in order to yield the complete Green's function.

The possibility to reduce the number of variables that enter the Green's function and the explicit subtraction of singular terms depend on the specific Green's function. A general scheme for the efficient evaluation of Green's functions that contain both the Coulomb singularity and resonances is proposed in Fig. 3.18.

In the following we will specialize the scheme of Fig. 3.18 to the Green's function $\overline{\boldsymbol{G}}^A(\boldsymbol{r}, \boldsymbol{r}')$ of a rectangular cavity. According to (2.171) this dyadic Green's function is diagonal and consists of the three nonvanishing components $G_{xx}^A(\boldsymbol{r}, \boldsymbol{r}')$, $G_{yy}^A(\boldsymbol{r}, \boldsymbol{r}')$, and $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')$. We will focus on the component $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')$, the remaining components can be treated in the same way.

(a) Reduction from six to three variables

Let us consider the ray representation (2.256) of $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')$ which can be rewritten according to

$$
\begin{aligned}
G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}') &= \frac{1}{4\pi} \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} \frac{e^{-jkR_{i,mnp}(\boldsymbol{r},\boldsymbol{r}')}}{R_{i,mnp}(\boldsymbol{r},\boldsymbol{r}')} \qquad (3.116)\\
&= G(|x-x'|, |y-y'|, |z-z'|) + G(|x-x'|, |y-y'|, z+z') \\
&\quad - G(|x-x'|, y+y', |z-z'|) - G(|x-x'|, y+y', z+z') \\
&\quad - G(x+x', |y-y'|, |z-z'|) - G(x+x', |y-y'|, z+z') \\
&\quad + G(x+x', y+y', |z-z'|) + G(x+x', y+y', z+z') \qquad (3.117)
\end{aligned}
$$

with

$$
G(u,v,w) := \frac{1}{4\pi} \sum_{m,n,p=-\infty}^{\infty} \frac{e^{-jkR_{mnp}(u,v,w)}}{R_{mnp}(u,v,w)} \qquad (3.118)
$$

and

$$
R_{mnp} := \sqrt{(u-2ml_x)^2 + (v-2nl_y)^2 + (w-2pl_z)^2}\,. \qquad (3.119)
$$
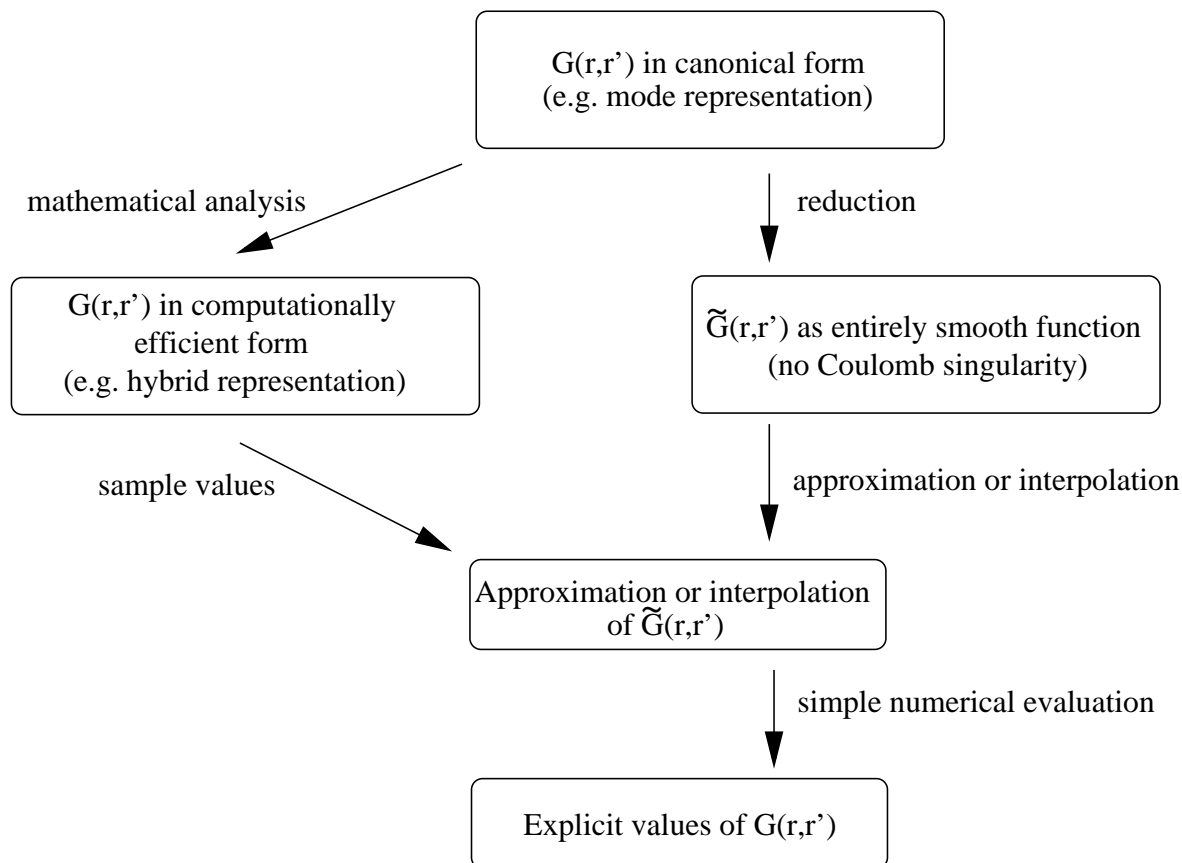
Figure 3.18: Proposed scheme for the efficient evaluation of a cavities Green's function $G(\boldsymbol{r}, \boldsymbol{r}')$.

The mapping $G(u, v, w) : \mathbb{R}^3 \rightarrow \mathbb{C}$ depends on three variables rather than on six variables. Accordingly, it is advantageous to approximate or interpolate $G(u, v, w)$ and calculate $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')$ from (3.117). For the calculation of sample values of $G(u, v, w)$ the following Ewald representation, which follows from (3.113), can be used,

$$G(u, v, w) = \underbrace{G_1(u, v, w)}_{\text{"mode part"}} + \underbrace{G_2(u, v, w)}_{\text{"ray part"}}, \tag{3.120}$$

$$G_1(u, v, w) = \frac{1}{8 l_x l_y l_z} \sum_{m,n,p=-\infty}^{\infty} \frac{e^{-\frac{k_{mnp}^2 - k^2}{4E^2}}}{k_{mnp}^2 - k^2} e^{j(k_x u + k_y v + k_z w)} , \tag{3.121}$$

$$G_2(u, v, w) = \frac{1}{8\pi} \sum_{m,n,p=-\infty}^{\infty} \Big[\frac{e^{jkR_{i,mnp}} \mathrm{erfc}(R_{mnp}E + jk/2E)}{R_{mnp}}$$
$$+ \frac{e^{-jkR_{i,mnp}} \mathrm{erfc}(R_{mnp}E - jk/2E)}{R_{mnp}} \Big] . \tag{3.122}$$

(b) Subtraction of Coulomb singularities and discontinuous derivatives

As is evident from (3.118) the function $G(u, v, w)$ encounters a Coulomb singularity whenever $R_{mnp} = 0$. This condition is fulfilled if and only if

$$u = 2ml_x, \quad v = 2nl_y, \quad w = 2pl_z. \tag{3.123}$$

Since $\boldsymbol{r}$, $\boldsymbol{r}'$ are located inside the rectangular cavity it follows from (3.117) that the variables u, v, and w are defined in the range $0 \le u \le 2l_x$, $0 \le v \le 2l_y$, and $0 \le w \le 2l_z$. Therefore, the condition (3.123) can be met if and only if $m, n, p \in \{0, 1\}$. It follows that in the summation (3.118) there are eight terms which contain a Coulomb singularity. Any of these singular terms is of the form $e^{-jkR_{mnp}}/R_{mnp}$ where R_{mnp} may approach zero. A Taylor expansion yields

$$\frac{e^{-jkR_{mnp}}}{R_{mnp}} = \frac{1}{R_{mnp}} - jk - \frac{k^2 R_{mnp}}{2} + \frac{jk^3 R_{mnp}^2}{6} + \dots \tag{3.124}$$

If the first term $1/R_{mnp}$ is subtracted we obtain a function which is continuous at $R_{mnp} = 0$. Due to the third term $-k^2 R_{mnp}/2$ the first derivative of this function is not continuous [12]. To achieve a better approximation by polynomial functions we also subtract this term. As a result, it is suggested to approximate the function

$$\tilde{G}(u, v, w) = G(u, v, w) - \frac{1}{4\pi} \sum_{m,n,p=0}^{1} \left(\frac{1}{R_{mnp}} - \frac{k^2 R_{mnp}}{2} \right) \tag{3.125}$$

rather than $G(u, v, w)$. From the approximated function $\tilde{G}(u, v, w)$ it is easy to calculate $G(u, v, w)$ by adding the eight terms that are subtracted in (3.125). Finally, the Green's function $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')$ is obtained from $G(u, v, w)$ via (3.117).

3.4 Explicit calculation of current distributions and antenna impedances

In this section we finally turn to the explicit solution of antenna problems. For a given electromagnetic excitation the solution of an antenna problem consists of the calculation of the electric current on one or several antennas. Once the antenna current is known the electromagnetic field can be calculated via the Green's function and integration. Other quantities of physical interest follow from the knowledge of the electric current and the electromagnetic field.

As described in Sec. 3.2, the mathematical equations that need to be solved are integral equations that have been obtained from the formulation of a general antenna problem in terms of an electromagnetic scattering problem. If N antennas are present

we arrive at N coupled integral equations of the form

$$\sum_{j=1}^{N} \left(\mathcal{L}_D^{-1} I_j\right)(\boldsymbol{r}_i) = -E_{\text{tan}}^{\text{inc}}(\boldsymbol{r}_i), \qquad i = 1, \dots, N, \qquad (3.126)$$

where $\mathcal{L}_D^{-1}$ denotes the relevant integral or integro-differential operator, I_j is the current on antenna j, and $\boldsymbol{r}_i$ indicates a position on antenna i. These coupled integral equations have to be solved for the unknown antenna currents I_j.

3.4.1   Choice of integral equations, basis functions, and weighting functions

We focus on linear thin-wire antennas that are located within a rectangular cavity. The choice of a rectangular cavity is natural and canonical since both linear antennas and a rectangular cavity are adapted to a Cartesian coordinate system. Accordingly, we have to decide which of the four standard integral equations of Sec. 3.2 are appropriate in this case. Leaving aside for the moment the possibility of approximate, analytical solutions and concentrating on method of moments solutions we take into account the following points:

- The integral kernel of Pocklington's equation is highly singular and appears to be not very attractive for numerical evaluations. However, in spite of possible numerical instabilities a solution of Pocklington's equation by the method of moments can straightforwardly be implemented. In this case the choice of piecewise sinusoidal basis functions together with delta functions as weighting functions ("point matching") leads to matrix elements (2.272) that have an analytic solution and it follows that no numerical integrations are necessary.

- Among the four integral equations Hallén's equation has the simplest kernel and can straightforwardly be solved by the method of moments. In particular, the choice of piecewise pulse functions as basis functions and delta functions as weighting functions allows to approximately calculate the corresponding matrix elements (2.272) in analytic form. The disadvantage of Hallén's equation is the cumbersome determination of integration constants. However, as long as we consider elementary antenna configurations this determination is manageable.

- The kernel of the mixed-potential integral equation involves a spatial derivative of the unknown antenna current. For a method of moment solution this implies that the basis functions of the unknown current must be differentiable. This criterion excludes the common piecewise basis functions and requires whole domain basis functions. It follows that the calculation of the matrix elements (2.272) requires integrations that extend over the complete antenna. In case of an antenna that is

located in free space the Green's function (2.136) that enters the kernel of the mixed potential is elementary such that it is feasible to numerically perform the required integrations. Within a cavity the Green's function is much more complicated. In this case numerical integration will still be possible but considerably more time-consuming.

- Schelkunoff's equation cannot be applied to the solution of antenna problems within a cavity since it requires the condition (3.90) which is not met within a cavity.

These arguments suggest to make use of Hallén's and Pocklington's equation. They provide the possibility to avoid numerical integration which, in our case, would involve to numerically integrate a series Green's function. This does not mean that such numerical integrations are necessarily inefficient. It just seems more reasonable to first use analytical results before invoking numerical tools. Details of the method of moment solution of Hallén's and Pocklington's equation are described in the following. Essentially, these are standard procedures that are also discussed in the relevant literature [147, 174, 187].

(a) Method of moment solution of coupled Pocklington's equation using point matching and piecewise sinusoidal functions as basis functions

We write Pocklington's equation in the general form (3.71)

$$ j\omega\mu \left[\int_{\Gamma} \overline{\boldsymbol{G}}^{E}(\boldsymbol{r}, \boldsymbol{r}') \boldsymbol{J}_s(\boldsymbol{r}') \, d^2 r' \right] \cdot \boldsymbol{e}_t(\boldsymbol{r}) = E_t^{\text{inc}}(\boldsymbol{r}) \,. \tag{3.127} $$

For N coupled antennas the condition (3.127) must be enforced on each antenna surface such that in this case it constitutes N coupled integral equations. More explicitly, we denote by $\boldsymbol{r}^{(i)}$, $\boldsymbol{r}^{(j)}$ a position vector that points to the surface of antenna i, j, respectively, with $i, j = 1, \ldots, N$. Together with a thin-wire approximation this yields

$$ j\omega\mu \sum_{j=1}^{N} \int_{\text{antenna}\,j} \boldsymbol{e}_t^T(\boldsymbol{r}^{(i)}) \overline{\boldsymbol{G}}^{E}(\boldsymbol{r}^{(i)}, \boldsymbol{r}^{(j)}) \boldsymbol{e}_t(\boldsymbol{r}^{(j)}) I(\boldsymbol{r}^{(j)}) \, d\boldsymbol{r}^{(j)} = E_t^{\text{inc}}(\boldsymbol{r}^{(i)}) \,. \tag{3.128} $$

To simplify notation we assume in the following that the antennas are aligned with one rectangular coordinate axis, say, the z-axis. The general case is obtained from linear combinations of this kind of specialization since the expression $\boldsymbol{e}_t^T(\boldsymbol{r}^{(i)}) \overline{\boldsymbol{G}}^{E}(\boldsymbol{r}^{(i)}, \boldsymbol{r}^{(j)}) \boldsymbol{e}_t(\boldsymbol{r}^{(j)})$ within (3.128) is a linear combination of nine terms of the form $e_{tx_i} G_{x_i x_j}^{E} e_{tx_j}$ where $x_i \in \{x, y, z\}$. With $\boldsymbol{e}_t(\boldsymbol{r}^{(i)}) = \boldsymbol{e}_t(\boldsymbol{r}^{(j)}) = \boldsymbol{e}_z$ the coupled integral equations (3.128) acquire the form

$$ j\omega\mu \sum_{j=1}^{N} \int_{\text{antenna}\,j} G_{zz}^{E}(\boldsymbol{r}^{(i)}, \boldsymbol{r}^{(j)}) I(\boldsymbol{r}^{(j)}) \, d\boldsymbol{r}^{(j)} = E_z^{\text{inc}}(\boldsymbol{r}^{(i)}) \,, \tag{3.129} $$

or

$$-\frac{1}{j\omega\varepsilon}\sum_{j=1}^{N}\int_{\text{antenna}\,j}\left(\frac{\partial^2}{\partial z^2}+k^2\right)G_{zz}^A(\boldsymbol{r}^{(i)},\boldsymbol{r}^{(j)})I(\boldsymbol{r}^{(j)})\,d\boldsymbol{r}^{(j)}=E_z^{\text{inc}}(\boldsymbol{r}^{(i)})\,. \qquad (3.130)$$

In order to solve this Pocklington's integral equation system by the method of moments each current is expanded as a linear combination of basis functions, $I(\boldsymbol{r}^{(i)})=\sum_k\alpha_k^{(i)}I_k(\boldsymbol{r}^{(i)})$. Then we divide each antenna into $2M$ intervals of length $h^{(j)}=L^{(j)}/(2M)$, choose the $2M-1$ points between adjacent intervals as matching points $\boldsymbol{r}_k^{(i)}$, and further take piecewise sinusoidal functions as basis functions, $I_k=S_k$,

$$S_k(\boldsymbol{r}^{(j)})=\begin{cases}\frac{\sin k(z^{(j)}-z_k^{(j)})}{\sin kh^{(j)}}, & \text{if}\quad z_{k-1}^{(j)}\leq z^{(j)}\leq z_k^{(j)}\\[2mm]\frac{\sin k(z_{k+1}^{(j)}-z^{(j)})}{\sin kh^{(j)}}, & \text{if}\quad z_k^{(j)}\leq z^{(j)}\leq z_{k+1}^{(j)}\\[2mm]0, & \text{else}\end{cases}\qquad k=1,\ldots 2M-1\,. \quad (3.131)$$

This yields the algebraic system of equations

$$\sum_{j=1}^{N}\sum_{k=1}^{2M-1}\alpha_k^{(j)}Z_k^{(j)E}(\boldsymbol{r}_l^{(i)})=-j\omega\varepsilon E_z^{\text{inc}}(\boldsymbol{r}_l^{(i)})\,,\quad i=1,\ldots,N,\quad l=1,\ldots,2M-1\quad (3.132)$$

which provides $N\times(2M-1)$ equations for $N\times(2M-1)$ unknowns $\alpha_k^{(j)}$. The matrix elements $Z_k^{(j)E}(\boldsymbol{r}_l^{(i)})$ turn out to be

$$Z_k^{(j)E}(\boldsymbol{r}_l^{(i)})=\frac{k}{\sin kh^{(j)}}\left[G_{zz}^A(\boldsymbol{r}_l^{(i)},\boldsymbol{r}_{k+1}^{(j)})+G_{zz}^A(\boldsymbol{r}_l^{(i)},\boldsymbol{r}_{k-1}^{(j)})-2\cos kh^{(j)}G_{zz}^A(\boldsymbol{r}_l^{(i)},\boldsymbol{r}_k^{(j)})\right]\,. \tag{3.133}$$

Therefore, the calculation of the matrix elements $Z_k^{(j)E}(\boldsymbol{r}_l^{(i)})$ requires no integration. It only involves to compute the Green's function $G_{zz}^A(\boldsymbol{r},\boldsymbol{r}')$ for various arguments.

(b) Method of moment solution of coupled Hallén's equations using point matching and piecewise pulse functions as basis functions

We write Hallén's equation in the general form (3.76)

$$\mu\left[\int_\Gamma\overline{\boldsymbol{G}}^A(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{J}_s(\boldsymbol{r}')\,d^2r'\right]\cdot\boldsymbol{e}_t(\boldsymbol{r})=-A_t^{\text{inc}}(\boldsymbol{r})\,. \tag{3.134}$$

Again, as in the case of Pocklington's equation, we denote by $\boldsymbol{r}^{(i)}$, $\boldsymbol{r}^{(j)}$ a position vector that points to the surface of antenna i, j, respectively, with $i,j=1,\ldots,N$. We also adopt a thin-wire approximation and assume that the antennas are aligned with one rectangular coordinate axis, say, the z-axis. In analogy to (3.129) we arrive at

$$\mu\sum_{j=1}^{N}\int_{\text{antenna}\,j}G_{zz}^A(\boldsymbol{r}^{(i)},\boldsymbol{r}^{(j)})I(\boldsymbol{r}^{(j)})\,d\boldsymbol{r}^{(j)}=-A_z^{\text{inc}}(\boldsymbol{r}^{(i)})\,, \tag{3.135}$$

with $I(\boldsymbol{r}^{(i)})$ the N unknown antenna currents. Again, in order to apply the methods of moments to (3.135) each current is expanded as a linear combination of basis functions, $I(\boldsymbol{r}^{(i)}) = \sum_k \alpha_k^{(i)} I_k(\boldsymbol{r}^{(i)})$. Now we divide each antenna into $2M - 1$ intervals of length $h^{(j)} = L^{(j)}/(2M - 1)$ with $L^{(j)}$ the length of antenna j, choose the midpoints of the intervals as matching points $\boldsymbol{r}_k^{(i)}$, and further take pulse functions as basis functions, $I_k = P_k$, with

$$P_k(\boldsymbol{r}^{(j)}) = \begin{cases} 1 & \text{if} \quad z^{(j)} \in \left[z_k^{(j)} - \frac{h^{(j)}}{2}, z_k^{(j)} + \frac{h^{(j)}}{2} \right] \\ 0 & \text{else} \end{cases}, \qquad k = 1, ..., 2M - 1 . \quad (3.136)$$

This yields the algebraic system of equations

$$\mu \sum_{j=1}^{N} \sum_{k=1}^{2M-1} \alpha_k^{(j)} Z_k^{(j)A}(\boldsymbol{r}_l^{(i)}) = -A_z^{\text{inc}}(\boldsymbol{r}_l^{(i)}), \qquad i = 1, \dots, N, \quad l = 1, \dots, 2M-1 \quad (3.137)$$

which provides $N \times (2M - 1)$ equations for $N \times (2M - 1)$ unknowns $\alpha_k^{(j)}$. The matrix elements $Z_k^{(j)A}(\boldsymbol{r}_l^{(i)})$ are given by

$$Z_k^{(j)A}(\boldsymbol{r}_l^{(i)}) = \int_{z_k^{(j)} - \frac{h^{(j)}}{2}}^{z_k^{(j)} + \frac{h^{(j)}}{2}} G_{zz}^A(\boldsymbol{r}_l^{(i)}, \boldsymbol{r}^{(j)}) \, d\boldsymbol{r}^{(j)} \quad (3.138)$$

$$\approx G_{zz}^A(\boldsymbol{r}_l^{(i)}, \boldsymbol{r}_k^{(j)})|\boldsymbol{r}_l^{(i)} - \boldsymbol{r}_k^{(j)}| \int_{z_k^{(j)} - \frac{h^{(j)}}{2}}^{z_k^{(j)} + \frac{h^{(j)}}{2}} \frac{1}{|\boldsymbol{r}_l^{(i)} - \boldsymbol{r}^{(j)}|} d\boldsymbol{r}^{(j)} . \quad (3.139)$$

In the last line we isolated the Coulomb singularity of the Green's function. Due to the thin-wire approximation the distance $|\boldsymbol{r}_l^{(i)} - \boldsymbol{r}^{(j)}|$ is bounded from below by the antenna radius ρ. The integral of (3.139) can be calculated analytically,

$$\int_{z_k^{(j)} - \frac{h^{(j)}}{2}}^{z_k^{(j)} + \frac{h^{(j)}}{2}} \frac{1}{|\boldsymbol{r}_l^{(i)} - \boldsymbol{r}^{(j)}|} d\boldsymbol{r}^{(j)} = \int_{z_k^{(j)} - \frac{h^{(j)}}{2}}^{z_k^{(j)} + \frac{h^{(j)}}{2}} \frac{1}{\sqrt{\rho^2 + (z_l^{(i)} - z^{(j)})^2}} \, dz^{(j)} \quad (3.140)$$

$$= \ln \left[\frac{z_k^{(j)} - z_l^{(i)} + \frac{h^{(j)}}{2} + \sqrt{\rho^2 + (z_k^{(j)} - z_l^{(i)} + \frac{h^{(j)}}{2})^2}}{z_k^{(j)} - z_l^{(i)} - \frac{h^{(j)}}{2} + \sqrt{\rho^2 + (z_k^{(j)} - z_l^{(i)} - \frac{h^{(j)}}{2})^2}} \right] .$$

$$(3.141)$$

Similar to (3.133) the calculation of the matrix elements $Z_k^{(j)A}(\boldsymbol{r}_l^{(i)})$ requires no integration. It only involves to compute the Green's function $G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}')$ for various arguments.

3.4.2   Calculation of self impedances

In this subsection we calculate self impedances of dipole antennas within a rectangular cavity. To calculate antenna impedances is not just a special aspect of antenna theory, it rather requires the *complete* solution of an antenna problem which has an unknown antenna current as solution. Such a complete solution also allows to calculate other quantities of interest, like the electromagnetic field at some point in space, for example. We focus on the calculation of antenna impedances since these quantities concisely characterize properties of both antennas and their electromagnetic environment.

(a) Cavity properties

We choose a cavity of dimensions $l_x = 6$m, $l_y = 7$m, and $l_z = 3$m. This cavity geometry corresponds to a mode-stirred chamber at the University of Magdeburg. The quality factor of this chamber was experimentally measured and the result compared to a simple but well-established model [113]. On the basis of this study we describe ohmic losses in the cavity walls by the approximate formula $Q(f) = 0.1\sqrt{f/\text{Hz}}$; this yields $Q(100\text{MHz}) = 1000$, for example. According to the examples of Sec. 2.3.1, cavity resonances are characterized by three integers m, n, and p via the wavenumber

$$k_{mnp}^2 = k_x^2 + k_y^2 + k_z^2 \tag{3.142}$$

$$= \left(\frac{m\pi}{l_x}\right)^2 + \left(\frac{n\pi}{l_y}\right)^2 + \left(\frac{p\pi}{l_z}\right)^2 \tag{3.143}$$

with corresponding resonance frequencies

$$f_{mnp} = \frac{1}{2\sqrt{\varepsilon\mu}}\sqrt{\left(\frac{m}{l_x}\right)^2 + \left(\frac{n}{l_y}\right)^2 + \left(\frac{p}{l_z}\right)^2}. \tag{3.144}$$

We assume that the cavity is filled with air. The corresponding values of the lowest thirty resonance frequencies f_{mnp} of the cavity are displayed in Tab. 3.2.

(b) Antenna properties

We consider perfectly conducting, cylindrical, and straight dipole antenna of wire radius $\rho = 10^{-3}$m. The antenna lengths $L_1 = L_2 = L$ will be chosen as $L = 1$m, $L = 2$m, or $L = 0.2$m such that the thin-wire approximation can be applied. Furthermore, we assume that the antennas are aligned to one of the coordinate axis, say, the z-axis, compare Fig. 3.19. In principle, arbitrary orientations can be considered as well. Due to linearity, an antenna problem which involves arbitrary orientations can be reduced to antenna problems with canonical antenna orientations that are in parallel to one of the coordinate axes. Here we are mainly interested to work out the basic physical mechanisms of antenna coupling and thus focus on canonical cases.

mnp	f_{mnp}	mnp	f_{mnp}	mnp	f_{mnp}
110	32.90 MHz	201	70.66 MHz	311	92.59 MHz
120	49.58 MHz	211	73.83 MHz	231	95.50 MHz
011	54.36 MHz	310	77.95 MHz	330	98.71 MHz
210	54.36 MHz	031	81.38 MHz	041	99.16 MHz
101	55.86 MHz	230	81.38 MHz	240	99.16 MHz
111	59.82 MHz	221	82.63 MHz	321	99.73 MHz
021	65.81 MHz	131	85.13 MHz	012	102.20 MHz
220	65.81 MHz	320	86.32 MHz	410	102.20 MHz
130	68.93 MHz	140	89.22 MHz	141	102.26 MHz
121	70.39 MHz	301	90.08 MHz	102	103.00 MHz

Table 3.2: The lowest thirty resonances and their corresponding frequencies of a cavity of dimensions $l_x = 6$m, $l_y = 7$m, and $l_z = 3$m.
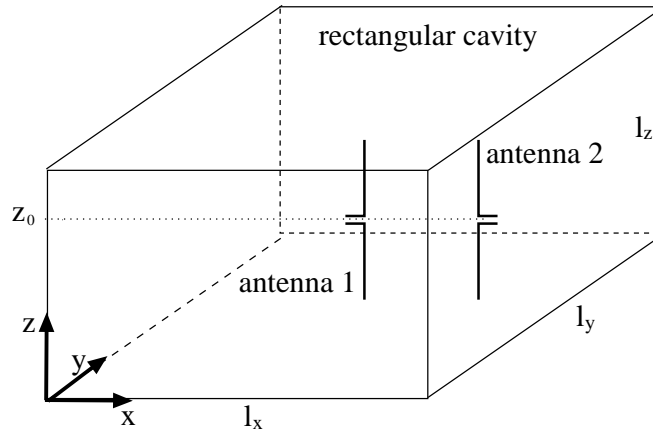
Figure 3.19: Two straight dipole antenna which are aligned to the z-axis and placed within a rectangular cavity of dimensions l_x, l_y, and l_z.

(c) Coupling between antennas and cavity

It follows from (3.40) or (3.55) that the electromagnetic coupling of an antenna current $\boldsymbol{J}$ to the electromagnetic field within a cavity is determined from the coupling coefficients

$$\alpha_{mnp}^A = \mu \frac{\langle \boldsymbol{J}, \boldsymbol{f}_{mnp}^A \rangle}{k^2 - k_{mnp}^2} \,. \tag{3.145}$$

or

$$\alpha_{mnp}^{E\parallel} = \frac{j\omega\mu}{k^2} \langle \boldsymbol{J}, \boldsymbol{L}_{mnp} \rangle \,, \qquad \alpha_{mnp}^{E\perp} = j\omega\mu \frac{\langle \boldsymbol{J}, \boldsymbol{F}_{mnp} \rangle}{k^2 - k_{mnp}^2} \,. \tag{3.146}$$

Up to a constant factor, each coefficient is the projection of an antenna current onto an eigenfunction which is weighted by $\frac{1}{k^2 - k_{mnp}^2}$ or $\frac{1}{k^2}$. If, in particular, this projection is zero

the antenna will not couple to the corresponding eigenmode.

Let us consider, as an example, an antenna 1 which is aligned to the z-axis, see Fig. 3.19. The most symmetric positioning is obtained if the center of the antenna is placed at $\boldsymbol{r}_0 = (l_x/2, l_y/2, l_z/2) = (3, 3.5, 1.5)$m. In this case we find for the projection of the antenna current $\boldsymbol{J}(\boldsymbol{r}) = I(z)\delta(x - x_0)\delta(y - y_0)\boldsymbol{e}_z$ onto the eigenfunction $\boldsymbol{f}_{mnp}^A$ the expression

$$\langle \boldsymbol{J}, \boldsymbol{f}_{mnp}^A \rangle = \langle J_z, f_{mnp}^{zA} \rangle \tag{3.147}$$

$$= \int_{z_0-L/2}^{z_0+L/2} I(z') f_{mnp}^{zA}(x_0, y_0, z') \, dz' \tag{3.148}$$

$$= \sqrt{\frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z}} \int_{z_0-L/2}^{z_0+L/2} I(z') \sin\left(\frac{m\pi}{l_x}x_0\right) \sin\left(\frac{n\pi}{l_y}y_0\right) \cos\left(\frac{p\pi}{l_z}z'\right) \, dz' \tag{3.149}$$

$$= \sqrt{\frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z}} \sin\left(\frac{m\pi}{2}\right) \sin\left(\frac{n\pi}{2}\right) \int_{z_0-L/2}^{z_0+L/2} I(z') \cos\left(\frac{p\pi}{l_z}z'\right) \, dz' . \tag{3.150}$$

This projection vanishes for even m and n. It also vanishes if both p is odd and the antenna is symmetrically excited around its center. In this case $\cos(\frac{p\pi}{l_z}z')$ is an odd function with respect to $z_0 = l_z/2$ and, due to the symmetry of the problem, $I(z)$ is an even function with respect $z_0 = l_z/2$ such that the integral of (3.150) vanishes. It follows, for example, that for a symmetric excitation and in the frequency range 25 MHz to 100 MHz the antenna will only couple to the modes 110 (32.9 MHz), 130 (68.9 MHz), 310 (78.0 MHz), and 330 (98.7 MHz), compare Tab. 3.2.

As a second example we consider an antenna 2 which is aligned to the z-axis with the center of the antenna placed at $\boldsymbol{r}_0 = (3l_x/4, l_y/2, l_z/2) = (4.5, 3.5, 1.5)$m, see Fig. 3.19. Then we find for the projection of the antenna current $\boldsymbol{J}(\boldsymbol{r}) = I(z)\delta(x - x_0)\delta(y - y_0)\boldsymbol{e}_z$ onto the eigenfunction $\boldsymbol{f}_{mnp}^A$ the expression

$$\langle \boldsymbol{J}, \boldsymbol{f}_{mnp}^A \rangle = \sqrt{\frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z}} \sin\left(\frac{3m\pi}{4}\right) \sin\left(\frac{n\pi}{2}\right) \int_{z_0-L/2}^{z_0+L/2} I(z') \cos\left(\frac{p\pi}{l_z}z'\right) \, dz' . \tag{3.151}$$

This projection vanishes for m an integer multiple of 4 and n even. It also vanishes if both p is odd and the antenna is symmetrically excited around its center. It follows for a symmetric excitation and in the frequency range 25 MHz to 100 MHz that the antenna will couple to the modes 110 (32.9 MHz), 210 (54.4 MHz), 130 (68.9 MHz), 310 (78.0 MHz), 230 (81.4 MHz), and 330 (98.7 MHz), compare Tab. 3.2.

(d) Self impedances from method of moment solutions of Pocklington's and Hallén's equation

We now calculate the self impedance of antenna 1 and antenna 2, respectively. To this end we choose the antenna lengths as $L = 1$m and proceed along the lines of Sec. 3.4.1

to numerically solve both Pocklington's and Hallén's equation within the rectangular cavity. The antennas will be excited by a delta-gap generator which is placed at the center of each antenna. Mathematically, a delta-gap generator is modeled by an incident electromagnetic field of the form

$$\boldsymbol{E}_t^{\mathrm{inc}}(z) = V_0 \delta(z - z_0) \qquad (3.152)$$

where z_0 indicates the position of the delta-gap generator. It has been shown by Rynne that the model of a delta-gap generator leads to Pocklington's and Hallén's equations that mathematically are well-posed [186]. From a physical point of view a distributional incident field represents a drastic idealization, though. In a numerical scheme the delta peak can be approximated by a pulse function $P_{k_{z0}}(z)$ where k_{z0} labels the interval of length h where the delta peak is located, compare (3.136). We then have the approximation

$$\delta(z - z_0) \approx \frac{P_{k_{z0}}(z)}{h} \, . \qquad (3.153)$$

With a fixed excitation of the form (3.152) the solution of Pocklington's and Hallén's equation is reduced to the calculation of the dyadic Green's function component G_{zz}^A, see (3.133) and (3.139). To calculate the various values of this component we use the interpolation scheme of Sec. 3.3.3. As interpolating functions we take cubic splines. The required sample values are obtained from the Ewald representation (3.120).

