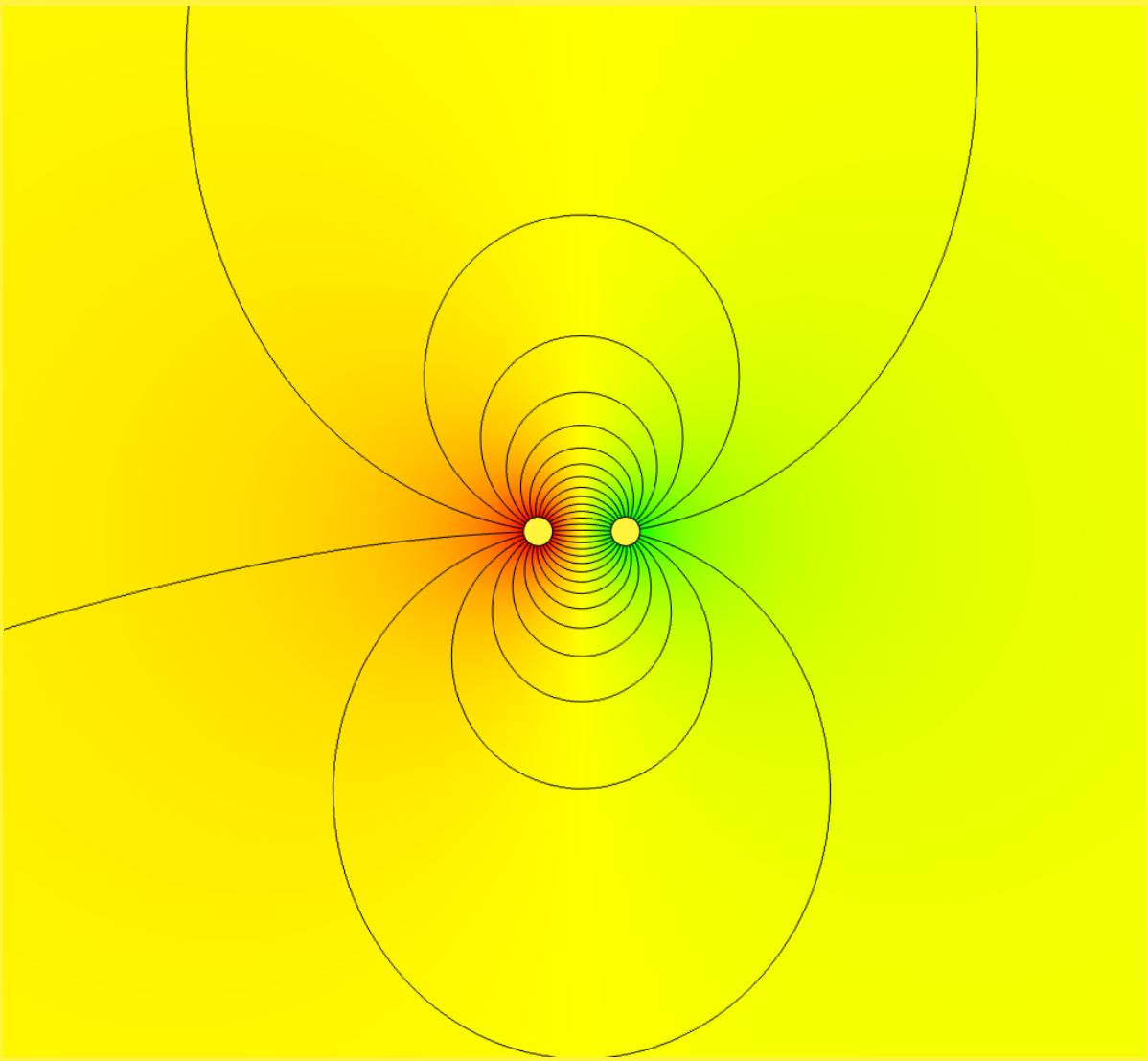
Microwave theory

Karlsson, Anders; Kristensson, Gerhard

2015

ANDERS KARLSSON
and
GERHARD KRISTENSSON

MICROWAVE THEORY



Rules for the ∇-operator

(1) $\nabla(\varphi+\psi)=\nabla\varphi+\nabla\psi$

(2) $\nabla(\varphi\psi)=\psi\nabla\varphi+\varphi\nabla\psi$

(3) $\nabla(\boldsymbol{a}\cdot\boldsymbol{b})=(\boldsymbol{a}\cdot\nabla)\boldsymbol{b}+(\boldsymbol{b}\cdot\nabla)\boldsymbol{a}+\boldsymbol{a}\times(\nabla\times\boldsymbol{b})+\boldsymbol{b}\times(\nabla\times\boldsymbol{a})$

(4) $\nabla(\boldsymbol{a}\cdot\boldsymbol{b})=-\nabla\times(\boldsymbol{a}\times\boldsymbol{b})+2(\boldsymbol{b}\cdot\nabla)\boldsymbol{a}+\boldsymbol{a}\times(\nabla\times\boldsymbol{b})+\boldsymbol{b}\times(\nabla\times\boldsymbol{a})+\boldsymbol{a}(\nabla\cdot\boldsymbol{b})-\boldsymbol{b}(\nabla\cdot\boldsymbol{a})$

(5) $\nabla\cdot(\boldsymbol{a}+\boldsymbol{b})=\nabla\cdot\boldsymbol{a}+\nabla\cdot\boldsymbol{b}$

(6) $\nabla\cdot(\varphi\boldsymbol{a})=\varphi(\nabla\cdot\boldsymbol{a})+(\nabla\varphi)\cdot\boldsymbol{a}$

(7) $\nabla\cdot(\boldsymbol{a}\times\boldsymbol{b})=\boldsymbol{b}\cdot(\nabla\times\boldsymbol{a})-\boldsymbol{a}\cdot(\nabla\times\boldsymbol{b})$

(8) $\nabla\times(\boldsymbol{a}+\boldsymbol{b})=\nabla\times\boldsymbol{a}+\nabla\times\boldsymbol{b}$

(9) $\nabla\times(\varphi\boldsymbol{a})=\varphi(\nabla\times\boldsymbol{a})+(\nabla\varphi)\times\boldsymbol{a}$

(10) $\nabla\times(\boldsymbol{a}\times\boldsymbol{b})=\boldsymbol{a}(\nabla\cdot\boldsymbol{b})-\boldsymbol{b}(\nabla\cdot\boldsymbol{a})+(\boldsymbol{b}\cdot\nabla)\boldsymbol{a}-(\boldsymbol{a}\cdot\nabla)\boldsymbol{b}$

(11) $\nabla\times(\boldsymbol{a}\times\boldsymbol{b})=-\nabla(\boldsymbol{a}\cdot\boldsymbol{b})+2(\boldsymbol{b}\cdot\nabla)\boldsymbol{a}+\boldsymbol{a}\times(\nabla\times\boldsymbol{b})+\boldsymbol{b}\times(\nabla\times\boldsymbol{a})+\boldsymbol{a}(\nabla\cdot\boldsymbol{b})-\boldsymbol{b}(\nabla\cdot\boldsymbol{a})$

(12) $\nabla\cdot\nabla\varphi=\nabla^2\varphi=\Delta\varphi$

(13) $\nabla\times(\nabla\times\boldsymbol{a})=\nabla(\nabla\cdot\boldsymbol{a})-\nabla^2\boldsymbol{a}$

(14) $\nabla\times(\nabla\varphi)=\boldsymbol{0}$

(15) $\nabla\cdot(\nabla\times\boldsymbol{a})=0$

(16) $\nabla^2(\varphi\psi)=\varphi\nabla^2\psi+\psi\nabla^2\varphi+2\nabla\varphi\cdot\nabla\psi$

(17) $\nabla r=\hat{\boldsymbol{r}}$

(18) $\nabla\times\boldsymbol{r}=\boldsymbol{0}$

(19) $\nabla\times\hat{\boldsymbol{r}}=\boldsymbol{0}$

(20) $\nabla\cdot\boldsymbol{r}=3$

(21) $\nabla\cdot\hat{\boldsymbol{r}}=\dfrac{2}{r}$

(22) $\nabla(\boldsymbol{a}\cdot\boldsymbol{r})=\boldsymbol{a},\quad \boldsymbol{a}$ constant vector

(23) $(\boldsymbol{a}\cdot\nabla)\boldsymbol{r}=\boldsymbol{a}$

(24) $(\boldsymbol{a}\cdot\nabla)\hat{\boldsymbol{r}}=\dfrac{1}{r}\left(\boldsymbol{a}-\hat{\boldsymbol{r}}(\boldsymbol{a}\cdot\hat{\boldsymbol{r}})\right)=\dfrac{\boldsymbol{a}_\perp}{r}$

(25) $\nabla^2(\boldsymbol{r}\cdot\boldsymbol{a})=2\nabla\cdot\boldsymbol{a}+\boldsymbol{r}\cdot(\nabla^2\boldsymbol{a})$

(26) $\nabla u(f)=(\nabla f)\dfrac{du}{df}$

(27) $\nabla\cdot\boldsymbol{F}(f)=(\nabla f)\cdot\dfrac{d\boldsymbol{F}}{df}$

(28) $\nabla\times\boldsymbol{F}(f)=(\nabla f)\times\dfrac{d\boldsymbol{F}}{df}$

(29) $\nabla=\hat{\boldsymbol{r}}(\hat{\boldsymbol{r}}\cdot\nabla)-\hat{\boldsymbol{r}}\times(\hat{\boldsymbol{r}}\times\nabla)$

Important vector identities

(1) $(\boldsymbol{a}\times\boldsymbol{c})\times(\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{c}\,((\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c})$

(2) $(\boldsymbol{a}\times\boldsymbol{b})\cdot(\boldsymbol{c}\times\boldsymbol{d}) = (\boldsymbol{a}\cdot\boldsymbol{c})(\boldsymbol{b}\cdot\boldsymbol{d}) - (\boldsymbol{a}\cdot\boldsymbol{d})(\boldsymbol{b}\cdot\boldsymbol{c})$

(3) $\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{a}\cdot\boldsymbol{c}) - \boldsymbol{c}(\boldsymbol{a}\cdot\boldsymbol{b})$

(4) $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{b}\cdot(\boldsymbol{c}\times\boldsymbol{a}) = \boldsymbol{c}\cdot(\boldsymbol{a}\times\boldsymbol{b})$

Integration formulas

Stoke's theorem and related theorems

(1) $$\iint_S (\nabla\times\boldsymbol{A})\cdot\hat{\boldsymbol{n}}\,\mathrm{d}S = \int_C \boldsymbol{A}\cdot\,d\boldsymbol{r}$$

(2) $$\iint_S \hat{\boldsymbol{n}}\times\nabla\varphi\,\mathrm{d}S = \int_C \varphi\,d\boldsymbol{r}$$

(3) $$\iint_S (\hat{\boldsymbol{n}}\times\nabla)\times\boldsymbol{A}\,\mathrm{d}S = \int_C d\boldsymbol{r}\times\boldsymbol{A}$$

Gauss' theorem (divergence theorem) and related theorems

(1) $$\iiint_V \nabla\cdot\boldsymbol{A}\,\mathrm{d}v = \iint_S \boldsymbol{A}\cdot\hat{\boldsymbol{n}}\,\mathrm{d}S$$

(2) $$\iiint_V \nabla\varphi\,\mathrm{d}v = \iint_S \varphi\hat{\boldsymbol{n}}\,\mathrm{d}S$$

(3) $$\iiint_V \nabla\times\boldsymbol{A}\,\mathrm{d}v = \iint_S \hat{\boldsymbol{n}}\times\boldsymbol{A}\,\mathrm{d}S$$

Green's formulas

(1) $$\iiint_V (\psi\nabla^2\varphi - \varphi\nabla^2\psi)\,\mathrm{d}v = \iint_S (\psi\nabla\varphi - \varphi\nabla\psi)\cdot\hat{\boldsymbol{n}}\,\mathrm{d}S$$

(2) $$\iiint_V (\psi\nabla^2\boldsymbol{A} - \boldsymbol{A}\nabla^2\psi)\,\mathrm{d}v$$
$$= \iint_S (\nabla\psi\times(\hat{\boldsymbol{n}}\times\boldsymbol{A}) - \nabla\psi(\hat{\boldsymbol{n}}\cdot\boldsymbol{A}) - \psi(\hat{\boldsymbol{n}}\times(\nabla\times\boldsymbol{A})) + \hat{\boldsymbol{n}}\psi(\nabla\cdot\boldsymbol{A}))\,\mathrm{d}S$$

Microwave theory

Anders Karlsson and Gerhard Kristensson

Lund, 16 February 2016

Contents

Preface

The book is about wave propagation along guiding structures, eg., transmission lines, hollow waveguides and optical fibers. There are numerous applications for these structures. Optical fiber systems are crucial for internet and many communication systems. Although transmission lines are replaced by optical fibers optical systems and wireless systems in telecommunication, they are still very important at short distance communication, in measurement equipment, and in high frequency circuits. The hollow waveguides are used in radars and instruments for very high frequencies. They are also important in particle accelerators where they transfer microwaves at high power. We devote one chapter in the book to the electromagnetic fields that can exist in cavities with metallic walls. Such cavities are vital for modern particle accelerators. The cavities are placed along the pipe where the particles travel. As a bunch of particles enters the cavity it is accelerated by the electric field in the cavity.

The electromagnetic fields in waveguides and cavities are described by Maxwell's equations. These equations constitute a system of partial differential equations (PDE). For a number of important geometries the equations can be solved analytically. In the book the analytic solutions for the most important geometries are derived by utilizing the method of separation of variables. For more complicated waveguide and cavity geometries we determine the electromagnetic fields by numerical methods. There are a number of commercial software packages that are suitable for such evaluations. We chose to refer to COMSOL Multiphysics, which is based on the finite element method (FEM), in many of our examples. The commercial software packages are very advanced and can solve Maxwell's equations in most geometries. However, it is vital to understand the analytical solutions of the simple geometries in order to evaluate and understand the numerical solutions of more complicated geometries.

The book requires basic knowledge in vector analysis, electromagnetic theory and circuit theory. The nabla operator is frequently used in order to obtain results that are coordinate independent.

Every chapter is concluded with a problem section. The more advanced problems are marked with an asterisk ($*$). At the end of the book there are answers to most of the problems.

Chapter 1

The Maxwell equations

The Maxwell equations constitute the fundamental mathematical model for all theoretical analysis of macroscopic electromagnetic phenomena. James Clerk Maxwell[1] published his famous equations in 1864. An impressive amount of evidences for the validity of these equations have been gathered in different fields of applications. Microscopic phenomena require a more refined model including also quantum effects, but these effects are out of the scope of this book.

The Maxwell equations are the cornerstone in the analysis of macroscopic electromagnetic wave propagation phenomena.[2] In SI-units (MKSA) they read

$$\nabla \times \boldsymbol{E}(\boldsymbol{r},t) = -\frac{\partial \boldsymbol{B}(\boldsymbol{r},t)}{\partial t} \tag{1.1}$$

$$\nabla \times \boldsymbol{H}(\boldsymbol{r},t) = \boldsymbol{J}(\boldsymbol{r},t) + \frac{\partial \boldsymbol{D}(\boldsymbol{r},t)}{\partial t} \tag{1.2}$$

The equation (1.1) (or the corresponding integral formulation) is the *Faraday's law of induction* [3], and the equation (1.2) is the *Ampère's (generalized) law.*[4] The vector fields in the Maxwell equations are[5]:

$\boldsymbol{E}(\boldsymbol{r},t)$	Electric field [V/m]
$\boldsymbol{H}(\boldsymbol{r},t)$	Magnetic field [A/m]
$\boldsymbol{D}(\boldsymbol{r},t)$	Electric flux density [As/m^2]
$\boldsymbol{B}(\boldsymbol{r},t)$	Magnetic flux density [Vs/m^2]
$\boldsymbol{J}(\boldsymbol{r},t)$	Current density [A/m^2]

All of these fields are functions of the space coordinates $\boldsymbol{r}$ and time t. We often suppress these arguments for notational reasons. Only when equations or expressions can be misinterpreted we give the argument.

[1] James Clerk Maxwell (1831–1879), Scottish physicist and mathematician.

[2] A detailed derivation of these macroscopic equations from a microscopic formulation is found in [8, 16].

[3] Michael Faraday (1791–1867), English chemist and physicist.

[4] André Marie Ampère (1775–1836), French physicist.

[5] For simplicity we sometimes use the names $\boldsymbol{E}$-field, $\boldsymbol{D}$-field, $\boldsymbol{B}$-field and $\boldsymbol{H}$-field.

The electric field $\boldsymbol{E}$ and the magnetic flux density $\boldsymbol{B}$ are defined by the force on a charged particle

$$\boldsymbol{F} = q\,(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \tag{1.3}$$

where q is the electric charge of the particle and $\boldsymbol{v}$ its velocity. The relation is called the *Lorentz' force*.

The motion of the free charges in materials, eg., the conduction electrons, is described by the current density $\boldsymbol{J}$. The current contributions from all bounded charges, eg., the electrons bound to the nucleus of the atom, are included in the time derivative of the electric flux density $\dfrac{\partial \boldsymbol{D}}{\partial t}$. In Chapter 2 we address the differences between the electric flux density $\boldsymbol{D}$ and the electric field $\boldsymbol{E}$, as well as the differences between the magnetic field $\boldsymbol{H}$ and the magnetic flux density $\boldsymbol{B}$.

One of the fundamental assumptions in physics is that electric charges are indestructible, i.e., the sum of the charges is always constant. The conservation of charges is expressed in mathematical terms by the continuity equation

$$\nabla \cdot \boldsymbol{J}(\boldsymbol{r}, t) + \frac{\partial \rho(\boldsymbol{r}, t)}{\partial t} = 0 \tag{1.4}$$

Here $\rho(\boldsymbol{r}, t)$ is the charge density (charge/unit volume) that is associated with the current density $\boldsymbol{J}(\boldsymbol{r}, t)$. The charge density ρ models the distribution of free charges. As alluded to above, the bounded charges are included in the electric flux density $\boldsymbol{D}$ and the magnetic field $\boldsymbol{H}$.

Two additional equations are usually associated with the Maxwell equations.

$$\nabla \cdot \boldsymbol{B} = 0 \tag{1.5}$$

$$\nabla \cdot \boldsymbol{D} = \rho \tag{1.6}$$

The equation (1.5) tells us that there are no magnetic charges and that the magnetic flux is conserved. The equation (1.6) is usually called *Gauss law*. Under suitable assumptions, both of these equations can be derived from (1.1), (1.2) and (1.4). To see this, we take the divergence of (1.1) and (1.2). This implies

$$\begin{cases} \nabla \cdot \dfrac{\partial \boldsymbol{B}}{\partial t} = 0 \\ \nabla \cdot \boldsymbol{J} + \nabla \cdot \dfrac{\partial \boldsymbol{D}}{\partial t} = 0 \end{cases}$$

since $\nabla \cdot (\nabla \times \boldsymbol{A}) \equiv 0$. We interchange the order of differentiation and use (1.4) and get

$$\begin{cases} \dfrac{\partial (\nabla \cdot \boldsymbol{B})}{\partial t} = 0 \\ \dfrac{\partial (\nabla \cdot \boldsymbol{D} - \rho)}{\partial t} = 0 \end{cases}$$

These equations imply

$$\begin{cases} \nabla \cdot \boldsymbol{B} = f_1 \\ \nabla \cdot \boldsymbol{D} - \rho = f_2 \end{cases}$$

where f_1 and f_2 are two functions that do not explicitly depend on time t (they can depend on the spatial coordinates $\boldsymbol{r}$). If the fields $\boldsymbol{B}$, $\boldsymbol{D}$ and ρ are identically zero before a fixed, finite time, i.e.,

$$\begin{cases} \boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{0} \\ \boldsymbol{D}(\boldsymbol{r},t) = \boldsymbol{0} \qquad t < \tau \\ \rho(\boldsymbol{r},t) = 0 \end{cases} \tag{1.7}$$

then (1.5) and (1.6) follow. Static or time-harmonic fields do not satisfy these initial conditions, since there is no finite time τ before the fields are all zero.[6] We assume that (1.7) is valid for time-dependent fields and then it is sufficient to use the equations (1.1), (1.2) and (1.4).