For the calculation we have to fix the parameter M which determines the fineness of the discretization. We first choose $M = 60$, i.e., for the solution of Pocklington's equation the antenna is divided into $2M = 120$ intervals and for the solution of Hallén's equation the antenna is divided into $2M - 1 = 119$ intervals. This yields approximate solutions for the unknown antenna current. Then the exciting voltage V_0, which can be taken as 1V, is divided by the value of the antenna current at the antenna center to yield the self impedance. The results for the real and imaginary parts of the self impedances Z_{self} of the single antennas 1 and 2 are shown in Figs. 3.20, 3.21, 3.22, and 3.23, respectively.
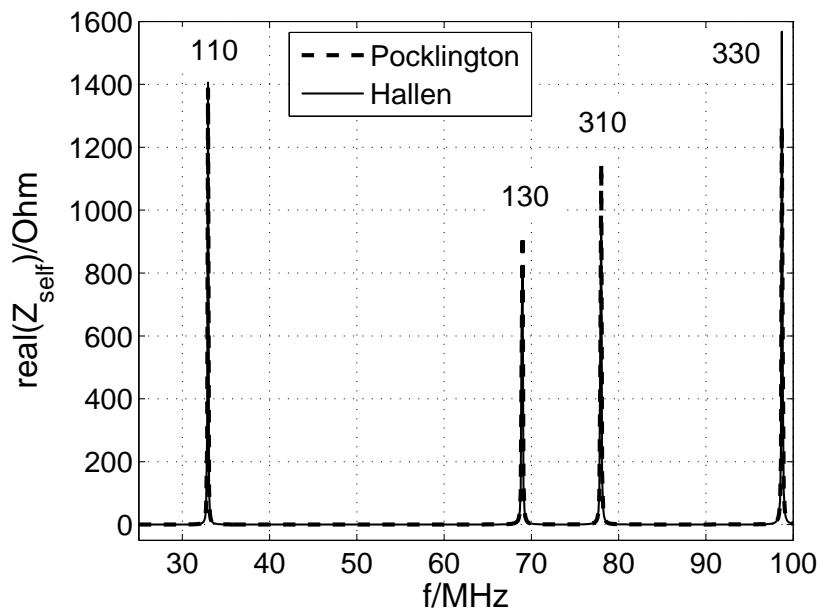
Figure 3.20: Real part of the self impedance of antenna 1. The characteristics of a resonance curve are recognized whenever the antenna couples to a cavity mode.
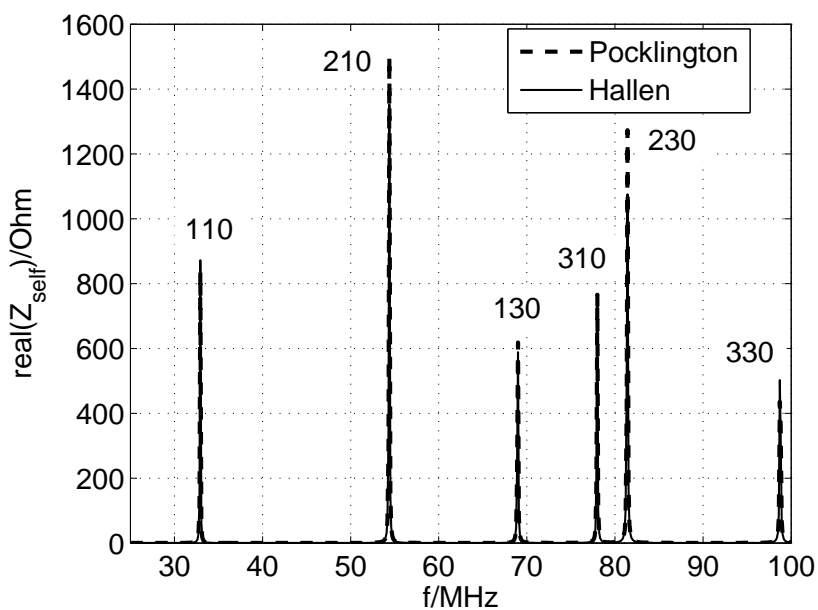
Figure 3.21: Real part of the self impedance of antenna 2. Similar to Fig. 3.20 we recognize resonance peaks whenever antenna 2 couples to a cavity mode. Compared to antenna 1 there are additional couplings to the modes 210 and 230.
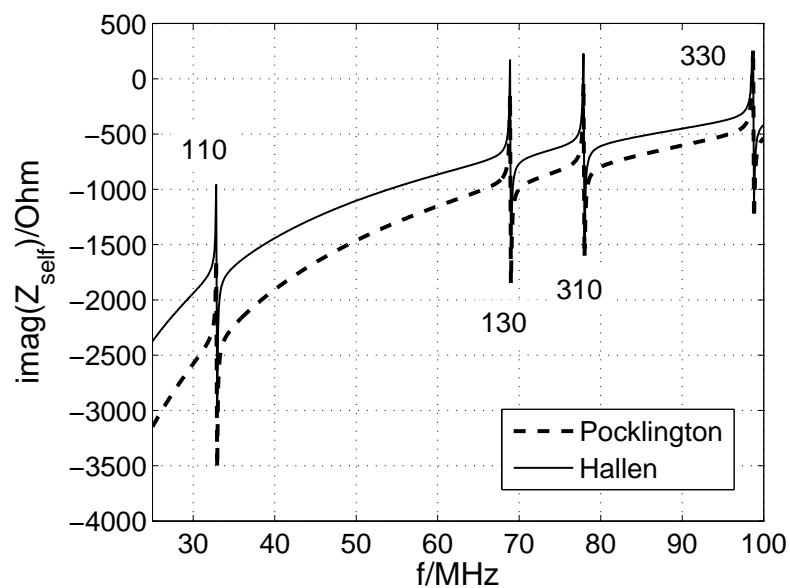
Figure 3.22: Imaginary part of the self impedance of antenna 1. Again, the characteristics of a resonance curve are recognized whenever the antenna couples to a cavity resonance.
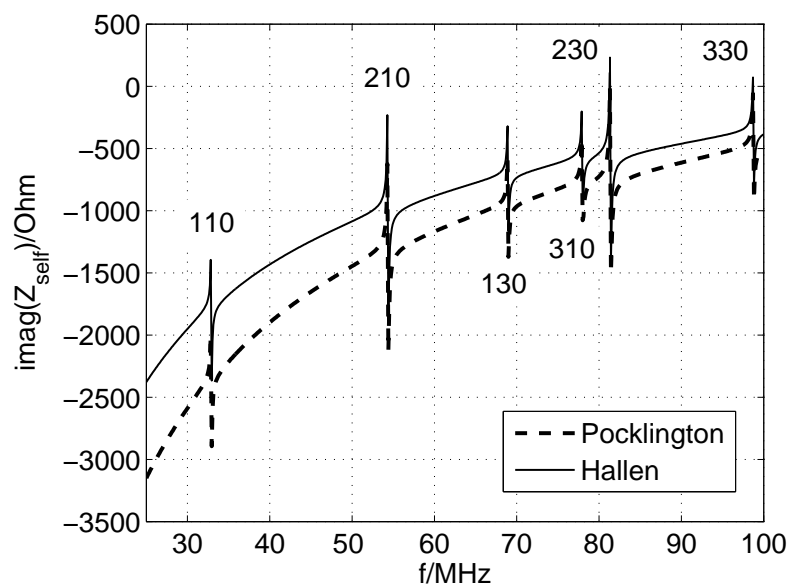
Figure 3.23: Imaginary part of the self impedance of antenna 2. It is analogous to that of antenna 1 in Fig. 3.22 except that antenna 2 couples to the additional modes 210 and 230.

The real part of the self impedances is characterized by sharp resonance peaks that occur if and only if the antenna couples to a cavity resonance. This follows from the discussion of Sec. 3.1.4: If we consider the expansion (3.46) of the self impedance we find, in our example, for the nominator of each summand

$$\langle \boldsymbol{J}, \boldsymbol{f}_{mnp}^A \rangle \langle \boldsymbol{J}, \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{f}_{mnp}^A) + k^2 \boldsymbol{f}_{mnp}^A \rangle = \langle J_z, f_{mnp}^{zA} \rangle \langle J_z, \frac{\partial^2 f_{mnp}^{zA}}{\partial z^2} + k^2 f_{mnp}^{zA} \rangle \qquad (3.154)$$

$$= (k^2 - k_z^2) \langle J_z, f_{mnp}^{zA} \rangle^2 \qquad (3.155)$$

Then the expression (3.46) for the self-impedance reduces to

$$Z_{\text{self}} = \frac{j\omega\mu(k_z^2 - k^2)}{I^2 k^2} \sum_n \frac{\langle J_z, f_{mnp}^{zA} \rangle^2}{k^2 - k_{mnp}^2}, \qquad (3.156)$$

that is, we obtain pole contributions for $k \to k_{mnp}$ as long as $\langle \boldsymbol{J}, \boldsymbol{f}_{mnp}^A \rangle = \langle J_z, f_{mnp}^{zA} \rangle$ is nonvanishing. The shapes of the resonance peaks are recognized as the real part of a resonance curve, see Fig. 3.6. The small bandwidth of the resonance peaks of Figs. 3.20 and 3.21 is due to the comparatively high quality factor of the cavity. It is also seen that the solutions of Pocklington's and Hallén's equation agree well.

Also the imaginary part of the self impedances exhibits sharp resonance peaks that are of the same shape than the imaginary part of the resonance curve in Fig. 3.6. Inbetween two separate resonances the value of the self impedance approaches that of free space. It is seen that the solutions of Pocklington's and Hallén's equation exhibit the same qualitative behavior, however, the values of the two solutions do not agree well. This disagreement is not a consequence of cavity effects. It also occurs in free space and is rooted in the strong singularity of the kernel of Pocklington's equation which requires a comparatively fine discretization in order to correctly quantify the Coulomb interaction within the method of moments. This is a known feature, see [42], for example, which is exemplified by Fig. 3.24 where the imaginary part of the input impedance of antenna 1 and antenna 2 in free space is plotted, as obtained by the solution of Hallén's equation and Pocklington's equation for various values of the parameter M. It is recognized that for finer discretizations the solution of Pocklington's equation approaches that of Hallén's equation.

This situation is somehow unfortunate: In the application of Pocklington's equation an accurate modeling of the Coulomb singularity requires a fine discretization and, in turn, to calculate within the method of moments an impedance matrix with a large number of entries. However, due to the resonance effects within the cavity the calculation of the matrix elements is time consuming and it is desirable to keep the impedance matrix as small as possible. As a consequence, it is desirable to separate in the calculation Coulomb singularity and resonance effects. This is exactly what the method of analytical regularization can achieve which has been introduced in Sec. 3.2.6.
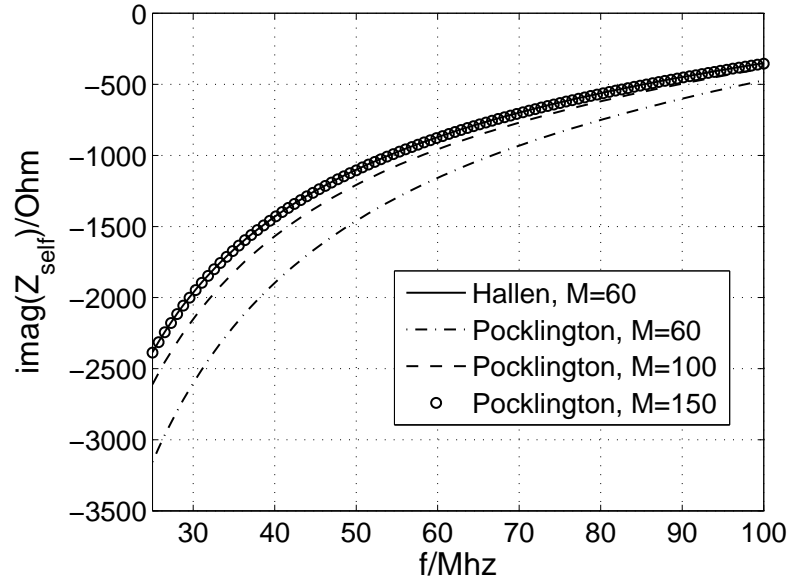
Figure 3.24: Imaginary part of the input impedance of the identical antennas 1 and 2 in free space. For finer discretizations the solution of Pocklington's equation approaches that of Hallén's equation.

(e) Regularization of Pocklington's equation

The discretization of Pocklington's equation leads to a system of algebraic equations that is of the form (3.132). Symbolically, we write this system of equations as

$$[Z^E][I] = -[E] \tag{3.157}$$

with $[Z^E]$ the known impedance matrix, $[I]$ the unknown sample values of the antenna current, and $[E]$ the known sample values of the electromagnetic excitation. This is the discrete analogue of an operator equation of the form (3.98) which is the starting point for the method of regularization. The impedance matrix $[Z^E]$ can be split into a free space part $[Z_0^E]$ which is calculated from the Green's function of free space and a remainder $[Z_1^E]$ which is calculated from the difference between the cavities Green's function and the Green's function of free space,

$$[Z^E] = [Z_0^E] + [Z_1^E]. \tag{3.158}$$

The free space solution $[I_0]$ is simply obtained by the inversion of $[Z_0^E]$,

$$[I_0] = -[Z_0^E]^{-1}[E]. \tag{3.159}$$

Similar to (3.103) it follows that the unknown antenna current $[I]$ is determined from the algebraic equation

$$[I] = [I_0] - [Z_0^E]^{-1}[Z_1^E][I] \tag{3.160}$$

with solution

$$[I] = \left([1] + [Z_0^E]^{-1}[Z_1^E]\right)^{-1}[I_0].$$ (3.161)

Here we introduced the unit matrix $[1]$.

The benefit of this procedure is obvious: The singularity of the Coulomb interaction is accounted for in the free space solution (3.159). The time to compute $[Z_0^E]$ is negligible if compared to the time to compute $[Z_1^E]$. Therefore, it poses no problems to obtain $[I_0]$ for a fine discretization. Then the unknown antenna current $[I]$ can be calculated from (3.161) with a more coarse discretization.

As an example we repeat the calculation of the self impedance of antenna 1 that led to Fig. 3.22 but now regularize Pocklington's equation as outlined above. To obtain the free space solution I_0 from (3.159) we choose a fine discretization with $M = 160$ and to subsequently calculate I via (3.161) we reduce the fineness to $M = 40$. The result is displayed in Fig. 3.25 and compared to the solution of Hallén's equation. It is obvious that the agreement between the different solution procedures is much better now.
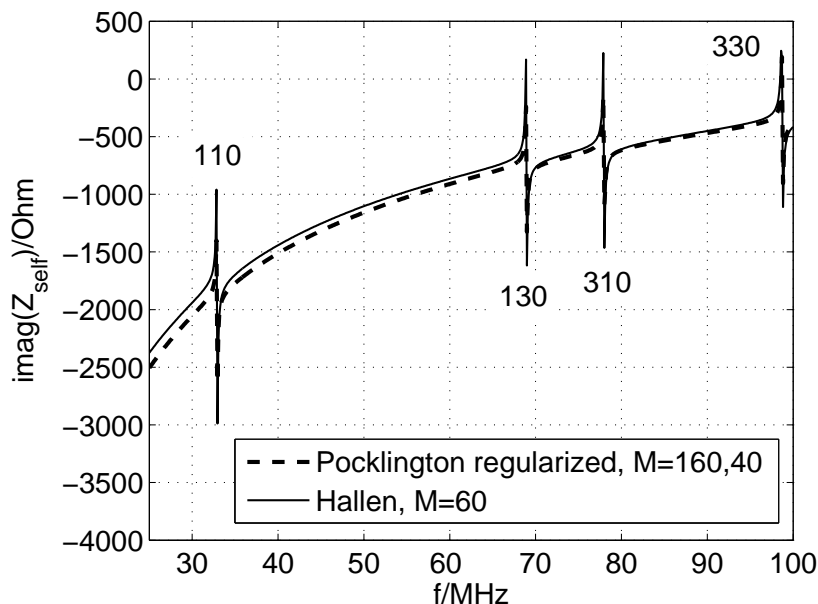
Figure 3.25: Imaginary part of the self impedance of antenna 1. The curve that shows the solution of Hallén's equation is the same as in Fig. 3.22. Now the solution of Pocklington's equation agrees much better since the method of analytical regularization was used.

In this connection it is worth mentioning that a criterion for the stability of a method of moments calculation is provided by the so-called *condition number* cond($[Z]$) of the impedance matrix $[Z]$ [79, 90]. The condition number can be introduced via the infinity

	$[Z^A]$ (M=60)	$[Z^E]$ (M=60)	$[Z_0^E]$ (M=160)	$[1] + [Z_0^E]^{-1}[Z_1^E]$ (M=40)
cond$_\infty$ at 25.0 MHz	7.311	3743	2107	1.009
cond$_\infty$ at 32.9 MHz	262.1	4790	2157	5.661

Table 3.3: Various condition numbers of matrices that need to be inverted to solve Hallén's, Pocklington's, and the regularized Pocklington's equation.

norm of an $n \times n$ matrix. This norm is defined by

$$\|[Z]\|_\infty := \max_{i=1,2,\ldots,n} \sum_{j=1}^{n} |z_{ij}|, \tag{3.162}$$

that is, this norm is the maximum of the sums obtained from adding the absolute values of the elements of $[Z]$ in each row. Then, in turn, the condition number is defined by

$$\text{cond}_\infty([Z]) := \|[Z]\|_\infty \|[Z]^{-1}\|_\infty \tag{3.163}$$

where we assigned the index ∞ to the condition number in order to indicate that it refers to the infinity norm (3.162)[7]. The condition number is bounded below by 1. High condition numbers indicate that the solution of the associated linear system of equations is sensitive towards errors of the given data.

For illustration we display in Tab. 3.3 condition numbers of the matrices that have to be inverted to yield solutions of Hallén's and Pocklington's equation. We choose from the results that are displayed in Figs. 3.22 and 3.25 the frequencies 25.0 MHz (off resonance) and 32.9 MHz (at resonance) and compute the condition numbers of the matrices $[Z^A]$ (Hallén), $[Z^E]$ (Pocklington), $[Z_0^E]$ (regularized Pocklington), and $[1] + [Z_0^E]^{-1}[Z_1^E]$ (regularized Pocklington). It is seen that the matrix $[Z^A]$ is characterized by low condition numbers if compared to $[Z^E]$, even though the condition numbers increase at resonance. The matrix $[Z_0^E]$ is calculated from the Green's function of free space and thus insensitive towards cavity resonances. The condition numbers of the regularized matrix $[1] + [Z_0^E]^{-1}[Z_1^E]$ are low and indicate the high stability of the regularized Pocklington's equation.

3.4.3 Calculation of mutual impedances between dipole antennas

The same methods that have been used for the calculation of antenna self impedances apply to the calculation of mutual impedances between different antennas. Of particular

---

[7]It is also common to define condition numbers with respect to other matrix norms. Common choices are the 1-norm, 2-norm, or the Frobenius norm [79].

importance are the integral expressions for antenna impedances that have been introduced in Sec. 3.1.3. We will take advantage of the integral expression for the mutual impedance Z_{12} that is given by (3.31). It requires to calculate electromagnetic quantities in the "a-situation" and the "b-situation" that are displayed in Fig. 3.4. In the "a-situation" we have to solve coupled Pocklington's and Hallén's equations that are defined on the surface of two antennas while in the "b-situation" only a single Pocklington's or Hallén's equation needs to be solved. The accuracy of the solution of Pocklington's equation will be improved by the method of regularization, as described above.

(a) Parallel antennas

We first investigate antennas that are in parallel and focus on the configuration that is shown in Fig. 3.19. Antenna 1, antenna 2, and the cavity are chosen as in the previous subsection with the exception that the lengths of the antennas now are enlarged to $L = 2$m in order to have larger mutual coupling effects. We recall that antenna 1 couples to the cavity modes 110, 130, 310, and 330, while, antenna 2 couples to the same modes and, additionally, to the modes 210 and 230.

In Fig. 3.26 the absolute value of the mutual impedance Z_{12} is shown on a logarithmic scale. The two curves that have independently been obtained by the solution of regularized Pocklington's and Hallén's equations agree well. It is observed that the mutual impedance becomes large whenever both antennas simultaneously couple to a common mode. This seems plausible from physical intuition and follows from the expansion (3.45) of the mutual impedance. The nominator of each summand has, in our example, the form

$$\langle \boldsymbol{J}^b, \boldsymbol{f}_{mnp}^A \rangle \langle \boldsymbol{J}^a, \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{f}_{mnp}^A) + k^2 \boldsymbol{f}_{mnp}^A \rangle = \langle J_z^b, f_{mnp}^{zA} \rangle \langle J_z^a, \frac{\partial^2 f_{mnp}^{zA}}{\partial z^2} + k^2 f_{mnp}^{zA} \rangle \quad (3.164)$$

$$= (k^2 - k_z^2) \langle J_z^b, f_{mnp}^{zA} \rangle \langle J_z^a, f_{mnp}^{zA} \rangle \quad (3.165)$$

It follows that the expression (3.45) for the self-impedance reduces to

$$Z_{12} = \frac{j\omega\mu(k^2 - k_z^2)}{I_2^a I_1^b k^2} \sum_n \frac{\langle J_z^b, f_{mnp}^{zA} \rangle \langle J_z^a, f_{mnp}^{zA} \rangle}{k^2 - k_{mnp}^2}, \quad (3.166)$$

such that dominating pole contributions are obtained if both $\langle J_z^a, f_{mnp}^{zA} \rangle$ and $\langle J_z^b, f_{mnp}^{zA} \rangle$ are nonvanishing. Hence, the primary resonance peaks 110, 130, 310, and 330 of Fig. 3.26 are explained by the sum of (3.166) which, up to a constant factor, constitutes the reaction between both antennas. However, the reaction does not explain the minor resonance peaks 210 and 230 that are observed if only antenna 2 couples to a mode. To explain these peaks we plot in Fig. 3.27 the absolute value of the electric current $I_2^a(z)$ on antenna 2 for three different frequencies that are in the vicinity of the resonance frequency of mode 210. The first, 51 MHz, is below resonance, the second, 54 MHz, is at resonance, and the third, 57 MHz is above resonance. While for 51 MHz and

57 MHz the current distribution is similar to a corresponding free space solution the current distribution drastically changes as the antenna couples to the cavity mode 210. One should note that the first natural antenna resonance (not cavity resonance!) of antenna 2 in free space occurs at 150 MHz, since in this case $\lambda = L_2 = 2$m. Within the cavity we observe a current distribution which resembles that of a natural antenna resonance at a much lower frequency. That is, due to the coupling of antenna 2 to a cavity mode the value of the input current I_2^a, which enters the formula (3.166) of the mutual impedance, drastically changes, and this is the reason for the secondary resonance peak 210 that is observed in the plot of the mutual impedance. The same conclusion is valid for the secondary peak 230.

The real and imaginary parts of the mutual impedance are displayed in Fig. 3.28 and Fig. 3.29, respectively. The shape of the resonance peaks is explained by the resonance curves of Fig. 3.5 and also familiar from the calculated self-impedances of the last subsection. In contrast to the real part of the self impedance, which must be strictly positive, the real part of the mutual impedance can become negative. This is due to the oscillations of the electromagnetic Green's function and a phenomenon which is also present in free space. The sign reversal is also reflected in the resonance peaks of the imaginary part of the mutual impedance. On the linear scale the secondary resonance peaks of the modes 210 and 230 can hardly be seen.
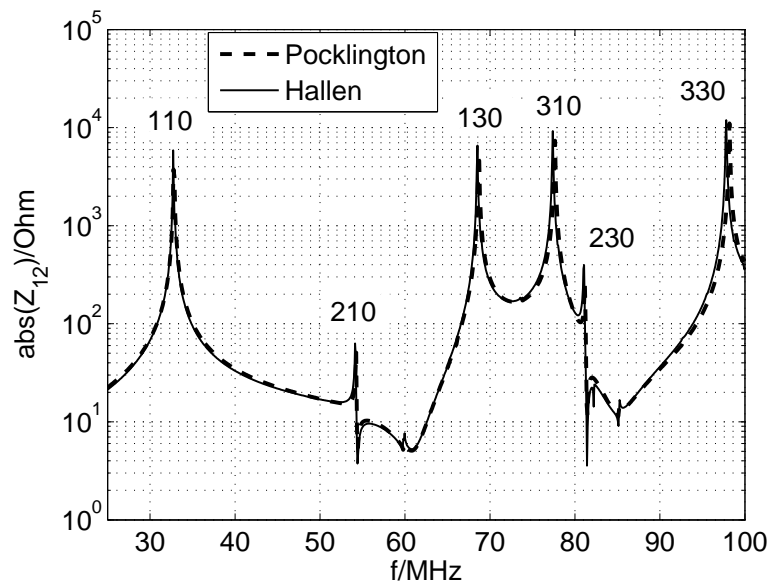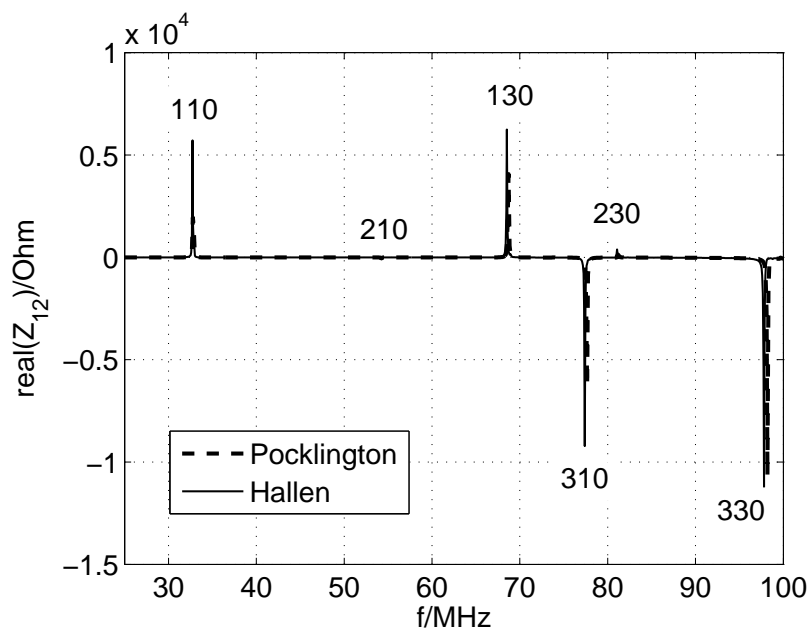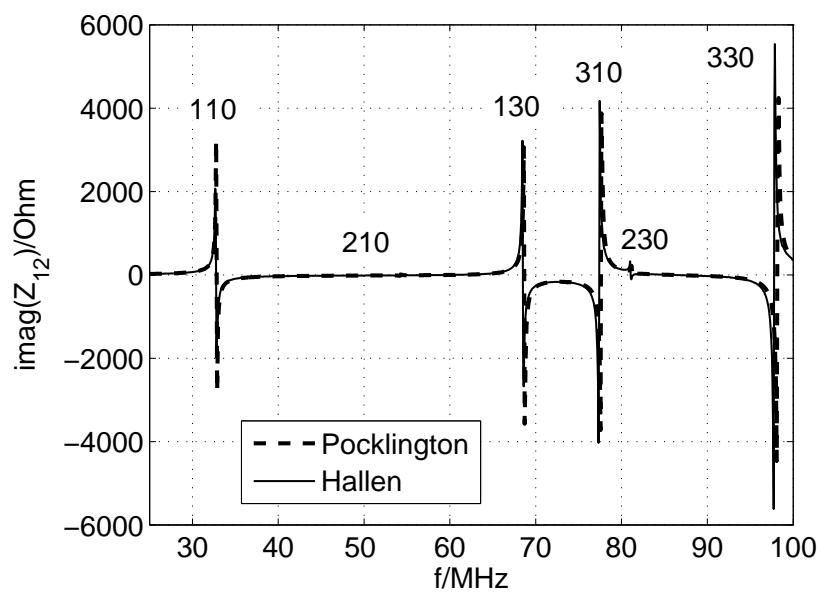

Figure 3.26: Absolute value of the mutual impedance between antenna 1 and antenna 2 in the frequency range of 25 MHz to 100 MHz. The antennas are in parallel.
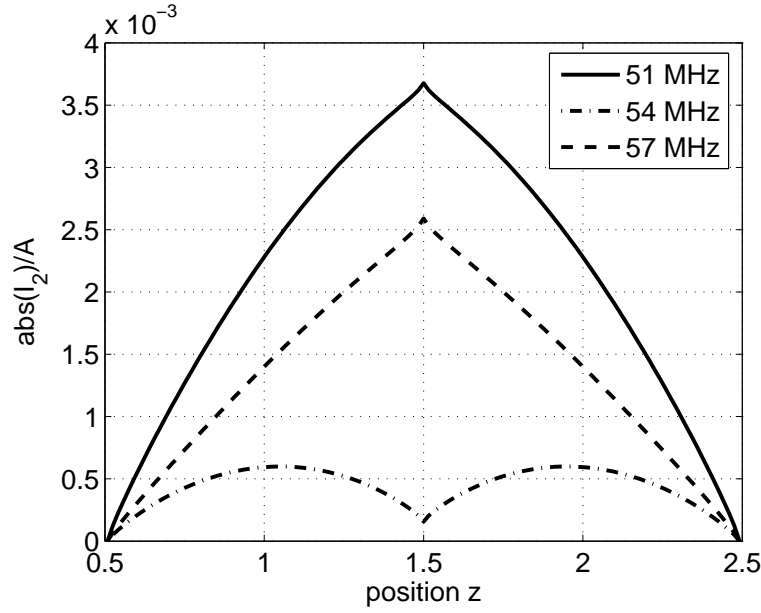
Figure 3.27: Absolute value of the current distribution $I_2^a(z)$ below (51 MHz), at (54 MHz), and above (57 MHz) the cavity resonance 210. While the resonance is passed the current distribution drastically changes and assumes the shape of a natural antenna resonance in free space.

(b) Cross polarized antennas

For further illustration we now consider cross polarized antennas. We start from the configuration Fig. 3.19 and turn antenna 1 by ninety degrees in the yz-plane. This yields the configuration of Fig. 3.30. Since antenna 2 is unchanged it still couples to the modes with m odd, n odd, and p even, i.e., it still couples to the modes 110 (32.9 MHz), 210 (54.4 MHz), 130 (68.9 MHz), 310 (78.0 MHz), 230 (81.4 MHz), and 330 (98.8 MHz). Antenna 1 is now aligned to the y-axis. Similar to (3.150) the projection of the antenna current $\boldsymbol{J}(\boldsymbol{r}) = I(y)\delta(x - x_0)\delta(z - z_0)\boldsymbol{e}_y$ onto the eigenfunction $\boldsymbol{f}_{mnp}^A$ yields the expression

$$\langle \boldsymbol{J}, \boldsymbol{f}_{mnp}^A \rangle = \langle J_y, f_{mnp}^{yA} \rangle \tag{3.167}$$

$$= \int_{y_0-L/2}^{y_0+L/2} I(y') f_{mnp}^{yA}(x_0, y', z_0)\, dy' \tag{3.168}$$

$$= \sqrt{\frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z}} \int_{y_0-L/2}^{y_0+L/2} I(y') \sin\left(\frac{m\pi}{l_x}x_0\right) \cos\left(\frac{n\pi}{l_y}y'\right) \sin\left(\frac{p\pi}{l_z}z_0\right)\, dz' \tag{3.169}$$

$$= \sqrt{\frac{\epsilon_{0m}\epsilon_{0n}\epsilon_{0p}}{l_x l_y l_z}} \sin\left(\frac{m\pi}{2}\right) \sin\left(\frac{p\pi}{2}\right) \int_{y_0-L/2}^{z_y+L/2} I(y') \cos\left(\frac{n\pi}{l_y}y'\right)\, dy'. \tag{3.170}$$

Figure 3.28: Real part of the mutual impedance between antenna 1 and antenna 2 in the frequency range of 25 MHz to 100 MHz.

Figure 3.29: Imaginary part of the mutual impedance between antenna 1 and antenna 2 in the frequency range of 25 MHz to 100 MHz.

This projection vanishes for even m, p, and, for a symmetric excitation, odd n. Therefore, antenna 1 now couples to modes with m odd, n even, and p odd. It follows that up to 100 MHz antenna 1 couples to the modes 101 (55.9 MHz), 121 (70.4 MHz), 301 (90.1 MHz), and 321 (99.7 MHz). As a consequence, there will be no more simultaneous coupling of both antennas to a common mode.
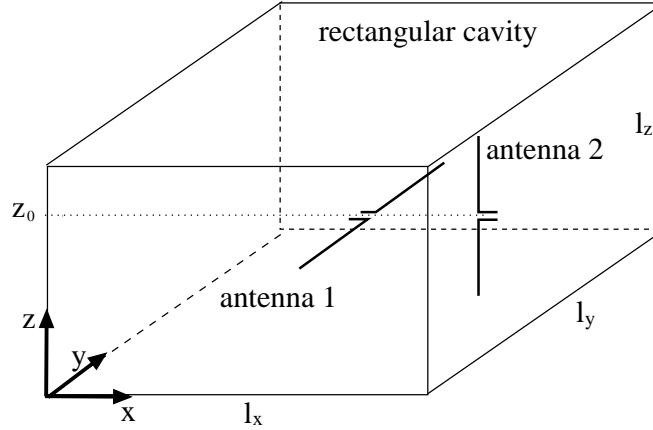
Figure 3.30: Two cross polarized antennas. Antenna 1 is aligned to the y-axis and antenna 2 is aligned to the z axis.