The vector equations (1.1) and (1.2) contain six different equations—one for each vector component. Provided the current density $\boldsymbol{J}$ is given, the Maxwell equations contain 12 unknowns—the four vector fields $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{D}$ and $\boldsymbol{H}$. We lack six equations in order to have as many equations as unknowns. The lacking six equations are called the *constitutive relations* and they are addressed in the next Chapter.

In vacuum $\boldsymbol{E}$ is parallel with $\boldsymbol{D}$ and $\boldsymbol{B}$ is parallel with $\boldsymbol{H}$, such that

$$\begin{cases} \boldsymbol{D} = \epsilon_0 \boldsymbol{E} \\ \boldsymbol{B} = \mu_0 \boldsymbol{H} \end{cases} \tag{1.8}$$

where ϵ_0 and μ_0 are the permittivity and the permeability of vacuum. The numerical values of these constants are: $\epsilon_0 \approx 8.854 \cdot 10^{-12}$ As/Vm and $\mu_0 = 4\pi \cdot 10^{-7}$ Vs/Am $\approx 1.257 \cdot 10^{-6}$ Vs/Am.

Inside a material there is a difference between the field $\epsilon_0 \boldsymbol{E}$ and the electric flux density $\boldsymbol{D}$, and between the magnetic flux density $\boldsymbol{B}$ and the field $\mu_0 \boldsymbol{H}$. These differences are a measure of the interaction between the charges in the material and the fields. The differences between these fields are described by the polarization $\boldsymbol{P}$, and the magnetization $\boldsymbol{M}$. The definitions of these fields are

$$\boldsymbol{P} = \boldsymbol{D} - \epsilon_0 \boldsymbol{E} \tag{1.9}$$

$$\boldsymbol{M} = \frac{1}{\mu_0} \boldsymbol{B} - \boldsymbol{H} \tag{1.10}$$

The polarization $\boldsymbol{P}$ is the volume density of electric dipole moment, and hence a measure of the relative separation of the positive and negative bounded charges in the material. It includes both *permanent* and *induced polarization*. In an analogous manner, the magnetization $\boldsymbol{M}$ is the volume density of magnetic dipole moment and hence a measure of the net (bounded) currents in the material. The origin of $\boldsymbol{M}$ can also be both *permanent* or *induced*.

The polarization and the magnetization effects of the material are governed by the constitutive relations of the material. The constitutive relations constitute the six missing equations.

[6] We will return to the derivation of equations (1.5) and (1.6) for time-harmonic fields in Chapter 2 on Page 15.

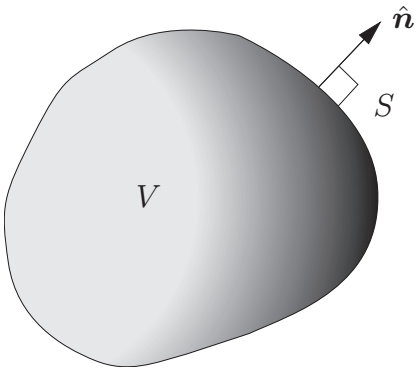


Figure 1.1: Geometry of integration.

1.1 Boundary conditions at interfaces

At an interface between two different materials some components of the electromagnetic field are discontinuous. In this section we give a derivation of these boundary conditions. Only surfaces that are fixed in time (no moving surfaces) are treated.

The Maxwell equations, as they are presented in equations (1.1)–(1.2), assume that all field quantities are differentiable functions of space and time. At an interface between two media, the fields, as already mentioned above, are discontinuous functions of the spatial variables. Therefore, we need to reformulate the Maxwell equations such that they are also valid for fields that are not differentiable at all points in space.

We let V be an arbitrary (simply connected) volume, bounded by the surface S with unit outward normal vector $\hat{\boldsymbol{n}}$, see Figure 1.1. We integrate the Maxwell equations, (1.1)–(1.2) and (1.5)–(1.6), over the volume V and obtain

$$\begin{aligned}
\iiint_V \nabla \times \boldsymbol{E}\,\mathrm{d}v &= -\iiint_V \frac{\partial \boldsymbol{B}}{\partial t}\,\mathrm{d}v \\
\iiint_V \nabla \times \boldsymbol{H}\,\mathrm{d}v &= \iiint_V \boldsymbol{J}\,\mathrm{d}v + \iiint_V \frac{\partial \boldsymbol{D}}{\partial t}\,\mathrm{d}v \\
\iiint_V \nabla \cdot \boldsymbol{B}\,\mathrm{d}v &= 0 \\
\iiint_V \nabla \cdot \boldsymbol{D}\,\mathrm{d}v &= \iiint_V \rho\,\mathrm{d}v
\end{aligned} \tag{1.11}$$

where $\mathrm{d}v$ is the volume measure ($\mathrm{d}v = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$).

The following two integration theorems for vector fields are useful:

$$\iiint_V \nabla \cdot \boldsymbol{A} \, \mathrm{d}v = \iint_S \boldsymbol{A} \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S$$
$$\iiint_V \nabla \times \boldsymbol{A} \, \mathrm{d}v = \iint_S \hat{\boldsymbol{n}} \times \boldsymbol{A} \, \mathrm{d}S$$

Here $\boldsymbol{A}$ is a continuously differentiable vector field in V, and $\mathrm{d}S$ is the surface element of S. The first theorem is usually called the *divergence theorem* or the *Gauss theorem*[7] and the other *Gauss analogous theorem* (see Problem 1.1).

After interchanging the differentiation w.r.t. time t and integration (volume V is fixed in time and we assume all field to be sufficiently regular) (1.11) reads

$$\iint_S \hat{\boldsymbol{n}} \times \boldsymbol{E} \, \mathrm{d}S = -\frac{d}{dt} \iiint_V \boldsymbol{B} \, \mathrm{d}v \tag{1.12}$$

$$\iint_S \hat{\boldsymbol{n}} \times \boldsymbol{H} \, \mathrm{d}S = \iiint_V \boldsymbol{J} \, \mathrm{d}v + \frac{d}{dt} \iiint_V \boldsymbol{D} \, \mathrm{d}v \tag{1.13}$$

$$\iint_S \boldsymbol{B} \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S = 0 \tag{1.14}$$

$$\iint_S \boldsymbol{D} \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S = \iiint_V \rho \, \mathrm{d}v \tag{1.15}$$

In a domain V where the fields $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{D}$ and $\boldsymbol{H}$ are continuously differentiable, these integral expressions are equivalent to the differential equations (1.1) and (1.6). We have proved this equivalence one way and in the other direction we do the analysis in a reversed direction and use the fact that the volume V is arbitrary.

The integral formulation, (1.12)–(1.15), has the advantage that the fields do not have to be differentiable in the spatial variables to make sense. In this respect, the integral formulation is more general than the differential formulation in (1.1)–(1.2). The fields $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{D}$ and $\boldsymbol{H}$, that satisfy the equations (1.12)–(1.15) are called weak solutions to the Maxwell equations, in the case the fields are not continuously differentiable and (1.1)–(1.2) lack meaning.

The integral expressions (1.12)–(1.15) are applied to a volume V_h that cuts the interface between two different media, see Figure 1.2. The unit normal $\hat{\boldsymbol{n}}$ of the interface S is directed from medium 2 into medium 1. We assume that all electromagnetic fields $\boldsymbol{E}$, $\boldsymbol{B}$, $\boldsymbol{D}$ and $\boldsymbol{H}$, and their time derivatives, have finite values in the limit from both sides of the interface. For the electric field, these limit values in medium 1 and 2 are denoted $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$, respectively, and a similar notation, with index 1 or 2, is adopted for the other fields. The current density $\boldsymbol{J}$ and the charge density ρ can adopt infinite values at the interface for perfectly conducting (metal)

[7] Distinguish between the Gauss law, (1.6), and the Gauss theorem.

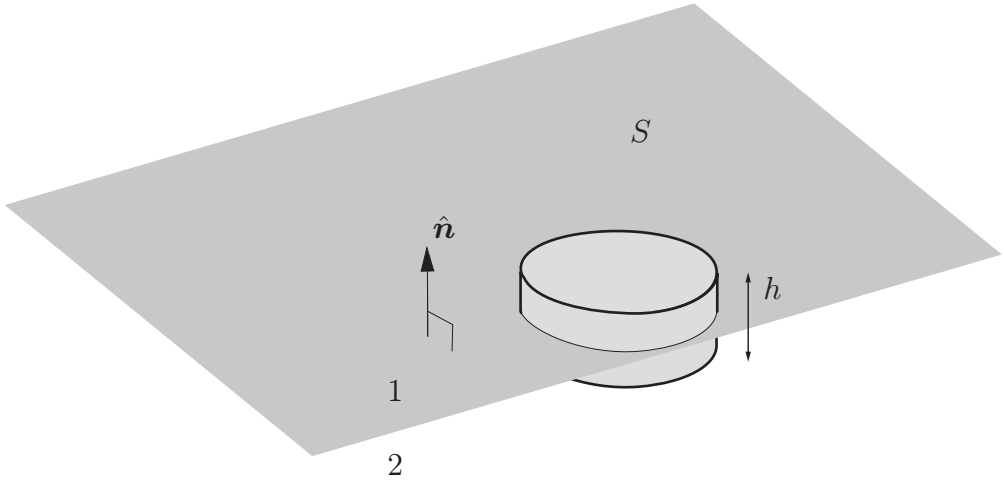


Figure 1.2: Interface between two different media 1 and 2.

surfaces.[8] It is convenient to introduce a *surface current density* $\boldsymbol{J}_S$ and *surface charge density* ρ_S as a limit process

$$\begin{cases} \boldsymbol{J}_S = h\boldsymbol{J} \\ \rho_S = h\rho \end{cases}$$

where h is the thickness of the layer that contains the charges close to the surface. We assume that this thickness approaches zero and that $\boldsymbol{J}$ and ρ go to infinity in such a way that $\boldsymbol{J}_S$ and ρ_S have well defined values in this process. The surface current density $\boldsymbol{J}_S$ is assumed to be a tangential field to the surface S. We let the height of the volume V_h be h and the area on the upper and lower part of the bounding surface of V_h be a, which is small compared to the curvature of the surface S and small enough such that the fields are approximately constant over a.

The terms $\frac{d}{dt}\iiint_{V_h} \boldsymbol{B}\,\mathrm{d}v$ and $\frac{d}{dt}\iiint_{V_h} \boldsymbol{D}\,\mathrm{d}v$ approach zero as $h \to 0$, since the fields $\boldsymbol{B}$ and $\boldsymbol{D}$ and their time derivatives are assumed to be finite at the interface. Moreover, the contributions from all side areas (area $\sim h$) of the surface integrals in (1.12)–(1.15) approach zero as $h \to 0$. The contribution from the upper part (unit normal $\hat{\boldsymbol{n}}$) and lower part (unit normal $-\hat{\boldsymbol{n}}$) are proportional to the area a, if the area is chosen sufficiently small and the mean value theorem for integrals are used. The contributions from the upper and the lower parts of the surface integrals in the limit $h \to 0$ are

$$\begin{aligned} a\,[\hat{\boldsymbol{n}} \times (\boldsymbol{E}_1 - \boldsymbol{E}_2)] &= \boldsymbol{0} \\ a\,[\hat{\boldsymbol{n}} \times (\boldsymbol{H}_1 - \boldsymbol{H}_2)] &= ah\boldsymbol{J} = a\boldsymbol{J}_S \\ a\,[\hat{\boldsymbol{n}} \cdot (\boldsymbol{B}_1 - \boldsymbol{B}_2)] &= 0 \\ a\,[\hat{\boldsymbol{n}} \cdot (\boldsymbol{D}_1 - \boldsymbol{D}_2)] &= ah\rho = a\rho_S \end{aligned}$$

[8] This is of course an idealization of a situation where the density assumes very high values in a macroscopically thin layer.

We simplify these expressions by dividing with the area a. The result is

$$\begin{cases} \hat{\boldsymbol{n}} \times (\boldsymbol{E}_1 - \boldsymbol{E}_2) = \boldsymbol{0} \\ \hat{\boldsymbol{n}} \times (\boldsymbol{H}_1 - \boldsymbol{H}_2) = \boldsymbol{J}_S \\ \hat{\boldsymbol{n}} \cdot (\boldsymbol{B}_1 - \boldsymbol{B}_2) = 0 \\ \hat{\boldsymbol{n}} \cdot (\boldsymbol{D}_1 - \boldsymbol{D}_2) = \rho_S \end{cases} \tag{1.16}$$

These boundary conditions prescribe how the electromagnetic fields on each side of the interface are related to each other(unit normal $\hat{\boldsymbol{n}}$ is directed from medium 2 into medium 1). We formulate these boundary conditions in words.

- The tangential components of the electric field are continuous across the interface.
- The tangential components of the magnetic field are discontinuous over the interface. The size of the discontinuity is $\boldsymbol{J}_S$. If the surface current density is zero, eg., when the material has finite conductivity[9], the tangential components of the magnetic field are continuous across the interface.
- The normal component of the magnetic flux density is continuous across the interface.
- The normal component of the electric flux density is discontinuous across the interface. The size of the discontinuity is ρ_S. If the surface charge density is zero, the normal component of the electric flux density is continuous across the interface.

In Figure 1.3 we illustrate the typical variations in the normal components of the electric and the magnetic flux densities as a function of the distance across the interface between two media.

A special case of (1.16) is the case where medium 2 is a perfectly conducting material, which often is a good model for metals and other materials with high conductivity. In material 2 the fields are zero and we get from (1.16)

$$\begin{cases} \hat{\boldsymbol{n}} \times \boldsymbol{E}_1 = \boldsymbol{0} \\ \hat{\boldsymbol{n}} \times \boldsymbol{H}_1 = \boldsymbol{J}_S \\ \hat{\boldsymbol{n}} \cdot \boldsymbol{B}_1 = 0 \\ \hat{\boldsymbol{n}} \cdot \boldsymbol{D}_1 = \rho_S \end{cases} \tag{1.17}$$

where $\boldsymbol{J}_S$ and ρ_S are the surface current density and surface charge density, respectively.

[9]This is an implication of the assumption that the electric field $\boldsymbol{E}$ is finite close to the interface. We have $\boldsymbol{J}_S = h\boldsymbol{J} = h\sigma\boldsymbol{E} \to \boldsymbol{0}$, as $h \to 0$.

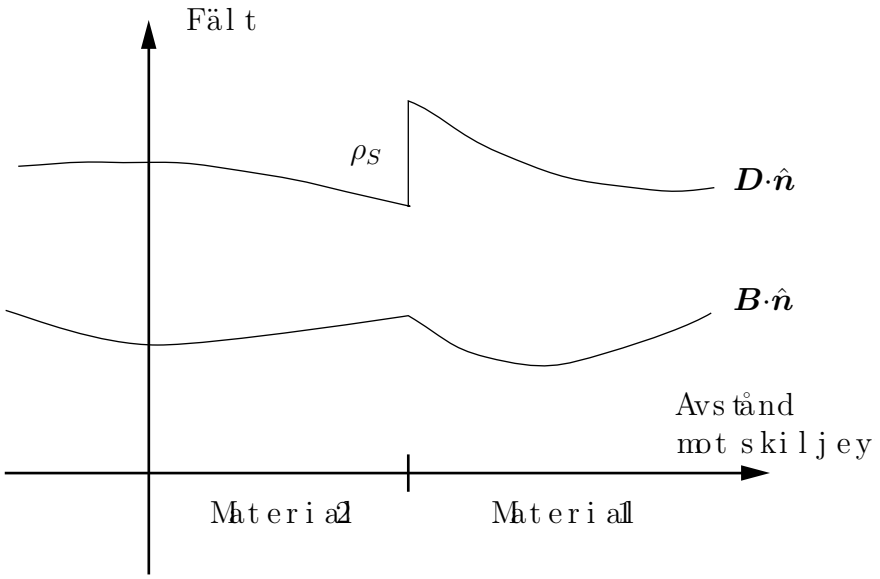


Figure 1.3: The variation of the normal components B_n and D_n at the interface between two different media.

1.1.1 Impedance boundary conditions

At an interface between a non-conducting medium and a metal, the boundary condition in (1.17) is often a good enough approximation. When there is a need for more accurate evaluations there are two ways to go. We can treat the two media as two regions and simply use the exact boundary conditions in (1.16). A disadvantage is that we have to solve for the electric and magnetic field in both regions. If we use FEM both regions have to be discretized. The wavelength in a conductor is considerably much smaller than the wavelength in free space, c.f., section 5.9. Since the mesh size is proportional to the wavelength a much finer mesh is needed in the metal than in the non-conducting region and that increases the computational time and required memory. There is a third alternative and that is to use an impedance boundary condition. This condition is derived in section 5.9. We let $\boldsymbol{E}$ and $\boldsymbol{H}$ be the electric and magnetic fields at the surface, but in the non-conducting region, and $\hat{\boldsymbol{n}}$ the normal unit vector directed out from the metal. Then the condition is

$$\begin{aligned} &\boldsymbol{E} - \hat{\boldsymbol{n}}(\boldsymbol{E}\cdot\hat{\boldsymbol{n}}) = -\eta_s \hat{\boldsymbol{n}}\times\boldsymbol{H} \\ &\eta_s = (1-\mathrm{i})\sqrt{\frac{\omega\mu_0}{2\sigma}} = \frac{1-\mathrm{i}}{\sigma\delta} \end{aligned} \tag{1.18}$$

Here η_s is the wave impedance of the metal, and $\delta = \sqrt{2/(\omega\mu_0\sigma)}$ the skin depth of the metal, c.f., section 5.9. Notice that $\boldsymbol{E} - \hat{\boldsymbol{n}}(\boldsymbol{E}\cdot\hat{\boldsymbol{n}})$ is the tangential component of the electric field.

Most commercial simulation programs, like COMSOL Multiphysics, have the impedance boundary condition as an option.

1.2 Energy conservation and Poynting's theorem

Energy conservation is shown from the Maxwell equations (1.1) and (1.2).

$$\begin{cases} \nabla \times \boldsymbol{E} = -\dfrac{\partial \boldsymbol{B}}{\partial t} \\ \nabla \times \boldsymbol{H} = \boldsymbol{J} + \dfrac{\partial \boldsymbol{D}}{\partial t} \end{cases}$$

We make a scalar multiplication of the first equation with $\boldsymbol{H}$ and the second with $\boldsymbol{E}$ and subtract. The result is

$$\boldsymbol{H} \cdot (\nabla \times \boldsymbol{E}) - \boldsymbol{E} \cdot (\nabla \times \boldsymbol{H}) + \boldsymbol{H} \cdot \frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{E} \cdot \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{E} \cdot \boldsymbol{J} = 0$$

By using the differential rule $\nabla \cdot (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{b} \cdot (\nabla \times \boldsymbol{a}) - \boldsymbol{a} \cdot (\nabla \times \boldsymbol{b})$ we obtain *Poynting's theorem*.

$$\nabla \cdot \boldsymbol{S} + \boldsymbol{H} \cdot \frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{E} \cdot \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{E} \cdot \boldsymbol{J} = 0 \tag{1.19}$$

We have here introduced the *Poynting's vector*,[10] $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$, which gives the power flow per unit area in the direction of the vector $\boldsymbol{S}$. The energy conservation is made visible if we integrate equation (1.19) over a volume V, bounded by the surface S with unit outward normal vector $\hat{\boldsymbol{n}}$, see Figure 1.1, and use the divergence theorem. We get

$$\begin{aligned} \iint_S \boldsymbol{S} \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S &= \iiint_V \nabla \cdot \boldsymbol{S} \, \mathrm{d}v \\ &= -\iiint_V \left[\boldsymbol{H} \cdot \frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{E} \cdot \frac{\partial \boldsymbol{D}}{\partial t} \right] \mathrm{d}v - \iiint_V \boldsymbol{E} \cdot \boldsymbol{J} \, \mathrm{d}v \end{aligned} \tag{1.20}$$

The terms are interpreted in the following way:

- The left hand side:
$$\iint_S \boldsymbol{S} \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S$$
This is the total power radiated out of the bounding surface S.

- The right hand side: The power flow through the surface S is compensated by two different contributions. The first volume integral on the right hand side
$$\iiint_V \left[\boldsymbol{H} \cdot \frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{E} \cdot \frac{\partial \boldsymbol{D}}{\partial t} \right] \mathrm{d}v$$

[10] John Henry Poynting (1852–1914), English physicist.

gives the power bounded in the electromagnetic field in the volume V. This includes the power needed to polarize and magnetize the material in V. The second volume integral in (1.20)

$$\iiint_V \boldsymbol{E} \cdot \boldsymbol{J} \, \mathrm{d}v$$

gives the work per unit time, i.e., the power, that the electric field does on the charges in V.

To this end, (1.20) expresses energy balance or more correctly power balance in the volume V, i.e.,

$$\begin{aligned} &\text{Through } S \text{ radiated power + power consumption in } V \\ &\qquad = - \text{ power bounded to the electromagnetic field in } V \end{aligned}$$

In the derivation above we assumed that the volume V does not cut any surface where the fields are discontinuous, eg., an interface between two media. We now prove that this assumption is no severe restriction and the assumption can easily be relaxed. If the surface S is an interface between two media, see Figure 1.2, Poynting's vector in medium 1 close to the interface is

$$\boldsymbol{S}_1 = \boldsymbol{E}_1 \times \boldsymbol{H}_1$$

and Poynting's vector close to the interface in medium 2 is

$$\boldsymbol{S}_2 = \boldsymbol{E}_2 \times \boldsymbol{H}_2$$

The boundary condition at the interface is given by (1.16).

$$\begin{aligned} \hat{\boldsymbol{n}} \times \boldsymbol{E}_1 &= \hat{\boldsymbol{n}} \times \boldsymbol{E}_2 \\ \hat{\boldsymbol{n}} \times \boldsymbol{H}_1 &= \hat{\boldsymbol{n}} \times \boldsymbol{H}_2 + \boldsymbol{J}_S \end{aligned}$$

We now prove that the power transported by the electromagnetic field is continuous across the interface. Stated differently, we prove

$$\iint_S \boldsymbol{S}_1 \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S = \iint_S \boldsymbol{S}_2 \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S - \iint_S \boldsymbol{E}_2 \cdot \boldsymbol{J}_S \, \mathrm{d}S \tag{1.21}$$

where the surface S is an arbitrary part of the interface. Note that the unit normal $\hat{\boldsymbol{n}}$ is directed from medium 2 into medium 1. The last surface integral gives the work per unit time done by the electric field on the charges at the interface. If there are no surface currents at the interface the normal component of Poynting's vector is continuous across the interface. It is irrelevant which electric field we use in the last surface integral in (1.21) since the surface current density $\boldsymbol{J}_S$ is parallel to the interface S and the tangential components of the electric field are continuous across the interface, i.e.,

$$\iint_S \boldsymbol{E}_1 \cdot \boldsymbol{J}_S \, \mathrm{d}S = \iint_S \boldsymbol{E}_2 \cdot \boldsymbol{J}_S \, \mathrm{d}S$$

Equation (1.21) is easily proved by a cyclic permutation of the vectors and the use of the boundary conditions.

$$\begin{aligned}\hat{\boldsymbol{n}} \cdot \boldsymbol{S}_1 &= \hat{\boldsymbol{n}} \cdot (\boldsymbol{E}_1 \times \boldsymbol{H}_1) = \boldsymbol{H}_1 \cdot (\hat{\boldsymbol{n}} \times \boldsymbol{E}_1) = \boldsymbol{H}_1 \cdot (\hat{\boldsymbol{n}} \times \boldsymbol{E}_2) \\ &= -\boldsymbol{E}_2 \cdot (\hat{\boldsymbol{n}} \times \boldsymbol{H}_1) = -\boldsymbol{E}_2 \cdot (\hat{\boldsymbol{n}} \times \boldsymbol{H}_2 + \boldsymbol{J}_S) \\ &= \hat{\boldsymbol{n}} \cdot (\boldsymbol{E}_2 \times \boldsymbol{H}_2) - \boldsymbol{E}_2 \cdot \boldsymbol{J}_S = \hat{\boldsymbol{n}} \cdot \boldsymbol{S}_2 - \boldsymbol{E}_2 \cdot \boldsymbol{J}_S\end{aligned}$$

By integration of this expression over the interface S we obtain power conservation over the surface S as expressed in equation (1.21).

Problems in Chapter 1

1.1 Show the following analogous theorem of Gauss theorem:

$$\iiint_V \nabla \times \boldsymbol{A} \, \mathrm{d}v = \iint_S \hat{\boldsymbol{n}} \times \boldsymbol{A} \, \mathrm{d}S$$

Apply the theorem of divergence (Gauss theorem) to the vector field $\boldsymbol{B} = \boldsymbol{A} \times \boldsymbol{a}$, where $\boldsymbol{a}$ is an arbitrary constant vector.

1.2 A finite volume contains a magnetic material with magnetization $\boldsymbol{M}$. In the absence of current density (free charges), $\boldsymbol{J} = \boldsymbol{0}$, show that the static magnetic field, $\boldsymbol{H}$, and the magnetic flux density, $\boldsymbol{B}$, satisfy

$$\iiint \boldsymbol{B} \cdot \boldsymbol{H} \, \mathrm{d}v = 0$$

where the integration is over all space.

Ampère's law $\nabla \times \boldsymbol{H} = \boldsymbol{0}$ implies that there exists a potential Φ such that

$$\boldsymbol{H} = -\nabla \Phi$$

Use the divergence theorem to prove the problem.

1.3 An infinitely long, straight conductor of circular cross section (radius a) consists of a material with finite conductivity σ. In the conductor a static current I is flowing. The current density $\boldsymbol{J}$ is assumed to be homogeneous over the cross section of the conductor. Compute the terms in Poynting's theorem and show that power balance holds for a volume V, which consists of a finite portion l of the conductor.

On the surface of the conductor we have $\boldsymbol{S} = -\hat{\boldsymbol{\rho}} \frac{1}{2} a \sigma E^2$ where the electric field on the surface of the conductor is related to the current by $I = \pi a^2 \sigma E$. The terms in Poynting's theorem are

$$\iint_S \boldsymbol{S} \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S = -\pi a^2 l \sigma E^2$$

$$\iiint_V \boldsymbol{E} \cdot \boldsymbol{J} \, \mathrm{d}v = \pi a^2 l \sigma E^2$$

Chapter 2

Time harmonic fields and Fourier transform

Time harmonic fields are common in many applications. We obtain the time harmonic formulation from the general results in the previous section by a Fourier transform in the time variable of all fields (vector and scalar fields).

The Fourier transform in the time variable of a vector field, eg., the electric field $\boldsymbol{E}(\boldsymbol{r}, t)$, is defined as

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \int_{-\infty}^{\infty} \boldsymbol{E}(\boldsymbol{r}, t) e^{\mathrm{i}\omega t} \, \mathrm{d}t$$

with its inverse transform

$$\boldsymbol{E}(\boldsymbol{r}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \boldsymbol{E}(\boldsymbol{r}, \omega) e^{-\mathrm{i}\omega t} \, \mathrm{d}\omega$$

The Fourier transform for all other time dependent fields are defined in the same way. To avoid heavy notation we use the same symbol for the physical field $\boldsymbol{E}(\boldsymbol{r}, t)$, as for the Fourier transformed field $\boldsymbol{E}(\boldsymbol{r}, \omega)$—only the argument differs. In most cases the context implies whether it is the physical field or the Fourier transformed field that is intended, otherwise the time argument t or the (angular)frequency ω is written out to distinguish the fields.

All physical quantities are real, which imply constraints on the Fourier transform. The field values for negative values of ω are related to the values for positive values of ω by a complex conjugate. To see this, we write down the criterion for the field $\boldsymbol{E}$ to be real.

$$\int_{-\infty}^{\infty} \boldsymbol{E}(\boldsymbol{r}, \omega) \mathrm{e}^{-\mathrm{i}\omega t} \, d\omega = \left\{ \int_{-\infty}^{\infty} \boldsymbol{E}(\boldsymbol{r}, \omega) \mathrm{e}^{-\mathrm{i}\omega t} \, d\omega \right\}^*$$

where the star (*) denotes the complex conjugate. For real ω, we have

$$\int_{-\infty}^{\infty} \boldsymbol{E}(\boldsymbol{r}, \omega) \mathrm{e}^{-\mathrm{i}\omega t} \, d\omega = \int_{-\infty}^{\infty} \boldsymbol{E}^*(\boldsymbol{r}, \omega) \mathrm{e}^{\mathrm{i}\omega t} \, d\omega = \int_{-\infty}^{\infty} \boldsymbol{E}^*(\boldsymbol{r}, -\omega) \mathrm{e}^{-\mathrm{i}\omega t} \, d\omega$$

where we in the last integral have substituted ω for $-\omega$. Therefore, for real ω we have

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{E}^*(\boldsymbol{r}, -\omega)$$

Band	**Frequency**	**Wave length**	**Application**
ELF	< 3 KHz	> 100 km	
VLF	3–30 KHz	100–10 km	Navigation
LV	30–300 KHz	10–1 km	Navigation
MV	300–3000 KHz	1000–100 m	Radio
KV (HF)	3–30 MHz	100–10 m	Radio
VHF	30–300 MHz	10–1 m	FM, TV
UHF	300–1000 MHz	100–30 cm	Radar, TV, mobile communication
†	1–30 GHz	30–1 cm	Radar, satellite communication
†	30–300 GHz	10–1 mm	Radar
	4.2–$7.9 \cdot 10^{14}$ Hz	0.38–0.72 μm	Visible light

This shows that when the physical field is constructed from its Fourier transform, it suffices to integrate over the non-negative frequencies. By the change in variables, $\omega \to -\omega$, and the use of the condition above, we have

$$\begin{aligned}
\boldsymbol{E}(\boldsymbol{r},t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \boldsymbol{E}(\boldsymbol{r},\omega)e^{-\mathrm{i}\omega t}\,\mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{0} \boldsymbol{E}(\boldsymbol{r},\omega)\mathrm{e}^{-\mathrm{i}\omega t}\,d\omega + \frac{1}{2\pi}\int_{0}^{\infty} \boldsymbol{E}(\boldsymbol{r},\omega)\mathrm{e}^{-\mathrm{i}\omega t}\,d\omega \\
&= \frac{1}{2\pi}\int_{0}^{\infty} \left[\boldsymbol{E}(\boldsymbol{r},\omega)\mathrm{e}^{-\mathrm{i}\omega t} + \boldsymbol{E}(\boldsymbol{r},-\omega)\mathrm{e}^{\mathrm{i}\omega t}\right]\,d\omega \\
&= \frac{1}{2\pi}\int_{0}^{\infty} \left[\boldsymbol{E}(\boldsymbol{r},\omega)\mathrm{e}^{-\mathrm{i}\omega t} + \boldsymbol{E}^{*}(\boldsymbol{r},\omega)\mathrm{e}^{\mathrm{i}\omega t}\right]\,d\omega = \frac{1}{\pi}\,\mathrm{Re}\int_{0}^{\infty} \boldsymbol{E}(\boldsymbol{r},\omega)\mathrm{e}^{-\mathrm{i}\omega t}\,d\omega
\end{aligned} \tag{2.1}$$

where $\mathrm{Re}\, z$ denotes the real part of the complex number z. A similar result holds for all other Fourier transformed fields that we are using.

Fields that are purely time harmonic are of special interests in many applications, see Table 2. Purely time harmonic fields have the time dependence

$$\cos(\omega t - \alpha)$$

A simple way of obtaining purely time harmonic waves is to use phasors. Then the complex field $\boldsymbol{E}(\boldsymbol{r},\omega)$ is related to the time harmonic field $\boldsymbol{E}(\boldsymbol{r},t)$ via the rule

$$\boldsymbol{E}(\boldsymbol{r},t) = \mathrm{Re}\left\{\boldsymbol{E}(\boldsymbol{r},\omega)\mathrm{e}^{-\mathrm{i}\omega t}\right\} \tag{2.2}$$

If we write $\boldsymbol{E}(\boldsymbol{r},\omega)$ as

$$\begin{aligned}
\boldsymbol{E}(\boldsymbol{r},\omega) &= \hat{\boldsymbol{x}}E_x(\boldsymbol{r},\omega) + \hat{\boldsymbol{y}}E_y(\boldsymbol{r},\omega) + \hat{\boldsymbol{z}}E_z(\boldsymbol{r},\omega) \\
&= \hat{\boldsymbol{x}}|E_x(\boldsymbol{r},\omega)|\mathrm{e}^{\mathrm{i}\alpha(\boldsymbol{r})} + \hat{\boldsymbol{y}}|E_y(\boldsymbol{r},\omega)|\mathrm{e}^{\mathrm{i}\beta(\boldsymbol{r})} + \hat{\boldsymbol{z}}|E_z(\boldsymbol{r},\omega)|\mathrm{e}^{\mathrm{i}\gamma(\boldsymbol{r})}
\end{aligned}$$

we obtain the same result as in the expression above. This way of constructing purely time harmonic waves is convenient and often used.

2.1 The Maxwell equations

As a first step in our analysis of time harmonic fields, we Fourier transform the Maxwell equations (1.1) and (1.2) ($\frac{\partial}{\partial t} \to -\mathrm{i}\omega$)

$$\nabla \times \boldsymbol{E}(\boldsymbol{r},\omega) = \mathrm{i}\omega \boldsymbol{B}(\boldsymbol{r},\omega) \tag{2.3}$$
$$\nabla \times \boldsymbol{H}(\boldsymbol{r},\omega) = \boldsymbol{J}(\boldsymbol{r},\omega) - \mathrm{i}\omega \boldsymbol{D}(\boldsymbol{r},\omega) \tag{2.4}$$

The explicit harmonic time dependence $\exp\{-\mathrm{i}\omega t\}$ has been suppressed from these equations, i.e., the physical fields are

$$\boldsymbol{E}(\boldsymbol{r},t) = \mathrm{Re}\left\{\boldsymbol{E}(\boldsymbol{r},\omega)\mathrm{e}^{-\mathrm{i}\omega t}\right\}$$

This convention is applied to all purely time harmonic fields. Note that the electromagnetic fields $\boldsymbol{E}(\boldsymbol{r},\omega)$, $\boldsymbol{B}(\boldsymbol{r},\omega)$, $\boldsymbol{D}(\boldsymbol{r},\omega)$ and $\boldsymbol{H}(\boldsymbol{r},\omega)$, and the current density $\boldsymbol{J}(\boldsymbol{r},\omega)$ are complex vector fields.

The continuity equation (1.4) is transformed in a similar way

$$\nabla \cdot \boldsymbol{J}(\boldsymbol{r},\omega) - \mathrm{i}\omega\rho(\boldsymbol{r},\omega) = 0 \tag{2.5}$$

The remaining two equations from Chapter 1, (1.5) and (1.6), are transformed into

$$\nabla \cdot \boldsymbol{B}(\boldsymbol{r},\omega) = 0 \tag{2.6}$$
$$\nabla \cdot \boldsymbol{D}(\boldsymbol{r},\omega) = \rho(\boldsymbol{r},\omega) \tag{2.7}$$

These equations are a consequence of (2.3) and (2.4) and the continuity equation (2.5) (c.f., Chapter 1 on Page 2). To see this we take the divergence of the Maxwell equations (2.3) and (2.4), and get ($\nabla \cdot (\nabla \times \boldsymbol{A}) = 0$)

$$\mathrm{i}\omega \nabla \cdot \boldsymbol{B}(\boldsymbol{r},\omega) = 0$$
$$\mathrm{i}\omega \nabla \cdot \boldsymbol{D}(\boldsymbol{r},\omega) = \nabla \cdot \boldsymbol{J}(\boldsymbol{r},\omega) = \mathrm{i}\omega\rho(\boldsymbol{r},\omega)$$

Division by $\mathrm{i}\omega$ (provided $\omega \neq 0$) gives (2.6) and (2.7).

In a homogenous non-magnetic source free medium we obtain the Helmholtz equation for the electric field by eliminating the magnetic field from (2.3) and (2.4). This is done by taking the rotation of (2.3) and utilizing (2.4). The result is

$$\nabla^2 \boldsymbol{E}(\boldsymbol{r},\omega) + k(\omega)^2 \boldsymbol{E}(\boldsymbol{r},\omega) = \boldsymbol{0} \tag{2.8}$$

where

$$k(\omega) = \omega\sqrt{\epsilon_0\mu_0(\epsilon + \mathrm{i}\sigma/(\omega\epsilon_0))}$$

is the wavenumber. The magnetic field satisfies the same equation

$$\nabla^2 \boldsymbol{H}(\boldsymbol{r},\omega) + k(\omega)^2 \boldsymbol{H}(\boldsymbol{r},\omega) = \boldsymbol{0} \tag{2.9}$$

To this end, in vacuum, the time-harmonic Maxwell field equations are

$$\begin{cases} \nabla \times \boldsymbol{E}(\boldsymbol{r},\omega) = \mathrm{i}k_0 \left(c_0 \boldsymbol{B}(\boldsymbol{r},\omega)\right) \\ \nabla \times \left(\eta_0 \boldsymbol{H}(\boldsymbol{r},\omega)\right) = -\mathrm{i}k_0 \left(c_0\eta_0 \boldsymbol{D}(\boldsymbol{r},\omega)\right) \end{cases} \tag{2.10}$$

where $\eta_0 = \sqrt{\mu_0/\epsilon_0}$ is the intrinsic impedance of vacuum, $c_0 = 1/\sqrt{\epsilon_0\mu_0}$ the speed of light in vacuum, and $k_0 = \omega/c_0$ the wave number in vacuum. In (2.10) all field quantities have the same units, i.e., that of the electric field. This form is the standard form of the Maxwell equations that we use in this book.

2.2 Constitutive relations

The constitutive relations are the relations between the fields $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{B}$ and $\boldsymbol{H}$. In this book we restrict ourselves to materials that are linear and isotropic. That covers most solids, liquids and gases. The constitutive relations then read

$$\begin{aligned}\boldsymbol{D}(\boldsymbol{r},\omega) &= \epsilon_0\epsilon(\omega)\boldsymbol{E}(\boldsymbol{r},\omega)\\ \boldsymbol{B}(\boldsymbol{r},\omega) &= \mu_0\mu(\omega)\boldsymbol{H}(\boldsymbol{r},\omega)\end{aligned}$$

The parameters ϵ and μ are the (relative) permittivity and permeability of the medium, respectively.