As in the previous example of two parallel antennas we consider again the expansion (3.45) of the mutual impedance. The nominator of each summand now has the form

$$\langle \boldsymbol{J}^b, \boldsymbol{f}^A_{mnp} \rangle \langle \boldsymbol{J}^a, \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{f}^A_{mnp}) + k^2 \boldsymbol{f}^A_{mnp} \rangle = \langle J^b_y, f^{yA}_{mnp} \rangle \langle J^a_z, \frac{\partial^2 f^{yA}_{mnp}}{\partial z \partial y} \rangle \tag{3.171}$$

$$= -k_z k_y \langle J^b_y, f^{yA}_{mnp} \rangle \langle J^a_z, f^{zA}_{mnp} \rangle. \tag{3.172}$$

The expression (3.45) for the self-impedance reduces to

$$Z_{12} = -\frac{j\omega\mu k_z k_y}{I^a_2 I^b_1 k^2} \sum_n \frac{\langle J^b_y, f^{yA}_{mnp} \rangle \langle J^a_z, f^{zA}_{mnp} \rangle}{k^2 - k^2_{mnp}}, \tag{3.173}$$

and it follows that contributions to the mutual impedance are obtained if both $\langle J^a_z, f^{zA}_{mnp} \rangle$ and $\langle J^b_y, f^{yA}_{mnp} \rangle$ are nonvanishing. But this does not happen since for any n or p either $\langle J^a_z, f^{zA}_{mnp} \rangle$ or $\langle J^b_y, f^{yA}_{mnp} \rangle$ vanishes. It follows that in this example the mutual impedance vanishes,

$$Z_{12} = 0. \tag{3.174}$$

This does not mean that the two antennas do not interact at all. But it means that the interaction is such that it has no effect at the antenna input terminals which are symmetrically positioned at the center of each antenna.

3.4.4   Electrically short antennas

The analysis of this chapter is mainly based on the Green's function formalism and the spectral properties of the electromagnetic field. It yields analytic formulas for quantities like the self or mutual impedance of dipole antennas. However, these formulas involve the, a priori, unknown antenna current. The antenna current is determined from an integral equation that can be solved by the method of moments. Usually, the method of moments is taken as a numerical solution procedure, but it can be turned into an approximative, analytical solution procedure if the unknown function, here the antenna current, is approximated by only a few basis functions. Then the resulting linear system of equations can analytically be solved.

The approximation of an antenna current by only a few terms is physically meaningful if the antenna considered is electrically small, $kL \ll 1$. An often made approximation follows from the assumption that the antenna current is sinusoidal. In this case the antenna current is modeled by only one piecewise sinusoidal basis function $S_k(\boldsymbol{r})$, compare (3.131).

Example: We consider the antenna-cavity configuration of Fig. 3.19 and choose the antenna lengths as $L = 0.2$m. To calculate the mutual impedance between both antennas we concentrate on the expression (3.31),

$$Z_{12} = -\frac{\langle \boldsymbol{E}^b, \boldsymbol{J}^a \rangle_p}{I_2^a I_1^b} \, . \tag{3.175}$$

With reference to Sec. 3.4.1 (a) we choose a discretization with $M = 1$ and expand each antenna current according to

$$I_1^b(\boldsymbol{r}^{(1)}) = \alpha^{(1)} S_1(\boldsymbol{r}^{(1)}) \, , \tag{3.176}$$

$$I_2^a(\boldsymbol{r}^{(2)}) = \alpha^{(2)} S_1(\boldsymbol{r}^{(2)}) \, . \tag{3.177}$$

The input currents of the antennas are simply given by the (unknown) expansion coefficients,

$$I_1^b = \alpha^{(1)} \, , \tag{3.178}$$

$$I_2^a = \alpha^{(2)} \, . \tag{3.179}$$

To evaluate the inner product of (3.175) we first evaluate the electric field by a calculation which is similar to the one which led to (3.133),

$$E_z^b(\boldsymbol{r}^{(2)}) = -j\omega\mu \int_{\text{antenna 1}} G_{zz}^E(\boldsymbol{r}^{(2)}, \boldsymbol{r}^{(1)}) \alpha^{(1)} S_1(\boldsymbol{r}^{(1)}) \, d\boldsymbol{r}^{(1)} \tag{3.180}$$

$$= -\frac{j\omega\mu\alpha^{(1)}}{k \sin(kh)} \left[G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1 + L/2\boldsymbol{e}_z) - G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1 - L/2\boldsymbol{e}_z) \right.$$

$$\left. -2\cos(kL/2) G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1) \right] \, . \tag{3.181}$$

Here, the position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2$ point to the centers of the antennas. With (3.178), (3.179), and (3.181) the expression for the mutual impedance (3.175) becomes

$$
Z_{12} = \frac{j\omega\mu}{k\sin(kL/2)} \left[G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1 + L/2\boldsymbol{e}_z) - G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1 - L/2\boldsymbol{e}_z) \right.
$$
$$
\left. -2\cos(kL/2)G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1) \right] \int_{\mathrm{antenna2}} S_1(\boldsymbol{r}^{(2)}) \, d\boldsymbol{r}^{(2)} . \quad (3.182)
$$

The last integral can be approximated according to

$$
\int_{\mathrm{antenna2}} S_1(\boldsymbol{r}^{(2)}) \, d\boldsymbol{r}^{(2)} = \frac{2}{\sin(kL/2)} \left(1 - \cos(kL/2) \right) \quad (3.183)
$$
$$
\approx \frac{L}{2} \quad (3.184)
$$

and we obtain the final result

$$
Z_{12} = \frac{j\omega\mu L}{2k\sin(kL/2)} \left[G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1 + L/2\boldsymbol{e}_z) - G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1 - L/2\boldsymbol{e}_z) \right.
$$
$$
\left. -2\cos(kL/2)G_{zz}^A(\boldsymbol{r}_2, \boldsymbol{r}_1) \right] . \quad (3.185)
$$

In Fig. 3.31 we compare the mutual impedance as calculated by the approximate, analytic result (3.185) to a method of moment solution. It is recognized that the approximate solution reproduces well the major resonance peaks that occur whenever both antennas couple to a common mode. Differences are observed inbetween these resonances, but the qualitative behavior of the curves agree fairly well. Of course, the advantage of the approximate solution is that it is calculated much faster than the method of moment solution and also provides a final answer in terms of the formula (3.185).

Clearly, as we turn to the simplified models of electrically short antennas we generally will no longer be able to take into account detailed features of the antenna current. For example, the approximation by a simple sinusoidal function will fail to yield any reliable value for the input impedance of an antenna. Therefore, turning to electrically short antennas means turning away from a number of details and subtleties of antenna theory. If antennas are drastically reduced in size to represent localized current elements that do not exhibit their own degrees of freedom we are left to study the behavior of the electromagnetic field. Then an antenna theory in resonating systems reduces to the conventional microwave theory, as displayed in Fig. 3.1.

Models of electrically short antennas nevertheless are useful to get an overall understanding of antenna coupling in resonating environments and also allow to estimate the order of coupling effects that are of practical interest. This has been exemplified in particular by Tkachenko and his coworkers [219, 153, 220].
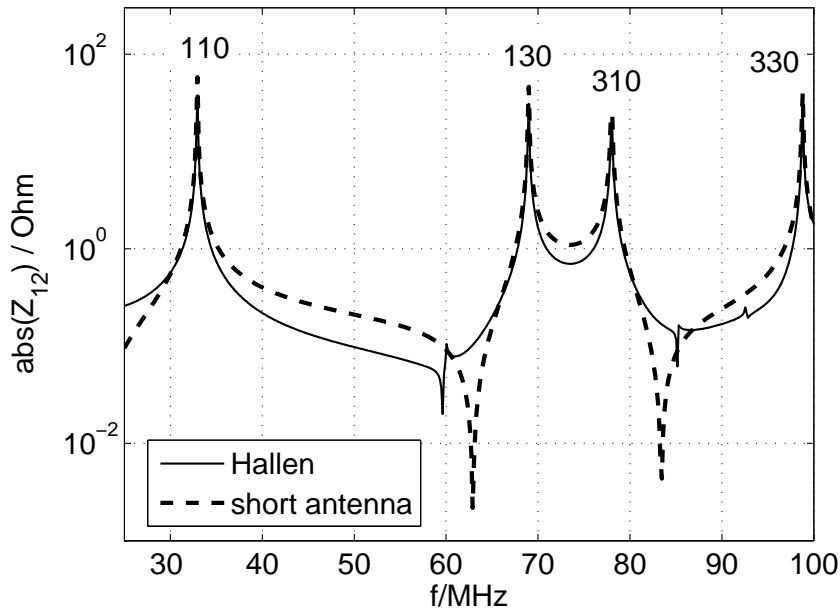
Figure 3.31: Mutual impedance between electrically short antennas, calculated by a method of moments solution of Hallén's equation and the approximate, analytic result (3.185) which is based on the assumption of electrically short antennas.

3.5 Remarks on antenna analysis in free space and resonating systems

We repeat that the main tasks of antenna analysis are

1. the solution of field integral equations for antenna currents and

2. the calculation of electromagnetic fields from antenna currents.

From a formal perspective, these two points constitute a simple road map. However, in practice they can be almost arbitrarily difficult to deal with. Both points require the knowledge of the electromagnetic Green's function. The Green's function establishes the integral kernel of the field integral equations to be solved and also relates the electric current to the electromagnetic field. It constitutes the solution of the Maxwell equations for prescribed boundary conditions.

In three-dimensional free space the Green's function represents outgoing spherical waves, see (2.136), that incorporate both Coulomb fields and radiation fields. The electromagnetic field that can be generated by an antenna in free space will be a superposition of such spherical waves that combine to yield a characteristic radiation pattern.

For a resonating environment the Green's function might not be known from the beginning. If it is not possible to construct (or, at least, approximate) this Green's function it follows that an antenna analysis, as outlined above, cannot be done and it is necessary to resort to numerical techniques. But if the Green's function is known also the spectral properties of the electromagnetic field within the resonating environment are known. Linear superpositions of the corresponding resonances (or, equivalently, eigenmodes) will make up the electromagnetic field that is generated by an antenna. That is, while in free space an antenna will always generate a superposition of outgoing spherical waves it will generate in a resonating environment a superposition of various resonances which yields both the appropriate Coulomb and radiation fields. The latter superposition is determined from coupling coefficients of the form (3.145) or (3.146). A complete set of known coupling coefficients can be seen as an antenna-cavity characteristic, comparable to an antenna radiation characteristic in free space. This set of coefficients not only depends on the geometry of the antenna but also on the positioning of the antenna in the resonating environment and on the resonating environment itself.

The calculation of antenna couplings within resonating environments is, in principle, straightforward. The mutual impedance between antennas is determined from formulas of the form (3.45) and (3.56). It critically depends on the *absolute* antenna positions within the resonating environment. This is in contrast to free space where the *relative* position of the antennas is decisive. In particular, within a cavity the mutual coupling does not necessarily decay in case of increasing distance between the two antennas since the cavity modes do not decay with increasing distance, as well. For strong antenna coupling at resonance it is not the relative distance but the simultaneous coupling to a resonant mode that is decisive. In view of applications to Electromagnetic Compatibility it is obvious that the electromagnetic coupling between unintentionally coupled antennas within a resonating environment can considerably be reduced not only by avoiding resonance frequencies or decreasing the quality factor of the cavity. An equally important aspect is the positioning of the antennas within the cavity that should be chosen such that a simultaneous coupling to single modes is avoided.

Chapter 4

Nonlinearly Loaded Antennas

Most of the concepts we have used so far, be this linear operator theory, the Green's function method, or the method of moments, for example, rest on the assumption that electromagnetic field theory is a linear theory. We know that this assumption is valid as long as the constitutive relations, which relate the electromagnetic excitations $\mathbf{D}$, $\mathbf{H}$ to the electromagnetic field strengths $\mathbf{E}$, $\mathbf{B}$, are linear. Examples of linear constitutive relations are provided by (1.36), (1.37) for the case of vacuum and by (1.38), (1.39) for the case of a general linear magnetoelectric medium.

But it is an inevitable truth that modern electric and electronic devices contain nonlinear elements like diodes, transistors, or other semiconductor elements. As a consequence, we need to have an understanding of how these nonlinearities might influence electric currents and electromagnetic fields. These effects of nonlinearities can be wanted or unwanted. Of course, nonlinear semiconductor elements are designed to serve some purpose, and this purpose usually will exploit nonlinear properties that lead to wanted effects. But unwanted effects occur as well and these effects need to be studied in the framework of Electromagnetic Compatibility.

To account for nonlinear effects might, at first sight, require to work with nonlinear constitutive relations and turn to nonlinear electromagnetic field theory [108]. In fact, many field theories are nonlinear. Examples are the three fundamental interactions gravity, weak, and strong interaction. But the resulting nonlinear field equations are difficult to solve [36, 200]. Often one relies on perturbation theory and attempts to calculate nonlinear corrections to linear first order solutions. Fortunately, the semiconductor elements that we just mentioned in the previous paragraph usually can be considered as concentrated, lumped elements. It is then not necessary to apply nonlinear field theory but rather appropriate to apply *nonlinear circuit theory*.

The reason why it is of interest to discuss nonlinearities in the context of antenna theory in resonating systems is best illustrated by an example. To this end we consider in Fig. 4.1 a cavity which contains a nonlinearity. In practice, this could be the model of a computer with a metallic housing or the model of some metallic compartment
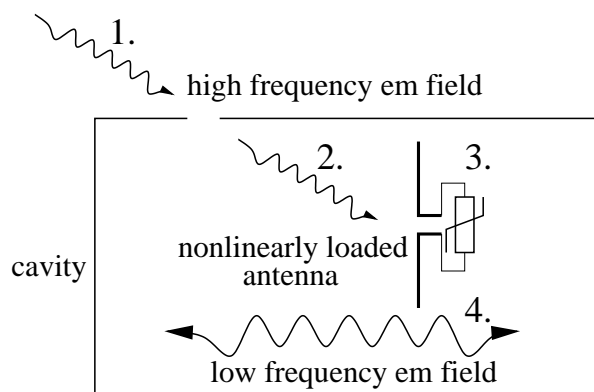
Figure 4.1: A high frequency electromagnetic field (1.) couples through a small opening into the interior of an electric or electronic system (2.). If the electromagnetic field interacts with a nonlinearity, which is modeled by a nonlinearly loaded antenna (3.), intermodulation will occur which can result in the generation of low frequency electromagnetic fields (4.). This is an example of "high to low frequency conversion" by a nonlinearly loaded antenna within a cavity.

within an aircraft or vehicle, for example. The cavity acts like a shield which prevents exterior electromagnetic fields to couple to the interior of the cavity. However, in realistic situations the cavity will not be completely closed but have small openings or slits. High frequency electromagnetic fields might penetrate through these openings and couple to the interior of the cavity. If nonlinearities are present in the interior it can happen that *intermodulation* occurs. This phenomenon describes the generation of secondary frequencies from the electromagnetic excitation of nonlinearities by primary frequencies and will be discussed below. Intermodulation can produce low frequency fields from high frequency fields[1]. These low frequency fields might disturb the system within the cavity, e.g., by the excitation of cavity resonances. This general example is not only of purely academic interest. Experiments on realistic smart defense systems have shown that even though a system is shielded against low frequency electromagnetic fields there can be a surprisingly high level of low frequency noise inside the system which disturbs enclosed analog electronic devices that operate in the low frequency region [152]. In this chapter we intend to give an impression of methods that are suitable to model this kind of problems of Electromagnetic Compatibility.

Nonlinearly loaded antennas and wires in free space already have been investigated in some detail. Schuman [194] combined a method of moment solution with a Newton-Raphson iterative technique to calculate the load current of a nonlinearly loaded thin-wire scatterer. Sarkar & Weiner [188] proposed an analytic, iterative solution scheme

---

[1]The notions "low frequency fields" and "high frequency fields" are not strictly defined and mainly used for linguistic convenience. Approximately, we think in the present context of the range from 10^3Hz to 10^7Hz as low frequency regime and 10^8Hz to 10^{10}Hz as high frequency regime.

which uses the Volterra series technique in the frequency domain and obtained power levels of the scattered field of a nonlinearly loaded antenna. Two different solution methods were proposed by Liu & Tesche [131]. The first one is a direct time-domain integral approach which is different from the Schuman method, the second one utilizes equivalent circuits in the frequency domain to split the problem in a linear and a nonlinear part, where the solution of the linear part is obtained from the method of moments and the solution of the nonlinear part is obtained by the iterative solution of a nonlinear integral equation. Similar analytical and numerical techniques were used by Kanda [100] to analyze an electrically short dipole with a nonlinear load. Landt et al. [116] numerically solved a time-domain integral equation to exhibit the behavior of nonlinearly loaded antennas. In their analysis they included a number of interesting examples that illustrated effects caused by short electromagnetic pulses such as generated from lightning or sparks. Another method of moment study, which involves dipoles loaded with passive and active diodes, was conducted by Janaswamy & Lee [94]. Finally, Huang & Chu [91] investigated wire scatterers with nonlinear or time-harmonic loads in the frequency domain. Similar to Sarkar and Liu & Tesche they transformed the electromagnetic field problem to a circuit problem and, eventually, solved for the unknown current of the nonlinear circuit by means of the harmonic balance technique to calculate scattered field spectra.

The analysis of nonlinear antenna problems within a resonating system follows the same patterns as the corresponding analysis in free space. Clearly, it is necessary to take into account the effects that result from the coupling between nonlinearly loaded antennas and electromagnetic resonances. Up to now there is relatively few literature on the subject of nonlinearly loaded antenna in resonating environments. In particular, Lee et al. have studied mutual coupling mechanisms within arrays of nonlinear antennas [118, 119, 120, 121]. These arrays form two-dimensional periodic structures and the corresponding Green's function exhibits similar features as that of a three-dimensional rectangular cavity. In the papers of Lee et al. the electromagnetic field problem is reduced to an equivalent circuit problem which, in turn, is solved by one of several different techniques. An equivalent circuit was also used in the investigation of a nonlinearly loaded antenna within a three-dimensional rectangular cavity where intermodulation effects were explicitly computed by means of the harmonic balance technique [70].

The strategy to reduce for the solution of nonlinear antenna problems within a resonating system the electromagnetic field problem to an equivalent circuit problem seems to be natural and convenient. In this case resonance effects will be incorporated in lumped antenna impedances and we know from Chapter 3, at least for canonical configurations, how to model and calculate these impedances in an efficient way.

4.1   Important aspects of the analysis of nonlinearly loaded antennas

This section introduces two basic but important aspects that are essential for the analysis of nonlinearly loaded antennas. The first is the representation of an antenna configuration by an equivalent circuit, the second is the generation of new frequencies by nonlinear, lumped circuit elements.

4.1.1   Equivalent circuit of a nonlinearly loaded antenna

We consider in Fig. 4.2 a nonlinearly loaded antenna. The antenna has been drawn as a dipole antenna, but it could also be of another type. The antenna is subject to an incident electromagnetic field. This field will cause, in general, a current $i_{\mathrm{nl}}(t)$ and a voltage $v_{\mathrm{nl}}(t)$ at the nonlinear load.
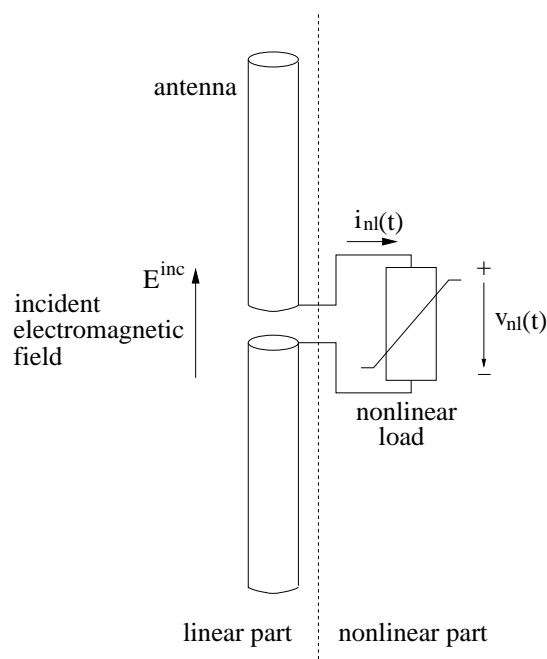
Figure 4.2: A linear antenna with a nonlinear load. The nonlinear load is connected to the antenna input terminal. As indicated, the configuration splits into a linear and a nonlinear part.

The frequency behavior of the linear part of the excited antenna can be represented by a Norton's or Thévenin's equivalent circuit looking toward the scatterer at the antenna input terminal, as shown in Fig. 4.3 [131]. The equivalent circuits involve an equivalent current source $I_{\mathrm{eq}}(\omega)$ or an equivalent voltage source $V_{\mathrm{eq}}(\omega)$, respectively, that are due to the electromagnetic source field. They also involve the antenna input admittance $Y_{\mathrm{in}}(\omega)$
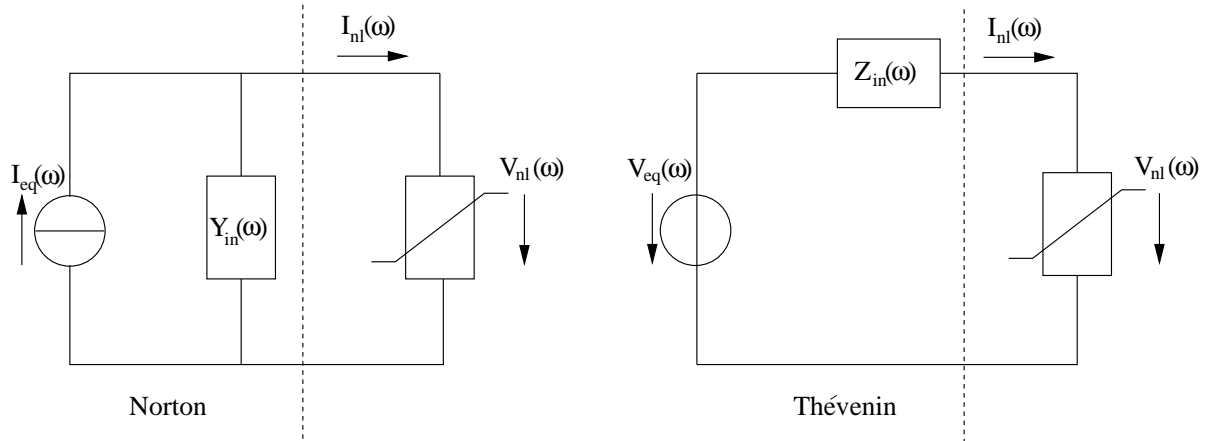
Figure 4.3: Equivalent circuit of the nonlinearly loaded antenna of Fig. 4.2. Both the Norton and Thévenin equivalent are shown. In the notation used the variable ω represents all frequencies that occur at a certain circuit element.

or input impedance $Z_{\text{in}}(\omega)$, respectively. These frequency domain functions need to be determined from the associated field-theoretical, linear boundary value problem, e.g., by means of the method of moments. Once this is done it remains to determine the current through and voltage across the nonlinear load by means of the methods of nonlinear circuit theory.

4.1.2 Intermodulation frequencies

Probably the most important feature of nonlinear circuits is the generation of secondary frequencies from primary frequencies that are provided by excitation sources. An intuitive approach to this feature is provided if the current-voltage relation $i = f(v)$ of a nonlinear element is expressed as a power series and subject to an excitation voltage which has multiple frequency components. This is a well-known approach which is described in a number of textbooks, see [133, §1], e.g..

Example 1: Let us consider the simple nonlinear circuit of Fig. 4.4 (a) and assume that the nonlinearity is characterized by the current-voltage relation

$$i(t) = a_1 v(t) + a_2 v^2(t) + a_3 v^3(t) \tag{4.1}$$

with constant factors a_1, a_2, and a_3. We further assume that the voltage source $v_{\text{eq}}(t)$ is given by a "two-tone excitation" of the form

$$v_{\text{eq}}(t) = v_1 \cos(\omega_1 t) + v_2 \cos(\omega_2 t). \tag{4.2}$$

With this excitation we want to solve for the unknown current $i(t)$ in (4.1). This is trivial since

$$v(t) = v_{\text{eq}}(t) \tag{4.3}$$
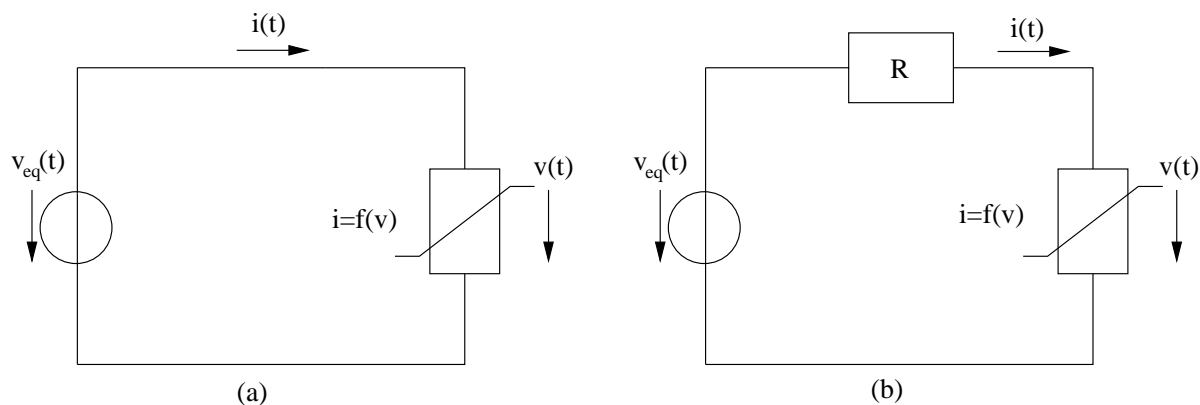
Figure 4.4: Two simple nonlinear circuits that exhibit the generation of intermodulation frequencies.

and, thus,

$$i(t) = a_1 \left(v_1 \cos(\omega_1 t) + v_2 \cos(\omega_2 t)\right) + a_2 \left(v_1 \cos(\omega_1 t) + v_2 \cos(\omega_2 t)\right)^2$$
$$+ a_3 \left(v_1 \cos(\omega_1 t) + v_2 \cos(\omega_2 t)\right)^3 . \quad (4.4)$$

The quadratic and cubic term can be expanded and rewritten by means of trigonometric identities such as

$$\cos^2(x) = \frac{1}{2}\left(1 + \cos(2x)\right) , \quad (4.5)$$

$$\cos^3(x) = \frac{1}{4}\left(3\cos(x) + \cos(3x)\right) , \quad (4.6)$$

in order to reduce powers of trigonometric functions to linear combinations of single trigonometric functions. This yields

$$\begin{aligned}
i(t) = & a_1 \left(v_1 \cos(\omega_1 t) + v_2 \cos(\omega_2 t)\right) \\
& + \frac{a_2}{2}\left(v_1^2 + v_2^2 + v_1^2 \cos(2\omega_1 t)\right. \\
& \qquad \left. + v_2^2 \cos(2\omega_2 t) + 2v_1 v_2[\cos((\omega_1 + \omega_2)t) + \cos((\omega_1 - \omega_2)t)]\right) \\
& + \frac{a_3}{4}\left(v_1^3 \cos(3\omega_1 t) + v_2^3 \cos(3\omega_1 t)\right. \\
& \qquad + 3v_1^2 v_2[\cos((2\omega_1 + \omega_2)t) + \cos((2\omega_1 - \omega_2)t)] \\
& \qquad + 3v_1 v_2^2[\cos((\omega_1 + 2\omega_2)t) + \cos((\omega_1 - 2\omega_2)t)] \\
& \qquad \left. + 3(v_1^3 + 2v_1 v_2^2)\cos(\omega_1 t) + 3(v_2^3 + 2v_1^2 v_2)\cos(\omega_2 t)\right) . \quad (4.7)
\end{aligned}$$

It follows that $i(t)$ contains besides the primary frequencies ω_1 and ω_2 new frequencies which are of the form

$$\omega_{m,n} = m\omega_1 + n\omega_2 \quad (4.8)$$

with $m, n = \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$. These frequencies are called *mixing frequencies* or *intermodulation frequencies*.

Example 2: As a second example we consider the nonlinear circuit of Fig. 4.4 (b) which includes a resistor R and the same nonlinearity as circuit (a). In this case we cannot directly obtain $i(t)$ from $v_{\text{eq}}(t)$ since

$$v_{\text{eq}}(t) = Ri(t) + v(t) \tag{4.9}$$

and this provides only one equation for the two unknowns $i(t)$ and $v(t)$. In order to eliminate in (4.9) the voltage $v(t)$ by the current $i(t)$ we invert the series (4.1). This yields [1, Eq. 3.6.25]

$$v(t) = \frac{1}{a_1}i(t) - \frac{a_2}{a_1^3}i^2(t) + \frac{2a_2^2 - a_1a_3}{a_1^5}i^3(t) - \frac{5(a_2^3 - a_1a_2a_3)}{a_1^7}i^4(t) + \ldots. \tag{4.10}$$

This series is, in fact, infinite and displayed up to the fourth term. With (4.9) we find a relation between $v_{\text{eq}}(t)$ and $i(t)$,

$$v_{\text{eq}}(t) = \left(\frac{1}{a_1} + R\right)i(t) - \frac{a_2}{a_1^3}i^2(t) + \frac{2a_2^2 - a_1a_3}{a_1^5}i^3(t) - \frac{5(a_2^3 - a_1a_2a_3)}{a_1^7}i^4(t) + \ldots. \tag{4.11}$$

From the inversion of this series we obtain the current

$$i(t) = \frac{1}{1/a_1 + R}v_{\text{eq}}(t) + \frac{a_2/a_1^3}{(1/a_1 + R)^3}v_{\text{eq}}^2(t) + \frac{2a_2^2/a_1^6 - (1/a_1 + R)(2a_2^2 - a_1a_3)/a_1^5}{(1/a_1 + R)^5}v_{\text{eq}}^3(t)$$

$$-5\frac{(1/a_1 + R)a_2(2a_2^2 - a_1a_3)/a_1^8 - (1/a_1 + R)^2(a_2^3 - a_1a_2a_3)/a_1^7 - a_2^3/a_1^9}{(1/a_1 + R)^7}v_{\text{eq}}^4(t) + \ldots, \tag{4.12}$$

which, again, is represented by an infinite series that is displayed up to the fourth term. In case of a two-tone excitation we may insert (4.2) into (4.12) and work out the powers of $v_{\text{eq}}(t)$ in an analogous way as we did to arrive at (4.7). In principle, for the complete series expansion of $i(t)$ this results in an infinite number of intermodulation frequencies of the form (4.8). However, in practice the series expansion (4.12) will only be meaningful if the parameters a_1, a_2, a_3, and the amplitudes of the exciting voltages v_1, v_2 are such that the series converges. As a general rule, this will be fulfilled if the nonlinearity is weak and the exciting voltages are small. Then the series terms of (4.12) will decay with increasing order and, as a result, the influence of high intermodulation frequencies that exceed a certain order can be neglected.

In the introduction to this chapter we mentioned the phenomenon of "high to low frequency conversion" which occurs if two frequencies of a two-tone (or multiple-tone) excitation are close to each other. Accordingly, if we assume two exciting frequencies ω_1, ω_2 with

$$\omega_2 = \omega_1 + \Delta\omega, \qquad \omega_1, \omega_2 \gg \Delta\omega \tag{4.13}$$

then the intermodulation frequencies $\omega_{m,n}$ will be of the form

$$\omega_{m,n} = (m+n)\omega_1 + n\Delta\omega \tag{4.14}$$

and the lowest intermodulation frequency is given by

$$\omega_{-1,1} = |\omega_{1,-1}| = \Delta\omega \,. \tag{4.15}$$

This is a second order intermodulation frequency which is exhibited by the quadratic term of the solution (4.7), for example. Further intermodulation frequencies with $m+n = 0$ are provided by

$$\omega_{-n,n} = |\omega_{n,-n}| = n\Delta\omega \,, \qquad n > 1 \,. \tag{4.16}$$

These are, obviously, integer multiples of $\Delta\omega$ which are, for n not too large, also low in comparison to ω_1, ω_2. Interpolation frequencies with $|m+n| = 1$ are located around ω_1, ω_2. Explicitly, we have

$$\omega_{1-n,n} = |\omega_{n-1,-n}| = \omega_1 + n\Delta\omega \,, \tag{4.17}$$
$$\omega_{1+n,-n} = |\omega_{-1-n,n}| = \omega_1 - n\Delta\omega \,. \tag{4.18}$$

In Fig. 4.5 the distribution of some of these intermodulation frequencies is shown.
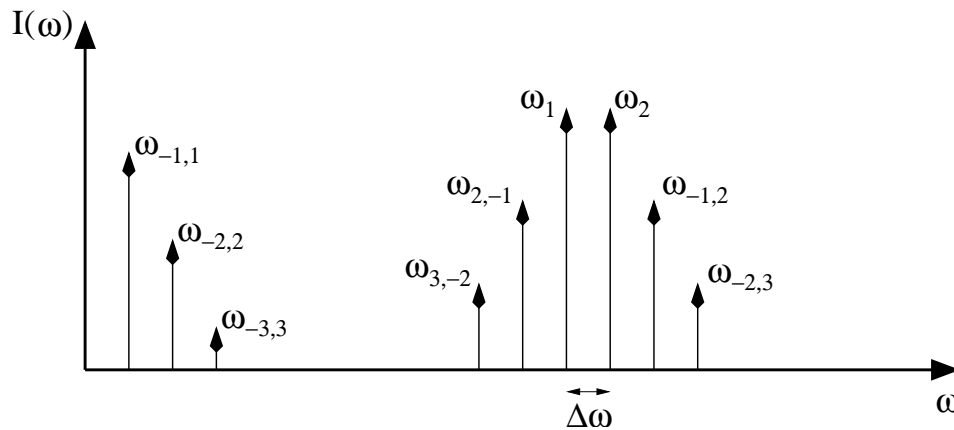
Figure 4.5: Selected intermodulation frequencies that result from a two-tone excitation if the exciting frequencies ω_1, ω_2 are close to each other. It is assumed that the amplitudes decrease for increasing order of intermodulation. The lowest intermodulation frequency is given by $\omega_{-1,1} = |\omega_{1,-1}| = \Delta\omega$. Not drawn are possible dc contributions at zero frequency.

4.2 Methods of nonlinear circuit theory

In the previous section we displayed in Fig. 4.3 the equivalent circuit of a nonlinearly loaded antenna and indicated by means of two elementary examples that the solution for the current $I_{nl}(\omega)$ (and, hence, for the voltage $V_{nl}(\omega)$) will contain a variety of intermodulation frequencies. It follows that in case of a two-tone excitation a nonlinear circuit, and thus a nonlinearly loaded antenna, will exhibit the phenomenon of "high to low frequency conversion". This is a statement of general validity. To explicitly calculate this phenomenon we have to be able to calculate from a given equivalent circuit with known sources $I_{eq}(\omega)$, $V_{eq}(\omega)$, known parameters $Y_{in}(\omega)$, $Z_{in}(\omega)$, and a known nonlinearity with current-voltage characteristic $i = f(v)$ the unknown current $I_{nl}(\omega)$.

In the following we will shortly introduce methods of nonlinear circuit theory that are suitable to calculate nonlinear networks of the form (4.3). Since the antenna characteristics usually will be given in the frequency domain, as exemplified by the antenna admittance or impedance, pure time-domain methods cannot directly be applied. Popular methods that *can* directly be applied include the *method of successive approximation* which is also known as *Picard iteration* [122, 34], the *Volterra series technique* [193, 184, 231, 133], and the *harmonic balance technique* [148, 109, 133].
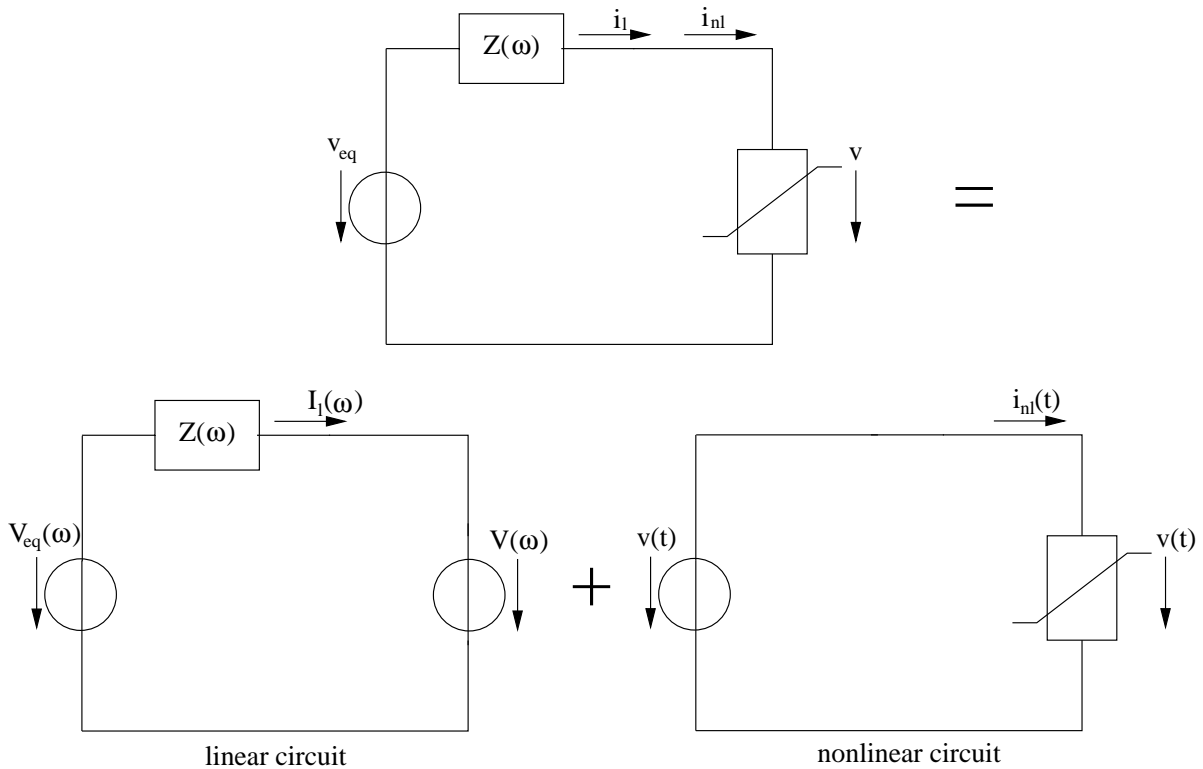
Figure 4.6: Division of a nonlinear circuit in two equivalent circuits.

In oder to attain an understanding of these methods it will be useful to divide a nonlinear circuit, as shown in Fig. 4.6, into a pair of equivalent circuits. One equivalent circuit represents the linear part and can be described in the frequency domain, the other equivalent circuit represents the nonlinear part and will be described in the time domain. The methods mentioned above take advantage of this division.

4.2.1   Successive approximation - Picard iteration

The method of successive approximation is described first because it is both an analytic method and comparatively simple to use. It requires that the nonlinearity is weak and the exciting signal is small such that the relation between voltage and current can be expressed by a converging power series. The method of successive approximation consists of the following steps:

1. Start from the linear circuit of Fig. 4.6 and set $V(\omega) = 0$. It follows that the computation of the current $I_l(\omega)$ is equivalent to the solution of a linear circuit problem. Solve this linear circuit problem. The result constitutes a first order solution and contains the exciting frequencies.

2. Analytically convert the current $I_l(\omega)$ into the time domain to obtain $i_l(t)$.

3. Turn to the nonlinear circuit of Fig. 4.6 and set $i_{nl}(t) = i_l(t)$. Characterize the nonlinearity not by a current-voltage relation but by a voltage-current relation of the form[2]

$$v(t) = F\left(i_{\mathrm{nl}}(t)\right) = \sum_{n=2}^{N} b_n i_{\mathrm{nl}}^n(t).$$

   (4.19)

   Insert the explicit expression for the current $i_{\mathrm{nl}}(t)$ into this relation and work out the powers $i_{\mathrm{nl}}^n(t)$ by means of trigonometric identities of the type (4.5), (4.6). This yields the voltage $v(t)$ as a superposition of single frequency contributions.

4. Obtain $V(\omega)$ from $v(t)$. This is immediate since $v(t)$ is given by a sum of single frequency contributions.

5. Turn back to the linear circuit and calculate with $V(\omega)$ an improved value of $I_l(\omega)$. Finally, return to step 2 and continue the process until $I_l(\omega)$ has converged.

During this process the most cumbersome part is to work out in step 3 the powers $i_{\mathrm{nl}}^n(t)$. In order to do this in a systematic way it is convenient to employ an exponential notation: If $i_{\mathrm{nl}}(t)$ is given as a linear combination of M frequencies,

$$i_{\mathrm{nl}}(t) = \sum_{m=1}^{M} i_m \cos(\omega_m t + \varphi_m),$$

   (4.20)

---

[2]We assume that the linear part of the nonlinearity has been absorbed by the impedance $Z(\omega)$ of the linear circuit.

it can be rewritten as

$$
\begin{aligned}
i_{\mathrm{nl}}(t) &= \frac{1}{2} \sum_{m=1}^{M} \left[i_m e^{j(\omega_m t + \varphi_m)} + i_m e^{-j(\omega_m t + \varphi_m)} \right] \\
&= \frac{1}{2} \sum_{m=1}^{M} \left[\hat{i}_m e^{j\omega_m t} + \hat{i}_m^* e^{-j\omega_m t} \right] \\
&= \frac{1}{2} \sum_{\substack{m=-M \\ m \neq 0}}^{M} \hat{i}_m e^{j\omega_m t} \, .
\end{aligned}
\tag{4.21}
$$

Here we defined $\hat{i}_m := i_m e^{j\varphi_m}$, $\hat{i}_{-m} := \hat{i}_m^*$, and $\omega_{-m} := -\omega_m$. It follows that a power of $i_{\mathrm{nl}}(t)$ is expressed according to

$$
i_{\mathrm{nl}}^n(t) = \left[\frac{1}{2} \sum_{\substack{m=-M \\ m \neq 0}}^{M} \hat{i}_m e^{j\omega_m t} \right]^n
\tag{4.22}
$$

$$
= \frac{1}{2^n} \sum_{\substack{m_1=-M \\ m_1 \neq 0}}^{M} \sum_{\substack{m_2=-M \\ m_2 \neq 0}}^{M} \cdots \sum_{\substack{m_n=-M \\ m_n \neq 0}}^{M} \hat{i}_1 \hat{i}_2 \ldots \hat{i}_m e^{j(\omega_{m_1} + \omega_{m_2} + \ldots + \omega_{m_n})t} \, ,
\tag{4.23}
$$

and this expression is easier to evaluate than powers of trigonometric functions.

Example: We consider the nonlinear circuit of Fig. 4.6 which involves a general impedance $Z(\omega)$. Let us further assume that the exciting voltage is given by a two-tone excitation of the form

$$
v_{\mathrm{eq}}(t) = \sum_{m=1}^{2} v_m \cos(\omega_1 t + \varphi_m)
\tag{4.24}
$$

$$
= \frac{1}{2} \sum_{\substack{m=-2 \\ m \neq 0}}^{2} \hat{v}_m e^{j\omega_m t} \, .
\tag{4.25}
$$

In the following we work out the first order Picard iteration, where we presuppose a third order nonlinearity of the form

$$
v(t) = \sum_{n=2}^{3} b_n i_{\mathrm{nl}}^n(t)
\tag{4.26}
$$

$$
= b_2 i_{\mathrm{nl}}^2(t) + b_3 i_{\mathrm{nl}}^3(t) \, .
\tag{4.27}
$$

We first set $V(\omega) = 0$ and obtain

$$
I_{\mathrm{l}}^{(1)}(\omega) = \sum_{m=1}^{2} \frac{\hat{v}_m}{Z(\omega_m)}
\tag{4.28}
$$

as solution of the linear circuit problem. Conversion to time domain leads to

$$i_1^{(1)}(t) = i_{\mathrm{nl}}^{(1)}(t) \tag{4.29}$$

$$= \frac{1}{2} \sum_{\substack{m=-2 \\ m \neq 0}}^{2} \frac{\hat{v}_m}{Z(\omega_m)} e^{j\omega_m t} . \tag{4.30}$$

This is inserted into (4.27) to yield the first order iterated voltage at the nonlinearity,

$$v^{(1)}(t) = \sum_{n=2}^{3} \frac{b_n}{2^n} \left[\sum_{\substack{m=-2 \\ m \neq 0}}^{2} \frac{\hat{v}_m}{Z(\omega_m)} e^{j\omega_m t} \right]^n \tag{4.31}$$

$$= \frac{b_2}{4} \sum_{\substack{m_1=-2 \\ m_1 \neq 0}}^{2} \sum_{\substack{m_2=-2 \\ m_2 \neq 0}}^{2} \frac{\hat{v}_{m_1}}{Z(\omega_{m_1})} \frac{\hat{v}_{m_2}}{Z(\omega_{m_2})} e^{j(\omega_{m_1}+\omega_{m_2})t}$$

$$+ \frac{b_3}{8} \sum_{\substack{m_1=-2 \\ m_1 \neq 0}}^{2} \sum_{\substack{m_2=-2 \\ m_2 \neq 0}}^{2} \sum_{\substack{m_3=-2 \\ m_3 \neq 0}}^{2} \frac{\hat{v}_{m_1}}{Z(\omega_{m_1})} \frac{\hat{v}_{m_2}}{Z(\omega_{m_2})} \frac{\hat{v}_{m_3}}{Z(\omega_{m_3})} e^{j(\omega_{m_1}+\omega_{m_2}+\omega_{m_3})t} . \tag{4.32}$$

The double sum of (4.32) already consists of $(2 \cdot 2)^2 = 16$ summands with corresponding frequencies

$$
\begin{array}{cccc}
-\omega_2 - \omega_2 & -\omega_2 - \omega_1 & -\omega_2 + \omega_1 & -\omega_2 - \omega_2 \\
-\omega_1 - \omega_2 & -\omega_1 - \omega_1 & -\omega_1 + \omega_1 & -\omega_1 - \omega_2 \\
+\omega_1 - \omega_2 & +\omega_1 - \omega_1 & +\omega_1 + \omega_1 & +\omega_1 + \omega_2 \\
-\omega_2 - \omega_2 & +\omega_2 - \omega_1 & +\omega_2 + \omega_1 & +\omega_2 + \omega_2 .
\end{array} \tag{4.33}
$$

These 16 terms show that $v^{(1)}(t)$ contains dc contributions and the frequencies $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$, and $|\omega_1 - \omega_2|$. The triple sum of (4.32) already consists of $(2 \cdot 2)^3 = 64$ terms that add contributions with frequencies ω_1, ω_2, $2\omega_1 + \omega_2$, $|2\omega_1 - \omega_2|$, $\omega_1 + 2\omega_2$, $|\omega_1 - 2\omega_2|$, $3\omega_1$, and $3\omega_2$.

The first order voltage (4.32) can also be written in terms of cosine functions. To this end we express the impedance $Z(\omega_m)$ by its absolute value $|Z(\omega_m)|$ and its argu-

ment φ_{Zm}. Then we find

$$
\begin{aligned}
v^{(1)}(t) = \frac{b_2}{2} \Bigg[& \left(\frac{V_1}{|Z(\omega_1)|} \right)^2 + \left(\frac{V_2}{|Z(\omega_2)|} \right)^2 + \left(\frac{V_1}{|Z(\omega_1)|} \right)^2 \cos(2(\omega_1 t + \varphi_1 - \varphi_{Z1})) \\
& + \left(\frac{V_2}{|Z(\omega_2)|} \right)^2 \cos(2(\omega_2 t + \varphi_2 - \varphi_{Z2})) \\
& + 2 \frac{V_1}{|Z(\omega_1)|} \frac{V_2}{|Z(\omega_2)|} \Big(\cos((\omega_1 + \omega_2)t + \varphi_1 - \varphi_{Z1} + \varphi_2 - \varphi_{Z2}) \\
& \hspace{5cm} + \cos((\omega_1 - \omega_2)t + \varphi_1 - \varphi_{Z1} - \varphi_2 + \varphi_{Z2}) \Big) \Bigg] \\
+ \frac{b_3}{4} \Bigg[& \left(\frac{V_1}{|Z(\omega_1)|} \right)^3 \cos(3(\omega_1 t + \varphi_1 - \varphi_{Z1})) + \left(\frac{V_2}{|Z(\omega_2)|} \right)^3 \cos(3(\omega_2 t + \varphi_2 - \varphi_{Z2})) \\
& + 3 \left(\frac{V_1}{|Z(\omega_1)|} \right)^2 \frac{V_2}{|Z(\omega_2)|} \Big(\cos(2(\omega_1 t + \varphi_1 - \varphi_{Z1}) + \omega_2 t + \varphi_2 - \varphi_{Z2}) \\
& \hspace{5cm} + \cos(2(\omega_1 t + \varphi_1 - \varphi_{Z1}) - \omega_2 t - \varphi_2 + \varphi_{Z2}) \Big) \\
& + 3 \frac{V_1}{|Z(\omega_1)|} \left(\frac{V_2}{|Z(\omega_2)|} \right)^2 \Big(\cos(\omega_1 t + \varphi_1 - \varphi_{Z1} + 2(\omega_2 t + \varphi_2 - \varphi_{Z2})) \\
& \hspace{5cm} + \cos(\omega_1 t + \varphi_1 - \varphi_{Z1} - 2(\omega_2 t + \varphi_2 - \varphi_{Z2})) \Big) \\
& + 3 \left(\left(\frac{V_1}{|Z(\omega_1)|} \right)^3 + 2 \frac{V_1}{|Z(\omega_1)|} \left(\frac{V_2}{|Z(\omega_2)|} \right)^2 \right) \cos(\omega_1 t + \varphi_1 - \varphi_{Z1}) \\
& + 3 \left(\left(\frac{V_2}{|Z(\omega_2)|} \right)^3 + 2 \left(\frac{V_1}{|Z(\omega_1)|} \right)^2 \frac{V_2}{|Z(\omega_2)|} \right) \cos(\omega_2 t + \varphi_2 - \varphi_{Z2}) \Bigg]. \quad (4.34)
\end{aligned}
$$

To calculate the second order voltage $v^{(2)}(t)$ it is necessary to repeat the procedure from (4.25), but with $v_{\text{eq}}(t)$ replaced by $v_{\text{eq}}(t) - v^{(1)}(t)$. This will result in a twelve-tone excitation. In this case the double and triple sum of (4.32) will consist of $(2 \cdot 12)^2 = 576$ and $(2 \cdot 12)^3 = 13824$ terms, respectively.

This example shows that the calculation of higher order Picard iterations quickly becomes laborious. Fortunately, it is rather straightforward to use computer algebra packages, such as *Maple* [54, 144] or *MATHEMATICA* [238], in order to perform the necessary analytic calculations with the help of a computer.

4.2.2 Volterra series analysis

The Volterra series analysis constitutes a generalized power series approach. As in the case of the method of successive approximation it requires that the nonlinearity is weak and the exciting signal is small such that it is meaningful to consider a current-voltage

relation of the form

$$i(t) = \sum_{n=1}^{N} a_n v^n(t) \,. \tag{4.35}$$

A power series approach has been used in the Example 1 of Sec. 4.1.1. In this example the exciting voltage $v_{\mathrm{eq}}(t)$ is equal to the voltage $v(t)$ at the nonlinearity and, as a result, we are able to immediately obtain the unknown current in terms of a power series, as displayed in (4.7). In order to generalize this example we assume that the exciting voltage consists of a multi-tone excitation which includes M different frequencies,

$$v_{\mathrm{eq}}(t) = \frac{1}{2} \sum_{\substack{m=-M \\ m\neq 0}}^{M} \hat{v}_m e^{j\omega_m t} \,. \tag{4.36}$$

If the exciting voltage $V_{\mathrm{eq}}(\omega)$ is related to the voltage $V(\omega)$ at the nonlinearity by a linear transfer function $H(\omega)$,

$$V(\omega) = H(\omega)V_{\mathrm{eq}}(\omega) \tag{4.37}$$

we can write

$$v(t) = \frac{1}{2} \sum_{\substack{m=-M \\ m\neq 0}}^{M} \hat{v}_m H(\omega_m) e^{j\omega_m t} \tag{4.38}$$

where $H(\omega_{-m}) = H^*(\omega_m)$. It follows that the current $i_{\mathrm{nl}}(t)$ is given by the expression

$$i_{\mathrm{nl}}(t) = \sum_{n=1}^{N} a_n \left[\frac{1}{2} \sum_{\substack{m=-M \\ m\neq 0}}^{M} \hat{v}_m H(\omega_m) e^{j\omega_m t} \right]^n \tag{4.39}$$

$$= \sum_{n=1}^{N} \frac{a_n}{2^n} \sum_{\substack{m_1=-M \\ m_1\neq 0}}^{M} \sum_{\substack{m_2=-M \\ m_2\neq 0}}^{M} \cdots \sum_{\substack{m_n=-M \\ m_n\neq 0}}^{M} \hat{v}_{m_1}\hat{v}_{m_2}\ldots\hat{v}_{m_n} \times$$

$$H(\omega_{m_1})H(\omega_{m_2})\ldots H(\omega_{m_n}) e^{j(\omega_{m_1}+\omega_{m_2}+\ldots+\omega_{m_n})t} \,. \tag{4.40}$$

This is a simple closed-form result which, unfortunately, can only be obtained if a linear relation of the form (4.38) is valid. This already excludes the equivalent circuits of a nonlinearly loaded antenna in Fig. 4.3.

In Volterra series analysis it is shown that in case of an excitation of the form (4.36) the response $i_{\mathrm{nl}}(t)$ of *any* weakly nonlinear circuit can be expressed as

$$i_{\mathrm{nl}}(t) = \sum_{n=1}^{N} \frac{1}{2^n} \sum_{\substack{m_1=-M \\ m_1\neq 0}}^{M} \sum_{\substack{m_2=-M \\ m_2\neq 0}}^{M} \cdots \sum_{\substack{m_n=-M \\ m_n\neq 0}}^{M} \hat{v}_{m_1}\hat{v}_{m_2}\ldots\hat{v}_{m_n} \times$$

$$H_n(\omega_{m_1},\omega_{m_2}\ldots\omega_{m_n}) e^{j(\omega_{m_1}+\omega_{m_2}+\ldots+\omega_{m_n})t} \,. \tag{4.41}$$

Here we introduced the functions $H_n(\omega_{m_1}, \omega_{m_2} \ldots \omega_{m_n})$ that are known as *nth-order nonlinear transfer functions*. The expression (4.41) is a generalization of (4.40). Clearly, for the special case

$$H_n(\omega_{m_1}, \omega_{m_2} \ldots \omega_{m_n}) = a_n H(\omega_{m_1}) H(\omega_{m_2}) \ldots H(\omega_{m_n}) \tag{4.42}$$

both expressions become equivalent.

In practice, the main task of Volterra series analysis is to determine the nonlinear transfer functions $H_n(\omega_{m_1}, \omega_{m_2} \ldots \omega_{m_n})$. Appropriate analytic methods include the *harmonic input method* and the *method of nonlinear currents* [133, §4], see also [15, 24]. These methods are reminiscent of the method of successive approximation which has been introduced in the last section since higher order contributions to the solution are obtained by working out powers of lower order contributions. In fact, since both the Volterra series analysis and the method of successive approximation are analytic methods which yield the solution in the form of a generalized power series it is expected that they produce identical results. A comparison of both methods, which employs iteration techniques, exhibits the equivalence of both methods [34]. However, the Volterra series analysis yields the solution in the form (4.41) which is advantageous for the analysis and design of analog circuits [231, 133].

In the context of nonlinearly loaded antennas the Volterra series analysis has been used to calculate for a specific example the frequency spectrum of the scattered electromagnetic field [188], and to this end nonlinear transfer functions were analytically worked out up to the third order. Similar calculations have been performed in the context of distortion analysis [48, 231]. In general, the third order appears to be an upper limit for the analytic calculation of transfer functions by hand. To calculate higher order transfer functions in this way simply becomes too laborious. Thus, similar as for the case of the method of successive approximation, it is tempting to use computer algebra to program the required calculations. It turns out that symbolic network analysis programs can be suitable in this respect [56, 57, 231]. But these programs are much more complex than a few dozens lines of computer algebra code that are sufficient to work out Picard iterations.

4.2.3   Harmonic balance technique

The harmonic balance technique is not restricted to weakly nonlinear circuits. However, it is not a purely analytic method and usually requires the use of numerical procedures. The main steps of the harmonic balance technique are based on the division of a nonlinear circuit into two linear and nonlinear circuits, as displayed in Fig. 4.6.

1. Start from the linear circuit of Fig. 4.6 and provide an initial guess for $V(\omega)$, i.e., for all harmonics that are represented by $V(\omega)$.

2. Compute from $V(\omega)$ the current $I_1(\omega)$. This is equivalent to the solution of a linear circuit problem.

3. Apply an inverse Fourier transformation to the voltage $V(\omega)$ to calculate the time domain current $v(t)$.

4. Turn to the nonlinear circuit of Fig. 4.6. and obtain $i_{\mathrm{nl}}(t)$ from the nonlinear voltage to current relation

$$i(t) = f\left(v(t)\right) . \tag{4.43}$$

   Then apply a Fourier transformation to calculate $I_{\mathrm{nl}}(\omega)$.

5. Compare the linear current $I_{\mathrm{l}}(\omega)$ of step 2 to the nonlinear current $I_{\mathrm{nl}}(\omega)$ of step 4. If both currents are (approximately) equal a solution is obtained. If the error

$$e_{I_{\mathrm{l}}-I_{\mathrm{nl}}} := I_{\mathrm{l}}(\omega) - I_{\mathrm{nl}}(\omega) \tag{4.44}$$

   appears to be not acceptable modify the voltage $V(\omega)$ "in an appropriate way". With this modified voltage return to step 2 and repeat the process until the error (4.44) is below an acceptable limit and, thus, the currents $I_{\mathrm{l}}(\omega)$ and $I_{\mathrm{nl}}(\omega)$ are balanced.

To modify in step 5 the voltage $V(\omega)$ "in an appropriate way" usually is reduced to finding zeros of a set of nonlinear equations. Corresponding iterative solution methods are part of the mathematical literature [165, 101]. These can also be used to solve the nonlinear equation in step 4. The standard way to numerically perform the required Fourier transformations in step 3 and step 4 is to apply the Fast Fourier Transform [13].

Harmonic balance techniques[3] have been used to study scattering properties of non-linearly loaded antennas and antenna arrays [91, 118]. The examples considered involved two-tone excitations and in these cases intermodulation phenomena were observed on the grounds of numerical calculations. But it is a disadvantage of the common harmonic balance techniques that in case of a two-tone or multi-tone excitation the various frequencies must be pairwise commensurable, i.e., the frequencies must be harmonics of a common fundamental frequency. This is the actual reason for the name *harmonic* balance technique. It should be mentioned, however, that methods have been proposed to generalize the harmonic balance technique in order to overcome this limitation [133, §3.6].

4.3   Calculation of intermodulation phenomena

We have introduced three frequency domain methods of nonlinear circuit theory in order to be able to calculate intermodulation phenomena at nonlinearly loaded antennas within resonators. We summarize the properties of these methods:

---

[3]Here the harmonic balance technique is in the plural since different variations of this technique exist [133, §3].

- The method of successive approximation is an analytic method which is suitable to analyze weak nonlinearities that are driven by small signals. It is laborious to use if higher order iterations are calculated by hand. However, it is rather straightforward to program the required analytic calculations by means of computer algebra.

- The Volterra series analysis also is an analytic method which is appropriate for the analysis of weak nonlinearities and small signals. The computation of higher order nonlinear transfer functions is laborious. Unlike the method of successive approximation, the Volterra series analysis is not easily transformed into a computer algorithm.

- The harmonic balance technique is a numerical method which is appropriate to analyze a variety of nonlinearities. However, it cannot directly be applied to general multi-tone excitations.

Considering these points it seems appropriate to apply in the following the method of successive approximation.

4.3.1 Intermodulation at a nonlinearly loaded antenna within a cavity

In order to study nonlinearly loaded antennas within cavities we need to know the input impedance at the antenna port where the nonlinear load is located. To become specific we focus on single dipole antennas that are nonlinearly loaded at the antenna center. Then an input impedance is represented by a self impedance, as calculated along the lines of Sec. 3.4.2. To illustrate this point we now return to the antenna-cavity configuration of Sec. 3.4.2 and reconsider antenna 1.

(a) Equivalent circuit of an antenna within a cavity

The real and imaginary part of the self impedance of antenna 1 are shown in Figs. 3.20 and 3.22, respectively. In order to model this self impedance by an equivalent circuit we recall from Sec. 3.1.4 (c) that the self impedance of an antenna element within a cavity can be composed of contributions of longitudinal and transverse eigenfunctions. The transverse contributions are represented as a collection of parallel RLC-circuits. Accordingly, we consider a circuit as displayed in Fig. 4.7. This circuit contains four parallel RLC circuits in series, representing the four modes 110, 130, 310, and 330 that couple to antenna 1 in the frequency range up to 100 MHz. The circuit also contains a capacitor C_{free} and an inductor L_{free} to take into account the antenna properties in free space. These two lumped element constitute a simplified pole/zero expansion of the corresponding input impedance in free space [192, §9.3], [8]. The values of C_{free} and L_{free} can be fitted to match the method of moments solution of the input impedance of
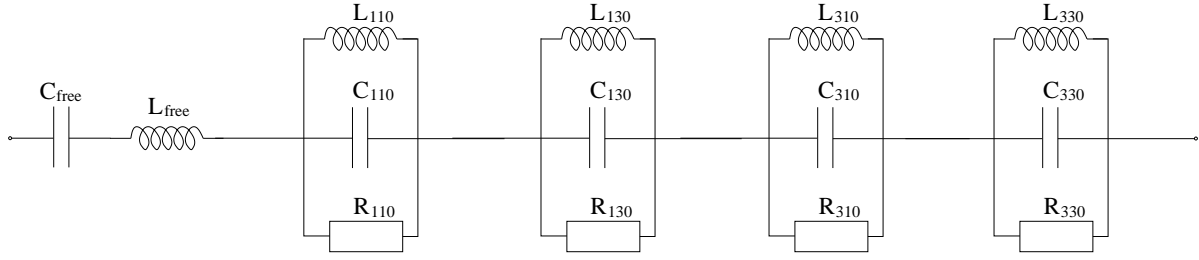
Figure 4.7: Equivalent circuit of the self impedance of antenna 1 which is located within a rectangular cavity. This equivalent circuit is valid in the frequency range up to 100 MHz where antenna 1 couples to four cavity resonances. For higher frequencies the coupling to additional resonances needs to be taken into account. However, this is trivial since in this case it is only required to add more parallel RLC circuits.

free space	resonance 110	resonance 130	resonance 310	resonance 330
$L_{\text{free}} = 286.4\text{nH}$	$L_{110} = 12.65\text{nH}$	$L_{130} = 4.171\text{nH}$	$L_{310} = 3.469\text{nH}$	$L_{330} = 2.434\text{nH}$
$C_{\text{free}} = 2.652\text{pF}$	$C_{110} = 1.850\text{nF}$	$C_{130} = 1.278\text{nF}$	$C_{310} = 1.202\text{nF}$	$C_{330} = 1.068\text{nF}$
———	$R_{110} = 1500\Omega$	$R_{130} = 1500\Omega$	$R_{310} = 1500\Omega$	$R_{330} = 1500\Omega$

Table 4.1: Explicit values of the lumped elements that constitute the equivalent circuit of Fig. 4.7.

free space. The values of the lumped elements L_{mnp}, C_{mnp} and R_{mnp} are determined from the method of moments solution that yields Figs. 3.20 and 3.22 and from the relations (3.64)–(3.66). As a result we obtain the values that are displayed in Tab. 4.1. With these values we calculate the impedance Z_{circ} of the equivalent circuit. The real and imaginary part of the result are shown in Fig. 4.8. Both parts agree well with the method of moments solution. One should note that the height of a plotted resonance peak depends on how close the actual resonance frequency is approached by one of the sampled frequencies.

We now have specified the equivalent circuit of an antenna impedance which can be implemented in the Thévenin equivalent of Fig. 4.3. In order to complete the problem it remains to choose a nonlinearity and an excitation. We will assume that the nonlinearity is characterized by a voltage-current relation of the form

$$v_{\text{nl}}(t) = b_1 i(t) + b_2 i^2(t) + b_3 i^3(t) \tag{4.45}$$

with constant parameters b_1, b_2 and b_3. As excitation we choose a two-tone excitation as

$$v_{\text{eq}}(t) = v_1 \cos(\omega_1 t) + v_2 \cos(\omega_2 t) \tag{4.46}$$

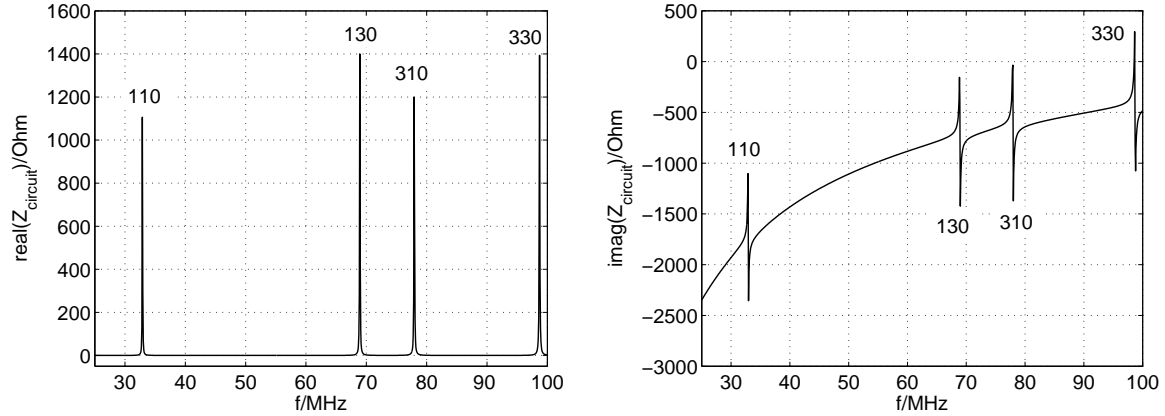with constant amplitudes V_1 and V_2. This yields a particular example of an equivalent

Figure 4.8: Real and imaginary part of the impedance $Z_{\text{circ}}(f)$ of the equivalent circuit of Fig. 4.7. They represent approximations to the method of moments solution of Figs. 3.20 and 3.22.

circuit which models the voltage and current at the nonlinear load of an antenna within a cavity.

(b) Application of the method of successive approximation

We now apply the method of successive approximation. The general algorithm has been described in Sec. 4.2.1 and the example of the same section exhibited the first iteration in detail. One should note that the method of successive approximation requires that a possible linear part of the nonlinearity gets absorbed in the circuit impedance $Z(\omega)$. In our specific case the voltage-current relation (4.27) exhibits a linear part which is determined by the constant b_1. It follows that the impedance $Z(\omega)$ of the general example in Sec. 4.2.1 now assumes the form $Z_{\text{circ}}(\omega) + b_1$.

The solution of the first order iteration is displayed in (4.32) and (4.34). It is clear that the impedance $Z(\omega)$ has no effect on the values of intermodulation frequencies but determines the amplitudes of the resulting spectrum. For further illustration we fix the parameters b_1, b_2, b_3, V_1, V_2 and choose

$$b_1 = 1\,k\Omega\,, \qquad b_2 = -6\,\frac{k\Omega}{A}\,, \qquad b_3 = 3\,\frac{k\Omega}{A^2}\,, \tag{4.47}$$

$$v_1 = v_2 = 100\,\text{V}\,. \tag{4.48}$$

For the given value of b_1 we plot in Fig. 4.9 the absolute value and the argument of $Z_{\text{circ}}(\omega) + b_1$. Similar to Fig. 4.8 resonance peaks are observed whenever antenna 1 couples to a cavity resonance.