We also note that a material with a conductivity that satisfies Ohm's law $\boldsymbol{J}(\boldsymbol{r},\omega) = \sigma(\omega)\boldsymbol{E}(\boldsymbol{r},\omega)$, always can be included in the constitutive relations by redefining the permittivity .

$$\epsilon_{\text{new}} = \epsilon_{\text{old}} + \mathrm{i}\frac{\sigma}{\omega\epsilon_0}$$

The right hand side in Ampère's law (2.4) is

$$\boldsymbol{J} - \mathrm{i}\omega\boldsymbol{D} = \sigma\boldsymbol{E} - \mathrm{i}\omega\epsilon_0\epsilon_{\text{old}}\cdot\boldsymbol{E} = -\mathrm{i}\omega\epsilon_0\epsilon_{\text{new}}\cdot\boldsymbol{E}$$

2.3 Poynting's theorem

In Chapter 1 we derived Poynting's theorem, see (1.19) on Page 9.

$$\nabla\cdot\boldsymbol{S}(t) + \boldsymbol{H}(t)\cdot\frac{\partial\boldsymbol{B}(t)}{\partial t} + \boldsymbol{E}(t)\cdot\frac{\partial\boldsymbol{D}(t)}{\partial t} + \boldsymbol{E}(t)\cdot\boldsymbol{J}(t) = 0$$

The equation describes conservation of power and contains products of two fields. In this section we study time harmonic fields, and the quantity that is of most interest for us is the time average over one period[1]. We denote the time average as $<\cdot>$ and for Poynting's theorem we get

$$<\nabla\cdot\boldsymbol{S}(t)> + <\boldsymbol{H}(t)\cdot\frac{\partial\boldsymbol{B}(t)}{\partial t}> + <\boldsymbol{E}(t)\cdot\frac{\partial\boldsymbol{D}(t)}{\partial t}> + <\boldsymbol{E}(t)\cdot\boldsymbol{J}(t)>= 0$$

[1]The time average of a product of two time harmonic fields $f_1(t)$ and $f_2(t)$ is easily obtained by averaging over one period $T = 2\pi/\omega$.

$$\begin{aligned}<f_1(t)f_2(t)> &= \frac{1}{T}\int_0^T f_1(t)f_2(t)\,\mathrm{d}t = \frac{1}{T}\int_0^T \mathrm{Re}\left\{f_1(\omega)\mathrm{e}^{-\mathrm{i}\omega t}\right\}\mathrm{Re}\left\{f_2(\omega)\mathrm{e}^{-\mathrm{i}\omega t}\right\}\,\mathrm{d}t\\ &= \frac{1}{4T}\int_0^T \left\{f_1(\omega)f_2(\omega)\mathrm{e}^{-2\mathrm{i}\omega t} + f_1^*(\omega)f_2^*(\omega)\mathrm{e}^{2\mathrm{i}\omega t} + f_1(\omega)f_2^*(\omega) + f_1^*(\omega)f_2(\omega)\right\}\,\mathrm{d}t\\ &= \frac{1}{4}\left\{f_1(\omega)f_2^*(\omega) + f_1^*(\omega)f_2(\omega)\right\} = \frac{1}{2}\mathrm{Re}\left\{f_1(\omega)f_2^*(\omega)\right\}\end{aligned}$$

The different terms in this quantity are

$$<\boldsymbol{S}(t)>=\frac{1}{2}\,\mathrm{Re}\,\{\boldsymbol{E}(\omega)\times\boldsymbol{H}^*(\omega)\} \tag{2.11}$$

and

$$\begin{aligned}
<\boldsymbol{H}(t)\cdot\frac{\partial\boldsymbol{B}(t)}{\partial t}>&=\frac{1}{2}\,\mathrm{Re}\,\{\mathrm{i}\omega\boldsymbol{H}(\omega)\cdot\boldsymbol{B}^*(\omega)\}\\
<\boldsymbol{E}(t)\cdot\frac{\partial\boldsymbol{D}(t)}{\partial t}>&=\frac{1}{2}\,\mathrm{Re}\,\{\mathrm{i}\omega\boldsymbol{E}(\omega)\cdot\boldsymbol{D}^*(\omega)\}\\
<\boldsymbol{E}(t)\cdot\boldsymbol{J}(t)>&=\frac{1}{2}\,\mathrm{Re}\,\{\boldsymbol{E}(\omega)\cdot\boldsymbol{J}^*(\omega)\}
\end{aligned}$$

Poynting's theorem (balance of power) for time harmonic fields, averaged over a period, becomes ($<\nabla\cdot\boldsymbol{S}(t)>=\nabla\cdot<\boldsymbol{S}(t)>$):

$$\begin{aligned}
\nabla\cdot<\boldsymbol{S}(t)>&+\frac{1}{2}\,\mathrm{Re}\,\{\mathrm{i}\omega\,[\boldsymbol{H}(\omega)\cdot\boldsymbol{B}^*(\omega)+\boldsymbol{E}(\omega)\cdot\boldsymbol{D}^*(\omega)]\}\\
&+\frac{1}{2}\,\mathrm{Re}\,\{\boldsymbol{E}(\omega)\cdot\boldsymbol{J}^*(\omega)\}=0
\end{aligned} \tag{2.12}$$

Of special interest is the case without currents[2] $\boldsymbol{J}=\boldsymbol{0}$. Poynting's theorem is then simplified to

$$\begin{aligned}
\nabla\cdot<\boldsymbol{S}(t)>&=-\frac{1}{2}\,\mathrm{Re}\,\{\mathrm{i}\omega\,[\boldsymbol{H}(\omega)\cdot\boldsymbol{B}^*(\omega)+\boldsymbol{E}(\omega)\cdot\boldsymbol{D}^*(\omega)]\}\\
&=-\frac{\mathrm{i}\omega}{4}\Big\{\boldsymbol{H}(\omega)\cdot\boldsymbol{B}^*(\omega)-\boldsymbol{H}^*(\omega)\cdot\boldsymbol{B}(\omega)\\
&\qquad+\boldsymbol{E}(\omega)\cdot\boldsymbol{D}^*(\omega)-\boldsymbol{E}(\omega)^*\cdot\boldsymbol{D}(\omega)\Big\}
\end{aligned}$$

where we used $\mathrm{Re}\,z=(z+z^*)/2$.

Problems in Chapter 2

2.1 Find two complex vectors, $\boldsymbol{A}$ and $\boldsymbol{B}$, such that $\boldsymbol{A}\cdot\boldsymbol{B}=0$ and

$$\begin{aligned}
\boldsymbol{A}'\cdot\boldsymbol{B}'&\neq 0\\
\boldsymbol{A}''\cdot\boldsymbol{B}''&\neq 0
\end{aligned}$$

where $\boldsymbol{A}'$ and $\boldsymbol{B}'$ are the real parts of the vectors, respectively, and where the imaginary parts are denoted $\boldsymbol{A}''$ and $\boldsymbol{B}''$, respectively.

$$\begin{cases}
\boldsymbol{A}=\hat{\boldsymbol{x}}+\mathrm{i}\hat{\boldsymbol{y}}\\
\boldsymbol{B}=(\hat{\boldsymbol{x}}+\xi\hat{\boldsymbol{y}})+\mathrm{i}(-\xi\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}})
\end{cases}$$

where ξ is an arbitrary real number.

[2]Conducting currents can, as we have seen, be included in the permittivity dyadic $\boldsymbol{\epsilon}$.

2.2 For real vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ we have

$$\boldsymbol{B} \cdot (\boldsymbol{B} \times \boldsymbol{A}) = 0$$

Prove that this equality also holds for arbitrary complex vectors $\boldsymbol{A}$ and $\boldsymbol{B}$.

Chapter 3

Transmission lines

When we analyze signals in circuits we have to know their frequency band and the size of the circuit in order to make appropriate approximations. We exemplify by considering signals with frequencies ranging from dc up to very high frequencies in a circuit that contains linear elements, i.e., resistors, capacitors, inductors and sources.

Definition: A circuit is discrete if we can neglect wave propagation in the analysis of the circuit. In most cases the circuit is discrete if the size of the circuit is much smaller than the wavelength in free space of the electromagnetic waves, $\lambda = c/f$.

- We first consider circuits at zero frequency, i.e., dc circuits. The wavelength $\lambda = c/f$ is infinite and the circuits are discrete. Capacitors correspond to an open circuit and inductors to a short circuit. The current in a wire with negligible resistance is constant in both time and space and the voltage drop along the wire is zero. The voltages and currents are determined by the Ohm's and Kirchhoff's laws. These follow from the static equations and relations

$$\begin{aligned} \nabla \times \boldsymbol{E}(\boldsymbol{r}) &= \boldsymbol{0} \\ \boldsymbol{J}(\boldsymbol{r}) &= \sigma \boldsymbol{E}(\boldsymbol{r}) \\ \nabla \cdot \boldsymbol{J}(\boldsymbol{r}) &= \boldsymbol{0} \end{aligned}$$

- We increase the frequency, but not more than that the wavelength $\lambda = c/f$ is still much larger than the size of the circuit. The circuit is still discrete and the voltage v and current i for inductors and capacitors are related by the induction law (1.1) and the continuity equation (1.4), that imply

$$\begin{aligned} i &= C\frac{dv}{dt} \\ v &= L\frac{di}{dt} \end{aligned}$$

  where C is the capacitance and L the inductance. These relations, in combination with the Ohm's and Kirchhoff's laws, are sufficient for determining

the voltages and currents in the circuit. In most cases the wires that connect circuit elements have negligible resistance, inductance and capacitance. This ensures that the current and voltage in each wire are constant in space, but not in time.

- We increase the frequency to a level where the wavelength is not much larger than the size of the circuit. Now wave propagation has to be taken into account. The phase and amplitude of the current and voltage along wires vary with both time and space. We have to abandon circuit theory and switch to transmission line theory, which is the subject of this chapter. The theory is based upon the full Maxwell equations but is phrased in terms of currents and voltages.
- If we continue to increase the frequency we reach the level where even transmission line theory is not sufficient to describe the circuit. This happens when components and wires act as antennas and radiate electromagnetic waves. We then need both electromagnetic field theory and transmission line theory to describe the circuit.

Often a system can be divided into different parts, where some parts are discrete while others need transmission line theory, or the full Maxwell equations. An example is an antenna system. The signal to the antenna is formed in a discrete circuit. The signal travels to the antenna via a transmission line and reaches the antenna, which is a radiating component.

3.1 Time and frequency domain

It is often advantageous to analyze signals in linear circuits in the frequency domain. We repeat some of the transformation rules between the time and frequency domains given in Chapter 2 and also give a short description of transformations based on Fourier series and Laplace transform. In the frequency domain the algebraic relations between voltages and currents are the same for all of the transformations described here. In the book we use either phasors or the Fourier transform to transform between time domain and frequency domain.

3.1.1 Phasors (jω method)

For time harmonic signals we use phasors. The transformation between the time and frequency domain is as follows:

$$v(t) = V_0 \cos(\omega t + \phi) \leftrightarrow V = V_0 e^{\mathrm{j}\phi}$$

where V is the complex voltage. This is equivalent to the transformation $v(t) = \mathrm{Re}\{V e^{\mathrm{j}\omega t}\}$, used in Chapter 2. An alternative is to use $\sin \omega t$ as reference for the phase and then the transformation reads

$$v(t) = V_0 \sin(\omega t + \phi) \leftrightarrow V = V_0 e^{\mathrm{j}\phi} \tag{3.1}$$

From circuit theory it is well-known that the relations between current and voltage are

$$\begin{cases} V = RI & \text{resistor} \\ V = \mathrm{j}\omega LI & \text{inductor} \\ V = \dfrac{I}{\mathrm{j}\omega C} & \text{capacitor} \end{cases}$$

In general the relationship between the complex voltage and current is written $V = ZI$ where Z is the impedance. This means that the impedance for a resistor is R, for an inductor it is $\mathrm{j}\omega L$ and for a capacitor it is $1/\mathrm{j}\omega C$. The admittance $Y = 1/Z$ is also used frequently in this chapter.

3.1.2 Fourier transformation

If the signal $v(t)$ is absolutely integrable, i.e., $\int_{-\infty}^{\infty} |v(t)|\,\mathrm{d}t < \infty$, it can be Fourier transformed

$$\begin{aligned} V(\omega) &= \int_{-\infty}^{\infty} v(t) e^{-\mathrm{j}\omega t}\,\mathrm{d}t \\ v(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} V(\omega) e^{\mathrm{j}\omega t}\,\mathrm{d}\omega \end{aligned} \tag{3.2}$$

The Fourier transform here differs from the one in Chapter 2 in that $e^{-\mathrm{i}\omega t}$ is exchanged for $e^{\mathrm{j}\omega t}$, see the comment below. As seen in Chapter 2 the negative values of the angular frequency is not a problem since they can be eliminated by using

$$V(\omega) = V^*(-\omega)$$

In the frequency domain the relations between current and voltage are identical with the corresponding relations obtained by the $\mathrm{j}\omega$-method, i.e.,

$$\begin{cases} V(\omega) = RI(\omega) & \text{resistor} \\ V(\omega) = \mathrm{j}\omega LI(\omega) & \text{inductor} \\ V(\omega) = \dfrac{I(\omega)}{\mathrm{j}\omega C} & \text{capacitor} \end{cases}$$

Comment on j **and** i

The electrical engineering literature uses the time convention $e^{\mathrm{j}\omega t}$ in the phasor method and the Fourier transformation, while physics literature uses $e^{-\mathrm{i}\omega t}$. We can transform expressions from one convention to the other by complex conjugation of all expressions and exchanging i and j. In this chapter we use $e^{\mathrm{j}\omega t}$ whereas in the rest of the book we use $e^{-\mathrm{i}\omega t}$. The reason is that transmission lines are mostly treated in the literature of electrical engineering while hollow waveguides and dielectric waveguides are more common in physics literature.

3.1.3 Fourier series

A periodic signal with the period T satisfies $f(t) = f(t+T)$ for all times t. We introduce the fundamental angular frequency $\omega_0 = 2\pi/T$. The set of functions $\{e^{\mathrm{j}n\omega_0 t}\}_{n=-\infty}^{n=\infty}$ is a complete orthogonal system of functions on an interval of length T and we may expand $f(t)$ in a Fourier series as

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{\mathrm{j}n\omega_0 t}$$

We obtain the Fourier coefficients c_m if we multiply with $e^{-\mathrm{j}m\omega_0 t}$ on the left and right hand sides and integrate over one period

$$c_m = \frac{1}{T}\int_0^T f(t) e^{-\mathrm{j}m\omega_0 t}\,\mathrm{d}t$$

An alternative is to use the expansion in the system $\{1, \cos(n\omega_0 t), \sin(n\omega_0 t)\}_{n=1}^{n=\infty}$

$$f(t) = a_0 + \sum_{n=1}^{\infty}[a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t)]$$

Also this set of functions is complete and orthogonal. The Fourier coefficients are obtained by multiplying with 1, $\cos(m\omega_0 t)$, and $\sin(m\omega_0 t)$, respectively, and integrate over one period

$$\begin{aligned}
a_0 &= \frac{1}{T}\int_0^T f(t)\,\mathrm{d}t \\
a_m &= \frac{2}{T}\int_0^T f(t)\cos(m\omega_0 t)\,\mathrm{d}t, \qquad m > 0 \\
b_m &= \frac{2}{T}\int_0^T f(t)\sin(m\omega_0 t)\,\mathrm{d}t
\end{aligned}$$

We see that $a_0 = c_0$ is the dc part of the signal. The relations for $n > 0$ are $c_n = 0.5(a_n - \mathrm{j}b_n)$ and $c_{-n} = c_n^*$, as can be seen from the Euler identity.

If we let the current and voltage have the expansions

$$\begin{aligned}
i(t) &= \sum_{n=-\infty}^{\infty} I_n e^{\mathrm{j}n\omega_0 t} \\
v(t) &= \sum_{n=-\infty}^{\infty} V_n e^{\mathrm{j}n\omega_0 t}
\end{aligned}$$

the relations between the coefficients V_n and I_n are

$$\begin{cases}
V_n = RI_n & \text{resistor} \\
V_n = \mathrm{j}n\omega_0 L I_n & \text{inductor} \\
V_n = \dfrac{I_n}{\mathrm{j}n\omega_0 C} & \text{capacitor}
\end{cases}$$

Thus it is straightforward to determine the Fourier coefficients for the currents and voltages in a circuit. In this chapter we will not use the expansions in Fourier series.

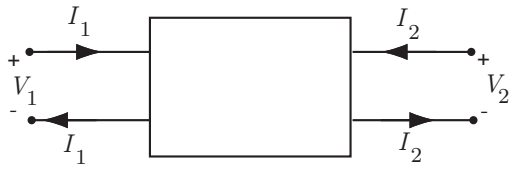


Figure 3.1: A two-port. Notice that the total current entering each port is always zero.

3.1.4 Laplace transformation

If the signal $v(t)$ is defined for $t \geq 0$ we may use the Laplace transform

$$V(s) = \int_{0^-}^{\infty} v(t)e^{-st}\,\mathrm{d}t$$

In most cases we use tables of Laplace transforms in order to obtain $v(t)$ from $V(s)$. If we exchange s for $\mathrm{j}\omega$ in the frequency domain we get the corresponding expression for the $\mathrm{j}\omega$-method and Fourier transformation. The Laplace transform is well suited for determination of transients and for stability and frequency analysis. The relations for the Laplace transforms of current and voltage read

$$\begin{cases} V(s) = RI(s) & \text{resistor} \\ V(s) = sLI(s) & \text{inductor} \\ V(s) = \dfrac{I(s)}{sC} & \text{capacitor} \end{cases}$$

3.2 Two-ports

A two-port is a circuit with two ports, c.f., figure 3.1. We only consider passive linear two-ports in this book. Passive means that there are no independent sources in the two-port. The sum of the currents entering a port is always zero. In the frequency domain the two-port is represented by a matrix with four complex elements. The matrix elements depends on which combinations of I_1, I_2, V_1 and V_2 we use, as seen below.

3.2.1 The impedance matrix

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = [Z] \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} \tag{3.3}$$

The inverse of the impedance matrix is the admittance matrix, $[Y] = [Z]^{-1}$

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = [Y] \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

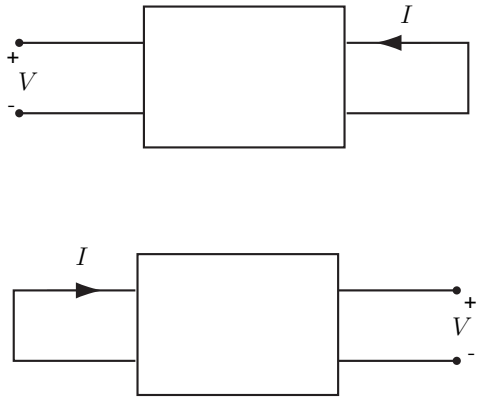


Figure 3.2: Reciprocal two-port. If voltage V at port 1 gives the shortening current I in port 2 then the voltage V at port 2 gives the shortening current I at port 1.

3.2.2 The cascade matrix ($ABCD$-matrix)

We introduce the $ABCD$ matrix as

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = [K] \begin{pmatrix} V_2 \\ -I_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_2 \\ -I_2 \end{pmatrix} \tag{3.4}$$

We have put a minus sign in front of I_2 in order to cascade two-ports in a simple manner. The relation can be inverted:

$$\begin{pmatrix} V_2 \\ I_2 \end{pmatrix} = [K'] \begin{pmatrix} V_1 \\ -I_1 \end{pmatrix} = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} \begin{pmatrix} V_1 \\ -I_1 \end{pmatrix}$$

We notice that the $[K']$ matrix is obtained from the $[K]^{-1}$ matrix by changing sign of the non-diagonal elements.

3.2.3 The hybrid matrix

$$\begin{pmatrix} V_1 \\ I_2 \end{pmatrix} = [H] \begin{pmatrix} I_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ V_2 \end{pmatrix}$$

The inverse hybrid matrix, $[G] = [H]^{-1}$, is given by

$$\begin{pmatrix} I_1 \\ V_2 \end{pmatrix} = [G] \begin{pmatrix} V_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} V_1 \\ I_2 \end{pmatrix}$$

3.2.4 Reciprocity

Assume a system where we place a signal generator at a certain point and measure the signal at another point. We then exchange the source and measurement points and measure the signal again. If the measured signal is the same in the two cases the system is reciprocal.

We use the following definition of reciprocity for two-ports: If V_2, I_2 give V_1, I_1 and V_2', I_2' give V_1', I_1' then the two-port is reciprocal if

$$V_1 I_1' - V_1' I_1 + V_2 I_2' - V_2' I_2 = 0$$

We insert the impedance matrix and get

$$(Z_{12} - Z_{21})(I_1' I_2 - I_2' I_1) = 0$$

for all I_1, I_2, I_1' and I_2'. Thus the two-port is reciprocal if and only if $[Z]$ is a symmetric matrix. The inverse of a symmetric matrix is symmetric and hence also $[Y]$ has to be symmetric in a reciprocal two-port. Reciprocity implies that if $I_1 = 0$, $I_2' = 0$ and $V_1 = V_2'$ then

$$V_1 I_1' = V_2' I_2 \Rightarrow I_1' = I_2$$

c.f., figure 3.2. If $V_1 = 0$, $V_2' = 0$ and $I_1 = I_2'$ then

$$V_2 I_2' = V_1' I_1 \Rightarrow V_1' = V_2$$

One can prove that all linear two-ports that do not have any dependent sources are reciprocal.