It remains to choose exciting frequencies $\omega_1 = 2\pi f_1$ and $\omega_2 = 2\pi f_2$. We first set $f_1 = 70\text{MHz}$ and $f_2 = 80\text{MHz}$. After four Picard iterations, that have been calculated
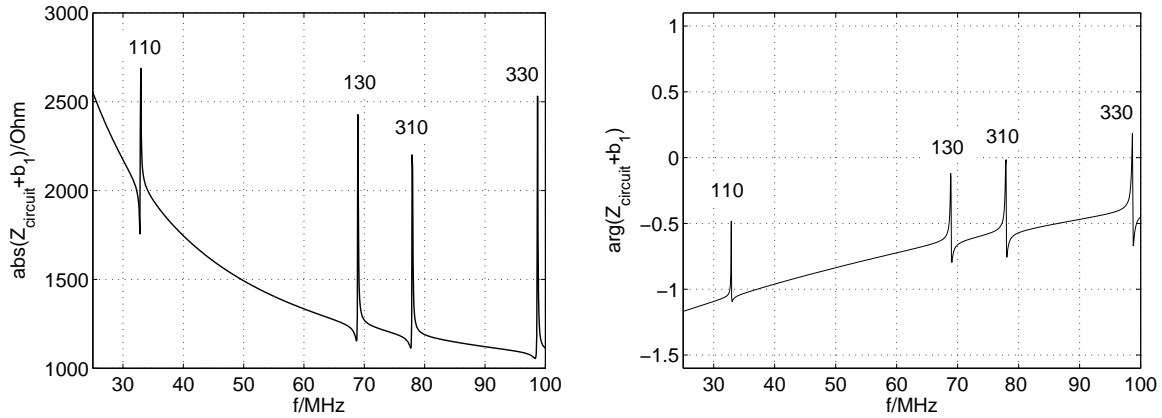
Figure 4.9: Absolute value and argument of the complex impedance $Z_{\text{circuit}}(\omega) + b_1$.

with the aid of the computer algebra program *Maple*, the amplitudes of the major frequencies in the emerging spectrum of the current through the nonlinear load do no longer significantly change and the corresponding result is shown in Fig. 4.10. The qualitative form of the spectrum is expected from the simple power series analysis of Sec. 4.1.2, it looks analogous to the spectrum of Fig. 4.5. We note, in particular, the emergence of the low intermodulation frequency at 10MHz.

The chosen frequencies $f_1 = 70$MHz and $f_2 = 80$MHz are slightly above the resonance frequencies f_{130} and f_{310}, respectively. The values of the antenna impedance at 70MHz and 80MHz are essentially those of free space. To change this situation we lower the exciting frequencies and choose $f_1 = 69$MHz and $f_2 = 78$MHz such that they approach resonance peaks with significantly increased absolute values of the antenna impedance, compare the left graph of Fig. 4.9. Since the absolute value of the antenna impedance enters the denominator of the terms that emerge from the Picard iteration, see (4.34), we expect that the new choice of exciting frequencies will lower the amplitudes of the current spectrum. This is verified by Fig. 4.11 which shows the current spectrum which results after four Picard iterations. An additional and obvious explanation of this phenomenon is that the cavity resonances, which are represented by the parallel RLC-circuits of Fig. 4.7, act like filters that suppress the exciting frequencies.

In the resonance region the absolute value of the impedance $Z_{\text{circuit}}(\omega) + b_1$ may also decrease if compared to the corresponding value that were obtained if the antenna were located in free space. This is observed in the left graph of Fig. 4.9: Towards lower frequencies any resonance peak attains a (not very pronounced) minimum. If we move the exciting frequencies to the vicinity of these minima the amplitudes of the resulting current spectrum should become (slightly) larger if compared to the first choice of exciting frequencies which resembled the situation of free space. Indeed, if we choose $f_1 = 68.7$MHz and $f_2 = 77.7$MHz we obtain after four Picard iterations the current spectrum of Fig. 4.12. Apparently, the amplitudes of the spectrum are larger than
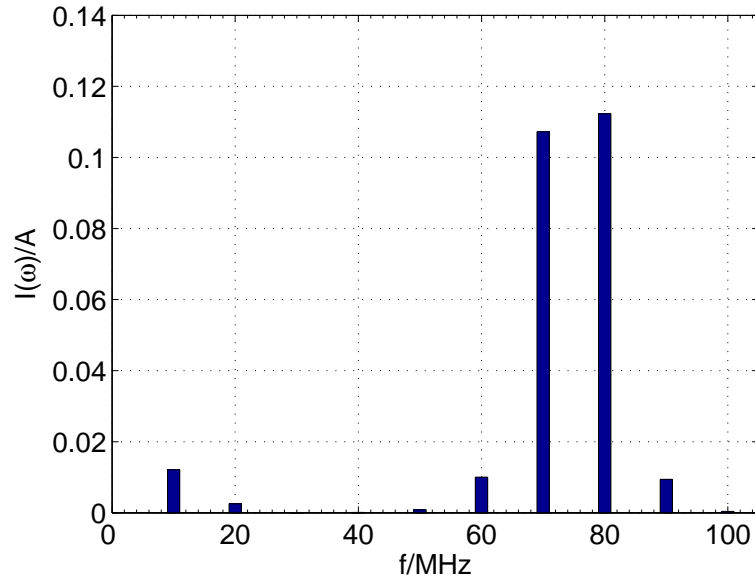
Figure 4.10: Resulting spectrum of the current $I(\omega)$ through the nonlinear antenna load. The chosen two-tone excitation consists of the frequencies $f_1 = 70\text{MHz}$ and $f_2 = 80\text{MHz}$. Significant contributions to the current at the intermodulation frequencies 10MHz, 60MHz, and 90 MHz are clearly visible.

those of Fig. 4.10. One should note that this is not a pure resonance effect. A pure resonance curve of a parallel RLC circuit, as exemplified by Fig. 3.6, exhibits exactly one maximum of the absolute value of the corresponding impedance but no minimum. The minima in the left graph of Fig. 4.9 are due to the fact that the free space part of the antenna impedance, which in the equivalent circuit of Fig. 4.7 is represented by the capacitance C_{free}, adds a large negative value to the imaginary part of the antenna impedance.

4.3.2   Remarks

It should be apparent by now that for the analysis of a nonlinearly loaded antenna by means of nonlinear circuit theory it is not really important if the antenna is located in a resonating environment or not. However, before we are able to invoke nonlinear circuit theory we have to know the equivalent circuit which characterizes the linear electromagnetic properties of the antenna that are present if the nonlinear load is removed. These linear electromagnetic properties, which are represented by antenna impedances, will reflect possible resonance effects. For the modeling and calculation of the relevant antenna impedances the methods of Chap. 3 are suitable. In particular, it turns out that within an equivalent circuit electromagnetic resonances are represented by parallel
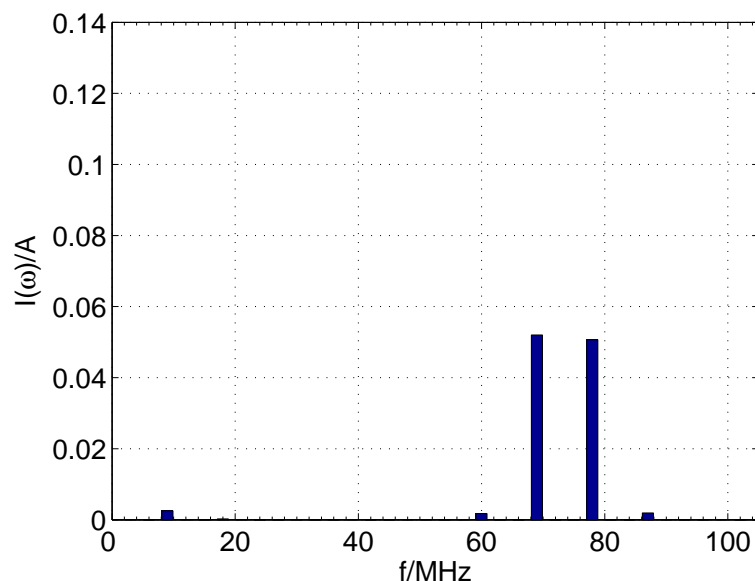
Figure 4.11: Resulting spectrum of the current $I(\omega)$ through the nonlinear antenna load if the two-tone excitation consists of the (resonant) frequencies $f_1 = 69$MHz and $f_2 = 78$MHz. If compared to the spectrum of Fig. 4.10 the amplitudes are considerably reduced.

RLC-subcircuits. The explicit values of the elements that constitute these subcircuits have to be determined from the complete solution of the corresponding linear antenna problem.

The nonlinear circuit theory provides voltage and current spectra as primary results. These spectra are composed of intermodulation frequencies. It follows from our analysis that the *values* of these intermodulation frequencies do depend on the excitation frequencies and on the nonlinear load, but they are independent of the linear antenna properties. In particular, they are independent of possible resonance effects. However, if resonance effects are present these can affect the *amplitudes* of the resulting spectra.

Therefore, the influence of resonance effects on the voltage and current spectra appear to be not drastic. But it would be rush in our judgment to assume that this excludes any drastic influence of resonance effects. We have calculated intermodulation frequencies such as the low intermodulation frequency $\Delta\omega = |\omega_1 - \omega_2|$ which results from a two-tone excitation. Once the values of these intermodulation frequencies coincide with a resonance frequency, drastic electromagnetic coupling effects can happen, as exemplified by the analysis of mutual antenna impedances in the presence of resonances in Chap. 3. We remark that strong intermodulation effects at nonlinearities within a resonator have also been experimentally verified for the special case of repetitive sinusoidal pulses as excitation [112].
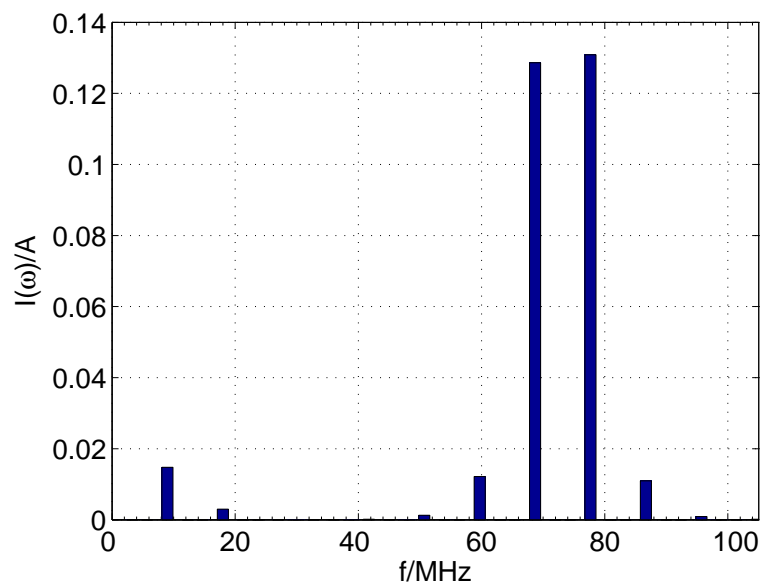
Figure 4.12: Resulting spectrum of the current $I(\omega)$ through the nonlinear antenna load if the two-tone excitation consists of the frequencies $f_1 = 68.7$MHz and $f_2 = 77.7$MHz. If compared to the spectrum of Fig. 4.10 the amplitudes are slightly increased.

We finally mention that analytic iterative techniques not only are suitable to predict intermodulation frequencies and their corresponding amplitudes. They can also be used to design filters, the so-called "pre-inverses" or "post-inverses", to suppress intermodulation frequencies in order to protect a system. To get an impression of this possibility we refer to [158, 231].

Chapter 5

Antenna Theory and Transmission Line Theories

Transmission lines consist of metallic structures that transmit electromagnetic signals and energy. In this respect they are similar to systems of transmitting and receiving antennas. However, the physical mechanisms that govern the electromagnetic transmission along transmission lines are quite different if compared to the electromagnetic transmission between pairs of antennas. This circumstance is illustrated in Fig. 5.1.
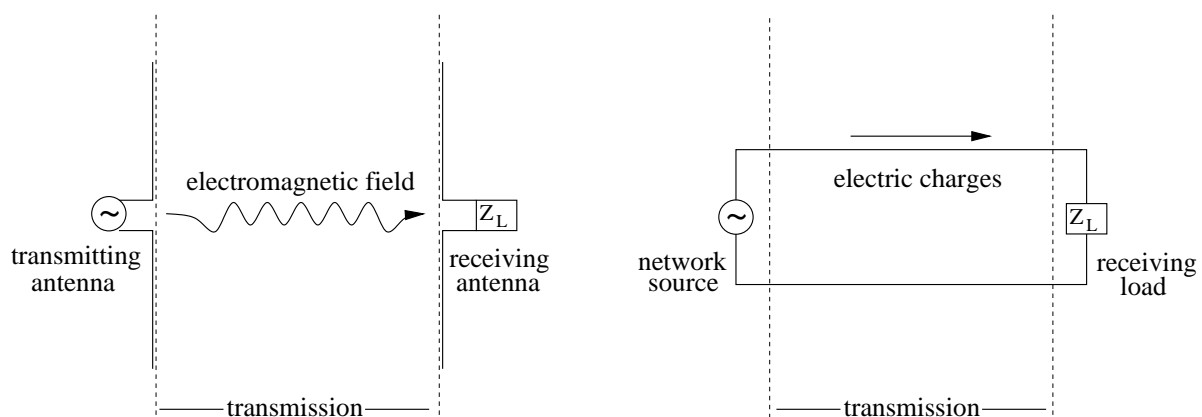

Figure 5.1: Electromagnetic transmission by means of a pair of antennas (left) and a transmission line (right). Inbetween the antennas an electromagnetic field mediates the actual transmission while the transmission line provides electric charges that mediate the transmission between the source and the load.

Inbetween a pair of antennas the electromagnetic transmission results from a propagating electromagnetic field which, for practical purposes, can often be approximated by a radiation field. This does not mean that in such a situation no Coulomb field is present. A Coulomb field will be related to the electric charges that move along the antennas and

will constitute their near-field. But in many cases the transmitting and receiving antennas are sufficiently far apart such that the main coupling is mediated by the radiation field which resembles a freely propagating electromagnetic field[1]. Electric charges are *not* involved in the actual electromagnetic transmission that happens inbetween the antennas. They only are required at the beginning and at the end of the transmission in order to, respectively, generate and receive the transmitting electromagnetic field.

The electromagnetic transmission along a transmission line *does* involve electric charges. These charges are located on the transmission line which normally consists of a highly conducting material. They are accompanied by a Coulomb field which dominate their mutual electromagnetic interaction at short distances. While the electric charges get accelerated, a radiation field will be produced. In particular, this will happen at high frequencies or if the transmission line is strongly curved or bent. Normally, the generation of a radiation field by electric charges on a transmission line is an unwanted effect which may affect the properties of a transmission line. For many situations this effect is small and negligible.

It follows that the dynamics along a transmission line is determined from the motion of electric charges and not from the degrees of freedom of an electromagnetic field. The fact that for the electromagnetic transmission along a transmission line the degrees of freedom of the electromagnetic field can often be neglected implies that the *classical transmission line theory* (classical TLT) has a much simpler structure than the Maxwell theory. The classical TLT is a limiting case of the Maxwell theory and contains the electric current $I(z)$, representing electric charges, and the electric voltage $V(z)$, representing the associated Coulomb fields, as main physical quantities [103, 169, 226, 217]. Clearly, these two quantities are not independent of each other. They are related by *Telegrapher equations* which, for a two-wire transmission line and in the frequency domain, are of the form

$$\frac{\partial V(z)}{\partial z} + (j\omega L' + R')I(z) = V_s'(z)\,, \tag{5.1}$$

$$\frac{\partial I(z)}{\partial z} + (j\omega C' + G')V(z) = I_s'(z)\,. \tag{5.2}$$

The primed quantities denote *per-unit-length parameters*. Explicitly, the quantities L', R', C', and G' are the per-unit-length inductance, resistance, capacitance, and conductance, respectively. They represent geometric and material properties of the transmission line. The quantities V_s' and I_s' denote *distributed voltage and current sources*, respectively, and represent the electromagnetic excitation of the line. For a given exciting

---

[1]We remind us at this point that in the presence of moving electric charges there will be an extended nonvanishing transverse current $\mathbf{J}_\perp$ which, in turn, will couple to the transverse parts of the electromagnetic field. In this general case there is no free propagation of an electromagnetic field since the wave equations (1.212) and (1.213) will always contain a non-vanishing source term. The propagation can only be *approximately* free at large distances to moving charges.

electromagnetic field their explicit form depends on the choice of one of several but equivalent coupling models which is used to calculate V_s' or I_s' [214, 3, 178].

Even though the classical TLT is approximate it is often preferred over the exact Maxwell theory because the Telegrapher equations (5.1), (5.2), which constitute a set of coupled first order differential equations, are much easier to solve than the complete set of the Maxwell equations. The price that has to be paid for this simplification is the limited scope of the classical TLT. It is obvious that prior to the application of classical TLT its limitations need to be understood.

If the conditions for an application of the classical TLT are not met it is suggestive to enlarge the range of applicability of the Telegrapher equations while keeping their mathematical structure. This leads to the subject of *generalized transmission line theories* (generalized TLTs). We have the relations

$$\text{Maxwell Theory} \quad \subseteq \quad \text{generalized TLTs} \quad \subset \quad \text{classical TLT} \,.$$

In order to arrive at a generalized TLT, two meaningful possibilities come to mind:

1. Supplement the classical TLT by corrective terms that take into account additional effects.

2. Derive more general Telegrapher equations from the exact Maxwell theory.

The second possibility appears to be the more systematic and logical one.

In this chapter we will discuss if the classical TLT can be applied in the presence of resonances. To this end we provide in Sec. 5.1 a derivation of the classical TLT from the full Maxwell theory, expressed in terms of integral equations of antenna theory, namely in terms of Pocklington's equation in the form (3.71) and the mixed potential integral equation (3.89). Based on this derivation we will comment in Sec. 5.2 on generalized TLTs and finally argue why the classical TLT can lead to satisfying results if applied inside a resonator.

5.1 Classical transmission line theory deduced from antenna theory

5.1.1 Coupled Pocklington's equations, antenna and transmission line mode

We consider a set of coupled Pocklington's equations (electric field integral equations) which models the electromagnetic coupling to a system of wires. The wires are supposed to represent a transmission line. For concreteness we consider two wires and assume a

thin-wire approximation. Then the corresponding coupled Pocklington's equations are analogous to those who model two coupled thin-wire antennas,

$$j\omega\mu \left[\int_{\text{wire 1}} \overline{\boldsymbol{G}}^E(\tau_1, \tau_1') \boldsymbol{I}_1(\tau_1') \, d\tau_1' + \int_{\text{wire 2}} \overline{\boldsymbol{G}}^E(\tau_1, \tau_2') \boldsymbol{I}_2(\tau_2') \, d\tau_2' \right] \cdot \boldsymbol{e}_{\tau_1} = E_{\text{tan}}^{\text{inc}}(\tau_1), \quad (5.3)$$

$$j\omega\mu \left[\int_{\text{wire 1}} \overline{\boldsymbol{G}}^E(\tau_2, \tau_1') \boldsymbol{I}_1(\tau_1') \, d\tau_1' + \int_{\text{wire 2}} \overline{\boldsymbol{G}}^E(\tau_2, \tau_2') \boldsymbol{I}_2(\tau_2') \, d\tau_2' \right] \cdot \boldsymbol{e}_{\tau_2} = E_{\text{tan}}^{\text{inc}}(\tau_2). \quad (5.4)$$

Here we introduced the variables τ_1, τ_2 that parameterize the length of wire 1 and wire 2, respectively. Fixed values of these variables represent fixed wire positions. The unit vectors $\boldsymbol{e}_{\tau_1}$, $\boldsymbol{e}_{\tau_1}$ are tangent to the line-like wires. The currents $\boldsymbol{I}_1(\tau_1)$, $\boldsymbol{I}_2(\tau_2)$ result from the thin-wire approximation and are defined by

$$\boldsymbol{I}_i(\tau_i) := I_i \boldsymbol{e}_{\tau_i} \qquad (5.5)$$

for $i = 1, 2$. The scalar I_i is the value of the electric current at the wire position τ_i.

If the wires form a transmission line we expect that they can be parameterized by a common coordinate ξ with value $\xi = \xi_0$ at the beginning and $\xi = \xi_L$ at the end of the line, compare Fig. 5.2. We take this coordinate as a common integration variable and write (5.3), (5.4) as

$$j\omega\mu \left[\int_{\xi_0}^{\xi_L} \left(\overline{\boldsymbol{G}}^E(\tau_1, \tau_1') \boldsymbol{I}_1(\tau_1') \frac{\partial \tau_1'}{\partial \xi'} + \overline{\boldsymbol{G}}^E(\tau_1, \tau_2') \boldsymbol{I}_2(\tau_2') \frac{\partial \tau_1'}{\partial \xi'} \right) d\xi' \right] \cdot \boldsymbol{e}_{\tau_1} = E_{\text{tan}}^{\text{inc}}(\tau_1), \quad (5.6)$$

$$j\omega\mu \left[\int_{\xi_0}^{\xi_L} \left(\overline{\boldsymbol{G}}^E(\tau_2, \tau_1') \boldsymbol{I}_1(\tau_1') \frac{\partial \tau_1'}{\partial \xi'} + \overline{\boldsymbol{G}}^E(\tau_2, \tau_2') \boldsymbol{I}_2(\tau_2') \frac{\partial \tau_2'}{\partial \xi'} \right) d\xi' \right] \cdot \boldsymbol{e}_{\tau_2} = E_{\text{tan}}^{\text{inc}}(\tau_2). \quad (5.7)$$

The variables τ_1, τ_2 are now understood as functions of the parameter ξ.

Next we introduce two currents $\boldsymbol{I}_A$ and $\boldsymbol{I}_{\text{TL}}$ as linear combinations of $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$,

$$\boldsymbol{I}_A := \frac{1}{2} \left(\boldsymbol{I}_1 + \boldsymbol{I}_2 \right), \qquad (5.8)$$

$$\boldsymbol{I}_{\text{TL}} := \frac{1}{2} \left(\boldsymbol{I}_1 - \boldsymbol{I}_2 \right). \qquad (5.9)$$

The inverse equations are

$$\boldsymbol{I}_1 = \boldsymbol{I}_A + \boldsymbol{I}_{\text{TL}}, \qquad (5.10)$$

$$\boldsymbol{I}_2 = \boldsymbol{I}_A - \boldsymbol{I}_{\text{TL}}. \qquad (5.11)$$

Identifications of this kind are familiar from the classical TLT with $\boldsymbol{I}_A$ being reminiscent of the so-called "antenna mode" or "common mode" current and $\boldsymbol{I}_{\text{TL}}$ being reminiscent of the "tranmission line mode" or "differential mode" current. In the present context these identifications are still formal. We note that neither $\boldsymbol{I}_A$ nor $\boldsymbol{I}_{\text{TL}}$ need to be tangent to
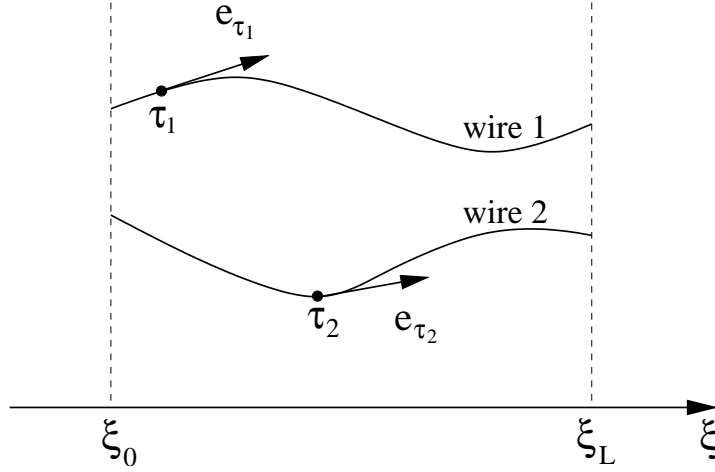
Figure 5.2: Introduction of a common variable ξ which parameterizes the wires of a transmission line.

one of the wires. But it is clear that we still may split $\boldsymbol{I}_A$ and $\boldsymbol{I}_{TL}$ into a component and a unit vector,

$$\boldsymbol{I}_A = \frac{1}{2}\left(I_1\boldsymbol{e}_{\tau_1} + I_2\boldsymbol{e}_{\tau_2}\right) =: I_A\boldsymbol{e}_{I_A}\,, \tag{5.12}$$

$$\boldsymbol{I}_{TL} = \frac{1}{2}\left(I_1\boldsymbol{e}_{\tau_1} - I_2\boldsymbol{e}_{\tau_2}\right) =: I_{TL}\boldsymbol{e}_{I_{TL}}\,. \tag{5.13}$$

If the relations (5.10), (5.11) are inserted into (5.6), (5.7) it is simple to find

$$j\omega\mu\left[\int_{\xi_0}^{\xi_L}\left(\left[\overline{\boldsymbol{G}}^E(\tau_1,\tau_1')\frac{\partial\tau_1'}{\partial\xi'} + \overline{\boldsymbol{G}}^E(\tau_1,\tau_2')\frac{\partial\tau_2'}{\partial\xi'}\right]I_A(\xi')\right.\right.$$
$$\left.\left. + \left[\overline{\boldsymbol{G}}^E(\tau_1,\tau_1')\frac{\partial\tau_1'}{\partial\xi'} - \overline{\boldsymbol{G}}^E(\tau_1,\tau_2')\frac{\partial\tau_2'}{\partial\xi'}\right]I_{TL}(\xi')\right)d\xi'\right]\cdot\boldsymbol{e}_{\tau_1} = E_{\tan}^{inc}(\tau_1)\,, \quad (5.14)$$

$$j\omega\mu\left[\int_{\xi_0}^{\xi_L}\left(\left[\overline{\boldsymbol{G}}^E(\tau_2,\tau_1')\frac{\partial\tau_1'}{\partial\xi'} + \overline{\boldsymbol{G}}^E(\tau_2,\tau_2')\frac{\partial\tau_2'}{\partial\xi'}\right]I_A(\xi')\right.\right.$$
$$\left.\left. + \left[\overline{\boldsymbol{G}}^E(\tau_2,\tau_1')\frac{\partial\tau_1'}{\partial\xi'} - \overline{\boldsymbol{G}}^E(\tau_2,\tau_2')\frac{\partial\tau_2'}{\partial\xi'}\right]I_{TL}(\xi')\right)d\xi'\right]\cdot\boldsymbol{e}_{\tau_2} = E_{\tan}^{inc}(\tau_2)\,. \quad (5.15)$$

We both add and subtract these equations and obtain

$$j\omega\mu\int_{\xi_0}^{\xi_L}\left(G_{+A}^E(\tau_1,\tau_2,\tau_1',\tau_2')I_A(\xi') + G_{+TL}^E(\tau_1,\tau_2,\tau_1',\tau_2')I_{TL}(\xi')\right)d\xi' = E_{\tan}^{inc}(\tau_1) + E_{\tan}^{inc}(\tau_2)\,,$$
$$(5.16)$$

$$j\omega\mu\int_{\xi_0}^{\xi_L}\left(G_{-A}^E(\tau_1,\tau_2,\tau_1',\tau_2')I_A(\xi') + G_{-TL}^E(\tau_1,\tau_2,\tau_1',\tau_2')I_{TL}(\xi')\right)d\xi' = E_{\tan}^{inc}(\tau_1) - E_{\tan}^{inc}(\tau_2)\,,$$
$$(5.17)$$

where we introduced the abbreviations

$$G_{+\mathrm{A}}^{E}(\tau_1, \tau_2, \tau_1', \tau_2') := \boldsymbol{e}_{\tau_1} \cdot \left(\overline{\boldsymbol{G}}^{E}(\tau_1, \tau_1')\frac{\partial \tau_1'}{\partial \xi'} + \overline{\boldsymbol{G}}^{E}(\tau_1, \tau_2')\frac{\partial \tau_2'}{\partial \xi'} \right) \cdot \boldsymbol{e}_{I_A}$$
$$+ \boldsymbol{e}_{\tau_2} \cdot \left(\overline{\boldsymbol{G}}^{E}(\tau_2, \tau_1')\frac{\partial \tau_1'}{\partial \xi'} + \overline{\boldsymbol{G}}^{E}(\tau_2, \tau_2')\frac{\partial \tau_2'}{\partial \xi'} \right) \cdot \boldsymbol{e}_{I_A}, \qquad (5.18)$$

$$G_{+\mathrm{TL}}^{E}(\tau_1, \tau_2, \tau_1', \tau_2') := \boldsymbol{e}_{\tau_1} \cdot \left(\overline{\boldsymbol{G}}^{E}(\tau_1, \tau_1')\frac{\partial \tau_1'}{\partial \xi'} - \overline{\boldsymbol{G}}^{E}(\tau_1, \tau_2')\frac{\partial \tau_2'}{\partial \xi'} \right) \cdot \boldsymbol{e}_{I_{TL}}$$
$$+ \boldsymbol{e}_{\tau_2} \cdot \left(\overline{\boldsymbol{G}}^{E}(\tau_2, \tau_1')\frac{\partial \tau_1'}{\partial \xi'} - \overline{\boldsymbol{G}}^{E}(\tau_2, \tau_2')\frac{\partial \tau_2'}{\partial \xi'} \right) \cdot \boldsymbol{e}_{I_{TL}}, \qquad (5.19)$$

$$G_{-\mathrm{A}}^{E}(\tau_1, \tau_2, \tau_1', \tau_2') := \boldsymbol{e}_{\tau_1} \cdot \left(\overline{\boldsymbol{G}}^{E}(\tau_1, \tau_1')\frac{\partial \tau_1'}{\partial \xi'} + \overline{\boldsymbol{G}}^{E}(\tau_1, \tau_2')\frac{\partial \tau_2'}{\partial \xi'} \right) \cdot \boldsymbol{e}_{I_A}$$
$$- \boldsymbol{e}_{\tau_2} \cdot \left(\overline{\boldsymbol{G}}^{E}(\tau_2, \tau_1')\frac{\partial \tau_1'}{\partial \xi'} + \overline{\boldsymbol{G}}^{E}(\tau_2, \tau_2')\frac{\partial \tau_2'}{\partial \xi'} \right) \cdot \boldsymbol{e}_{I_A}, \qquad (5.20)$$

$$G_{-\mathrm{TL}}^{E}(\tau_1, \tau_2, \tau_1', \tau_2') := \boldsymbol{e}_{\tau_1} \cdot \left(\overline{\boldsymbol{G}}^{E}(\tau_1, \tau_1')\frac{\partial \tau_1'}{\partial \xi'} - \overline{\boldsymbol{G}}^{E}(\tau_1, \tau_2')\frac{\partial \tau_2'}{\partial \xi'} \right) \cdot \boldsymbol{e}_{I_{TL}}$$
$$- \boldsymbol{e}_{\tau_2} \cdot \left(\overline{\boldsymbol{G}}^{E}(\tau_2, \tau_1')\frac{\partial \tau_1'}{\partial \xi'} - \overline{\boldsymbol{G}}^{E}(\tau_2, \tau_2')\frac{\partial \tau_2'}{\partial \xi'} \right) \cdot \boldsymbol{e}_{I_{TL}}. \qquad (5.21)$$

5.1.2 Decoupling of antenna and transmission line mode current: uniform transmission lines in free space

The expressions we obtained so far look more complicated than the original equations (5.3) and (5.4) that we started from. To nevertheless appreciate this form of coupled Pocklington's equations we specialize to the case of straight and parallel wires. We align a Cartesian coordinate system such that the z-axis is parallel to the wires and choose $\xi = z$. This leads to the simplifications

$$\boldsymbol{e}_{\tau_1} = \boldsymbol{e}_{\tau_2} = \boldsymbol{e}_{I_A} = \boldsymbol{e}_{I_{TL}}, \qquad (5.22)$$
$$\boldsymbol{e}_{\tau_{1,2}} \cdot \overline{\boldsymbol{G}}^{E} \cdot \boldsymbol{e}_{I_{\mathrm{A,TL}}} = G_{zz}^{E}, \qquad (5.23)$$
$$\frac{\partial \tau_1}{\partial \xi} = \frac{\partial \tau_2}{\partial \xi} = 1. \qquad (5.24)$$

Accordingly, equations (5.18) – (5.21) reduce to

$$G_{+\mathrm{A}}^{E}(z_1, z_2, z_1', z_2') = G_{zz}^{E}(z_1, z_1') + G_{zz}^{E}(z_1, z_2') + G_{zz}^{E}(z_2, z_1') + G_{zz}^{E}(z_2, z_2'), \qquad (5.25)$$
$$G_{+\mathrm{TL}}^{E}(z_1, z_2, z_1', z_2') = G_{zz}^{E}(z_1, z_1') - G_{zz}^{E}(z_1, z_2') + G_{zz}^{E}(z_2, z_1') - G_{zz}^{E}(z_2, z_2'), \qquad (5.26)$$
$$G_{-\mathrm{A}}^{E}(z_1, z_2, z_1', z_2') = G_{zz}^{E}(z_1, z_1') + G_{zz}^{E}(z_1, z_2') - G_{zz}^{E}(z_2, z_1') - G_{zz}^{E}(z_2, z_2'), \qquad (5.27)$$
$$G_{-\mathrm{TL}}^{E}(z_1, z_2, z_1', z_2') = G_{zz}^{E}(z_1, z_1') - G_{zz}^{E}(z_1, z_2') - G_{zz}^{E}(z_2, z_1') + G_{zz}^{E}(z_2, z_2'). \qquad (5.28)$$

Let us now further suppose that the transmission line is located in free space. Then the Green's function is that of free space, $G_{zz}^{E} = G_{0\,zz}^{E}$. It is, in particular, translation

invariant,

$$G^E_{0\,zz}(z, z') = G^E_{0\,zz}(|z - z'|).$$ (5.29)

For straight, parallel wires we have, compare Fig. 5.3,

$$|z_1 - z'_1| = |z_2 - z'_2| \quad \Longrightarrow \quad G^E_{0\,zz}(z_1, z'_1) = G^E_{0\,zz}(z_2, z'_2),$$ (5.30)

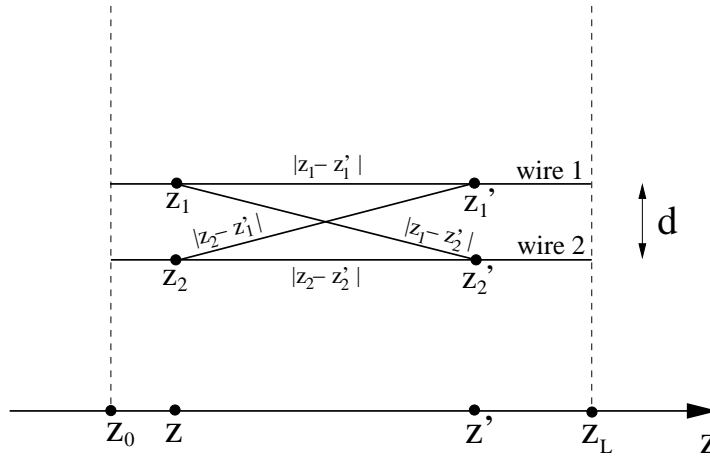$$|z_1 - z'_2| = |z_2 - z'_1| \quad \Longrightarrow \quad G^E_{0\,zz}(z_1, z'_2) = G^E_{0\,zz}(z_2, z'_1).$$ (5.31)

Figure 5.3: Geometry of a straight two-wire transmission line.

It follows that the relations (5.25)–(5.28) reduce to

$$G^E_{0+A}(z_1, z_2, z'_1, z'_2) = 2\left(G^E_{0\,zz}(z_1, z'_1) + G^E_{0\,zz}(z_1, z'_2)\right),$$ (5.32)

$$G^E_{0+\mathrm{TL}}(z_1, z_2, z'_1, z'_2) = 0,$$ (5.33)

$$G^E_{0-A}(z_1, z_2, z'_1, z'_2) = 0,$$ (5.34)

$$G^E_{0-\mathrm{TL}}(z_1, z_2, z'_1, z'_2) = 2\left(G^E_{0\,zz}(z_1, z'_1) - G^E_{0\,zz}(z_1, z'_2)\right),$$ (5.35)

and in view of the integral equation system (5.16), (5.17) it is recognized that the antenna mode I_A and the transmission line mode I_{TL} completely decouple,

$$j\omega\mu \int_{z_0}^{z_L} G^E_{0+A}(z_1, z_2, z'_1, z'_2) I_A(z')\, dz' = E^{\mathrm{inc}}_{\tan}(z_1) + E^{\mathrm{inc}}_{\tan}(z_2),$$ (5.36)

$$j\omega\mu \int_{z_0}^{z_L} G^E_{0-\mathrm{TL}}(z_1, z_2, z'_1, z'_2) I_{\mathrm{TL}}(z')\, dz' = E^{\mathrm{inc}}_{\tan}(z_1) - E^{\mathrm{inc}}_{\tan}(z_2).$$ (5.37)

5.1.3 Transmission line mode current and classical transmission line theory

The antenna mode current I_{A} in (5.36) vanishes at the beginning and at the end of the transmission line. This is analogous to the boundary conditions of the antenna current on a single wire antenna. Also the Green's function $G^E_{+\mathrm{A}}(z_1, z_2, z'_1, z'_2)$ exhibits the same qualitative behavior as the kernel of the Pocklington's equation for a single wire antenna since in (5.32) the terms $G^E_{0\,zz}(z_1, z'_1)$ and $G^E_{0\,zz}(z_1, z'_2)$ add up and only considerably differ if the distance $|z - z'|$ is smaller or of the order of the distance d between wire 1 and wire 2. It follows that (5.36) can be solved with the methods of antenna theory.

The classical TLT is contained in (5.37). To explicitly see this we first note that Pocklington's equation (3.72),

$$j\omega\mu \int_{z_0}^{z_L} G^E_{zz}(z, z')I(z')\,dz' = E^{\mathrm{inc}}_{\tan}(z)\,, \qquad (5.38)$$

is equivalent to the mixed potential integral equation (3.89),

$$\frac{1}{j\omega\varepsilon} \int_{z_0}^{z_L} \left[\frac{\partial G^\phi(z, z')}{\partial z} \frac{\partial I(z')}{\partial z'} + k^2 G^A_{zz}(z, z')I(z') \right] dz' = -E^{\mathrm{inc}}_z(z)\,, \qquad (5.39)$$

with $G^\phi(z, z')$ and $G^A_{zz}(z, z')$ indicating the Green's functions for the scalar potential ϕ and the magnetic vector potential $\boldsymbol{A}$ in the Lorenz gauge, respectively. We introduce a per-unit-length charge q' by the continuity equation

$$\frac{\partial I(z)}{\partial z} + j\omega q'(z) = 0 \qquad (5.40)$$

and define a potential $V^{q'}$ by

$$V^{q'}(z) := \frac{1}{\varepsilon} \int_{z_0}^{z_L} G^\phi(z, z')q'(z')\,dz'\,. \qquad (5.41)$$

Furthermore, in view of (5.35) and (5.37), we introduce the combinations

$$G^\phi_{0-\mathrm{TL}}(z, z') = 2\big(G^\phi_0(z_1, z'_1) - G^\phi_0(z_1, z'_2)\big)\,, \qquad (5.42)$$

$$G^A_{0-\mathrm{TL}}(z, z') = 2\big(G^A_{0\,zz}(z_1, z'_1) - G^A_{0\,zz}(z_1, z'_2)\big)\,. \qquad (5.43)$$

For the straight two-wire transmission line of Fig. 5.3 we explicitly have

$$G^\phi_{0-\mathrm{TL}}(z, z') = G^A_{0-\mathrm{TL}}(z, z') = \frac{1}{2\pi} \left[\frac{e^{-jk\sqrt{(z-z')^2+\rho^2}}}{\sqrt{(z-z')^2+\rho^2}} - \frac{e^{-jk\sqrt{(z-z')^2+d^2}}}{\sqrt{(z-z')^2+d^2}} \right]\,, \qquad (5.44)$$

where ρ denotes the wire radius which limits the Coulomb singularity in the thin-wire approximation. It is an elementary mathematical task to verify that $G^\phi_{0-\mathrm{TL}}$ and $G^A_{0-\mathrm{TL}}$,

as given by (5.44), are strongly localized functions that are characterized by a sharp peak in the domain where $k|z - z'|$ is of the order of the distance d. Since the spatial variation of $q(z)$ and $I(z)$ is of the order of the wavelength considered these quantities do only slightly change along an interval of length $k|z - z'| \approx d$ if we assume that $kd \ll 1$. This circumstance is essential in the derivation of the classical TLT, compare the appendix of [218], for example, since it leads to the simplifications

$$\int_{z_0}^{z_L} G_{0-\mathrm{TL}}^{\phi}(z, z') q'(z') \, dz' \approx q'(z) \int_{z_0}^{z_L} G_{0-\mathrm{TL}}^{\phi}(z, z') \, dz' \,, \tag{5.45}$$

$$\int_{z_0}^{z_L} G_{0-\mathrm{TL}}^{A}(z, z') I(z') \, dz' \approx I(z) \int_{z_0}^{z_L} G_{0-\mathrm{TL}}^{A}(z, z') \, dz' \,. \tag{5.46}$$

The integrals in these equations can be approximately evaluated according to

$$\frac{1}{2\pi} \int_{z_0}^{z_L} \left[\frac{e^{-jk\sqrt{(z-z')^2+\rho^2}}}{\sqrt{(z-z')^2+\rho^2}} - \frac{e^{-jk\sqrt{(z-z')^2+d^2}}}{\sqrt{(z-z')^2+d^2}} \right] dz'$$

$$\approx \frac{1}{2\pi} \int_{z_0}^{z_L} \left[\frac{1}{\sqrt{(z-z')^2+\rho^2}} - \frac{1}{\sqrt{(z-z')^2+d^2}} \right] dz' \quad \text{(5.47)}$$

$$\approx \frac{\ln(d/\rho)}{\pi} \,, \tag{5.48}$$

where the first approximation (5.47) requires the previously made assumption $kd \ll 1$ and the second approximation (5.48) requires $z_L - z \gg d$ and $z - z_0 \gg d$, i.e., the point z should not be close to the beginning or the end of the line such that boundary effects are negligible.

The mixed potential integral equation (5.39) and the continuity equation (5.40) can now be written as classical Telegrapher equations

$$\frac{\partial V^{q'}(z)}{\partial z} + j\omega L' I(z) = E_{\tan}^{\mathrm{inc}}(z_1) - E_{\tan}^{\mathrm{inc}}(z_2) \,, \tag{5.49}$$

$$\frac{\partial I(z)}{\partial z} + j\omega C' V^{q'}(z) = 0 \,, \tag{5.50}$$

with

$$C' := \frac{\varepsilon}{\int_{z_0}^{z_L} G_{0-\mathrm{TL}}^{\phi}(z, z') \, dz'} \tag{5.51}$$

$$\approx \frac{\pi\varepsilon}{\ln(d/\rho)} \,, \tag{5.52}$$

$$L' := \mu \int_{z_0}^{z_L} G_{0-\mathrm{TL}}^{A}(z, z') \, dz' \tag{5.53}$$

$$\approx \frac{\mu}{\pi} \ln(d/\rho) \,. \tag{5.54}$$

As a result, with the Telegrapher equations (5.49), (5.50) and the per-unit-length parameters (5.52), (5.54) we have arrived at the classical TLT. Since the source terms are given by

$$V'_s(z) = E_{\tan}^{\text{inc}}(z_1) - E_{\tan}^{\text{inc}}(z_2)\,, \tag{5.55}$$

$$I'_s(z) = 0\,, \tag{5.56}$$

we have obtained the field-to-line coupling model of Agrawal [3, 216].

5.1.4   Discussion

We summarize the assumptions that are necessary to derive the classical TLT: First, the coupled Pocklington's equations (5.3), (5.4) have been written in a form which presupposes

1. perfect conductivity of the wires;

2. a thin-wire approximation.

These two simplifying assumptions are also often made in antenna theory. Additionally, a decoupling of antenna mode current and transmission line mode current required

3. geometrical uniformity, i.e., the conductors have to be straight and in parallel;

4. translation invariance of the relevant Green's function along the transmission line.

Finally, a reduction of the integral equation for the transmission line current and the continuity equation to classical Telegrapher equations required that

5. the wavelength considered is large in comparison to the separation of the transmission line conductors, $kd \ll 1$.

The assumptions 1., 3., and 4. are clearly cut and represent precise statements. Also assumption 2., the thin-wire approximation, can be made precise [235]. But assumption 5. is not a precise statement. It is required to simplify the integro-differential equation (5.39) to a differential equation and this, in turn, requires to assume the per-unit-length charge $q'(z)$ and current $I(z)$ to be (approximately) constant along an (electrically small) range of integration. That is, we employ electrostatic and magnetostatic assumptions and, thus, neglect that the per-unit-length charge and current along the line actually are subject to change. As a result, we neglect a radiation field which, according to (1.243) and (1.244), *always* is present if electric charges are accelerated. Classical TLT would conceptually be much simpler if we had available a mathematical method to strictly decompose a given electromagnetic field into its constitutive velocity (Coulomb) and acceleration (radiation) field. But no such method is known and this is the reason why

we have to employ imprecise approximations, as represented by assumption 5, in order to isolate the dynamics of electric charges and their associated Coulomb field from the radiation field. As a result, the Telegrapher equations do not represent the complex dynamics of the complete electromagnetic field and its sources which is characterized by the inseparability of Coulomb field and radiation field.

5.2 Generalized transmission line theories

It has been mentioned that the Telegrapher equations (5.1), (5.2) of the classical TLT have a much simpler structure than the complete set of Maxwell equations. Indeed, Telegrapher equations can readily be solved: Suppose that Telegrapher equations are written in the combined form

$$\frac{\partial}{\partial \xi} \begin{pmatrix} V(\xi) \\ I(\xi) \end{pmatrix} = \begin{pmatrix} P_{11}(\xi) & P_{12}(\xi) \\ P_{21}(\xi) & P_{22}(\xi) \end{pmatrix} \begin{pmatrix} V(\xi) \\ I(\xi) \end{pmatrix} + \begin{pmatrix} V_s'(\xi) \\ I_s'(\xi) \end{pmatrix} \tag{5.57}$$

or, in a more condensed notation[2], as

$$\frac{\partial \boldsymbol{X}(\xi)}{\partial \xi} = \overline{\boldsymbol{P}}(\xi)\boldsymbol{X}(\xi) + \boldsymbol{X}_s'(\xi) \tag{5.58}$$

with a, possibly position dependent, per-unit-length parameter matrix $\overline{\boldsymbol{P}}(\xi)$. Then it is known from the general theory of differential equations that the general solution for the unknowns $V(\xi), I(\xi) \equiv \boldsymbol{X}(\xi)$ is given by [39, 53]

$$\boldsymbol{X}(\xi) = \mathcal{M}_{\xi_0}^{\xi}\{\overline{\boldsymbol{P}}\}\boldsymbol{X}(\xi_0) + \int_{\xi_0}^{\xi} \mathcal{M}_{\eta}^{\xi}\{\overline{\boldsymbol{P}}\}\boldsymbol{X}_s'(\eta)\,d\eta \,. \tag{5.59}$$

Here, the operator $\mathcal{M}_{\xi_0}^{\xi}\{\overline{\boldsymbol{P}}\}$ denotes the *product integral* or *matrizant* of the matrix $\overline{\boldsymbol{P}}(\xi)$,

$$\mathcal{M}_{\xi_0}^{\xi}\{\overline{\boldsymbol{P}}\} := \overline{\boldsymbol{I}} + \int_{\xi_0}^{\xi} \overline{\boldsymbol{P}}(\eta)\,d\eta + \int_{\xi_0}^{\xi} \overline{\boldsymbol{P}}(\eta) \int_{\xi_0}^{\eta} \overline{\boldsymbol{P}}(\tau)\,d\tau\,d\eta + \dots \tag{5.60}$$

$$= \prod_{\xi_0}^{\xi} e^{\overline{\boldsymbol{P}}(\eta)\,d\eta} \,. \tag{5.61}$$

In the classical TLT the parameter matrix $\overline{\boldsymbol{P}}$ is position independent and the product integral simplifies to a matrix exponential.

In generalized transmission line theories one or more of the necessary assumptions of the classical TLT, which are listed in Sec. 5.1.4, are dropped while the electromagnetic propagation still is described by Telegrapher equations of the form (5.58). Once those generalized Telegrapher equations are determined, it is possible to study on the basis of (5.59) their solutions.

---

[2]While in (5.57) the notation implies a two-wire transmission line, the notation in (5.58) is suitable to represent the Telegrapher equations of a multiconductor transmission line [7].

5.2.1 Generalized transmission line theories as methods for solving field integral equations of antenna theory

Most of the literature on extensions of the classical TLT deals with geometrically nonuniform transmission lines. This is discussed and reviewed in [74, 154, 7], for example. Here we will restrict ourselves to point out recent advantages in the framework of generalized transmission line theories which are based on rigorous derivations and take the field integral equations of antenna theory as starting point.

It already had been noted by King [103] that in the conversion of an electric field integral equation to a Telegrapher equation it is essential to pull, similar to (5.45), (5.46), the charge and current distributions out of the relevant integrals. Realizing this necessity, Haase et al. [73, 74, 75, 76] developed a new method to derive a generalized transmission line theory which they called *transmission line super theory* (TLST). Their derivation is based on the mixed-potential integral equation and the continuity equation. Both equations are written in a form which is appropriate to be applied to general transmission line structures. To rewrite the mixed potential integral equation as a Telegrapher equation it is assumed that the electric current is given by the general solution of a second order differential equation[3]. This general solution can be expressed in the form (5.59) and be inserted into the mixed potential integral equation. As a result, the functions of the charge and current distribution become independent of the integration variable and can be pulled in front of the integrals in order to yield generalized transmission line equations. A solution of these equations is not trivial since it requires, in particular, an iterative solution of an integral equation for the, a priori, unknown generalized transmission line parameters and source terms. This, in turn, requires mathematical procedures to efficiently evaluate expressions that contain product integrals [9, 74]. The TLST has been applied to a variety of examples, reproducing known analytic solutions as well as results from numerical method of moment calculations and experimental measurements.

Generalized transmission line theories add radiation effects to the classical Telegrapher equations and it turns out that there is a freedom in absorbing the corresponding additional terms by the per-unit-length parameters or the distributed sources. In particular, this complicates the physical interpretation of generalized per-unit-length parameters which, in the context of a generalized transmission line theory, usually are position dependent, frequency dependent, or gauge dependent. To a good degree these features have been elucidated by a number of articles by Nitsch & Tkachenko [156, 157, 159].

We also mention results of Tkachenko et al. which demonstrate the benefit of analytical techniques for the analysis of canonical but geometrically nonuniform transmission line configurations. The results are based on the solution of Pocklington's equation or the mixed potential integral equation [218, 155, 223]. In these studies, the canonical transmission line structures usually involve a perfectly conducting ground plane. This

---

[3]This assumption is backed by a study of Mei [139] where it is shown that the current distribution along a nonuniform transmission line generally is characterized by a second order differential equation.

makes it possible to discard the symmetric antenna mode current right from the beginning and, similar to the analysis of Sec. 5.1.3, concentrate on the evaluation of the transmission line mode current.

In summary, considerably progress in the formulation of generalized transmission line theories has been achieved during the last years. *Generalized transmission line theories offer an alternative way to solve integral equations of antenna theory* and are applicable as long as the structures considered resemble a transmission line which can be parameterized by a single spatial variable. In particular, the TLST produces results that are in good agreement with corresponding method of moments calculations and, additionally, are in a form which is suitable to embed transmission lines into network representations of larger systems [74].

5.2.2 Transmission lines in the presence of resonances

There have been investigations of transmission lines in cavities, but these are, by far, not extensive. Some investigations, such as [49, 151, 166], combine numerical methods with classical TLT and validate their results against (more time-consuming) purely numerical calculations and experimental measurements. Recently, by means of a semi-analytical approach, it has been more systematically investigated if the classical TLT still is valid within a cavity [207]. To this end, a classical TLT model of a specific transmission line inside a rectangular cavity has been combined with an analytically calculated incident electromagnetic cavity field. Then the results of this classical TLT model were compared to corresponding results of a numerical full-wave calculation and it was observed that the classical TLT reproduced most of the characteristic features of the full-wave calculation. However, the mechanism that explains *why* the classical TLT, for certain configurations, yields satisfying results inside cavities has not been discussed.

Any rigorous approach towards transmission lines in cavities will necessarily involve the integral equations of antenna theory in resonating systems. From the developments of Chap. 3 it is anticipated that the solution of any integral equation of this type will require numerical evaluations, as exemplified in [221]. Clearly, it is always possible to resort to method of moments calculations, such as presented in Sec. 3.4, in order to calculate the current on a transmission line in the same way as the current on an antenna – the integral equations to solve are the same. Similarly, it is expected that the TLST can also be adapted to cavities' Green's functions, even though this has not been shown in detail, yet.

But let us return to the integral equation based derivation of classical TLT in Sec. 5.1. To arrive at the Telegrapher equations (5.49), (5.50) we employed three times the properties of free space Green's functions:

1. To decouple for straight, parallel wires the antenna mode from the transmission line mode we used translation invariance of the free space Green's function $G_{0\,zz}^{E}$, compare (5.29).

2. To pull in (5.45), (5.46) the per-unit-length charge q' and the current I out of the integrals we took advantage of the strongly localized Coulomb singularity that is contained in the combinations $G^{\phi}_{0-\mathrm{TL}}$ and $G^{A}_{0+\mathrm{TL}}$ of free space Green's functions.

3. To calculate the per-unit-length parameter L' and C' we used in (5.47) the explicit mathematical expression of the free space combinations $G^{\phi}_{0-\mathrm{TL}}$ and $G^{A}_{0-\mathrm{TL}}$.

For the classical TLT to be applicable within a cavity it needs to be checked if these three steps are (approximately) valid if the Green's functions of free space are replaced by the corresponding cavities' Green's functions. To this end we recall that, according to Sec. 2.3.3, it is physically meaningful to split a cavities' Green's function $\overline{\boldsymbol{G}}_{\mathrm{cav}}$ into a free space contribution $\overline{\boldsymbol{G}}_{0}$ and a remainder $\widetilde{\overline{\boldsymbol{G}}}$ which takes into account the interaction with the cavity,

$$\overline{\boldsymbol{G}}_{\mathrm{cav}} = \overline{\boldsymbol{G}}_{0} + \widetilde{\overline{\boldsymbol{G}}}\,. \tag{5.62}$$

We now state that the steps 1.–3. can *approximately* be performed within a cavity.

1. The decoupling of antenna and transmission line mode requires the kernels $G^{E}_{+\mathrm{TL}}(z_1, z_2, z'_1, z'_2)$ and $G^{E}_{-\mathrm{A}}(z_1, z_2, z'_1, z'_2)$ to vanish. From the relations (5.26), (5.27), and Fig. 5.3 with $d = y_1 - y_2$ it follows that it is meaningful to consider the Taylor expansions

$$\tilde{G}^{E}_{zz}(z_1, z'_2) \approx \tilde{G}^{E}_{zz}(z_1, z'_1) - \frac{\partial \tilde{G}^{E}_{zz}}{\partial y}(z_1, z'_1)\, d\,, \tag{5.63}$$

$$\tilde{G}^{E}_{zz}(z_1, z'_2) \approx \tilde{G}^{E}_{zz}(z_2, z'_2) + \frac{\partial \tilde{G}^{E}_{zz}}{\partial y}(z_2, z'_2)\, d\,, \tag{5.64}$$

$$\tilde{G}^{E}_{zz}(z_2, z'_1) \approx \tilde{G}^{E}_{zz}(z_2, z'_2) + \frac{\partial \tilde{G}^{E}_{zz}}{\partial y}(z_2, z'_2)\, d\,, \tag{5.65}$$

$$\tilde{G}^{E}_{zz}(z_2, z'_1) \approx \tilde{G}^{E}_{zz}(z_1, z'_1) - \frac{\partial \tilde{G}^{E}_{zz}}{\partial y}(z_1, z'_1)\, d\,, \tag{5.66}$$

since then

$$G^{E}_{+\mathrm{TL}}(z_1, z_2, z'_1, z'_2) = G^{E}_{-\mathrm{A}}(z_1, z_2, z'_1, z'_2) \approx \left(\frac{\partial \tilde{G}^{E}_{zz}}{\partial y}(z_1, z'_1) + \frac{\partial \tilde{G}^{E}_{zz}}{\partial y}(z_2, z'_2) \right) d\,. \tag{5.67}$$

If the terms involving the derivative of the Green's function are small the kernels $G^{E}_{+\mathrm{TL}}(z_1, z_2, z'_1, z'_2)$ and $G^{E}_{-\mathrm{A}}(z_1, z_2, z'_1, z'_2)$ are small as well and, as a result, antenna and transmission line mode approximately decouple. For a qualitative discussion of this point one might represent the cavities Green's function $\overline{\boldsymbol{G}}^{E}_{\mathrm{cav}}$ by means of an expansion in transverse and longitudinal eigenfunctions, compare Sec. 2.3.1.

The *transverse* eigenfunctions are solutions of sourceless Helmholtz equations and the spatial variation of the dominant eigenfunctions is *of the order* of the wavelength considered. It follows that the contributions of the transverse eigenfunctions to the

derivative terms in (5.63)–(5.66) will be of the order of kd. Thus, according to the assumption $kd \ll 1$ of classical TLT, these contributions will be small.

The *longitudinal* eigenfunctions are solutions of electrostatic Poisson equations and constitute the Coulomb singularity. Their spatial variation diverges close to a Coulomb singularity and, otherwise, quickly decays. Since $\widetilde{\overline{G}} = \overline{G}_{\text{cav}} - \overline{G}_0$ contains no Coulomb singularity it is expected that, in general, the longitudinal eigenfunctions will lead to no significant spatial variation of $\widetilde{\overline{G}}$. However, this argument does not apply if a source point is close to a cavity wall. In this case there will be a dominant "mirrored" Coulomb interaction with the cavity wall which is embedded in $\widetilde{\overline{G}}$ and the spatial variation of $\widetilde{\overline{G}}$ might exceed the order of the parameter kd. However, this will be an electrostatic effect and no resonance effect.

2. We need to reconsider the approximations (5.45), (5.46) which are due to the dominance of the Coulomb interactions at short distances when source and observation point approach each other. This dominance is unaffected by the presence of a cavity.

3. The electrostatic calculation that led to the values of the per-unit-length parameters (5.52), (5.54) consists of the evaluation of integrals of the form $\int_{-L/2}^{L/2} G_{-\text{TL}}^{\phi,A}(z, z')\, dz'$. Within a cavity these parameters will only significantly change if the Coulomb field in the vicinity of the transmission lines is significantly perturbed by the presence of the cavity. In this case, the parameters L', C' have to be calculated from an electrostatic calculation which takes into account the cavity wall.

These arguments are of a general nature. They should be taken as guidelines rather than as mathematical proofs. However, they express the basic mechanism that underlies conventional TLT: The basic mechanism of conventional TLT is the *subtraction* of electromagnetic field contributions of two (or more) conductors such that the sharp, localized Coulomb fields remain and the smooth, extended radiation fields (approximately) cancel. This mechanism applies both in free space and within resonating systems.

Appendix A

Tensor Analysis, Integration, and Lie derivative

Within a theoretical formulation physical quantities are modeled as mathematical objects. The understanding and application of appropriate mathematics yields, in turn, the properties of physical quantities. In the development of the axiomatic approach, we made repeated use of integration, of the Poincaré lemma, and of the Stokes theorem. It is with these mathematical concepts that it is straightforward to derive the basics of electromagnetism from a small number of axioms.

Integration is an operation that yields coordinate independent values. It requires an integration measure, the dimension of which depends on the type of region that is integrated over. We want to integrate over one-dimensional curves, two-dimensional surfaces, or three-dimensional volumes that are embedded in three-dimensional space. Therefore, we have to define line-, surface-, and volume-elements as integration measures. Then we can think of suitable objects as integrands that can be integrated over to yield coordinate independent physical quantities.

A.1 Integration over a curve and covariant vectors as line integrands

We consider a one-dimensional curve $\boldsymbol{c} = \boldsymbol{c}(t)$ in three-dimensional space. In a specific coordinate system x^i, with indices $i = 1, 2, 3$, a parameterization of $\boldsymbol{c}$ is given by the vector

$$\boldsymbol{c}(t) = \left(c^1(t), c^2(t), c^3(t) \right). \tag{A.1}$$

The functions $c^i(t)$ define the shape of the curve. For small changes of the parameter t, with $t \rightarrow t + \Delta t$, the difference vector between $\boldsymbol{c}(t + \Delta t)$ and $\boldsymbol{c}(t)$ is given by

$$\Delta \boldsymbol{c}(t) = \left(\frac{\Delta c^1}{\Delta t}, \frac{\Delta c^2}{\Delta t}, \frac{\Delta c^3}{\Delta t} \right) \Delta t, \tag{A.2}$$

compare Fig.A.1. In the limit where Δt becomes infinitesimally we obtain the line element

$$d\boldsymbol{c}(t) = (dc^1(t), dc^2(t), dc^3(t))$$
$$:= \left(\frac{\partial c^1(t)}{\partial t}, \frac{\partial c^2(t)}{\partial t}, \frac{\partial c^3(t)}{\partial t} \right) dt \,. \tag{A.3}$$

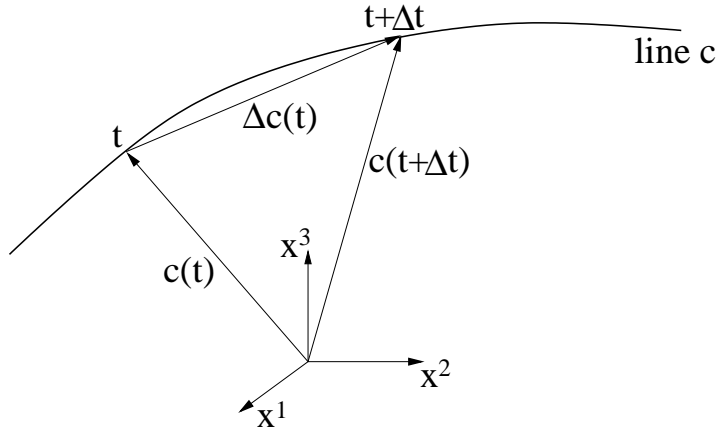It is characterized by an infinitesimal length and an orientation.

Figure A.1: Parameterization of a curve $\boldsymbol{c}(t)$. The difference vector $\Delta\boldsymbol{c}(t)$ between $\boldsymbol{c}(t + \Delta t)$ and $\boldsymbol{c}(t)$ yields, in the limit $\Delta t \to 0$, the line element $d\boldsymbol{c}(t)$.

We now construct objects that we can integrate over the curve $\boldsymbol{c}$ in order to obtain a coordinate invariant scalar. The line element $d\boldsymbol{c}$ contains three independent components dc^i. If we shift from old coordinates x^i to new coordinates $y^{j'} = y^{j'}(x^i)$ these components transform according to

$$dc^{j'} = \frac{\partial y^{j'}}{\partial x^i} dc^i \,. \tag{A.4}$$

Therefore, we can form an invariant expression if we introduce objects $\boldsymbol{\alpha} = \boldsymbol{\alpha}(x^i)$, with three independent components α_i, that transform in the opposite way,

$$\alpha_{j'} = \frac{\partial x^i}{\partial y^{j'}} \alpha_i \,. \tag{A.5}$$

This transformation behavior characterizes a vector or, more precisely, a covariant vector (a 1-form). It follows that the expression

$$\alpha_i \, dc^i = \alpha_{j'} \, dc^{j'} \tag{A.6}$$

yields the same value in each coordinate system.

Thus, we can now immediately define integration over a curve by the expression

$$\int \alpha_i \, dc^i = \int \alpha_1 \, dc^1 + \alpha_2 \, dc^2 + \alpha_3 \, dc^3$$

$$= \int \left(\alpha_1 \frac{\partial c^1}{\partial t} + \alpha_2 \frac{\partial c^2}{\partial t} + \alpha_3 \frac{\partial c^3}{\partial t} \right) dt \,. \tag{A.7}$$

The last line shows how to carry out explicitly the integration since α_i and c^i are functions of the parameter t.

A.2 Integration over a surface and contravariant vector densities as surface integrands

Now we consider a two-dimensional surface $\boldsymbol{a} = \boldsymbol{a}(t, s)$. Within a specific coordinate system x^i, a parameterization of $\boldsymbol{a}$ is of the form

$$\boldsymbol{a}(t, s) = (a^1(t, s), a^2(t, s), a^3(t, s)) \tag{A.8}$$

with parameters t, s and functions $a^i(t, s)$ that define the shape of the surface.
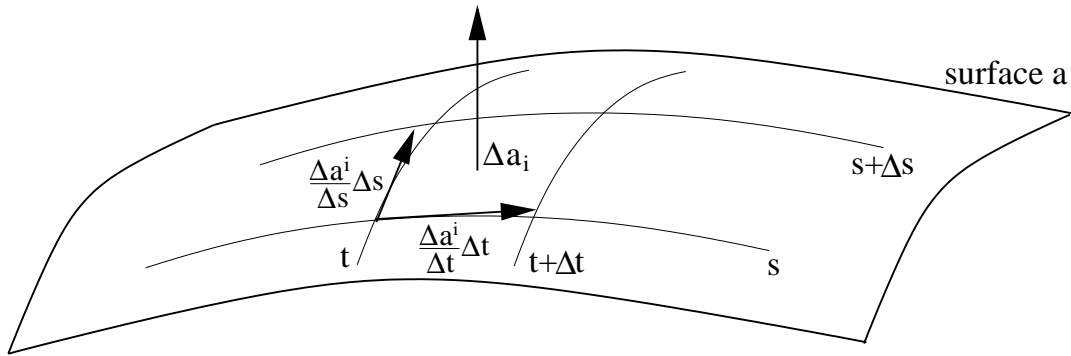

Figure A.2: Parameterization of a surface $\boldsymbol{a}(t, s)$. The lines $t =$const, $t + \Delta t =$const, $s =$const, and $s + \Delta s =$const circumscribe a surface Δa_i that is spanned by the edges $\frac{\Delta a^i}{\Delta t} \, dt$ and $\frac{\Delta a^i}{\Delta s} \, ds$. In the limit $\Delta t \to dt$, $\Delta s \to ds$, it becomes an elementary surface element da_i.

An elementary surface element is bound by lines $t =$const, $t + dt =$const, $s =$const, and $s + ds =$const, compare Fig.A.2. It is characterized by the two edges $\frac{\partial a^i}{\partial t} \, dt$ and $\frac{\partial a^i}{\partial s} \, ds$. These edges span an infinitesimal surface, the area and orientation of which is characterized by a covariant vector da_i that points normal to the infinitesimal surface. The vector da_i is given by the vector product of $\frac{\partial a^i}{\partial t} \, dt$ and $\frac{\partial a^i}{\partial s} \, ds$,

$$da_i = \epsilon_{ijk} \frac{\partial a^j}{\partial t} \frac{\partial a^k}{\partial s} \, dt \, ds \,. \tag{A.9}$$

In order to know how the components da_i transform under coordinate transformations $y^{j'} = y^{j'}(x^i)$, we have to know the transformation behavior of the symbol ϵ_{ijk}. Since in any coordinate system, ϵ_{ijk} assumes the values 0, 1, or -1 by definition, it is obvious that in general

$$\epsilon_{i'j'k'} \neq \frac{\partial x^i}{\partial y^{i'}} \frac{\partial x^j}{\partial y^{j'}} \frac{\partial x^k}{\partial y^{k'}} \epsilon_{ijk} . \tag{A.10}$$

This is because the determinant of the transformation matrix, i.e.,

$$\det\left(\partial x/\partial y\right) = \epsilon_{ijk} \frac{\partial x^i}{\partial y^{i'}} \frac{\partial x^j}{\partial y^{j'}} \frac{\partial x^k}{\partial y^{k'}} , \tag{A.11}$$

is, in general, not equal to one. But it follows from (A.11) that the correct transformation rule for ϵ_{ijk} is given by

$$\begin{aligned}
\epsilon_{i'j'k'} &= \frac{1}{\det(\partial x/\partial y)} \frac{\partial x^i}{\partial y^{i'}} \frac{\partial x^j}{\partial y^{j'}} \frac{\partial x^k}{\partial y^{k'}} \epsilon_{ijk} \\
&= \det(\partial y/\partial x) \frac{\partial x^i}{\partial y^{i'}} \frac{\partial x^j}{\partial y^{j'}} \frac{\partial x^k}{\partial y^{k'}} \epsilon_{ijk} .
\end{aligned} \tag{A.12}$$

With (A.9) this yields the transformation rule for the components da_i,

$$da_{j'} = \det(\partial y/\partial x) \frac{\partial x^i}{\partial y^{j'}} da_i . \tag{A.13}$$

Now we construct quantities that can be integrated over a surface. Since a surface element is determined from three independent components da_i we introduce an integrand with three independent components β^i that transform according to

$$\beta^{j'} = \frac{1}{\det(\partial y/\partial x)} \frac{\partial y^{j'}}{\partial x^i} \beta^i . \tag{A.14}$$

Transformation rules that involve the determinant of the transformation matrix characterize so-called *densities*. Densities are sensitive towards changes of the scale of elementary volumes. In physics they represent additive quantities, also called extensities, that describe how much of a quantity is distributed within a volume or over the surface of a volume. This is in contrast to intensities. The covariant vectors that we introduced as natural line integrals are intensive quantities that represent the strength of a physical field.