3.2.5 Transformation between matrices

The transformations between the matrices $[Z]$, $[K]$, and $[H]$ and between $[Y]$, $[G]$, and $[K']$ are given by the table below:

	$[Z]$	$[H]$	$[K]$
$[Z]$	$\begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}$	$\begin{pmatrix} \frac{\Delta_H}{h_{22}} & \frac{h_{12}}{h_{22}} \\ -\frac{h_{21}}{h_{22}} & \frac{1}{h_{22}} \end{pmatrix}$	$\begin{pmatrix} \frac{A}{C} & \frac{\Delta_K}{C} \\ \frac{1}{C} & \frac{D}{C} \end{pmatrix}$
$[H]$	$\begin{pmatrix} \frac{\Delta_Z}{Z_{22}} & \frac{Z_{12}}{Z_{22}} \\ -\frac{Z_{21}}{Z_{22}} & \frac{1}{Z_{22}} \end{pmatrix}$	$\begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$	$\begin{pmatrix} \frac{B}{D} & \frac{\Delta_K}{D} \\ -\frac{1}{D} & \frac{C}{D} \end{pmatrix}$
$[K]$	$\begin{pmatrix} \frac{Z_{11}}{Z_{21}} & \frac{\Delta_Z}{Z_{21}} \\ \frac{1}{Z_{21}} & \frac{Z_{22}}{Z_{21}} \end{pmatrix}$	$\begin{pmatrix} -\frac{\Delta_H}{h_{21}} & -\frac{h_{11}}{h_{21}} \\ -\frac{h_{22}}{h_{21}} & -\frac{1}{h_{21}} \end{pmatrix}$	$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$

	$[Y]$	$[G]$	$[K']$
$[Y]$	$\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}$	$\begin{pmatrix} \frac{\Delta_G}{g_{22}} & \frac{g_{12}}{g_{22}} \\ -\frac{g_{21}}{g_{22}} & \frac{1}{g_{22}} \end{pmatrix}$	$\begin{pmatrix} \frac{A'}{B'} & -\frac{1}{B'} \\ -\frac{\Delta'_K}{B'} & \frac{D'}{B'} \end{pmatrix}$
$[G]$	$\begin{pmatrix} \frac{\Delta_Y}{Y_{22}} & \frac{Y_{12}}{Y_{22}} \\ -\frac{Y_{21}}{Y_{22}} & \frac{1}{Y_{22}} \end{pmatrix}$	$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$	$\begin{pmatrix} \frac{C'}{D'} & -\frac{1}{D'} \\ \frac{\Delta'_K}{D'} & \frac{B'}{D'} \end{pmatrix}$
$[K']$	$\begin{pmatrix} -\frac{Y_{11}}{Y_{12}} & -\frac{1}{Y_{12}} \\ -\frac{\Delta Y}{Y_{12}} & -\frac{Y_{22}}{Y_{12}} \end{pmatrix}$	$\begin{pmatrix} -\frac{\Delta_G}{g_{12}} & -\frac{g_{22}}{g_{12}} \\ -\frac{g_{11}}{g_{12}} & -\frac{1}{g_{12}} \end{pmatrix}$	$\begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$

We use $\Delta_K = \det\{K\}$ to denote the determinant of the cascade matrix. From these transformations we see that a reciprocal two-port has a hybrid matrix that is anti symmetric, i.e., $h_{12} = -h_{21}$, since the impedance matrix is symmetric. We also notice that $\Delta_K = 1$ for a reciprocal two-port and that $[G]$ is anti symmetric and $\Delta_{K'} = 1$, since $[Y]$ is symmetric.

3.2.6 Circuit models for two-ports

We have seen that a general two-port is determined by four complex parameters. They can be substituted by an equivalent two-port with two impedances and two dependent sources. In figure 3.3 we see the two equivalent two-ports that can be obtained directly from the $Z-$ and $H-$matrices, respectively. A reciprocal two-port is determined by the three complex numbers Z_{11}, $Z_{12} = Z_{21}$, and Z_{22}. In this case we can still use the equivalent circuits in figure 3.3 but we can also find equivalent $T-$ and $\Pi-$circuits with passive components. A $T-$circuit, c.f., figure 3.4, has the following impedance matrix

$$[Z] = \begin{pmatrix} Z_a + Z_c & Z_c \\ Z_c & Z_b + Z_c \end{pmatrix}$$

This equals the impedance matrix for a reciprocal two-port if we let

$$\begin{aligned} Z_a &= Z_{11} - Z_{21} \\ Z_b &= Z_{22} - Z_{21} \\ Z_c &= Z_{21} \end{aligned}$$

The admittance matrix for a Π-coupling, c.f., figure 3.4, is obtained by shortening port 1 and port 2, respectively.

$$[Y] = \begin{pmatrix} Y_a + Y_c & -Y_c \\ -Y_c & Y_b + Y_c \end{pmatrix}$$

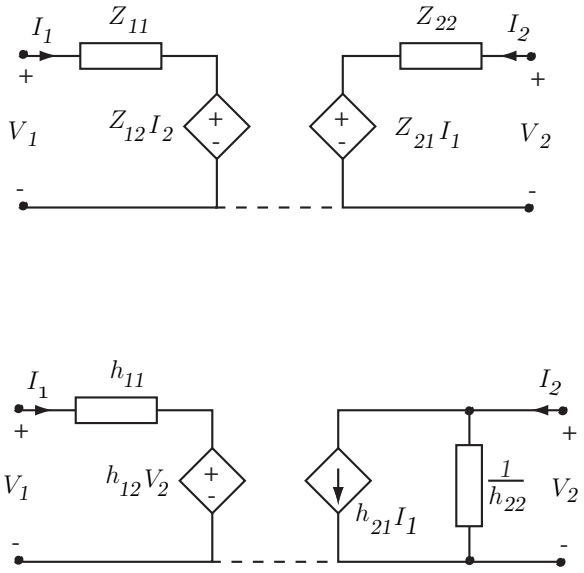


Figure 3.3: Equivalent circuits for a passive two-port. The upper corresponds to the impedance representation and the lower to the hybrid representation.

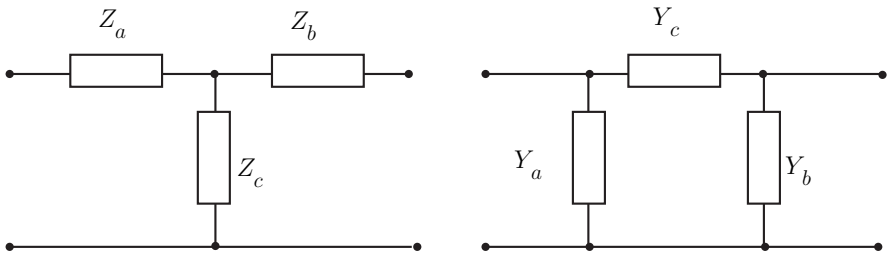


Figure 3.4: Equivalent $T-$ and Π-circuits for a reciprocal passive two-port.

We can always substitute a reciprocal two-port for a Π-coupling by using

$$Y_a = Y_{11} + Y_{21}$$
$$Y_b = Y_{22} + Y_{21}$$
$$Y_c = -Y_{21}$$

3.2.7 Combined two-ports

A two-port can be feedback coupled by another two-port in four different ways. We can use these different couplings when we create feedback amplifiers. The four couplings correspond to voltage amplifier, $V \to V$, current amplifier, $I \to I$, transimpedance amplifier, $I \to V$, and transadmittance amplifier $V \to I$. The input impedance should be as large as possible when voltage is the input and as low as possible when current is the input. The output impedance should be as small as possible when voltage is the output and as high as possible when current is the outpot.

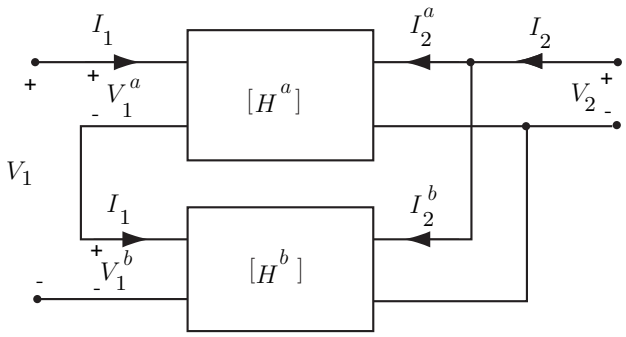


Figure 3.5: Series-parallel coupling: $[H] = [H^a] + [H^b]$

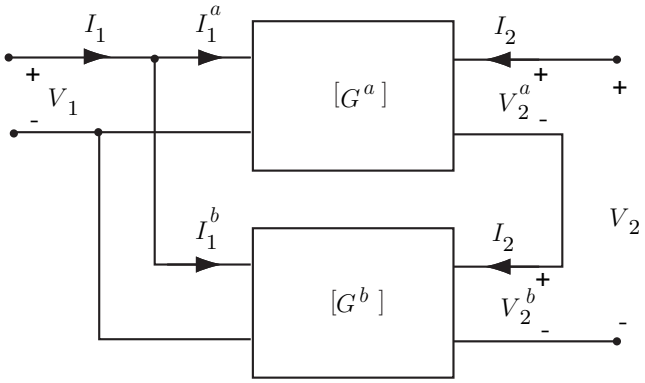


Figure 3.6: Parallel-series coupling: $[G] = [G]^a + [G^b]$

Series-parallel coupling ($V - V$-coupling)

According to figure 3.5 we get

$$\begin{pmatrix} V_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} V_1^a \\ I_2^a \end{pmatrix} + \begin{pmatrix} V_1^b \\ I_2^b \end{pmatrix} = ([H^a] + [H^b]) \begin{pmatrix} V_2 \\ I_1 \end{pmatrix}$$

The total hybrid matrix is then given by $[H] = [H^a] + [H^b]$. Series-parallel coupling of N two-ports with hybrid matrices $[H_n]$ result in the total hybrid matrix is $[H] = \sum_{n=1}^{N} [H_n]$.

Parallel-series coupling ($I - I$-coupling)

According to figure 3.6 we get

$$\begin{pmatrix} I_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} I_1^a \\ V_2^a \end{pmatrix} + \begin{pmatrix} I_1^b \\ V_2^b \end{pmatrix} = ([G^a] + [G^b]) \begin{pmatrix} V_1 \\ I_2 \end{pmatrix}$$

The total hybrid matrix is given by $[G] = [G^a] + [G^b]$. If we use N two-ports with hybrid matrices $[G_n]$ we get the total hybrid matrix $[G] = \sum_{n=1}^{N} [G_n]$.

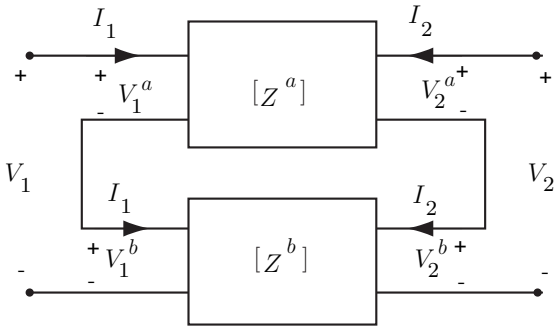


Figure 3.7: Series coupling: $[Z] = [Z^a] + [Z^b]$.

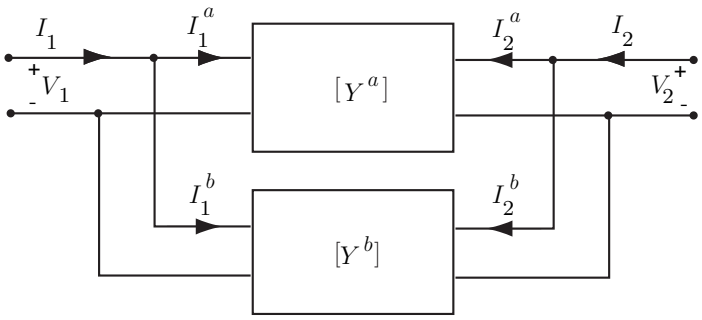


Figure 3.8: Parallel coupling: $[Y] = [Y^a] + [Y^b]$.

Series coupling ($V - I$-coupling)

According to figure 3.7 we get

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} V_1^a \\ V_2^a \end{pmatrix} + \begin{pmatrix} V_1^b \\ V_2^b \end{pmatrix} = ([Z^a] + [Z^b]) \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$$

The total impedance matrix is given by $[Z] = [Z^a] + [Z^b]$. With N two-ports with impedance matrices $[Z_n]$ in series the total impedance matrix is $[Z] = \sum_{n=1}^{N} [Z_n]$.

Parallel coupling ($I - V$-coupling)

According to figure 3.8 we get

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} I_1^a \\ I_2^a \end{pmatrix} + \begin{pmatrix} I_1^b \\ I_2^b \end{pmatrix} = ([Y^a] + [Y^b]) \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

The total admittance matrix is given by $[Y] = [Y^a] + [Y^b]$. With N two-ports with admittance matrices $[Y_n]$ in parallel the total admittance matrix is $[Y] = \sum_{n=1}^{N} [Y_n]$.

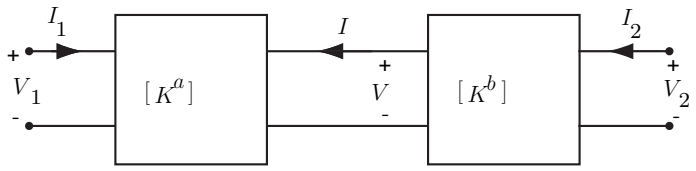


Figure 3.9: Cascade coupling: $[K] = [K^a][K^b]$.

3.2.8 Cascad coupled two-ports

We cascade two two-ports according to figure 3.9 and get the total cascade matrix in the following way

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = [K^a] \begin{pmatrix} V \\ -I \end{pmatrix} = [K^a][K^b] \begin{pmatrix} V_2 \\ -I_2 \end{pmatrix}$$

The total matrix is given by $[K] = [K^a][K^b]$. The two matrices do not commute, in general. The order of the two-ports is thus important. When N two-ports with cascade matrices $[K_n]$ are cascaded the total cascade matrix is

$$[K] = \prod_{n=1}^{N} [K_n]$$

3.3 Transmission lines in time domain

Transmission lines are wires that are not short compared to the wavelength. The wavelength is given by $\lambda = c/f$ where c is the speed of light in the medium surrounding the wires and f is the frequency.

The signals propagating along a line can be expressed in terms of the voltage between the wires and current in the wires. This leads to the scalar wave equation for the voltage (or current). The signal can also be described as electromagnetic waves that are bound to the wires. To find the electromagnetic fields we have to solve the Maxwell equations with appropriate boundary conditions. The two views lead to the same results but in most cases it is more convenient to use the equations for voltage and current rather than the Maxwell equations.

3.3.1 Wave equation

A transmission line consists of two conductors. We always consider the line to be straight and let it run along the $z-$direction. The voltage between the conductors, $v(z,t)$, and the current $i(z,t)$ are defined by figure 3.10.

A transmission line is defined by the four distributed line parameters

C : capacitance per unit length (F/m)
L : inductance per unit length (H/m)
G : conductance per unit length (1/(Ωm))
R : resistance per unit length (Ω/m)

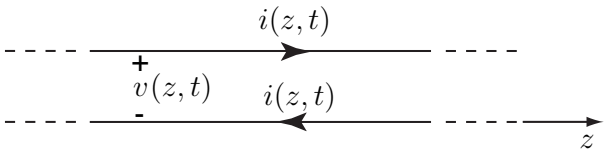


Figure 3.10: Voltage and current for the transmission line

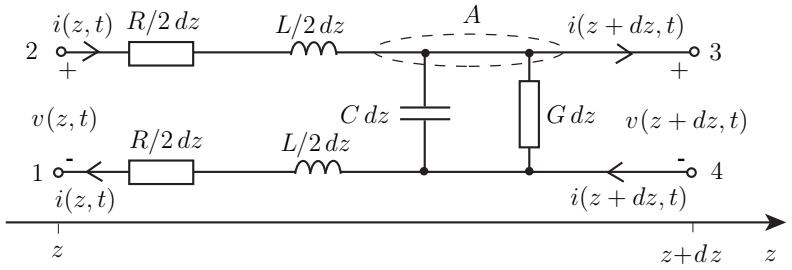


Figure 3.11: Circuit model for a transmission line

For a homogeneous transmission line the line parameters are independent of z and t. The voltage v between the conductors gives rise to line charges ρ_ℓ and $-\rho_\ell$ on the conductors. The capacitance per unit length is defined by $C = \rho_\ell(z,t)/v(z,t)$. The current i on the conductors gives rise to a magnetic flux Φ per unit length of the line. The inductance per unit length is defined by $L = \Phi(z,t)/i(z,t)$. The voltage v between the conductors may give rise to a leakage of current between the conductors. The conductance per unit length is defined by $G = i_{\text{leak}}(z,t)/v(z,t)$, where i_{leak} is the leakage current per unit length. The resistance of the conductors give a voltage drop along the line. If the voltage drop per unit length is v_{drop} then the resistance per unit length is $R = 2v_{\text{drop}}(z,t)/i(z,t)$.

We derive the equations for the transmission line by examining a very short piece of the line. The piece, which is shown in figure 3.11, is short enough to be a discrete circuit.

The Kirchhoff's voltage law for the loop $1 \to 2 \to 3 \to 4$ and the Kirchhoff's current law for the node A give the transmission line equations

$$v(z,t) = R\,dz\,i(z,t) + L\,dz\,\frac{\partial i(z,t)}{\partial t} + v(z+dz,t)$$
$$i(z,t) = G\,dz\,v(z,t) + C\,dz\,\frac{\partial v(z,t)}{\partial t} + i(z+dz,t)$$

We divide by dz and let dz $\to 0$ to get

$$-\frac{\partial v(z,t)}{\partial z} = Ri(z,t) + L\frac{\partial i(z,t)}{\partial t} \tag{3.5}$$

$$-\frac{\partial i(z,t)}{\partial z} = Gv(z,t) + C\frac{\partial v(z,t)}{\partial t} \tag{3.6}$$

We eliminate the current by operating with $\frac{\partial}{\partial z}$ on (3.5) and with $R + L\frac{\partial}{\partial t}$ on (3.6). We combine the two equations to get the scalar wave equation for the voltage

$$\frac{\partial^2 v(z,t)}{\partial z^2} - LC\frac{\partial^2 v(z,t)}{\partial t^2} - (LG + RC)\frac{\partial v(z,t)}{\partial t} - RGv(z,t) = 0 \tag{3.7}$$

Also the current $i(z,t)$ satisfies this equation.