The transformation behavior (A.14) of the components β^i characterizes a contravariant vector density. With this transformation behavior the surface integral

$$\int \beta^i da_i = \int \beta^i \epsilon_{ijk} \frac{\partial a^j}{\partial t} \frac{\partial a^k}{\partial s} \, dt \, ds \tag{A.15}$$

yields a scalar value that is coordinate independent.

A.3   Integration over a volume and scalar densities as volume integrands

We finally consider integration over a three-dimensional volume $\boldsymbol{v}$ in three-dimensional space. Again we choose a specific coordinate system x^i and specify a parameterization of $\boldsymbol{v}$ by

$$\boldsymbol{v}(t,s,r) = \left(v^1(t,s,r), v^2(t,s,r), v^3(t,s,r)\right), \tag{A.16}$$

with three parameters t, s, and r.

An elementary volume element dv is characterized by three edges $\frac{\partial v^i}{\partial t}dt$, $\frac{\partial v^i}{\partial s}ds$, and $\frac{\partial v^i}{\partial r}dr$. The volume, which is spanned by these edges, is given by the determinant

$$
\begin{aligned}
dv &= \det\left(\frac{\partial v^i}{\partial t}dt, \frac{\partial v^i}{\partial s}ds, \frac{\partial v^i}{\partial r}dr\right) \\
&= \epsilon_{ijk}\frac{\partial v^i}{\partial t}\frac{\partial v^j}{\partial s}\frac{\partial v^k}{\partial r}\, dt\, ds\, dr\,.
\end{aligned}
\tag{A.17}
$$

It is not coordinate invariant but transforms under coordinate transformations $y^{j'} = y^{j'}(x^i)$ according to

$$dv' = \det(\partial y/\partial x)\, dv\,. \tag{A.18}$$

Since the volume element dv constitutes one independent component, a natural object to integrate over a volume has one independent component as well. We denote such an integrand by γ. It transforms according to

$$\gamma' = \frac{1}{\det(\partial y/\partial x)}\,\gamma\,. \tag{A.19}$$

This transformation rule characterizes a scalar density and yields

$$\int \gamma\, dv = \int \gamma\, \epsilon_{ijk}\frac{\partial v^i}{\partial t}\frac{\partial v^j}{\partial s}\frac{\partial v^k}{\partial r}\, dt\, ds\, dr \tag{A.20}$$

as a coordinate independent value.

A.4   Poincaré Lemma

The axiomatic approach takes advantage of the Poincaré lemma. The Poincaré lemma states under which conditions a mathematical object can be expressed in terms of a derivative, i.e., in terms of a potential.

We consider integrands α_i, β^i, and γ of line-, surface-, and volume integrals, respectively, and assume that they are defined in an open and simply connected region of three-dimensional space. Then the Poincaré lemma yields the following conclusions:

1. If α_i is curl free, it can be written as the gradient of a scalar function f,

$$\epsilon^{ijk}\partial_j\alpha_k = 0 \qquad \Longrightarrow \qquad \alpha_i = \partial_i f\,. \tag{A.21}$$

2. If β^i is divergence free, it can be written as the curl of the integrand α_i of a line integral,

$$\partial_i\beta^i = 0 \qquad \Longrightarrow \qquad \beta^i = \epsilon^{ijk}\partial_j\alpha_k\,. \tag{A.22}$$

3. The integrand γ of a volume integral can be written as the divergence of an integrand β^i of a surface integral,

$$\gamma \text{ is a volume integrand} \quad \Longrightarrow \qquad \gamma = \partial_i\beta^i\,. \tag{A.23}$$

While conclusions (A.21), (A.22) are familiar from elementary vector calculus, this might not be the case for conclusion (A.23). However, (A.23) is rather trivial since, in Cartesian coordinates x, y, z, for a given volume integrand $\gamma = \gamma(x,y,z)$ the vector β^i with components $\beta^x = \int_0^x \gamma(t,y,z)/3\,dt$, $\beta^y = \int_0^y \gamma(x,t,z)/3\,dt$, and $\beta^z = \int_0^z \gamma(x,y,t)/3\,dt$ fulfills (A.23). Of course, the vector β^i is not uniquely determined from γ since any divergence free vector field can be added to β^i without changing γ. We further note that γ, as a volume integrand, constitutes a scalar density. It can be integrated as above to yield the components of β^i as components of a contravariant vector density. Therefore, the integration does not yield a coordinate invariant scalar such that γ cannot be considered as a natural integrand of a line integral.

A.5 Stokes Theorem

In our notation Stokes theorem, if applied to line integrands α_i or surface integrands β^i, yields the identities:

$$\int_V \partial_i\beta^i\,dv = \int_{\partial V} \beta^i\,da_i\,, \tag{A.24}$$

$$\int_S \epsilon^{ijk}\partial_j\alpha_k\,da_i = \int_{\partial S} \alpha_i\,dc^i\,, \tag{A.25}$$

where ∂V denotes the two-dimensional boundary of a simply connected volume V and ∂S denotes the one-dimensional boundary of a simply connected surface S.

A.6 Lie derivative

The Lie derivative l_v describes the change of an object T between two infinitesimally neighboring points p (with coordinates x^i) and $\tilde{p}$ (with coordinates $x^i + \varepsilon v^i(p)$), as noticed

by an observer who applies a coordinate system x' that is dragged along the vector field v^i [190, 225],

$$l_v T := \lim_{\tilde{p} \to p} \frac{T'(\tilde{p}) - T(p)}{||\tilde{p} - p||} \tag{A.26}$$

$$= \lim_{\varepsilon \to 0} \frac{T'(x^i + \varepsilon v^i) - T(x^i)}{\varepsilon} \, . \tag{A.27}$$

At the point $\tilde{p}$ the relation between the coordinate system x and the dragged coordinate system x' is given by the coordinate transformation

$$x^n = x'^n + \varepsilon v^n(x^i) \tag{A.28}$$

with transformation matrix

$$\frac{\partial x^n}{\partial x'^i} = \delta_i^n + \varepsilon \partial_i v^n \tag{A.29}$$

Example: We apply the definition (A.26) to a contravariant vector density β^n with transformation behavior

$$\beta'^n = \det\left(\frac{\partial x^n}{\partial x'^i}\right) \frac{\partial x'^n}{\partial x^i} \beta^i \, . \tag{A.30}$$

To evaluate (A.26) we only keep terms linear in ε and neglect higher order terms. We note that

$$\det\left(\frac{\partial x^n}{\partial x'^i}\right) = \det(\delta_i^n + \varepsilon \partial_i v^n) \tag{A.31}$$

$$= 1 + \varepsilon \underbrace{\text{Trace}(\partial_i v^n)}_{=\partial_j v^j} + \mathcal{O}(\epsilon^2) \, . \tag{A.32}$$

It follows

$$\beta'^n(\tilde{p}) = \det\left(\frac{\partial x^n}{\partial x'^i}\right) \frac{\partial x'^n}{\partial x^i} \beta^i(\tilde{p}) \tag{A.33}$$

$$= (1 + \varepsilon \, \partial_j v^j)(\delta_i^n - \varepsilon \, \partial_i v^n)(\beta^i(p) + \varepsilon \, v^j \partial_j \beta^n(p)) \tag{A.34}$$

$$= \beta^n(p) + \varepsilon \big(v^j \partial_j \beta^n(p) - \partial_i v^n \beta^i(p) + \partial_j v^j \beta^n(p)\big) \, , \tag{A.35}$$

and we obtain from (A.27) for the Lie derivative of β^n the expression

$$l_v \beta^n = v^j \partial_j \beta^n - \beta^i \partial_i v^n + \beta^n \partial_j v^j \, . \tag{A.36}$$

In a similar way we obtain the Lie derivative of a covariant vector α_n as

$$l_v \alpha_n = v^j \partial_j \alpha_n + \alpha_j \partial_n v^j \tag{A.37}$$

and the Lie derivative of a scalar density γ turns out to be

$$l_v \gamma = v^j \partial_j \gamma + \gamma \partial_j v^j \tag{A.38}$$

$$= \partial_j(\gamma v^j) \, . \tag{A.39}$$

Appendix B

Some Formulas of Vector and Dyadic Calculus in Three Dimensions

This appendix collects a few identities that are useful to work out expressions that involve vector and dyadic quantities. A number of these identities has been used in the previous chapters. A much more exhaustive collection of vector and dyadic identities is provided by the appendices of the classic book of Van Bladel [227].

B.1 Vector identities

The following vector identities involve three-component vector functions $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$, a scalar function ψ, and the differential operator $\boldsymbol{\nabla}$.

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b}) \tag{B.1}$$

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c} \tag{B.2}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d}) - (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c}) \tag{B.3}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla}\psi = 0 \tag{B.4}$$

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{a}) = 0 \tag{B.5}$$

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{a}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{a}) - \Delta\boldsymbol{a} \tag{B.6}$$

$$\boldsymbol{\nabla} \cdot (\psi\boldsymbol{a}) = \boldsymbol{a} \cdot \boldsymbol{\nabla}\psi + \psi\boldsymbol{\nabla} \cdot \boldsymbol{a} \tag{B.7}$$

$$\boldsymbol{\nabla} \times (\psi\boldsymbol{a}) = \boldsymbol{\nabla}\psi \times \boldsymbol{a} + \psi\boldsymbol{\nabla} \times \boldsymbol{a} \tag{B.8}$$

$$\boldsymbol{\nabla}(\boldsymbol{a} \cdot \boldsymbol{b}) = (\boldsymbol{a} \cdot \boldsymbol{\nabla})\boldsymbol{b} + (\boldsymbol{b} \cdot \boldsymbol{\nabla})\boldsymbol{a} + \boldsymbol{a} \times (\boldsymbol{\nabla} \times \boldsymbol{b}) + \boldsymbol{b} \times (\boldsymbol{\nabla} \times \boldsymbol{a}) \tag{B.9}$$

$$\boldsymbol{\nabla} \cdot (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{b} \cdot (\boldsymbol{\nabla} \times \boldsymbol{a}) - \boldsymbol{a} \cdot (\boldsymbol{\nabla} \times \boldsymbol{b}) \tag{B.10}$$

$$\boldsymbol{\nabla} \times (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{a}(\boldsymbol{\nabla} \cdot \boldsymbol{b}) - \boldsymbol{b}(\boldsymbol{\nabla} \cdot \boldsymbol{a}) + (\boldsymbol{b} \cdot \boldsymbol{\nabla})\boldsymbol{a} - (\boldsymbol{a} \cdot \boldsymbol{\nabla})\boldsymbol{b} \tag{B.11}$$

B.2   Dyadic identities

In the following $\overline{G}$ denotes a dyadic and $\overline{G}^{\mathrm{T}}$ its transpose.

$$a \cdot (\overline{G} \cdot b) = (a \cdot \overline{G}) \cdot b = a \cdot \overline{G} \cdot c \qquad (\text{B.12})$$

$$(a \cdot \overline{G}) \times b = a \cdot (\overline{G} \times b) = a \cdot \overline{G} \times b \qquad (\text{B.13})$$

$$(a \times \overline{G}) \cdot b = a \times (\overline{G} \cdot b) \qquad (\text{B.14})$$

$$(a \times b) \cdot \overline{G} = a \cdot (b \times \overline{G}) = -b \cdot (a \times \overline{G}) \qquad (\text{B.15})$$

$$(\overline{G} \times a) \cdot b = \overline{G} \cdot a \times b = -(\overline{G} \times b) \cdot a \qquad (\text{B.16})$$

$$a \times (b \times \overline{G}) = b(a \cdot \overline{G}) - (a \times b)\overline{G} \qquad (\text{B.17})$$

$$a \cdot \overline{G} = \overline{G}^{\mathrm{T}} \cdot a = (a \cdot \overline{G})^{\mathrm{T}} \qquad (\text{B.18})$$

B.3   Integral identities

The following integral identities involve volume integrals that extend over a regular volume $\Omega \subset \mathbb{R}^3$ which is bounded by a closed surface $\Gamma = \partial\Omega$. The unit normal vector e_n is defined to point inward from the surface Γ into the volume Ω.

Vector-dyadic Green's first theorem

$$\int_\Omega \left[(\boldsymbol{\nabla} \times a) \cdot (\boldsymbol{\nabla} \times \overline{G}) - a \cdot (\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \overline{G}) \right] d^3r = \oint e_n \times (a \times \boldsymbol{\nabla} \times \overline{G}) \, d^2r \quad (\text{B.19})$$

Vector-dyadic Green's second theorem

$$\int_\Omega \left[(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times a) \cdot \overline{G} - a \cdot (\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \overline{G}) \right] d^3r = \qquad (\text{B.20})$$

$$- \oint_\Gamma \left[(n \times a) \cdot (\boldsymbol{\nabla} \times \overline{G}) + (e_n \times \boldsymbol{\nabla} \times a) \cdot \overline{G} \right] d^2r$$

By means of vector and dyadic identities the vector-dyadic Green's second theorem can be put into the equivalent form

$$\int_\Omega \left[(\Delta a) \cdot \overline{G} - a \cdot \Delta\overline{G} \right] d^3r = \qquad (\text{B.21})$$

$$\oint_\Gamma \left[(e_n \times a) \cdot (\boldsymbol{\nabla} \times \overline{G}) - (\boldsymbol{\nabla} \times a) \cdot (e_n \times \overline{G}) + e_n \cdot a(\boldsymbol{\nabla} \cdot \overline{G}) - e_n \cdot \overline{G}(\boldsymbol{\nabla} \cdot a) \right] d^2r$$

Appendix C

Ewald Representation of the Dyadic Vector Potential's Green's Function

In this appendix we derive the Ewald representation (3.113) of the dyadic vector potential's Green's function of a rectangular cavity from the ray representation (2.256). Since the three nonvanishing components G_{xx}^A, G_{yy}^A, and G_{zz}^A are related to each other by cyclic exchange of x, y and z it is sufficient to concentrate on the component G_{zz}^A. The formulas that we will use are a consequence of the results that are provided by the original paper of Ewald [45], and all that needs to be done is to adapt these results to the geometry of a rectangular cavity.

We start from the ray representation (2.256) of the component G_{zz}^A which is repeated as

$$G_{zz}^A(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{4\pi} \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} \frac{e^{-jkR_{i,mnp}(\boldsymbol{r}, \boldsymbol{r}')}}{R_{i,mnp}(\boldsymbol{r}, \boldsymbol{r}')} \tag{C.1}$$

with

$$R_{i,mnp}(\boldsymbol{r}, \boldsymbol{r}') = \sqrt{(X_i + 2ml_x)^2 + (Y_i + 2nl_y)^2 + (Z_i + 2pl_z)^2}, \tag{C.2}$$

where the distances X_i, Y_i and Z_i depend on $\boldsymbol{r}$ and $\boldsymbol{r}'$, compare (2.259)–(2.261). We now consider the identity

$$\frac{e^{-jkR_{i,mnp}}}{R_{i,mnp}} = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-R_{i,mnp}^2 s^2 + \frac{k^2}{4s^2}} \, ds \tag{C.3}$$

which is valid if the complex path of integration is such that the integrand remains finite for $s \to 0$ and tends to zero for $s \to \infty$. With (C.3) the component G_{zz}^A can be split into

two parts G_{zz1}^A and G_{zz2}^A,

$$G_{zz1}^A = \frac{1}{2\pi^{3/2}} \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} \int_0^E e^{-R_{i,mnp}^2 s^2 + \frac{k^2}{4s^2}} ds, \tag{C.4}$$

$$G_{zz2}^A = \frac{1}{2\pi^{3/2}} \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} \int_E^{\infty} e^{-R_{i,mnp}^2 s^2 + \frac{k^2}{4s^2}} ds. \tag{C.5}$$

We first concentrate on the expression G_{zz1}^A. By means of (C.2) it can be rewritten as

$$G_{zz1}^A = \frac{1}{2\pi^{3/2}} \sum_{i=0}^{7} A_i^{zz} \int_0^E e^{\frac{k^2}{4s^2}} \times$$

$$\left[\sum_{m=-\infty}^{\infty} e^{-(X_i+2ml_x)^2 s^2} \sum_{n=-\infty}^{\infty} e^{-(Y_i+2nl_y)^2 s^2} \sum_{p=-\infty}^{\infty} e^{-(Z_i+2pl_z)^2 s^2} \right] ds \tag{C.6}$$

In a next step we transform the summations according to

$$\sum_{m=-\infty}^{\infty} e^{-(X_i+2ml_x)^2 s^2} = \frac{\sqrt{\pi}}{2l_x s} \sum_{m=-\infty}^{\infty} e^{-\left(\frac{m\pi}{2l_x s}\right)^2 + j\frac{m\pi}{l_x} X_i}, \tag{C.7}$$

$$\sum_{n=-\infty}^{\infty} e^{-(Y_i+2nl_y)^2 s^2} = \frac{\sqrt{\pi}}{2l_y s} \sum_{n=-\infty}^{\infty} e^{-\left(\frac{n\pi}{2l_y s}\right)^2 + j\frac{n\pi}{l_y} Y_i}, \tag{C.8}$$

$$\sum_{p=-\infty}^{\infty} e^{-(Z_i+2pl_z)^2 s^2} = \frac{\sqrt{\pi}}{2l_z s} \sum_{p=-\infty}^{\infty} e^{-\left(\frac{p\pi}{2l_z s}\right)^2 + j\frac{p\pi}{l_z} Z_i}. \tag{C.9}$$

At this point it is convenient to employ the notation introduced by (2.167) and (2.168),

$$k_x = \frac{m\pi}{l_x}, \qquad k_y = \frac{n\pi}{l_y}, \qquad k_z = \frac{p\pi}{l_z}, \qquad k_{mnp}^2 = k_x^2 + k_y^2 + k_z^2, \tag{C.10}$$

in order to express (C.6) in the form

$$G_{zz1}^A = \frac{1}{16l_x l_y l_z} \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} e^{j(k_x X_i + k_y Y_i + k_z Z_i)} \int_0^E \frac{e^{-\frac{k_{mnp}^2 - k^2}{4s^2}}}{s^3} ds. \tag{C.11}$$

The integration can be performed in closed form,

$$\int_0^E \frac{e^{-\frac{k_{mnp}^2 - k^2}{4s^2}}}{s^3} ds = 2\frac{e^{-\frac{k_{mnp}^2 - k^2}{4E^2}}}{k_{mnp}^2 - k^2}, \tag{C.12}$$

yielding the result (3.114)

$$G_{zz1}^A = \frac{1}{8l_x l_y l_z} \sum_{m,n,p=-\infty}^{\infty} \sum_{i=0}^{7} A_i^{zz} \frac{e^{-\frac{k_{mnp}^2 - k^2}{4E^2}}}{k_{mnp}^2 - k^2} e^{j(k_x X_i + k_y Y_i + k_z Z_i)}. \tag{C.13}$$

Next we consider the expression (C.5) for G_{xx2}^A. Ewald, in his paper [45], derived the identity

$$\frac{2}{\sqrt{\pi}} \int_E^\infty e^{-R_{i,mnp}^2 s^2 + \frac{k^2}{4s^2}} \, ds = \frac{1}{2R_{i,mnp}} \left[e^{jkR_{i,mnp}} \text{erfc}(R_{i,mnp}E + jk/2E) + \right.$$
$$\left. + e^{-jkR_{i,mnp}} \text{erfc}(R_{i,mnp}E - jk/2E) \right] \qquad (C.14)$$

which immediately leads to the result (3.115),

$$G_{zz2}^A = \frac{1}{8\pi} \sum_{m,n,p=-\infty}^\infty \sum_{i=0}^7 A_i^{zz} \left[\frac{e^{jkR_{i,mnp}} \text{erfc}(R_{i,mnp}E + jk/2E)}{R_{i,mnp}} \right.$$
$$\left. + \frac{e^{-jkR_{i,mnp}} \text{erfc}(R_{i,mnp}E - jk/2E)}{R_{i,mnp}} \right]. \qquad (C.15)$$

This completes the derivation of the Ewald representation (3.113) from the ray representation (2.256).

Bibliography

[1] Abramowitz, M. and Stegun. I.A.(eds.): *Pocketbook of Mathematical Functions*, (Harri Deutsch, Thun, 1984).

[2] Achieser, N.I. and Glasmann, I.M.: *Theorie der linearen Operatoren im Hilbert-Raum*, 8th ed., (Akademie-Verlag, Berlin, 1981).

[3] Agrawal, A.K. Price, H.J. and Gurbaxani, S.H.: "Transient response of multiconductor transmission-lines excited by a nonuniform electromagnetic field", *IEEE Trans. Electromagn. Compat.*, vol. 22, (May 1980), 119–129.

[4] Balanis, C.A.: *Advanced Engineering Electromagnetics*, (John Wiley & Sons, New York, 1989).

[5] Balanis, C.A.: *Antenna Theory*, 3rd ed., (John Wiley & Sons, New York, 2005).

[6] Balian, R. and Duplantier, B.: "Electromagnetic waves near perfect conductors. I. Multiple scattering expansions. Distribution of modes", *Annals of Physics*, vol. 104, (1977) 300–335.

[7] Baum, C.E., Nitsch, J., and Sturm, R.: "Analytical solution for uniform and nonuniform multiconductor transmission lines with sources", in *Review of Radio Science*, W.R. Stone (ed.), (Oxford University Press, Oxford, 1996).

[8] Baum. C.E.: "General Properties of Antennas", *IEEE Trans. Electromagn. Compat.*, vol. 44, (February 2002), 18–24.

[9] Baum, C.E. and Steinmetz, T.: "An interpolation technique for analyzing sections of nonuniform multiconductor transmission lines", in *Proc. of EMC Zurich 03*, Zurich, Switzerland, February 2003, 593–596.

[10] Bohm, D.: *Quantum Theory*, (Prentice Hall, New York, 1951).

[11] Booton, R.C.: *Computational Methods for Electromagnetics and Microwaves*, (Wiley, New York, 1992).

[12] Borji, A. and Safavi-Nacini, S.: "Rapid calculation of the Green's function in a rectangular enclosure with applications to conductor loaded cavity resonators", *IEEE Trans. Microw. Theory Tech.*, vol. 52, (July 2004) 1724–1731.

[13] Brigham, E.O.: *Fast Fourier Transform and Its Applications*, (Prentice Hall, Englewood Cliffs, 1988).

[14] Budó, A: *Theoretische Mechanik*, 9th ed., (VEB Verlag, Berlin, 1978).

[15] Bussgang, J., Ehrman, L., and Graham, J.: "Analysis of nonlinear systems with multiple inputs", *Proc. IEEE*, vol. 62, (August1974), 1088–1118.

[16] Butler, C.M. and Wilton, D.R.: "Analysis of Various Numerical Techniques Applied to Thin-Wire Scatterers", *IEEE Trans. Antennas Propagat.*, vol. 23, (July 1975), 534–540.

[17] Byron, F.W. and Fuller, R.W.: *Mathematics of Classical and Quantum Physics*, (Dover, New York, 1992).

[18] Carlberg, U., Kildal, P.-S., and Carlsson: "Study of Antennas in Reverberation Chamber Using Method of Moments with Cavity Green's Function Calculated by Ewald Summation" *IEEE Trans. Electromagn. Compat.*, vol. 47, (November 2005), 805–814.

[19] Carpes Jr., W.P., Ferreira, G.S., Raizer, A., Pichon, L., and Razek, A.: "TLM and FEM methods applied in the analysis of electromagnetic coupling", *IEEE Trans. Magn.*, vol. 36, (July 2000), 982–985.

[20] Carpes Jr., W.P., Pichon, L., and Razek, A.: "Analysis of the coupling of an incident wave with a wire inside a cavity using an FEM in frequency and time domains", *IEEE Trans. Electromagn. Compat.*, vol. 44, (August 2002), 470–475.

[21] Chen, T.-P. and Li, L.-F.: *Gauge theory of elementary particle physics*, (Clarendon Press, Oxford, 1984).

[22] Chew, W.C.: *Waves and Fields in Inhomogeneous Media*, (IEEE Press, New York, 1995).

[23] Chihara, T.S.: *An Introduction to Orthogonal Polynomials*, (Gordon and Breach, New York, 1978).

[24] Chua, L.O. and Ng, C.-Y.: "Frequency-domain analysis of nonlinear systems: formulation of transfer functions", *IEE J. Electronic Circuits and Systems*, vol. 3, (November 1979), 257–269.

[25] Chua, L.O., Desoer, C.A., and Kuh, E.S.: *Linear and Nonlinear Circuits*, (McGraw-Hill, New York, 1987).

[26] Cockrell, C.R: "The Input Admittance of the Rectangular Cavity-Backed Slot Antenna", *IEEE Trans. Antennas Propagat.* vol. 24, (May 1976), 288–294.

[27] Cohen, E.: "An Ewald Transformation of Frequency Domain Integral Formulations", *Electromagnetics*, vol. 15, (1995), 427–439.

[28] Cohen-Tannoudji, C., Dupont-Roc, J., and Grynberg G.: *Photons & Atoms – Introduction to Quantum Electrodynamics*, (John Wiley & Sons, New York, 1997).

[29] Collin, R.E. and Zucker, F.J.: *Antenna Theory*, Part I and II, (McGraw-Hill, New York, 1969).

[30] Collin, R.E.: *Field Theory of Guided Waves*, 2nd ed., (IEEE Press, New York, 1997).

[31] Colton, D. and Kress, R.: *Integral Equation Methods in Scattering Theory*, (John Wiley & Sons, New York, 1983).

[32] Courant, R. and Hilbert, D.: *Methoden der mathematischen Physik*, 4th ed., (Springer, Berlin, 1993).

[33] Crowley, B.J.B.: "Some generalizations of the Poisson summation formula" *J. Phys. A*, vol. 12, (1979), 1951–1955.

[34] Danilov, L.V.: "Volterra-Picard series in the theory of nonlinear circuits", (Radio i Svyaz, Moscow, 1987).

[35] Davis, E.B.: *Spectral Theory and Differential Operators*, (Cambridge University Press, Cambridge, 1995).

[36] Debnath, L.: *Nonlinear Partial Differential Equations for Scientists and Engineers*, (Birkhäuser, Boston, 1997).

[37] Debnath, L. and Mikusiński, P.: *Introduction to Hilbert Spaces with Applications*, (Academic Press, San Diego, 1999).

[38] Dienstfrey, A., Hang, F., and Huang, J.: "Lattice sums and the two-dimensional, periodic Green's function for the Helmholtz equation", *Proc. Royal Soc. London*, vol. A457, (2001) 67–86.

[39] Dollard, J.D. and Friedman, C.N.: *Product Integration with Application to Differential Equations*, (Addison Wesley, Reading, 1979).

[40] Dudley, D.G.: *Mathematical Foundations for Electromagnetic Theory*, (IEEE Press, New York, 1994).

[41] Dunford, N. and Schwartz, J.T.: *Linear Operators – Part I: General Theory, Part II: Spectral Theory, Part III: Spectral Operators*, (John Wiley & Sons, New York, 1988).

[42] Elliott, R.S.: *Antenna theory and design*, (Prentice–Hall, Englewood Cliffs, 1981).

[43] Elliott, R.S.: *Electromagnetics – History, Theory, and Applications* (IEEE Press, New York, 1992).

[44] Ewald, P.P.: "Theorie der Dispersion, Reflexion und Brechung", *Annalen der Physik*, vol. 49, (1916) 1–38.

[45] Ewald, P.P.: "Die Berechnung optischer und elektrostatischer Gitterpotentiale", *Annalen der Physik* **64** (1921) 253–287.

[46] Felsen, L.P., (ed.): *Hybrid Formulation of Wave Propagation and Scattering*, (Martinus Nijhoff, Dordrecht, 1984).

[47] Felsen, L.P.: "Progressing and Oscillatory Waves for Hybrid Synthesis of Source Excited Propagation and Diffraction", *IEEE Trans. Antennas Propagat.* vol. 32, (August 1984), 775–796.

[48] Fiori, F. and Crovetti, P.S.: "Prediction of EMI effects in operational amplifiers by a two-input Volterra series model", *IEE Proc.-Circuits Devices Syst.*, vol. 150, (June 2003), 185–193.

[49] Frei, S. and Jobava, R.: "Coupling of inhomogeneous fields into an automotive cable harness with arbitrary terminations", in *Proc. of EMC Zurich 01*, Zurich, Switzerland, February 2001, 87–92.

[50] Friedman, B.: "Principles and Techniques of Applied Mathematics", (John Wiley & Sons, New York, 1957).

[51] Friedman, A.: *Foundations of Modern Analysis*, (Dover, New York, 1982).

[52] Galejs, J.: "Admittance of a Rectangular Slot Which is Backed by a Rectangular Cavity", *IEEE Trans. Antennas Propagat.* vol. 11, (March 1963), 119–126.

[53] Gantmacher, F.R.: *The theory of matrices, vol. 2*, (Chelsea Publishing, New York, 1984).

[54] Garvan, F.: *The Maple Book*, (CRC Press, Boca Raton, 2001).

[55] Gelfand, I.M. and Shilov, G.E.: *Generalized functions, vols. 1-4*, (Academic Press, New York, 1964–1968).

[56] Gielen, G. and Sansen, W.: *Symbolic analysis for automated design of analog integrated circuits*, (Kluwer Academic Publishers, Boston, 1991).

[57] Gielen, G., Wambacq, P., and Sanseb, W.: "Symbolic analysis methods and applications for analog circuits: a tutorial overview", *Proc. IEEE*, vol. 82, (February 1994), 287–304.

[58] Gilmore R.: *Alice in Quantumland*, (Copernicus Books, New York, 1995).

[59] Gong, J.; Volakis, J.L., Woo, A.C., and Wang, H.T.G.: "A hybrid finite element-boundary integral method for the analysis of cavity-backed antennas of arbitrary shape", *IEEE Trans. Antennas Propagat.* vol. 42, (September 1994), 1233–1242.

[60] Griffel, D.H.: "Applied Functional Analysis", (John Wiley & Sons, New York, 1981).

[61] Griffiths, D.J.: *Introduction to Electrodynamics*, 3rd ed., (Prentice Hall, New Jersey, 1999).

[62] Gronwald, F.: "Metric–Affine Gauge Theory of Gravity I. Fundamental Structure and Field Equations", *Int. J. Mod. Phys. D*, vol. 6, (1997) 263–303.

[63] Gronwald, F. and Nitsch, J.: "A Geometric Scattering Expansion of the Current on Conducting Surfaces and Transmission Line Structures", *Proc. of the Int. IEEE Symposium on EMC*, Seattle, USA, (August 1999), 739–744.

[64] Gronwald, F. and Nitsch, J.: "The physical origin of gauge invariance in electrodynamics", *Electrical Engineering*, **81** (1999) 363–367.

[65] Gronwald, F. and Nitsch, J.: "Analytically derived radiation and reflection properties of high-frequency signals on arbitrarily curved transmission lines", *Proc. of EMC Europe 2000*, Brugge, Belgium, 125–130, September 2000.

[66] Gronwald, F. and Nitsch, J.: "The structure of the electrodynamic field as derived from first principles", *IEEE Antennas and Propagation Magazine*, vol. 43, (August 2001) 64–79.

[67] Gronwald, F. and Nitsch, J.: "Universeller Zusammenhang: Wie das elektromagnetische Feld die Welt verbindet", (in german) *Magdeburger Wissenschaftsjournal*, vol. 1/2 (2001), 19–28.

[68] Gronwald, F., Nitsch, J., and Tkachenko, S: "Hybrid representation methods for the efficient analysis of antenna coupling within cavities", *EMC 2002 – Proc.of the Sixteenth International Wroclaw Symposium on Electromagnetic Compatibility*, Wroclaw, Poland, (June 2002), 109–114.

[69] Gronwald, F.: "The influence of electromagnetic singularities on an active dipole antenna within a cavity", *Advances in Radio Science*, vol. 1, (2003), 57–61.

[70] Gronwald, F., Blume E., and Nitsch, J.: "Computation of the frequency response of a nonlinearly loaded antenna within a cavity", *Advances in Radio Research*, vol. 2 (2004) 57–62.

[71] Gronwald, F. and Blume, E.: "Reciprocity and mutual impedance formulas within lossy cavities", *Advances in Radio Science*, vol. 3 (2005) 91–97.

[72] Gronwald, F.: "Calculation of Mutual Antenna Coupling within Rectangular Enclosures", *IEEE Trans. on Electromagn. Compat.*, vol. 47 (November 2005) 1021–1025.

[73] Haase, H. and Nitsch, J.: "Full-wave transmission line theory for the analysis of three-dimensional wire-like structures", in *Proc. of EMC Zurich 01*, Zurich, Switzerland, February 2001, 235–240.

[74] Haase, H., Nitsch, J., and Steinmetz, T.: "Transmission-Line Super Theory: A New Approach to an Effective Calculation of Electromagnetic Interactions", *URSI Radio Science Bulletin (Review of Radio Science)*, vol. 307, (December 2003), 33–60.

[75] Haase, H., Steinmetz, T., and Nitsch, J.: "New propagation models for electromagnetic waves along uniform and nonuniform cables", *IEEE Trans. on Electromagn. Compat.*, vol. 46, (August 2004), 345–352.

[76] Haase, H.: "Full-Wave Field Interactions of Nonuniform Transmission Lines", PhD thesis, appeared in *Res Electricae Magdeburgenses, Magdeburger Forum zur Elektrotechnik*, Nitsch, J. and Styczynski, Z.A. (eds.), vol. 9, (Otto-von-Guericke-Universität Magdeburg, 2005).

[77] Hallén, E.: "Theoretical investigations into the transmitting and receiving qualities of antennae", *Nova Acta Regiae Soc. Sci. Upsaliensis*, Ser. IV, no. 4, (1938), 1-44.

[78] Hampe, M. and Dickmann, S.: "Damping of Cavity-Mode Resonances in PCB Power-Bus Structures Using Discrete Capacitors", *IEEE Trans. Electromagn. Compat.*, vol. 47, (November 2005), 880–888.

[79] Hanke-Bourgeois, M.: *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*, (Teubner, Stuttgart, 2002).

[80] Hansen, W.W.: "A New Type of Expansion in Radiation Problems", *Phys. Rev.*, vol. 47 (1935) 139–143.

[81] Hanson, G.W. and Yakovlev: *Operator Theory for Electromagnetics*, (Springer, New York, 2002).

[82] Harrington, R.F.: "Time-Harmonic Electromagnetic Fields", (McGraw-Hill, New York, 1961).

[83] Harrington, R.F.: *Field Computation by Moment Methods*, (IEEE Press, New York 1993).

[84] Hehl, F.W. and Obukhov, Y.: *Foundations of Classical Electrodynamics: Charge, Flux, and Metric*, (Birkhäuser, Boston, 2003).

[85] Hehl, F.W. and Obukhov, Yu.N.: "Linear media in classical electrodynamics and the Post constraint," *Phys. Lett. A*, vol. 334 (2005) 249–259.

[86] von Helmholtz, H.: "Über Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen", *Crelles Journal*, vol. 55, (1858), 25–55.

[87] Hertz, H.: "Ueber sehr schnelle elektrische Schwingungen" and "Nachtrag zu der Abhandlung ueber sehr schnelle elektrische Schwingungen", *Annalen der Physik und Chemie*, vol. 31 (1887) 421–448 and 543–544.

[88] Hill, D.A.: "Linear dipole response in a reverberation chamber", *IEEE Trans. Electromagn. Compat.*, vol 41., (November 1999), 365–368.

[89] Hoppe, D.J. and Rahmat-Samii, Y.: "Impedance Boundary Conditions in Electromagnetics", (Taylor & Francis, Washington, 1995).

[90] Hosking, R.J., Joe, S., Joyce, D.C., and Turner, J.C.: "First Steps in Numerical Analysis", 2nd ed., (Arnold, London, 1996).

[91] Huang C.-C. and Chu, T.-H.: "Analysis of Wire Scatterers with Nonlinear or Time-Harmonic Loads in the Frequency Domain", *IEEE Trans. Antennas Propagat.* vol. 41, January 1993, 25–30.

[92] Ilyinski, A.S., Slepyan, G.Y., and Slepyan, A.Y.: "Propagation, scattering and dissipation of electromagnetic waves", *IEE Electromagnetic Waves Series*, vol. 36 (IEE, London, 1993).

[93] Jackson, J.D.: *Classical Electrodynamics*, 3rd edition (John Wiley & Sons, New York, 1998).

[94] Janaswamy, R. and Lee, S.-W.: "Scattering from Dipoles Loaded with Diodes" *IEEE Trans. Antennas Propagat.*, vol. 36, (November 1988), 1649–1651.

[95] Jeng, G. and Wexler, A.: "Self-Adjoint Variational Formulation of Problems Having Non-self-adjoint Operator", *IEEE Trans. Microwave Theory Tech.*, vol. 26, (February 1978), 91–94.

[96] Jones, D.S.: *Methods in Electromagnetic Wave Propagation*, (IEEE Press, New York, 1995).

[97] Jordan, E.C. and Balmian, K.G.: "Electromagnetic Waves and Radiating Systems", 2nd ed., (Prentice Hall, Englewood Cliffs, 1968).

[98] Jordan, K.E., Richter, G.R., and Sheng, P.: "An Efficient Numerical Evaluation of the Green's Function for the Helmholtz Operator on Periodic Structures", *Journal of Computational Physics*, vol. 63, (1986) 222–235.

[99] Jorgensen, R.E. and Mittra, R.: "Efficient calculation of the free-space periodic Green's functions", *IEEE Trans. Antennas Propagat.*, vol. 38, (May 1990), 633–642.

[100] Kanda, M.: "Analytical and Numerical Techniques for Analyzing an Electrically Short Dipole with a Nonlinear Load", *IEEE Trans. Antennas Propagat.* vol. 28, (January 1980), 71–78.

[101] Kelley, C.T.: *Iterative Methods for Linear and Nonlinear Equations*, (SIAM, Philadelphia, 1995).

[102] Kerr, A.R.: "A Technique for Determining the Local Oscillator Waveforms in a Microwave Mixer", *IEEE Trans. Microwave Theory Tech.*, vol. 23, (1975), 828–831.

[103] King, R.W.P.: "Transmission Line Theory", (McGraw-Hill, New York, 1955).

[104] King, R.W.P.: "The Theory of Linear Antennas", (Harvard University Press, Cambridge, 1956).

[105] King, R.W.P.: "The Linear Antenna – Eighty Years of Progress", *Proc. of the IEEE*, **55**, (January 1967), 2–16.

[106] King, R.W.P. and Harrison, C.W.: *Antennas and Waves: A modern approach*, (The M.I.T. Press, Cambridge, 1969).

[107] Konefal, T., Dawson, J.F., Denton, A.C., Benson, T.M., Christopoulos, C., Marvin, A.C., Porter, S.J., and Thomas, D.W.P.: "Electromagnetic coupling between wires inside a rectangular cavity using multiple-mode-analogous-transmission-line circuit theory", *IEEE Trans. Electromagn. Compat.*, vol. 43, (August 2001), 273–281.

[108] Kose, V. (ed.): *Non-linear electromagnetic systems: advanced techniques and mathematical methods*, (IOS Press, Amsterdam, 1998).

[109] Kundert, K.S. and Sangiovanni-Vincentelli: "Simulation of nonlinear circuit in the frequency domain", *IEEE Trans. Computer-Aided Design*, vol. CAD-5, (October 1986) 521–535.

[110] Kustepeli, A. and Martin, A.Q.: "On the splitting parameter in the Ewald method", *IEEE Microwave Guided Wave Lett.*, vol. 10, (May 2000), 168–170.

[111] Krall, A.M.: *Hilbert Space, Boundary Value Problems and Orthogonal Polynomials*, (Birkhäuser, Basel, 2002).

[112] Krauthäuser, H. G., Tkachenko, S., and Nitsch, J.: "The Action of Non-Linear Effects in a Resonator", in *Proc. of the XXVIIth General Assembly of the International Union of Radio Science, URSI GA 2002*, Maastricht, The Netherlands, August 2002, 4 pages.

[113] Krauthäuser, H.G. and Nitsch, J.: "Effects of the Variation of the Excitation and Boundary Conditions of Mode-Stirred Chambers and Consequences for Calibration and Measurements", in *Proc. of EMC Zurich 03*, Zurich, Switzerland, February 2003, 615–620.

[114] Lämmerzahl, C. and Hehl, F.W.: "Riemannian light cone from vanishing birefringence in premetric vacuum electrodynamics," *Phys. Rev. D*, vol. 70, (2004), 105022 (7 pages).

[115] Lamb, H.: *Hydrodynamics*, 6th ed., (Cambridge University Press, Cambridge, 1936 and Dover, New York 1993).

[116] Landt, J.A., Miller, E.K., and Deadrick, F.J.: "Time Domain Modeling of Nonlinear Loads" *IEEE Trans. Antennas Propagat.*, vol. 22, (July 1974), 611–613.

[117] Lee, K.S.H., (ed.): *EMP Interaction: Principles, Techniques, and Reference Data*, revised printing, Taylor & Francis, Washington D.C. 1995.

[118] Lee, K.-C.: "Two Efficient Algorithms for the Analyses of a Nonlinearly Loaded Antenna and Antenna Array in the Frequency Domain", *IEEE Trans. Electromagn. Compat.* vol. 42, (November 2000), 339–346.

[119] Lee, K.-C.: "Genetic algorithms based analyses of nonlinearly loaded antenna arrays including mutual coupling effects", *IEEE Trans. Antennas Propag.* vol. 51, (April 2003), 776–781.

[120] Lee, K.-C. and Lin T.-N.: "Application of neural networks to analyses of nonlinearly loaded antenna arrays including mutual coupling effects", *IEEE Trans. Antennas Propag.* vol. 53, (March 2005), 1126–1132.

[121] Lee, K.-C. and Chu T.-H.: "Mutual Coupling Mechanisms Within Arrays of Nonlinear Antennas", *IEEE Trans. Electromagn. Compat.* vol. 47, (November 2005), 963–970.

[122] Leon, B.I. and Schaefer, D.I.: "Volterra series and Picard iteration for nonlinear circuits and systems", *IEEE Trans. CAS*, vol. 25, (1977), 941–948.

[123] Leone, M.: "The radiation of a rectangular power-bus structure at multiple cavity-mode resonances", *IEEE Trans. Electromagn. Compat.*, vol. 45, (August 2005), 486–492.

[124] Leontovich, M.A.: "Investigations of Radiowave Propagation, Part II", (Printing House of the Academy of Sciences, Moscow, 1948).

[125] Li, M.-Y., Hummer, K.A., and Chang, K.: "Theoretical and Experimental Study of the Input Impedance of the Cylindrical Cavity-Backed Rectangular Slot Antennas", *IEEE Trans. Antennas Propag.* vol. 39, (August 1991), 1158–1166.

[126] Li, M., Nuebel, J. Drewniak, J.L., DuBroff, R.E., Hubing, T.H., and Van Doren, T.P.: "EMI from Cavity Modes of Shielding Enclosures – FDTD Modeling and Measurements" *IEEE Trans. Electromagn. Compat.*, vol. 42, (February 2000), 29–38.

[127] Li, M., Drewniak, J.L., Radu, S., Nuebel, J., Hubing, T.H., DuBroff, R.E., and Van Doren, T.P.:"An EMI estimate for shielding-enclosure evaluation", *IEEE Trans. Electromagn. Compat.*, vol. 43, (August 2001), 295–304.

[128] Lindell, I.V.: *Differential Forms in Electromagnetics* (IEEE Press, Piscataway, NJ, and Wiley-Interscience, 2004).

[129] Ling, F., Jiao, D., and Jing, J.N.: "Efficient electromagnetic modeling of microstrip structures in multilayer media", *IEEE Trans. Microwave Theory Tech.*, vol. 47, (September 1999), 1810–1818.

[130] Linton, C.M.: "The Green's function for the two-dimensional Helmholtz equation in periodic domains", *J. Engng. Math.*, vol. 33, (2001), 377–402.

[131] Liu, T.K. and Tesche, F.M.: "Analysis of Antennas and Scatterers with Nonlinear Loads", *IEEE Trans. Antennas Propagat.*, vol. 24, (March 1976), 131–139.

[132] Ma, M.T.: *Theory and Application of Antenna Arrays* (John Wiley & Sons, New York, 1974).

[133] Maas, S.A.: *Nonlinear Microwave Circuits*, 2nd ed., (Artech House, Norwood, 2003).

[134] Mackey, G.W.: *The Mathematical Foundations of Quantum Mechanics*, (Benjamin, New York, 1963).

[135] Magnusson, P.C., Alexander, G.C., Tripathi, V.K., and Weisshaar, A.: *Transmission Lines and Wave Propagation*, (CRC Press, Boca Raton, 2001).

[136] Marcuvitz, N.: *Waveguide Handbook* vol. 10, MIT Radiation Laboratory Series, (McGraw-Hill, New York, 1951).

[137] Marliani, F. and Ciccolella, A.: "Computationally Efficient Expressions of the Dyadic Green's Function for Rectangular Enclosures", *Progress in Electromagnetics Research*, vol. 31, (2001), 195–223.

[138] Mathis, A.W. and Peterson A.F.: "A Comparison of Acceleration Procedures for the Two-Dimensional Periodic Green's Function", *IEEE Trans. Antennas Propagat.*, vol. 44, (1996) 567–571.

[139] Mei, K.K.: "Theory of Maxwellian Circuits", *URSI Radio Science Bulletin*, vol. 305, (September 2003), 3–13.

[140] Merzbacher, E.: *Quantum Mechanics*, 2nd ed., (John Wiley & Sons, New York, 1961).

[141] Michlin, S.G.: *Lehrgang der mathematischen Physik*, 2nd ed., (Akademie-Verlag, Berlin, 1975).

[142] Mittra, R. (ed.): *Computer Techniques for Electromagnetics*, (Springer, Berlin, 1973).

[143] Mittra, R. (ed.): *Numerical and Asymptotic Techniques in Electromagnetics*, (Springer, Berlin, 1975).

[144] Monagan, M.B., Geddes, K.O., and Heal, K.M.: *Maple V Programming Guide*, 2nd ed., (Springer, New York, 1996).

[145] Monk, P.: *Finite Element Methods for Maxwell's Equations*, (Clarendon Press, Oxford, 2003).

[146] Morse, P.M. and Feshbach, H.: *Methods of Theoretical Physics I+II*, (McGraw Hill, New York, 1953).

[147] Nakano, H.: "Antenna Analysis using Integral Equations", in *Analysis Methods for Electromagnetic Wave Problems, Volume Two*, E. Yamashita (ed.) (Artech House, Boston, 1996).

[148] Nakhla, M.S. and Vlach, J.: "A Piecewise Harmonic Balance Technique for Determination of Periodic Response of Nonlinear Systems", *IEEE Trans. CAS*, vol. 23, (1976), 85.

[149] Nauwelaers, B.K.J.C. and Van de Capelle, A.R.: "Integrals for the Mutual Coupling Between Dipoles or Between Slots: With or Without Complex Conjugate?", *IEEE Trans. Antennas Propagat.*, vol. 36, (1988) 1357–1381.

[150] Naylor, A.W. and Sell, G.R.: *Linear Operator Theory in Engineering and Science*, 2nd ed., (Springer, New York, 1982).

[151] Neumayer, R., Stelzer, A., Haslinger, F., Steinmair, G., Troscher, M., Held, J., Unger, B., and Weigel, R.: "Numerical EMC-simulation for automotive applications", in *Proc. of EMC Zurich 03*, Zurich, Switzerland, February 2003, 459–464.

[152] Nitsch, J., Bohl, J., Strähle, U., Kaiser, A., Meyer, L., and Vogel H.-J.: "High-Power Microwaves Effects on Smart Electronic Weapon Systems", *Proc. of the AGARD-Symposium, AGARD-CP-600*, vol. 2, ISBN 92-836-0048-7, (December 1997) B9-1 – B9-13.

[153] Nitsch, J., Gronwald, F., Pelster, R., and Steinmetz, T.: "EME-Schutz von fliegenden Waffensystemen – EMC zukünftiger Verteidigungssysteme" (in german) (Magdeburg, 1999), 129 pages.

[154] Nitsch, J. and Gronwald, F.: "Analytical solutions in multiconductor transmission line theory", *IEEE Trans. Electromagn. Compat.*, vol. 41, (November 1999), 469–479.

[155] Nitsch, J. and Tkachenko, S.: "Eine Transmission-Line Beschreibung für eine vertikale Halbschleife auf leitender Ebene", Feser, K. (ed.), in *Elektromagnetische Verträglichkeit, EMV 2004*, (VDE-Verlag, Berlin, 2004), 291–300.

[156] Nitsch, J. and Tkachenko, S.: "Telegrapher Equations for Arbitrary Frequencies and Modes-Radiation of an Infinite, Lossless Transmission Line", *Radio Science*, vol. 39, (2004), RS2026, doi: 10.1029/2002RS002817.

[157] Nitsch, J. and Tkachenko, S.: "Complex-Valued Transmission-Line Parameters and their Relation to the Radiation Resistance", *IEEE Trans. Electromagn. Compat.*, vol. 46, (August 2004), 477–487.

[158] Nitsch, J., Korovkin, N., Solovieva, E., and Gronwald, F.: "Characterization and Compensation of Beat-Induced Intermodulation Effects in Analog Circuits", in *ANALOG '05, 8. GMM/ITG-Diskussionssitzung*, Hannover, März 2005, p. 191–196.

[159] Nitsch, J. and Tkachenko, S.: "Global and Modal Parameters in the Generalized Transmission Line Theory and Their Physical Meaning", *URSI Radio Science Bulletin*, vol. 312, (March 2005), 21–31.

[160] Neumark, M.A.: "Lineare Differentialoperatoren", (Akademie Verlag, Berlin, 1960).

[161] Noether, E.: "Invariante Variationsprobleme", *Nachr. König. Gesell. Wiss. Göttingen, Math.-Phys. Kl.* (1918) 235–257; English translation in: *Transport Theory and Stat. Phys.*, vol. 1 (1971), 186-207.

[162] Nosich, A.I.: " The method of analytical regularization in wave-scattering and eigenvalue problems: foundations and review of solutions", *IEEE Antennas and Propagation Magazine*, vol. 41, (June 1999), 34–49.

[163] Obukhov, Yu.N. and Hehl, F.W.: "Measuring a piecewise constant axion field in classical electrodynamics," *Phys. Lett. A*, vol. 341, (2005), 357–365.

[164] Olyslager, F., Laermans, E., De Zutter, D., Criel, S., De Smedt, R., Lietaert, N., and De Clercq, A.:"Numerical and experimental study of the shielding effectiveness of metallic enclosures", *IEEE Trans. Electromagn. Compat.*, vol. 41, (August 1999), 202–213.

[165] Ortega, J.M. and Rheinboldt, W.C.: *Iterative Solution of Nonlinear Equations in Several Variables*, (SIAM, Philadelphia, 2000)

[166] Paletta, L., Parmantier, J., Issac, F., Dumas, P., and Alliot, J.: "Susceptibility analysis of wiring in a complex system combining a 3-D solver and a transmission-line network simulation", *IEEE Trans. Electromagn. Compat.*, vol. 44, (May 2002), 309–317.

[167] Park, M.-J., Park, J., and Nam, S.: "Efficient Calculation of the Green's Function for the Rectangular Cavity", *IEEE Microwave and Guided Wave Letters*, vol. 8, (March 1998), 124–126.

[168] Paul, C.R.: *Introduction to Electromagnetic Compatibility*, (John Wiley & Sons, New York, 1992).

[169] Paul, C.R.: *Analysis of Multiconductor Transmission Lines*, (John Wiley & Sons, New York, 1994).

[170] Poppe, G.P.M. and Wijers, C.M.J.: "More efficient computation of the complex error function", *ACM Trans. Math. Softw.*, vol. 16, (March 1990), 38–46.

[171] Post, E.J.: *Formal Structure of Electromagnetics: General Covariance and Electromagnetics*, (North Holland, Amsterdam, 1962 and Dover, New York, 1997).

[172] Pozar, D.M.: *Microwave Engineering*, 3rd ed., (John Wiley & Sons, New York, 2004)

[173] Olyslager, F. and Lindell, V.: "Electromagnetics and exotic media: A quest for the holy grail," *IEEE Antennas and Propagation Magazine*, vol. 44, No.2 (2002) 48-58.

[174] Peterson, A.F., Ray, S.L., and Mittra, R.: *Computational Methods for Electromagnetism*, (IEEE Press, New York, 1998).

[175] Pocklington, H.C.: "Electrical oscillations in wires", *Cambridge Philos. Soc. Proc.*, vol. 9, (1897) 324–332.

[176] Popović, B.D., Dragović, M.B. and Djordjecić, A.R.: *Analysis and Synthesis of Wire Antennas*, (Research Studies Press, Chichester, 1982).

[177] Popović, B.D. and Kolundžija: *Analysis of metallic antennas and scatterers*, (IEE, London,1994).

[178] Rachidi, F.: "Formulation of the Field-to-Transmission Line Coupling Equations in Terms of Magnetic Excitation Field", *IEEE Trans. Electromagn. Compat.*, vol. 35, (August 1993), 404–407.

[179] Rahman, H.: "Coupling characterization of multiple wires in a rectangular cavity: Part I – Thin wire approximation", *Journal of Electromagnetic Waves and Applications*, vol. 11, (1997), 77-88.

[180] Rahman, H.: "Coupling characterization of multiple wires in a rectangular cavity: Part II – Thick wire approximation", *Journal of Electromagnetic Waves and Applications*, vol. 11, (1997), 89-99.

[181] Ramo, S., Whinnery, J.R., and Van Duzer: *Fields and Waves in Communication Electronics*, 3rd ed., (John Wiley & Sons, New York, 1997).

[182] Rohrlich, F.: *Classical Charged Particles*, (Addison-Wesley, Reading, 1965).

[183] Rothwell, E.J. and Cloud, M.J.: *Electromagnetics*, (CRC Press, Boca Raton, 2001).

[184] Rugh, W.J.: *Nonlinear System Theory*, (The John Hopkins University Press, Baltimore, 1981).

[185] Ryder, L.: *Quantum Field Theory*, 2nd ed., (Cambridge University Press, Cambridge, 1996).

[186] Rynne, B.P.: "The Well-Posedness of the Integral Equations for Thin-Wire Antennas with Distributional Incident Fields", *Q. Jl Mech. appl. Math.*, vol. 52, (1999), 489–497.

[187] Sadiku, M.N.O.: *Numerical Techniques in Electromagnetics*, 2nd ed., (CRC Press, Boca Raton, 2000).

[188] Sarkar, T.K. and Weiner, D.D.: "Scattering Analysis of Nonlinearly Loaded Antennas", *IEEE Trans. Antennas Propagat.*, vol. 24, (March 1976), 125–131.

[189] Sarkar, T. K., Djordjević, A. R., and Kolundžija, B. M.: "Method of Moments Applied to Antennas", in *The Handbook of Antennas in Wireless Communications*, Godara, L. and Barroso, V. (eds.), (CRC Press, Boca Raton, 2001).

[190] Schouten, J.A.: *Tensor Analysis for Physicists*, 2nd ed. reprinted (Dover, New York, 1989).

[191] Schelkunoff, S.A.: "Concerning Hallén's Integral Equation for Cylindrical Antennas", *Proc. IRE*, (December 45), 872–878.

[192] Schelkunoff, S.A. and Friis, H.T.: *Antennas – Theory and Practice*, 3rd ed. (John Wiley & Sons, New York, 1966).

[193] Schetzen, M.: *The Volterra and Wiener Theories of Nonlinear Systems*, (John Wiley & Sons, New York, 1980).

[194] Schuman, H.: "Time-Domain Scattering from a Nonlinearly Loaded Wire" *IEEE Trans. Antennas Propagat.*, vol. 22, (July 1974), 611–613.

[195] Schreier, F.: "The Voigt and Complex Error Function: A Comparison of Computational Methods", *J. Quant. Spectros. Radiat. Transfer*, vol. 48, (1992) 743–762.

[196] Schumaker, L.: *Spline Functions: Basic theory*, (John Wiley & Sons, New York, 1981).

[197] Schwab, A.J.: *Elektromagnetische Verträglichkeit*, 4th. ed., (Springer, Berlin, 1996).

[198] Schwartz, L.: *Mathematics for the physical sciences*, (Addison-Wesley, Reading, 1966).

[199] Schwinger, J., DeRaad Jr., L.L., Milton, K.A., and Tsai, W.: *Classical Electrodynamics*, (Perseus Books, Reading, 1998).

[200] Scott, A.: *Nonlinear Science*, (Oxford University Press, Oxford, 1999).

[201] Senior, T.B.A. and Volakis,J.L.: *Approximate boundary conditions in electromagnetics*, (IEE, London, United Kingdom, 1995).

[202] Shi, S., Hirasawa, K., and Chen Z.N.: "Circularly Polarized Rectangularly Bent Slot Antennas Backed by a Rectangular Cavity", *IEEE Trans. Antennas Propagat.*, vol. 49, Nov. 2001, 1517–1524.

[203] Siah, E.S., Sertel, K., Volakis, J.L., Liepa, V.V., and Wiese, R.: "Coupling studies and shielding techniques for electromagnetic penetration through apertures on complex cavities and vehicular platforms", *IEEE Trans. Electromagn. Compat.*, vol. 45, (May 2003), 245–256.

[204] Singh, S., Richards, W.F., Zinecker, J.R., and Wilton, D.R.: "Accelerating the convergence of series representing the free space periodic Green's function", *IEEE Trans. Antennas Propagat.*, vol. 38, (December 1990), 633–642.

[205] Sobolev, S.L.: *Partial Differential Equations of Mathematical Physics*, (Dover, New York, 1989).

[206] Sommerfeld, A.: *Mechanics of Deformable Bodies*, (Academic Press, New York, 1950).

[207] Spadacini, G., Pignari, A., and Marliani, F.: "Closed-Form Transmission Line Model for Radiated Susceptibility in Metallic Enclosures" *IEEE Trans. Electromagn. Compat.*, vol. 47, (November 2005), 701–708.

[208] Stakgold, I.: *Green's Functions and Boundary Value Problems*, 2nd ed., (John Wiley & Sons, New York, 1998).

[209] Steele, C.W.: *Numerical Computation of Electric and Magnetic Fields*, 2nd ed., (Chapman & Hall, New York, 1997).

[210] Stratton, J.A.: *Electromagnetic Theory*, (McGraw-Hill, New York, 1941).

[211] Stutzman, W.L. and Thiele, G.A.: "Antenna Theory and Design", 2nd ed., (John Wiley & Sons, New York, 1997).

[212] Tai, C.-T. and Rozenfeld, P.: "Different Representations of Dyadic Green's Functions for a Rectangular Cavity", *IEEE Trans. Microwave Theory Tech.*, vol. 24, (September 1976), 597–601.

[213] Tai, C.-T.: *Dyadic Green Functions in Electromagnetic Theory*, (IEEE Press, New York, 1994).

[214] Taylor, C.D., Satterwhite, R.S., and Harrison C.H.: "The Response of a Terminated Two-Wire Transmission Line Excited by a Nonuniform Electromagnetic Field", *IEEE Trans. Antennas Propag.*, vol. 13, (November 1965), 987–989.

[215] Taylor, M.E.: "Partial Differential Equations II – Qualitative Studies of Linear Equations", (Springer, New York, 1996).

[216] Tesche, F.M.: "Principles and applications of EM field coupling to transmission lines", in *Proc. of EMC Zurich 95*, Zurich, Switzerland, March 1995, 21–31.

[217] Tesche, F.M., Ianoz, M.V., and Karlsson, T.: *EMC Analysis Methods and Computational Methods*, (John Wiley & Sons, New York, 1997).

[218] Tkachenko, S., Rachidi, F., and Ianoz, M.: "Electromagnetic Field Coupling to a Line of Finite Length: Theory and Fast Iterative Solutions in Frequency and Time Domains", *IEEE Trans. Electromagn. Compat.*, vol. 37, (November 1995), 509–518.

[219] Tkachenko, S., Vodopianov, G.V., and Martinov, L.V.: "Electromagnetic field coupling to an electrically small antenna in a rectangular cavity", in *Proc. of EMC Zurich 99*, Zurich, Switzerland, February 1999, 379–384.

[220] Tkachenko, S., Gronwald, F., Krauthäuser H.G., and Nitsch, J.: "Investigation of Electromagnetic Field Coupling to a Small Antenna within a Resonator" (in russian, english translation available), *Proc. of the Vth International Symposium on Electromagnetic Compatibility and Electromagnetic Ecology (EMC 2003)*, St. Petersburg, Russia, September 2003, p. 68–74.

[221] Tkachenko, S.; Gronwald, F.; Nitsch, J.; Steinmetz, T.: "On Electromagnetic Field Coupling to Transmission Lines in Resonators" (in russian, english translation available), *Proc. of the Vth International Symposium on Electromagnetic Compatibility and Electromagnetic Ecology (EMC 2003)*, St. Petersburg, Russia, September 2003, p. 74–78.

[222] Tkachenko, S., Gronwald, F., and Nitsch, J.: "Electromagnetic field coupling to a loaded transmission line inside a rectangular cavity", in *Proc. PIERS 2004 – Progress in Electromagnetics Research Symp.*, Pisa, Italy, (March 2004), p. 335.

[223] Tkachenko, S. and Nitsch, J.: "On the Electromagnetic Field Excitation of Smoothly Curved Wires", *Proc. of the VIth International Symposium on Electromagnetic Compatibility and Electromagnetic Ecology (EMC-2005)*, St. Petersburg, Russia, June, 2005, pp. 115-120.

[224] Trinkle, J., Cantoni, A., and Fynn, K: "Efficient Impedance Calculation of Loaded Power Ground Planes", in *Proc. of EMC Zurich 03*, Zurich, Switzerland, February 2003, 285–290.

[225] Truesdell, C. and Toupin, R.A.: *The classical field theories*, in *Handbuch der Physik*, vol. III/1, S. Flügge ed. (Springer, Berlin, 1960), 226–793.

[226] Unger, H.G.: *Elektromagnetische Wellen auf Leitungen*, 4th. ed., (Hüthig Buch Verlag, Heidelberg, 1996).

[227] Van Bladel, J.: *Electromagnetic Fields*, revised printing, (Hemisphere Publishing, New York, 1985).

[228] Van Bladel , J.: *Singular Electromagnetic Fields and Sources*, (Clarendon Press, Oxford, 1991).

[229] Vinogradov, S.S., Smith, P.D., and Vinogradova, E.D.: *Canonical Problems in Scattering and Potential Theory – Part I: Canonical Structures in Potential Theory*, (Chapman & Hall/CRC, Boca Raton, 2001).

[230] Vinogradov, S.S., Smith, P.D., and Vinogradova, E.D.: *Canonical Problems in Scattering and Potential Theory – Part II: Acoustic and Electromagnetic Diffraction by Canonical Structures*, (Chapman & Hall/CRC, Boca Raton, 2002).

[231] Wambacq, P. and Sansen, W.: *Distortion Analysis of Analog Integrated Circuits*, (Kluwer Academic Publishers, Boston, 1998).

[232] Warne, L.K. and Lee, K.S.H.: "Some remarks on antenna response in a reverberation chamber", *IEEE Trans. Electromagn. Compat.*, vol. 43, (May 2001), 239–240.

[233] Warne, L.K., Lee, K.S., Hudson, G., Johnson, W.A., Jorgenson, R.E., and Stronach, S.L.: "Statistical Properties of Linear Antenna Impedance in an Electrically Large Cavity", *IEEE Trans. Antennas Propagat.*, vol. 51, (2003), 978–992.

[234] Weideman, J.A.C.: "Computation of the Complex Error Function", *SIAM J. Numer. Anal.*, vol. 31, (1994) 1497–1518.

[235] Werner, D.H., Huffman, J.A., and Werner, P.L.: "Techniques for Evaluating the Uniform Current Vector Potential at the Isolated Singularity of the Cylindrical Wire Kernel", *IEEE Trans. Antennas Propagat.*, vol. 42, (1994), 1549–1553.

[236] Weyl, H.: "Electron and Gravitation", *Zeit. f. Phys.*, vol. 56, (1929) 330–352.

[237] Wilton, D.R. and Butler, C.M.: "Efficient Numerical Techniques for Solving Pocklington's Equation and Their Relationships to Other Methods", *IEEE Trans. Antennas Propagat.*, vol. 24, (January 1976), 83–86.

[238] Wolfram, S.: *The MATHEMATICA Book. Version 4.*, (Cambridge University Press, Cambridge, 1999).

[239] Wu, D.I. und Chang D.C.: "An Investigation of a Ray-Mode Representation of the Green's Function in a Rectangular Cavity", *National Bureau of Standards Technical Note* **1312**, (Sept. 1987).

[240] Wu, D.I. and Chang, D.C.: "A Hybrid Representation of the Green's Function in an Overmoded Rectangular Cavity", *IEEE Trans. Microwave Theory Tech.*, vol. 36, (1988), 1334–1342.

[241] Xia, L., Wang, C.-F., Li, L.-W., Kooi, P.S., and Leong, M.-S.: "Resonant Behaviors of Microstrip Antenna in Multilayered Media: An Efficient Full-Wave Analysis", *Progress in Electromagnetics Research*, vol. 31, (2001), 55–67.

[242] Zhao, K., Vouvakis, M.N., Fee, J.-F.: "The Adaptive Cross Approximation Algorithm for Accelerated Method of Moments Computations of EMC Problems", *IEEE Trans. Electromagn. Compat.*, vol. 47, (November 2005), 763–773.

[243] Zhou, P.: *Numerical Analysis of Electromagnetic Fields*, (Springer, Berlin, 1993).

Acknowledgment

In connection with this thesis it is a great pleasure to thank a number of people:

- Prof. Dr. Jürgen Nitsch for providing an outstanding research environment, for his constant advice, support, and interest in science

- Prof. Dr. Günter Wollenberg for many valuable discussions and instructive opinions

- Prof. Dr. Dr. Carl E. Baum for his continuous willingness to share his exceptionally deep scientific knowledge

- Prof. Friedrich W. Hehl for many enlightening discussions on fundamental field theory

- Prof. Dr. Achim Enders for his kind interest in this work

- Dr. Sergey Tkachenko for sharing his profound expertise and for his friendship

- Prof. Nikolay Korovkin and Prof. Elena Solovieva for helpful discussions on non-linear circuit analysis

- my present and former colleagues at the Institute for Fundamental Electrical Engineering, in particular Dr. Heiko Haase, Dr. Uwe Knauff, Dr. Hans Georg Krauthäuser, Dr. Hans-Jürgen Scheibe, and Dr. Torsten Steinmetz, for common research experiences over many years and a pleasant working climate

- the staff at the Institute for Fundamental Electrical Engineering for friendliness and helpfulness

- Justine Bell for english language advice and kind encouragement

- my parents, Wolfgang and Monika Gronwald, for their loving support

Curriculum Vitae

Frank Gronwald

born on July 15$^{\text{th}}$, 1968, in Aachen, Germany

Professional Experience

since 2000 **Otto-von-Guericke-University Magdeburg,**
Faculty of Electrical Engineering and Information Technology
Researcher at Assistant Professor level, pursuing Habilitation degree
Lecturer "Numerical Methods in Electromagnetic Field Computation"

1998 – 2000 **Otto-von-Guericke-University Magdeburg,**
Faculty of Electrical Engineering and Information Technology
Research Assistant, Electromagnetic Compatibility (EMC) and
Theoretical Electrical Engineering

1992 – 1998 **University of Cologne,**
Faculty of Mathematics and Natural Sciences
Research Assistant, Theoretical Physics
PhD in Theoretical Physics, 1996

Education

1987 – 1992 **University of Cologne,**
Faculty of Mathematics and Natural Sciences
Diploma in Physics, 1992

1978 – 1987 **Extended Secondary School "Humboldtgymnasium Köln"**
Graduated "Abitur", 1987

Stay Abroad

1993 – 1994 **University of Texas at Austin, USA, Center for Particle Physics**
Visiting Researcher, Theoretical Physics

Electrical Engineering Awards and Recognition

2006 Best HPE Paper Award - best basic paper in 2004–2005, SUMMA Foundation

2005 Elected Member of the International Union of Radio Science (U.R.S.I)

2005 Senior Member of IEEE

2002 U.R.S.I. Young Scientist Award

2000 Young Scientist Award, ISAP, Fukuoka, Japan

2000 Best HPE Paper Award - best basic paper in 1998–1999, SUMMA Foundation

1999 U.R.S.I. Young Scientist Award