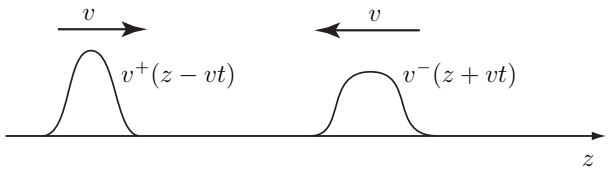


Figure 3.12: Wave propagation of pulses

3.3.2 Wave propagation in the time domain

For simplicity we start with a lossless transmission line. That means that both R and G are zero and the transmission line equations read

$$-\frac{\partial v(z,t)}{\partial z} = L\frac{\partial i(z,t)}{\partial t} \tag{3.8}$$

$$-\frac{\partial i(z,t)}{\partial z} = C\frac{\partial v(z,t)}{\partial t} \tag{3.9}$$

$$\frac{\partial^2 v(z,t)}{\partial z^2} - LC\frac{\partial^2 v(z,t)}{\partial t^2} = 0 \tag{3.10}$$

The general solution to this equation is

$$v(z,t) = f(z - v_\mathrm{p}t) + g(z + v_\mathrm{p}t)$$

where

$$v_\mathrm{p} = \frac{1}{\sqrt{LC}}$$

is the phase speed. To verify the solution we make the substitution of variables

$$\xi = z - v_\mathrm{p}t$$
$$\eta = z + v_\mathrm{p}t$$

that transforms the wave equation (3.10) to the equation

$$\frac{\partial^2 v}{\partial\eta\partial\xi} = 0 \quad \Rightarrow \frac{\partial v}{\partial\xi} = f_1(\xi) \quad \Rightarrow v = \int f_1(\xi)\,\mathrm{d}\xi + g(\eta) = f(\xi) + g(\eta)$$

The function $f(z - v_\mathrm{p}t)$ has constant argument when $z = v_\mathrm{p}t + \text{konstant}$. That means that $f(z - v_\mathrm{p}t)$ is a wave that propagates with the speed $v_\mathrm{p} = 1/\sqrt{LC}$ in the positive $z-$direction, see figure 3.12. In the same way we can argue that $g(z + v_\mathrm{p}t)$ is a wave that propagates with speed v_p in the negative $z-$direction. We indicate the direction of propagation by writing the solution as

$$v(z,t) = v^+(z - v_\mathrm{p}t) + v^-(z + v_\mathrm{p}t) \tag{3.11}$$

The shape of v^+ and v^- are determined by the sources and load impedances of the transmission line. We get the currents corresponding to v^+ and v^- by inserting (3.11) into (3.8) or (3.9)

$$i(z,t) = i^+(z - v_\mathrm{p}t) + i^-(z + v_\mathrm{p}t)$$

where

$$\frac{v^+(z-v_\mathrm{p}t)}{i^+(z-v_\mathrm{p}t} = -\frac{v^-(z+v_\mathrm{p}t)}{i^-(z+v_\mathrm{p}t)} = \sqrt{\frac{L}{C}} = Z_0$$

We have introduced the characteristic impedance of the transmission line, Z_0, which has the dimension Ω. It is important to understand that the characteristic impedance is the quotient of the voltage and current running in the positive $z-$direction. It is not the quotient of the total voltage and current. We can prove that the phase speed $v_\mathrm{p} = 1/\sqrt{LC}$ is equal to the wave speed of electromagnetic waves in the material surrounding the conductors, i.e., $v_\mathrm{p} = 1/\sqrt{\epsilon_0\epsilon\mu_0} = c/\sqrt{\epsilon}$ where ϵ_0 and μ_0 are the permittivity and permeability of vacuum, ϵ is the relative permittivity for the material surrounding the conductors and $c = 3\cdot 10^8$ m/s is the speed of light in vacuum. We summarize the results for wave propagation on a lossless transmission line

$$\begin{aligned}
&v(z,t) = v^+(z-v_\mathrm{p}t) + v^-(z+v_\mathrm{p}t)\\
&i(z,t) = i^+(z-v_\mathrm{p}t) + i^-(z+v_\mathrm{p}t) = \frac{1}{Z_0}(v^+(z-v_\mathrm{p}t) - v^-(z+v_\mathrm{p}t))\\
&v_\mathrm{p} = 1/\sqrt{LC}, \qquad Z_0 = \sqrt{L/C}
\end{aligned} \tag{3.12}$$

3.3.3 Reflection on a lossless line

We assume a lossless transmission line along $0 < z < \ell$. It has the characteristic impedance Z_0 and is terminated by a load resistance R_L at $z = \ell$. A wave $v_i(z,t) = v^+(z-v_\mathrm{p}t)$ has been generated at $z = 0$ and it generates a reflected wave $v_r(z,t) = v^-(z+v_\mathrm{p}t)$ once it reaches $z = \ell$. The total voltage and current on the line are

$$\begin{aligned}
v(z,t) &= v_i(z,t) + v_r(z,t)\\
i(z,t) &= i_i(z,t) + i_r(z,t) = \frac{1}{Z_0}\left(v_i(z,t) - v_r(z,t)\right)
\end{aligned} \tag{3.13}$$

We need the boundary condition at $z = \ell$

$$\frac{v(\ell,t)}{i(\ell,t)} = R_L$$

We insert (3.13) into the boundary condition

$$Z_0\frac{v_i(\ell,t) + v_r(\ell,t)}{v_i(\ell,t) - v_r(\ell,t)} = R_L$$

which gives us the reflected wave

$$v_r(\ell,t) = \frac{R_L - Z_0}{R_L + Z_0}v_i(\ell,t)$$

We see that the reflected wave is just a scaled version of the incident wave. The scaling constant is the dimensionless reflection coefficient

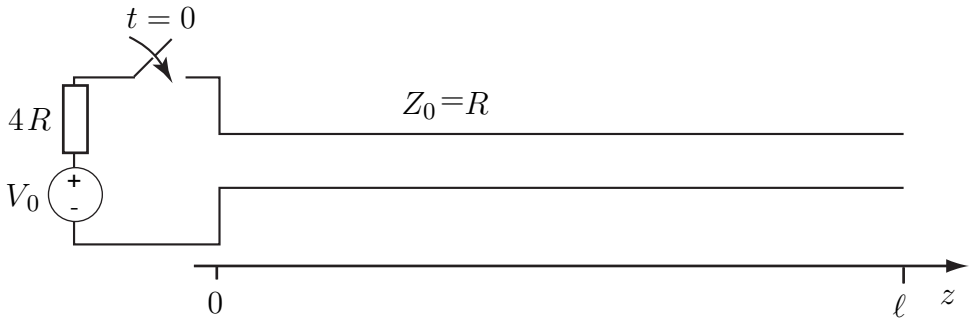
$$\Gamma = (R_L - Z_0)/(R_L + Z_0) \tag{3.14}$$

Example 3.1

- If $R_L = Z_0$ then $\Gamma = 0$ and $v_r(z,t) = 0$. The load impedance is then matched to the line. All power sent to the load is absorbed by it. In most cases we try to use matched loads.
- If the line is <u>open</u> at $z = \ell$ then $R_L = \infty$. We get $\Gamma = 1$ and $v_r(\ell,t) = v_i(\ell,t)$.
- If the line is <u>shortened</u> at $z = \ell$ then $R_L = 0$ and we get $\Gamma = -1$ and $v_r(\ell,t) = -v_i(\ell,t)$.

■

Example 3.2
A dc source with open circuit voltage V_0 and an inner resistance $R_i = 4R$ is at time $t = 0$



connected at $z = 0$ to a transmission line with characteristic impedance $Z_0 = R$. The transmission line has the length ℓ and is open at $z = \ell$. We determine the total voltage $v(0,t)$ for $t > 0$ by analyzing the process in a chronological order.

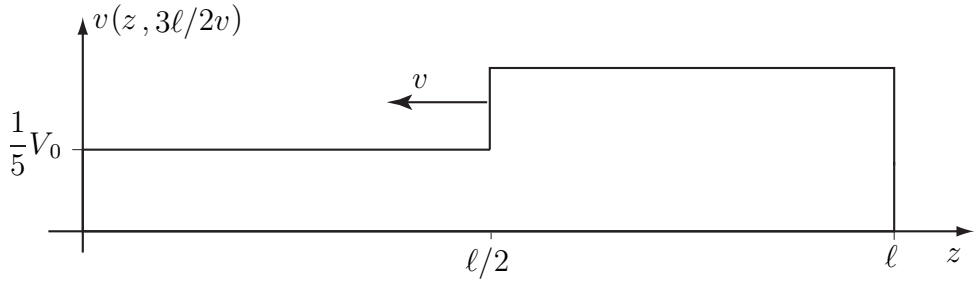
$t = 0^+$: A step pulse $v_i(z,t)$ starts to propagate along the line. The voltage at $z = 0$ is given by $v(0,t) = v_i(0,t) = Z_0 i_i(0,t) = Ri(0,t)$. The line is equivalent to a resistance R and the voltage division formula gives

$$v(0,t) = v^i(0,t) = \frac{R}{4R+R}V_0 = \frac{1}{5}V_0$$

$0 < t < \frac{\ell}{v_\mathrm{p}}$: The step with amplitude $v_1 = V_0/5$ is moving with speed v_p towards $z = \ell$.

$t = \frac{\ell}{v_\mathrm{p}}$: The step v_1 is reflected at $z = \ell$. At $z = \ell$ the line is open. Thus $Z_b = R_b = \infty$ and the reflection coefficient is $\Gamma_\ell = 1$. The reflected step has the amplitude $v_2 = \Gamma_\ell v_1 = V_0/5$. Still $v(0,t) = v_1 = V_0/5$ since the reflected wave has not reached $z = 0$.

$\frac{\ell}{v_\mathrm{p}} < t < \frac{2\ell}{v_\mathrm{p}}$: The step v_2 is moving like a $v^-(z + v_\mathrm{p}t)$ wave towards $z = 0$.



$t = \frac{2\ell}{v_\mathrm{p}}$: v_2 is reflected at $z = 0$. The load impedance at $z = 0$ is the inner resistance $4R$ and the reflection coefficient is

$$\Gamma_0 = \frac{4R - R}{4R + R} = \frac{3}{5}$$

The reflected wave is a step with amplitude $v_3 = \Gamma_0 v_2 = 3V_0/25$. The total voltage at $z = 0$ is

$$v(0, 2\ell/v_\mathrm{p}^+) = v_1 + v_2 + v_3 = V_0/5 + V_0/5 + 3V_0/25 = 13V_0/25$$

$t > \frac{2\ell}{v_\mathrm{p}}$: The reflections continue. After infinite time the voltage at $z = 0$ is a geometrical series

$$v(0, \infty) = v_1 + v_2 + v_3 \cdots = \frac{2V_0}{5} \sum_{k=0}^{\infty} \left(\frac{3}{5}\right)^k = \frac{2V_0}{5} \frac{1}{1 - 3/5} = V_0$$

This does not surprise us since after long time the circuit is a dc circuit. The time evolution of the voltage is very rapid since the travel time ℓ/v_p is very short.

■

3.4 Transmission lines in frequency domain

We now turn to time harmonic signals. The voltages are generated by a time harmonic source that was switched on early enough such that all transients have disappeared. The incident and reflected sinusiodal waves form a standing wave pattern along the line. We transform the voltages and currents to the frequency domain by the jω method. In the frequency domain the complex voltage $V(z)$ and current $I(z)$ satisfy the frequency domain versions of the transmission line equations, c.f., (3.5) and (3.6)

$$\begin{aligned} -\frac{dV(z)}{dz} &= (R + \mathrm{j}\omega L) I(z) \\ -\frac{dI(z)}{dz} &= (G + \mathrm{j}\omega C) V(z) \end{aligned} \tag{3.15}$$

We first consider the lossless case $R = 0$, $G = 0$. By differentiating the upper equation w.r.t. z and using the lower equation we get

$$\frac{d^2V(z)}{dz^2} + \beta^2 V(z) = 0$$

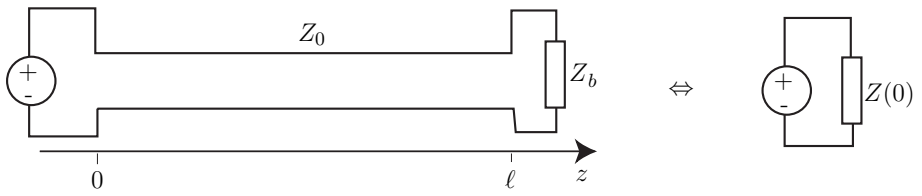


Figure 3.13: The input impedance $Z(0)$.

where $\beta = \omega\sqrt{LC} = \omega/v_{\mathrm{p}}$ is the phase coefficient. The equation has the two independent solutions

$$V(z) = V^+(z) + V^-(z) = V_p e^{-\mathrm{j}\beta z} + V_n e^{\mathrm{j}\beta z} \tag{3.16}$$

In general $V_p = |V_p|e^{\mathrm{j}\phi_p}$ and $V_n = |V_n|e^{\mathrm{j}\phi_n}$. The time dependent voltage is given by

$$\begin{aligned} v(z,t) &= \mathrm{Re}\left\{V(z)e^{\mathrm{j}\omega t}\right\} = v^+(z,t) + v^-(z,t) \\ &= |V_p|\cos(\omega t - \beta z + \phi_p) + |V_n|\cos(\omega t + \beta z + \phi_n) \end{aligned}$$

Since $\omega t - \beta z = \omega(t - z/v_{\mathrm{p}})$ we see that $v^+(z,t)$ is a wave propagating in the positive $z-$direction with speed v. In the same manner $v^-(z,t)$ is a wave propagating in the negative $z-$direction with speed v_{p}. The wavelength for the time harmonic waves is the shortest length $\lambda > 0$ for which $v^+(z,t) = v^+(z+\lambda,t)$ for all t. That gives $\beta\lambda = 2\pi$ or

$$\lambda = 2\pi/\beta = 2\pi v_{\mathrm{p}}/\omega = v_{\mathrm{p}}/f$$

The complex current satisfies the same equation as $V(z)$ and hence

$$I(z) = I^+(z) + I^-(z) = I_p e^{-\mathrm{j}\beta z} + I_n e^{\mathrm{j}\beta z} \tag{3.17}$$

We insert (3.16) and (3.17) in (3.15) and get

$$\frac{V^+(z)}{I^+(z)} = -\frac{V^-(z)}{I^-(z)} = Z_0$$

where $Z_0 = \sqrt{L/C}$ is the characteristic impedance of the transmission line, c.f., (3.12).

3.4.1 Input impedance

We assume a lossless line with length ℓ and characteristic impedance Z_0 with a load impedance Z_L at $z = \ell$. The line and load are equivalent to an input impedance $Z(0)$ at $\ell = 0$ where

$$Z(0) = \frac{V(0)}{I(0)} = Z_0\frac{V_p + V_n}{V_p - V_n} = Z_0\frac{1 + V_n/V_p}{1 - V_n/V_p} \tag{3.18}$$

To get the quotient V_n/V_p we use the boundary condition at $z = \ell$

$$Z_L = \frac{V(\ell)}{I(\ell)} = Z_0\frac{V_p e^{-\mathrm{j}\beta\ell} + V_n e^{\mathrm{j}\beta\ell}}{V_p e^{-\mathrm{j}\beta\ell} - V_n e^{\mathrm{j}\beta\ell}}$$

We get

$$\frac{V_n}{V_p} = \frac{Z_L - Z_0}{Z_L + Z_0}e^{-2\mathrm{j}\beta\ell}$$

and insert this into (3.18)

$$\begin{aligned} Z(0) &= Z_0 \frac{(Z_L + Z_0)e^{\mathrm{j}\beta\ell} + (Z_L - Z_0)e^{-\mathrm{j}\beta\ell}}{(Z_L + Z_0)e^{\mathrm{j}\beta\ell} - (Z_L - Z_0)e^{-\mathrm{j}\beta\ell}} \\ &= Z_0 \frac{Z_L \cos(\beta\ell) + \mathrm{j}Z_0 \sin(\beta\ell)}{Z_0 \cos(\beta\ell) + \mathrm{j}Z_L \sin(\beta\ell)} \end{aligned} \tag{3.19}$$

This is the complex input impedance in the frequency domain. Circuit theory would give $Z(0) = Z_L$ which in many cases is completely wrong, as will be seen from the examples below.

We see that at the load, $z = \ell$, the reflection coefficient in the frequency domain is in accordance with the corresponding coefficient in time domain, c.f., (3.14)

$$\Gamma = \frac{V^-(\ell)}{V^+(\ell)} = \frac{Z_L - Z_0}{Z_L + Z_0}$$

At a position $z < \ell$ the reflection coefficient gets a phase shift of $2\beta(\ell - z)$, i.e.,

$$\frac{V^-(z)}{V^+(z)} = \Gamma e^{-2\mathrm{j}\beta(\ell - z)} = \frac{Z_L - Z_0}{Z_L + Z_0} e^{-2\mathrm{j}\beta(\ell - z)}$$

Example 3.3
Matched line: If $Z_L = Z_0$ then $\Gamma = 0$ and $Z(0) = Z_0$ regardless of the length of the line and there are no waves propagating in the negative $z-$direction. We say that the impedance is matched to the line. ∎

Example 3.4
Shortened and open lines: An open line at $z = \ell$ has $Z_L = \infty$ and a shortened line at $z = \ell$ has $Z_L = 0$. The corresponding input impedances are

$$Z_L = \infty \quad \Rightarrow \quad Z(0) = -\mathrm{j}Z_0 \cot \beta\ell \tag{3.20}$$

$$Z_L = 0 \quad \Rightarrow \quad Z(0) = \mathrm{j}Z_0 \tan \beta\ell \tag{3.21}$$

The input impedance is purely reactive in both cases. This is expected since there is no dissipation of power in the line or in the load. ∎

Example 3.5
Quarter wave transformer: When the length of the line is a quarter of a wavelength long, $\ell = \lambda/4$, then $\beta\ell = \pi/2$ and

$$Z(0) = \frac{Z_0^2}{Z_L} \tag{3.22}$$

When $Z_L = \infty$ then $Z(0) = 0$ and when $Z_L = 0$ then $Z(0) = \infty$, which is opposite of what circuit theory predicts. ∎

Example 3.6
Matching a load by $\lambda/4$ transformer: Assume that we like to match a resistive load R_L to a lossless line with characteristic impedance Z_1. This is done by using a quarter wave transformer with characteristic impedance $Z_0 = \sqrt{Z_1 R_L}$. ∎

3.4.2 Standing wave ratio

At high frequencies it is difficult to determine the load impedance and the characteristic impedance by direct measurements. A convenient method to obtain these quantities is to measure the standing wave ratio (SWR). We then measure the amplitude, $|V(z)|$, along the line with an instrument that can register the rms voltages.

The standing wave ratio is the quotient between the largest and smallest value of $|V(z)|$ along the line

$$\text{SWR} = \frac{|V(z)|_{\max}}{|V(z)|_{\min}}$$

When the waves $V_p e^{-\mathrm{j}\beta z}$ and $V_n e^{\mathrm{j}\beta z}$ are in phase we get the maximum voltage and when they are out of phase we get the minimum voltage. Thus

$$\begin{aligned}
\text{SWR} &= \frac{|V_p| + |V_n|}{|V_p| - |V_n|} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \\
|\Gamma| &= \frac{\text{SWR} - 1}{\text{SWR} + 1} \\
|\Gamma| &= \frac{|V_n|}{|V_p|} = \left| \frac{Z_L - Z_0}{Z_L + Z_0} \right|
\end{aligned}$$

The distance Δz between two maxima is determined by

$$e^{2\mathrm{j}\beta\Delta z} = e^{\mathrm{j}2\pi}, \qquad \beta\Delta z = \pi$$

and hence $\Delta z = \lambda/2$.

3.4.3 Waves on lossy transmission lines in the frequency domain

When $R > 0$ and $G > 0$ the transmission line is lossy and some of the power we transport along the line is transformed to heat in the wires and in the material between the wires. Due to these power losses the waves are attenuated and decay exponentially along the direction of propagation. The losses are assumed to be quite small such that $R \ll \omega L$ and $G \ll 1/(\omega C)$.

From the general transmission line equations (3.15) we derive the equation for the voltage

$$\frac{d^2V(z)}{dz^2} - \gamma^2 V(z) = 0 \tag{3.23}$$

where

$$\gamma = \sqrt{(R + \mathrm{j}\omega L)(G + \mathrm{j}\omega C)} = \underline{\text{propagation constant}}. \tag{3.24}$$

The general solution to (3.23) is

$$V(z) = V_p e^{-\gamma z} + V_n e^{\gamma z}$$

The corresponding current is

$$I(z) = I_p e^{-\gamma z} + I_n e^{\gamma z} = \frac{1}{Z_0}(V_p e^{-\gamma z} - V_n e^{\gamma z})$$

where

$$Z_0 = \sqrt{\frac{R + \mathrm{j}\omega L}{G + \mathrm{j}\omega C}} \tag{3.25}$$

is the characteristic impedance. We can decompose the propagation constant in its real and imaginary parts

$$\gamma = \alpha + \mathrm{j}\beta$$

where α =attenuation constant and β =phase constant. With $\cos \omega t$ as phase reference the time domain expressions for a time harmonic wave are

$$v(z,t) = \mathrm{Re}\left\{V(z)e^{\mathrm{j}\omega t}\right\} = |V_p|e^{-\alpha z}\cos(\omega t - \beta z + \phi_p) + |V_n|e^{\alpha z}\cos(\omega t + \beta z + \phi_n)$$

where $V_p = |V_p|e^{\mathrm{j}\phi_p}$ and $V_n = |V_n|e^{\mathrm{j}\phi_n}$. Also in this case we can define a wave speed. In order to have a constant argument in $\cos(\omega t - \beta z + \phi_p)$ we must have $z = \omega t/\beta$+constant and this leads us to the definition of the phase speed

$$v_p = \frac{\omega}{\beta}$$

If we have a line with length ℓ and characteristic impedance Z_0, that is terminated by a load impedance Z_L, the input impedance is

$$Z(0) = Z_0 \frac{Z_L \cosh \gamma\ell + Z_0 \sinh \gamma\ell}{Z_L \sinh \gamma\ell + Z_0 \cosh \gamma\ell}$$

The derivation is almost identical to the one for the lossless line.

3.4.4 Distortion free lines

When we have losses the phase speed, attenuation constant and the characteristic impedance are all frequency dependent. If we send a pulse along such a transmission line the shape of the pulse changes. We say that the pulse gets distorted when it propagates. The distortion of pulses is a serious problem in all communication systems based on guided waves. Luckily enough we can get rid of the distortion if we can adjust L or C such that

$$\frac{R}{L} = \frac{G}{C}$$

Then

$$\gamma = \sqrt{LC}\sqrt{(R/L + \mathrm{j}\omega)(G/C + \mathrm{j}\omega)} = \sqrt{RG} + \mathrm{j}\omega\sqrt{LC}$$

and we get a line that is distortion free since the attenuation $\alpha = \sqrt{RG}$ and the phase speed $v_p = \omega/\beta = 1/\sqrt{LC}$ are frequency independent. The characteristic impedance of a distortion free line is the same as for a lossless line, i.e., $Z_0 = \sqrt{L/C}$. If we send a pulse along a distortion free line the amplitude of the pulse decreases exponentially with distance but its shape is unaffected.

Historical notes on distortion free lines

In the early times of telephone communication, distortion was a big problem. One could only transmit speech over short distances otherwise it would be too distorted. Also in telegraphy, where the Morse code was usually used, the transmission speed was limited by

the distorsion. It was Oliver Heaviside that realised that if one increases the inductance of the telephone lines the distortion is reduced. He wrote a paper on this in 1887 but the telegraphic companies ignored his results. It took some years before the american company AT&T rediscovered Heaviside's work and added inductances to their telephone lines. The inductance were coils that were placed with some distance apart along the lines. These coils are called Pupin coils due to their inventor M. I. Pupin. Today we often use optical fibers rather than copper wires for communication. In optical fibers distortion is also a major problem. It causes pulses to be broader when they propagate and this limits the bit rate of the cable.

The first atlantic cable for telegraphy was laid in 1858. After a month of operation the operator tried to increase the transmission speed by increasing the voltage. The cable was overheated and destroyed. In 1865 and 1866 two, more successful, cables were laid. The transmission rate was very limited for these early cables. The main reason was the resistance of the cables. The cable had only one wire since they used the sea water as the other conductor. The wire was made of copper and had a radius of approximately 1.6 mm. It was surrounded by an insulating cover. The total radius of the cable was 15 mm. Based on the parameters that is known for the cable we can estimate the line parameters to be $R = 2.2 \cdot 10^{-3}$ Ω/m, $L = 0.4$ μH/m, $C = 80$ pF/m and $G = 10^{-13}$ (Ωm^{-1}. With these parameters it is seen that the attenuation of a received signal increases exponentially with frequency. Already at 5 Hz the signal is attenuated 60 dB. Only the very low frequencies of the signal are transmitted and at such low frequencies $G \ll \omega C$ and $R \gg \omega L$. The voltage then satisfies the equation

$$\frac{\partial^2 v(z,t)}{\partial z^2} = RC\frac{\partial v(z,t)}{\partial t}$$

This is the diffusion equation. Assume that we apply a voltage at $z = 0$ that is a step function in time $v(0,t) = V_0 H(t)$, where $H(t)$ is the Heaviside step function and let $v(z,0) = 0$ for $z > 0$. The solution to this problem is well-known

$$v(z,t) = V_0 \int_0^t z\sqrt{\frac{RC}{4\pi(t-s)^3}}\exp\left(-\frac{z^2 RC}{4(t-s)}\right)\,\mathrm{d}s$$

The voltage at the receiving station in America is seen in Figure 3.14 for a unit step voltage at England. The problem for the receiver is that the current becomes very low, only 0.1 mA for $V_0 = 1$ V when the receiver is a shortage. Hence the signal is hard to detect unless the voltage at the transmitter is high. The communication speed was eight words per minute using Morse code which corresponds to approximately less than one sign per second.

3.5 Wave propagation in terms of $\boldsymbol{E}$ and $\boldsymbol{H}$

We have described the wave propagation in terms of voltage and current. We may treat the wave propagation in terms of the electromagnetic fields instead. In most cases the analysis is then more complicated since we have to use the vector fields $\boldsymbol{E}$ and $\boldsymbol{H}$ instead of the scalar quantities v and i. Another drawback is that it is hard to link the analysis to circuit theory concepts. Despite of this, it is of interest to analyze the wave propagation in terms of the electromagnetic fields. As we will see, there are cases where we are forced

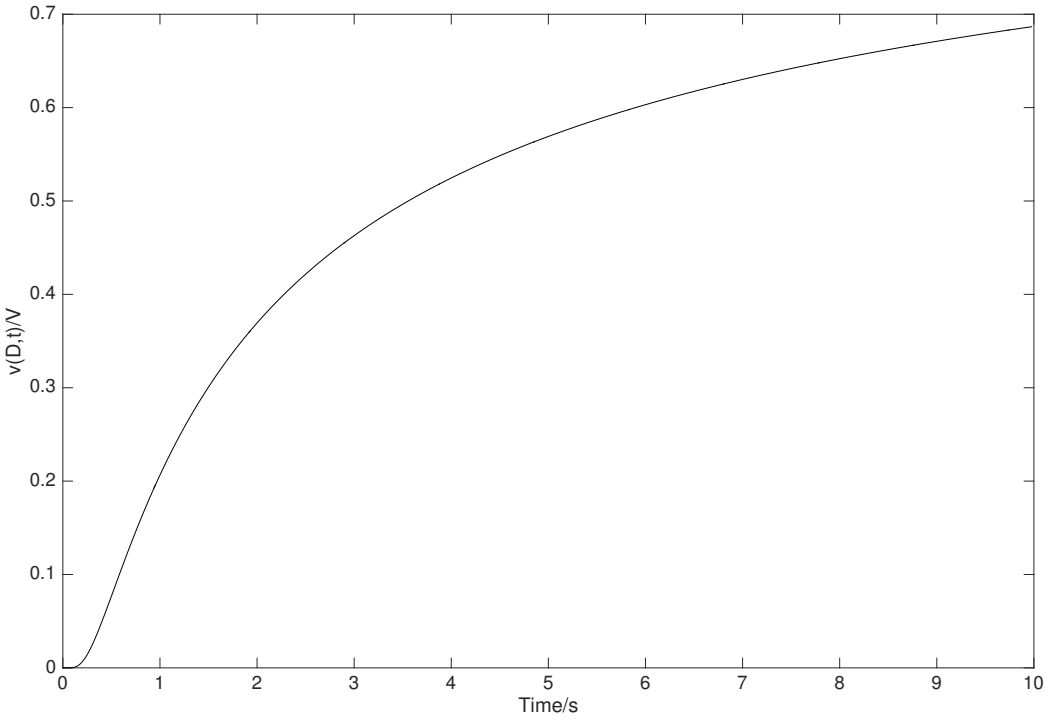


Figure 3.14: The received voltage $v(D,t)$ for a unit step voltage $v(0,t) = H(t)$ V, where $D = 4300$ km is the length of the Atlantic cable. The line parameters are $R = 2.2 \cdot 10^{-3}$ Ω/m, $L = 0.4$ μH/m, $C = 80$ pF/m and $G = 10^{-13}$ $(\Omega\text{m})^{-1}$.

to use electromagnetic theory. We first consider transmission lines for which the material between the conductors is homogenous. We also assume that the resistance per unit length is very small. The propagating wave is a transverse electromagnetic (TEM) wave, which means that the electric and magnetic fields lack $z-$component. In the frequency domain the current, voltage and the fields must have the same propagator $e^{\pm\gamma z}$.

$$\begin{aligned} \boldsymbol{E}(x,y,z) &= \boldsymbol{E}_T(x,y)e^{\pm\gamma z} \\ \boldsymbol{H}(x,y,z) &= \boldsymbol{H}_T(x,y)e^{\pm\gamma z} \end{aligned} \tag{3.26}$$

From the $z-$component of the induction law it follows that $\nabla \times \boldsymbol{E}_T(x,y) = \boldsymbol{0}$. It then exists a potential $\Psi(x,y)$ such that $\boldsymbol{E}_T(x,y) = -\nabla\Psi(x,y)$. In the region between conductors there is no charge density and Gauss law gives $0 = \nabla \cdot \boldsymbol{E}_T(x,y) = -\nabla^2\Psi(x,y)$. The Ampère's law shows that the relations between the electric and magnetic fields are the same as for a plane wave in free space. We conclude that the electric and magnetic fields are obtained from the following boundary value problem

$$\begin{aligned} &\nabla^2\Psi(x,y) = 0, \text{ between conductors} \\ &\Psi(x,y) = \begin{cases} V_1 & \text{on conductor 1} \\ V_2 & \text{on conductor 2} \end{cases} \end{aligned} \tag{3.27}$$

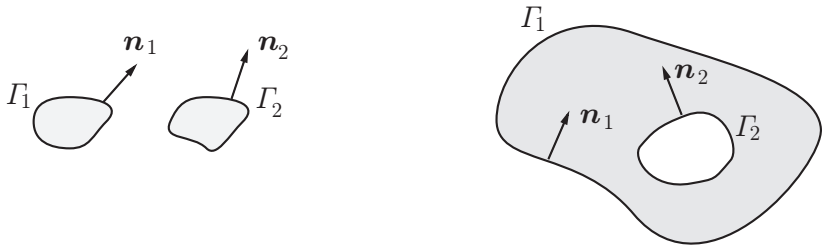


Figure 3.15: The cross-section of two types of transmission lines. The structure to the left is open and the structure to the right is of coaxial type where the electromagnetic field is confined to the region between the conductors. The curves Γ_1 and Γ_2 define the boundary of the conductors. The normal vectors $\hat{\boldsymbol{n}}_1$ and $\hat{\boldsymbol{n}}_2$ are directed out from the conductors.

We let $V = V_1 - V_2$ denote the voltage between the conductors. The electric and magnetic fields are obtained from

$$\boldsymbol{E}(x,y,z) = -e^{\pm\gamma z}\nabla\Psi(x,y)$$
$$\boldsymbol{H}(x,y,z) = \pm\mathrm{j}\frac{\gamma}{\omega\mu_0}\hat{\boldsymbol{z}}\times\boldsymbol{E}(x,y,z)$$

where the upper (lower) sign is for waves traveling in the negative (positive) $z-$direction.

3.6 Transmission line parameters

We first assume a transmission line with two conductors, with conductivity σ_c, imbedded in a homogenous dielectric material with relative permittivity ϵ and conductivity σ_d. The conductivity σ_d is assumed to satisfy $\sigma_d \ll \omega\epsilon_0\epsilon$, whereas $\sigma_c \gg \omega\epsilon_0$. The skin depth $\delta = \sqrt{\dfrac{2}{\omega\sigma_c\mu_0}}$ of the conductors is assumed to be much smaller than the diameters of the conductors. It is sufficient to solve the static problem (3.27) to obtain all four line parameters. We assume that conductor 1 has the larger potential, i.e., $V = V_1 - V_2 > 0$. From $\Psi(x,y)$ we get the charge per unit length as

$$\rho_\ell = \oint_{\Gamma_1}\rho_s\,\mathrm{d}\ell = -\oint_{\Gamma_2}\rho_s\,\mathrm{d}\ell = -\epsilon_0\epsilon\oint_{\Gamma_1}\hat{\boldsymbol{n}}\cdot\nabla\Psi(x,y)\,\mathrm{d}\ell$$

and the capacitance per unit length as

$$C = \frac{\rho_\ell}{V}$$

An alternative is to use the energy relation in (3.33). The inductance is obtained from the phase speed $v_\mathrm{p} = (\epsilon_0\epsilon\mu_0)^{-1/2} = (LC)^{-1/2}$

$$L = \frac{1}{Cv_\mathrm{p}^2} = \frac{\epsilon_0\epsilon\mu_0}{C}, \tag{3.28}$$

where we assume that the R and G are very small. An alternative is to obtain L from (3.33). We get the conductance from the relation

$$G = C\frac{\sigma_d}{\epsilon_0\epsilon} \tag{3.29}$$

To obtain the resistance we first relate the surface current density to the surface charge density ρ_S

$$\boldsymbol{J}_S = \hat{\boldsymbol{n}} \times \boldsymbol{H} = \eta^{-1}\hat{\boldsymbol{n}} \times (\hat{\boldsymbol{z}} \times \boldsymbol{E}) = \eta^{-1}\hat{\boldsymbol{z}}\,(\hat{\boldsymbol{n}} \cdot \boldsymbol{E}) = \hat{\boldsymbol{z}} v_{\mathrm{p}} \rho_S \tag{3.30}$$

where $v_{\mathrm{p}} = 1/\sqrt{\mu_0\epsilon_0\epsilon}$ is the phase speed and $\eta = \sqrt{\mu_0/\epsilon_0\epsilon}$ is the wave impedance. The dissipated power per unit length in the conductors is

$$P_d = \frac{1}{2} R |I|^2$$

We can express this power in terms of the surface currents and surface charge densities

$$P_d = \frac{1}{2} R_s \oint_{\Gamma_1+\Gamma_2} |\boldsymbol{J}_s|^2, \mathrm{d}ell = \frac{1}{2} v_{\mathrm{p}}^2 R_s \oint_{\Gamma_1+\Gamma_2} \rho_S^2 \,\mathrm{d}\ell$$

where R_S is the surface resistance

$$R_S = \frac{1}{\sigma_c \delta} = \sqrt{\frac{\omega\mu_0}{2\sigma_c}} \tag{3.31}$$

The expression for the surface resistance is derived in chapter 5. The resistance per unit length is thus given by

$$R = \frac{1}{|I|^2} R_s \oint_{\Gamma_1+\Gamma_2} |\boldsymbol{J}_s|^2 \,\mathrm{d}\ell = \frac{1}{|I|^2} v_{\mathrm{p}}^2 R_s \oint_{\Gamma_1+\Gamma_2} \rho_S^2 \,\mathrm{d}\ell \tag{3.32}$$

The current is either obtained from $|I| = Z_0^{-1}|V| = v_{\mathrm{p}}|\rho_\ell|$, or from integration over one of the conductors

$$|I| = \oint_{\Gamma_1} |\boldsymbol{J}_s| \,\mathrm{d}\ell = \oint_{\Gamma_2} |\boldsymbol{J}_s| \,\mathrm{d}\ell$$

Example 3.7
We exemplify by considering wave propagation in the positive $z-$direction in a coaxial cable. Consider a circular coaxial cable with metallic surfaces at $\rho = a$ and $\rho = b > a$. The region $a < \rho < b$ is filled with a homogenous material with relative permittivity ϵ. Due to the axial symmetry the potential only depends on the radial distance ρ and then

$$\begin{aligned}
&\nabla^2 \Psi(\rho) = \frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial\Psi(\rho)}{\partial\rho} = 0 \\
&\Psi(a) = V \\
&\Psi(b) = 0
\end{aligned}$$

We let the potential be zero on the outer conductor since in most applications the outer conductor of a coaxial cable is connected to ground. The solution is obtained by integrating the equation two times and using the boundary condition

$$\Psi(\rho) = V\frac{\ln(\rho/b)}{\ln(a/b)}$$

The corresponding electric and magnetic fields read

$$\begin{aligned}
\boldsymbol{E}(\rho) &= -\nabla\Psi(\rho) = -\frac{V}{\rho\ln(a/b)}\hat{\boldsymbol{\rho}} \\
\boldsymbol{H}(\rho) &= \frac{\beta}{\omega\mu_0}\hat{\boldsymbol{z}} \times \boldsymbol{E}(\rho) = -\frac{1}{\eta}\frac{V}{\rho\ln(a/b)}\hat{\boldsymbol{\phi}}
\end{aligned}$$

where $\eta = \sqrt{\mu_0/\epsilon_0\epsilon}$ is the wave impedance of the material between the conductors. We may also express the magnetic field in terms o the current along the inner conductor. The Ampère's law gives

$$\boldsymbol{H}(\rho) = \frac{I}{2\pi\rho}\hat{\boldsymbol{\phi}}$$

From these expressions we get the capacitance of the coaxial cable. Since the charge per unit length on the inner conductor is given by

$$\rho_\ell = \int_0^{2\pi} \hat{\boldsymbol{\rho}} \cdot \boldsymbol{D}(a) a \,\mathrm{d}\phi = -\epsilon_0\epsilon \int_0^{2\pi} \hat{\boldsymbol{\rho}} \cdot \nabla\Psi(a) a \,\mathrm{d}\phi = 2\pi \frac{\epsilon_0\epsilon V}{\ln(b/a)},$$

the capacitance per unit length is

$$C = 2\pi \frac{\epsilon_0\epsilon}{\ln(b/a)}$$

We can also obtain this expression from the energy relation

$$\begin{aligned} \frac{1}{4}C|V|^2 &= \frac{1}{4}\epsilon_0\epsilon \iint_\Omega |\boldsymbol{E}(\rho)|^2 \,\mathrm{d}S = \frac{\pi}{2}\epsilon_0\epsilon \int_a^b |\boldsymbol{E}(\rho)|^2 \rho \,\mathrm{d}\rho \\ &= \text{time average of electric energy density per unit length} \end{aligned} \tag{3.33}$$

The inductance L, conductance G and resistance R follow from (3.28), (3.29) and (3.32). The expressions are listed in subsection 3.6.1.

∎

Example 3.8

In the determination of the line parameters we assume that $E_z = 0$, but this is not really true when we have a resistance in the conductors. Due to Ohm's law we have $E_z = J_z/\sigma_c$ in the conductor where J_z is the current density. Since E_z is continuous across the surface it is not zero on the boundary, and hence not in the region between the conductors. In this example we show that the approximation $E_z = 0$ is relevant for most transmission lines.

We first observe that the surface current density on each conductor is given by (3.30) $\boldsymbol{J}_S = \hat{\boldsymbol{z}} v_\mathrm{p} \rho_S$. The current density $\boldsymbol{J}$ in the conductor is a function of the distance ξ to the surface i.e., locally it can be written as $\boldsymbol{J}(\xi)$. The current density at the surface $\xi = 0$ is given by[1]

$$J(0) = J_S/\delta = v_\mathrm{p}\rho_S/\delta$$

where $\delta = \sqrt{2/(\omega\mu_0\sigma_c)}$ is the skin depth of the conductors. The relation between $J(0)$ and the electric field on the surface of the conductor is

$$E_z = J(0)/\sigma_c$$

The relation between the normal component of the electric field on the surface of the conductor and the surface charge density is

$$E_n = \rho_s/(\epsilon_0\epsilon)$$

[1] This will be shown in chapter 5

Thus on the surface of the conductor the difference in magnitude between the transverse and $z-$components of the electric field is

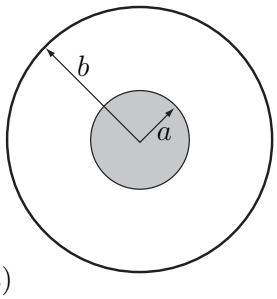
$$\frac{E_n}{E_z} = \sqrt{\frac{2\sigma_c}{\epsilon_0 \epsilon \omega}}$$

At 10 Ghz this value is on the order of 10^4 for a transmission line with metallic conductors. ∎

3.6.1 Explicit expressions

For the coaxial cable, the parallel plate line, and the circular two-wire line we can determine R, L, C and G analytically. The expressions are given below. For other types it is often convenient to use a numerical method, eg., the finite element method to obtain the parameters.

Coaxial cable



$$C = \frac{2\pi\epsilon_0\epsilon}{\ln(b/a)}$$

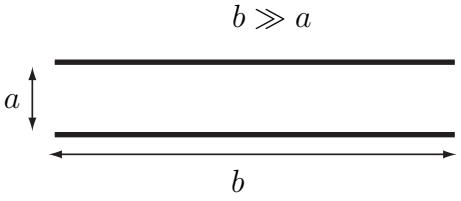
$$L = \frac{\mu_0}{2\pi} \ln \frac{b}{a}$$

$$Z_0 = \sqrt{\frac{L}{C}} = \frac{1}{2\pi}\sqrt{\frac{\mu_0}{\epsilon_0\epsilon}} \ln \frac{b}{a} \quad \text{(lossless)}$$

$$R = \frac{R_S}{2\pi}\left(\frac{1}{a} + \frac{1}{b}\right) \text{ where } R_S = 1/(\delta\sigma_c) = \sqrt{\frac{\omega\mu_0}{2\sigma_c}}$$

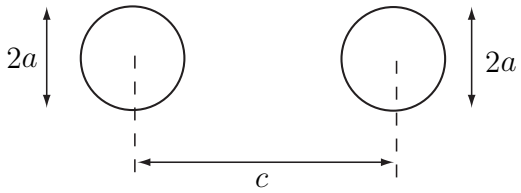
Parallel plate

Notice that $b \gg a$.



$$
\begin{aligned}
C &= \frac{\epsilon_0 \epsilon b}{a} \\
L &= \frac{\mu_0 a}{b} \\
Z_0 &= \sqrt{\frac{L}{C}} = \sqrt{\frac{\mu_0}{\epsilon_0 \epsilon}} \frac{a}{b} \quad \text{(lossless)} \\
R &= \frac{2R_S}{b}
\end{aligned}
\tag{3.34}
$$

Two-wire line



$$
C = \pi \epsilon_0 \epsilon \left(\ln \left(\frac{c}{2a} + \sqrt{\left(\frac{c}{2a}\right)^2 - 1} \right) \right)^{-1}
$$

$$
L = \frac{\mu_0}{\pi} \ln \left(\frac{c}{2a} + \sqrt{\left(\frac{c}{2a}\right)^2 - 1} \right)
$$

$$
Z_0 = \sqrt{\frac{L}{C}} = \frac{1}{\pi} \sqrt{\frac{\mu_0}{\epsilon_0 \epsilon}} \ln \left(\frac{c}{2a} + \sqrt{\left(\frac{c}{2a}\right)^2 - 1} \right) \quad \text{(lossless)}
$$

$$
R = \frac{R_S}{\pi a} \frac{c/(2a)}{\sqrt{(c/(2a))^2 - 1}}
$$

Notice that C and R for a two-wire line become infinite when $c \to 2a$, while L and $Z_0 \to 0$.

3.6.2 Determination of R, L, G, C with the finite element method

When the medium surrounding the wires is homogenous it is straightforward to determine the line parameters by numerical methods. In this book we use the commercial finite element method program COMSOL for the numerical simulations. In COMSOL there are two modules that we use, the ac/dc-module and the RF-module. The ac/dc-module is used in this section. We determine the line parameters for a two wire line and indicate all

steps that are to be done in COMSOL. The wires have radius $a = 1$ cm and the distance between their axes is $c = 3$ cm. The surrounding medium is air.

1. Start COMSOL and choose **2D> AC/DC-module>Electrostatics>Stationary**.
2. The default length unit in COMSOL is meter.
3. We must use a finite computational region and reduce the infinite region to a large rectangle. The rectangle has to be much larger than the size of the cross section of the wires. We first let the rectangle be a square 80×80 cm^2.
4. Next we draw the circular wires and put them in the center of the square, c.f., figure 3.16.
5. We make the domain between the large rectangle and the conductors our computational domain. In COMSOL this is done by creating the difference between the rectangle and the conductors.
6. We let the potential be zero on the large rectangle, 0.5 V on the left conductor and -0.5 V on the right.
7. Add **air** to **Materials** and let the computational region have air as material.
8. Start with the default mesh. Then check the result with a finer mesh to check the accuracy. Use **Mesh>size>fine** to get the finer mesh.
9. We obtain ρ_ℓ by integrating the surface charge density over the surface of the conductor with positive voltage. In COMSOL this is done by **Results>Derived results>Line integration>Surface charge density**.
10. Since we have one volt between the conductors the capacitance C equals the value of the integral. An alternative is to use the energy relation in (3.33). We then let COMSOL calculate the electric energy by using **Results>Derived results>Surface integral>Electric field norm**. Take the square of the norm and multiply with ε_0 (i.e., es.normE$^2 * 8.854e - 12$ in **expression**).
11. From C we get L and G from (3.28) and (3.29).
12. Finally we get the resistance by letting the FEM program numerically evaluate the integral in (3.32).

The geometry and solution are shown in figure 3.16. COMSOL gives the value $C = 2.885 \cdot 10^{-11}$ F/m with the default mesh. A finer resolution only improves the accuracy slightly. This is to be compared with the value obtained from the analytic expression which gives $C = 2.8901 \cdot 10^{-11}$ F/m. The numerical value is accurate and it shows that the finite element method is a suitable tool for this type of calculation. If we decrease the size of the surrounding square to 20×20 cm^2 we get the capacitance $C = 2.95 \cdot 10^{-11}$ F/m even with a refined mesh. We see from the example that it is important to check results obtained from COMSOL by comparing calculations with different sizes of the computational domain and different refinements of the mesh.

Example 3.9
In this example we let COMSOL calculate the capacitance of a plate transmission line with

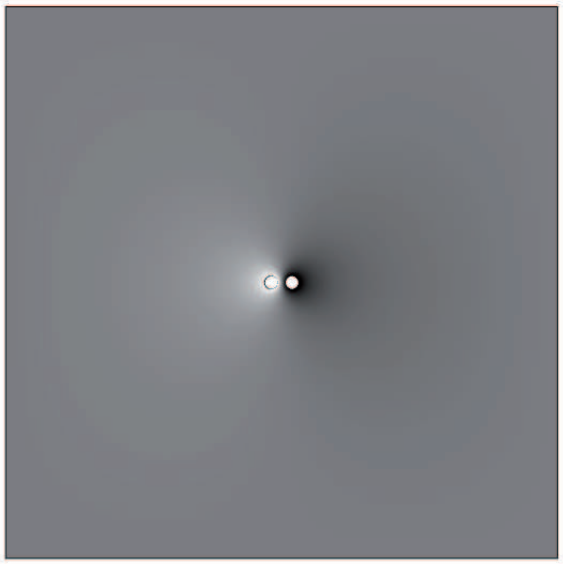


Figure 3.16: The geometry and voltage for the two-wire calculation in COMSOL. The right wire has the potential $V = -0.5$ V and the left the potential $V = 0.5$ V. The outer square has the size 80×80 cm^2 and potential 0 V.

different values of the quotient b/a between the width b of the plates and the distance a between the plates. The results obtained from the FEM calculations are presented in table 9. These are very accurate and we see that the error in the approximate analytic result is not negligible unless b/a is very large.

■

3.6.3 Transverse inhomogeneous region

If the material between the conductors is inhomogeneous in the transverse directions we need to modify the method above. We present an approximate method for determining the line parameters and the propagation constant. A more accurate method is given in Chapter 5. The two methods are compared and it is seen that the approximate method is sufficiently accurate for most applications.

It is assumed that the medium between the wires has an xy-dependent conductivity and permittivity. Such transmission lines are quite common in microwave circuits where microstrips on substrates are used. The waves are referred to as quasi TEM-waves since there are small $z-$components of the electric and magnetic fields. We first neglect the $z-$components and introduce a potential $\psi(x, y)$ for the electric field as $\boldsymbol{E}(x, y, z) = -\nabla_T \psi(\boldsymbol{\rho}) e^{-\gamma z}$. The potential satisfies the equation

$$\nabla_T \cdot \epsilon(\boldsymbol{\rho}) \nabla_T \psi(\boldsymbol{\rho}) = 0 \tag{3.35}$$

From this equation we obtain the capacitance and the conductance. Since the permeability is $\mu = 1$ everywhere, the magnetic field is very close to the magnetic field we would have

b/a	analytic	FEM
5	44.2 pF	59.2 pF
10	88.5 pF	105 pF
20	177 pF	195 pF
50	442 pF	464 pF
100	885 pF	909 pF

Table 3.1: The capacitance of a parallel plate line as a function of b/a. The approximate analytic values are from (3.34). The FEM values are from a 2D-solver and are very accurate.

if there is vacuum between the wires. Hence we can obtain the magnetic field and the inductance by using vacuum between the wires.

We use COMSOL to solve the problem. The line parameters are obtained from the following steps:

- We use **2D>AC/DC module>Electric currents (ec)>Frequency domain**.
- We draw the geometry of the wire. In order to have a finite computational domain we use an outer rectangle with length much larger than the distance between the wires and the diameter of the wires.
- We create the computational domain as the difference between the rectangle and the wires.
- The frequency is entered in **Study>Frequency domain**.
- The materials in the different regions are now defined. We may either find a predefined material or design our own. We design a material by **Materials>+ material** and add the values of conductivity and relative permittivity for the material.
- For convenience we set the potential $\psi = 0.5$ V on one of the conductors and -0.5 V on the other, while the rectangle is ground. There are cases when one of the conductors is the ground plane. Then it is convenient to let the bottom of the rectangle be the ground plane and let the other conductor have $\psi = V = 1$ V.
- We use two mesh sizes, eg., fine and finer, and compare the results to check the accuracy.
- The capacitance is obtained from the energy relation

$$\frac{1}{4}CV^2 = W_E = \text{time average of electric energy} = \frac{1}{4}\iint_\Omega \boldsymbol{E}\cdot\boldsymbol{D}\,\mathrm{d}S$$

where S is the computational domain and v is the voltage between the conductors. We set $V = 1$ V and calculate the electric energy by **Results>Derived results>Surface integral**. The result is the surface charge, but also the capacitance per unit length, since $V = 1$ V.

- The conductance per unit length G is obtained from the energy relation

$$\frac{1}{2}GV^2 = P_\sigma = \text{active Ohmic power loss} = \frac{1}{2}\iint_\Omega \frac{|\boldsymbol{J}(\boldsymbol{\rho})|^2}{\sigma(\boldsymbol{\rho})}\,\mathrm{d}S$$

  We let COMSOL calculate P_σ from **Results>Derived results>Surface integral> Resistive losses** (or **Total power dissipation density** or **Electromagnetic power loss density**). We now have both C and G and turn to L.

- We change the subdomain settings such that there is air in all regions in the computational domain, but keep the voltages. We solve the problem and calculate the capacitance in the same manner as above. Since there is vacuum between the wires we use the relation $c_0 = 1/\sqrt{LC}$ to get L, the inductance per unit length. Here c_0 is the speed of light in vacuum.

- The resistance per unit length R is obtained from the vacuum calculation in the following manner: The surface resistance of each conductor is $R_s = \sqrt{\omega\mu_0/(2\sigma_c)}$, where σ_c is the conductivity of the conductor. The resistance of each conductor is obtained from

$$R_k = \frac{R_S \oint_{\Gamma_k} |\boldsymbol{J}_S \cdot \hat{\boldsymbol{z}}|^2\,\mathrm{d}\ell}{(\oint_{\Gamma_k} \boldsymbol{J}_S \cdot \hat{\boldsymbol{z}}\,\mathrm{d}\ell)^2} = \frac{c_0^2 R_S \oint_{\Gamma_k} (\rho_S)^2\,\mathrm{d}\ell}{\left(c_0 \oint_{\Gamma_k} \rho_S\,\mathrm{d}\ell\right)^2} = R_S \frac{\oint_{\Gamma_k} (\rho_S)^2\,\mathrm{d}\ell}{\left(\oint_{\Gamma_k} \rho_S\,\mathrm{d}\ell\right)^2},\ k = 1, 2$$

  where we have used $\boldsymbol{J}_s \cdot \hat{\boldsymbol{z}} = |\boldsymbol{H}| = \eta_0^{-1}|\boldsymbol{E}| = c_0\rho_S$. The derivation of these relations is given in Chapter 5. The resistance R is the sum of the resistances of the two conductors, $R = R_1 + R_2$. The two integrals are evaluated by using the line integral of the normal component of the displacement field and the line integral of the square of the normal component of the displacement field, respectively.

- The propagation constant is calculated from $\gamma = \sqrt{(R + \mathrm{j}\omega L)(G + \mathrm{j}\omega C)}$ and the characteristic impedance from $Z = \sqrt{(R + \mathrm{j}\omega L)/(G + \mathrm{j}\omega C)}$.

Example 3.10

We consider a microstrip on a substrate. The substrate thickness is 0.5 mm, the permittivity is $\epsilon = 4$ and the conductivity is $\sigma = 0.01$ S/m. The width of the strip is 3 mm and the thickness 0.1 mm. The material in the strip is copper with conductivity $\sigma_c = 5.998 \cdot 10^7$ S/m. We use the methods described above to determine the line parameters, and from them the characteristic impedance and the propagation constant. The frequency is $f = 100$ MHz.

We choose the size 60 mm x 50 mm for the outer rectangle. The bottom side of the rectangle is the ground plane. The voltage of the microstrip is set to one volt in the computations. We first determine the inductance per unit length by solving Laplace equation when the substrate is replaced by vacuum. We let the microstrip have potential one volt and the other surfaces be ground. That gives the time average of the electric energy per unit length $W_e = 1.993 \cdot 10^{-11}$ N, which corresponds to the capacitance $C_0 = 4W_e/1^2 = 79.72$ pF/m and the distributed inductance $L = 1/(c_0^2 C_0) = 139.5$ nH/m. We then do the quasi-static calculation with the substrate. Again we use the potential of one volt on the microstrip and let the other surfaces be ground. The electric energy per unit

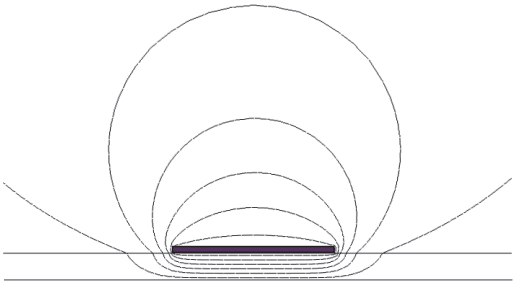


Figure 3.17: The voltage contour lines for the microstrip with substrate $\sigma = 0.01$ S/m and $\epsilon = 4$. Notice the discontinuity in the tangent of the lines at the surface of the substrate.

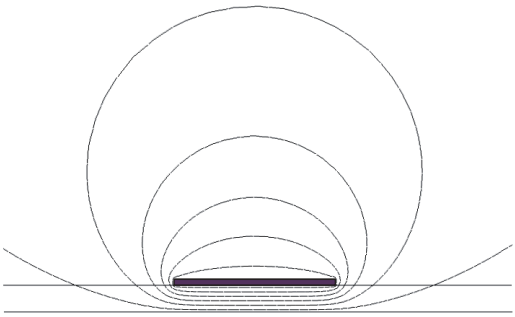


Figure 3.18: The voltage contour lines for the microstrip with vacuum substrate $\sigma = 0$ S/m and $\epsilon = 1$. The voltage contour lines are identical with the magnetic field lines.

length is in this case $W_e = 6.638 \cdot 10^{-11}$ N, corresponding to $C = 4W_e/1^2 = 265.5$ pF/m. The time average of the Ohmic losses in the substrate is $P_\sigma = 0.03471$ W/m and this gives the conductance per unit length $G = 2P_\sigma/1^2 = 0.0694$ S/m.

The surface resistance is $R_s = \sqrt{\omega\mu_0/(2\sigma_c)} = 0.00256\,\Omega$. The resistance of the line is obtained from the calculation where the substrate is replaced by vacuum. Then

$$R = \frac{R_S \oint |\boldsymbol{J}_S \cdot \hat{\boldsymbol{z}}|^2 \,\mathrm{d}\ell}{(\oint \boldsymbol{J}_S \cdot \hat{\boldsymbol{z}} \,\mathrm{d}\ell)^2} = \frac{c_0^2 R_S \oint_\Gamma (\rho_S)^2 \,\mathrm{d}\ell}{\left(c_0 \oint_\Gamma \rho_S \,\mathrm{d}\ell\right)^2} = R_S \frac{\oint_\Gamma (\rho_S)^2 \,\mathrm{d}\ell}{\left(\oint_\Gamma \rho_S \,\mathrm{d}\ell\right)^2}$$

where we have used $\boldsymbol{J}_s \cdot \hat{\boldsymbol{z}} = |\boldsymbol{H}| = \eta_0^{-1}|\boldsymbol{E}| = c_0\rho_S$. This gives $R = 0.601\,\Omega$/m. We can now calculate the propagation constant

$$\gamma = \sqrt{(R + \mathrm{j}\omega L)(G + \mathrm{j}\omega C)} = 0.7927 + \mathrm{j}3.90\,\mathrm{m}^{-1} \tag{3.36}$$

If we use the more accurate method presented in Chapter 5 we get $\gamma = 0.791 + \mathrm{j}3.908$. This indicates that the method presented here is accurate for microstrip problems. The

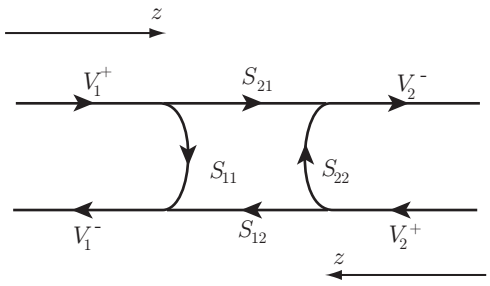


Figure 3.19: A flow diagram of the scattering matrix. The incident waves V_1^+ and V_2^+ are assumed to travel in the positive $z-$direction and hence two coordinate systems are used.

resistance is very small and can be neglected in the calculation of γ. If high accuracy is needed then the method in Chapter 5 is recommended. ∎

3.7 The scattering matrix S

The scattering matrix for a two-port is a very important quantity for measurements at high frequencies. It relates the voltages of a two-port. The voltages are decomposed into waves traveling to the ports and the waves traveling from the ports. When the characteristic impedances of the transmission lines at the two-ports are the same, the S-matrix is defined by

$$\begin{pmatrix} V_1^- \\ V_2^- \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} V_1^+ \\ V_2^+ \end{pmatrix} = [S] \begin{pmatrix} V_1^+ \\ V_2^+ \end{pmatrix} \tag{3.37}$$

where superindex $+$ indicates that the wave travels towards the two-port and $-$ that it travels from the two-port. Then S_{11} and S_{22} are the reflection coefficients at port 1 and 2, respectively, and the element S_{21} is the transmission coefficients from port 1 to port 2, and vice versa for S_{12}.

The transmitted power from an incident wave at port one to port two and the reflected power at port one are given by

$$\begin{aligned} P_t &= |S_{21}|^2 P_i \\ P_r &= |S_{11}|^2 P_i \end{aligned} \tag{3.38}$$

3.7.1 S-matrix when the characteristic impedance is not the same

Now assume a two-port where the transmission line connected to port 1 has characteristic impedance Z_1 and the transmission line that is connected to port 2 has characteristic impedance Z_2. We like to keep the expressions for the transmitted and reflected powers in (3.38) and for this reason we need to alter the definition of S_{21} and S_{12}

$$\begin{aligned} S_{21} &= \frac{V_2^+}{V_1^+}\sqrt{\frac{Z_1}{Z_2}} \\ S_{12} &= \frac{V_1^+}{V_2^+}\sqrt{\frac{Z_2}{Z_1}} \end{aligned} \tag{3.39}$$

The definitions of S_{11} and S_{22} are the same as in (3.37) and the expressions for the powers in (3.38) still hold.

3.7.2 Relation between S and Z

There is a relation between the scattering matrix and the impedance matrix. We let the two transmission lines that are connected to the two ports have the same characteristic impedance Z_0. We let $[U]$ denote the 2 by 2 unit matrix and utilize that

$$V = V^+ + V^- = ([U] + [S])V^+$$
$$I = I^+ + I^- = \frac{1}{Z_0}(V^+ - V^-) = \frac{1}{Z_0}([U] - [S])V^+$$

Since $V = [Z]I$ we get

$$\frac{1}{Z_0}[Z]([U] - [S]) = [U] + [S]$$

and

$$[Z] = Z_0([U] + [S])([U] - [S])^{-1}$$
$$[S] = (Z_0[U] + [Z])^{-1}([Z] - Z_0[U])$$

The impedance matrix for a reciprocal two-port is symmetric and that means that also the corresponding scattering matrix is symmetric. A lossless two-port satisfies

$$\begin{aligned}\mathrm{Re}\{V^t I^*\} &= \frac{1}{Z_0}\mathrm{Re}\{(V^+ + V^-)^t(V^+ - V^-)^*\} \\ &= \frac{1}{Z_0}\mathrm{Re}\{V^{+t}V^{+*} + V^{-t}V^{+*} - V^{+t}V^{-*} - V^{-t}V^{-*}\} = 0\end{aligned}$$

Since $V^{-t}V^{+*} - V^{+t}V^{-*}$ is purely imaginary and $V^{+t}V^{+*} - V^{-t}V^{-*}$ purely real we get

$$V^{+t}V^{+*} - V^{-t}V^{-*} = V^{+t}([U] - [S]^t[S]^*)V^{+*} = 0$$

This is valid for all input signals V^+ and then $[S]$ satisfies

$$[S]^t[S]^* = [U]$$

We see that the scattering matrix of a lossless two-port is a unitary matrix and hence its inverse is equal to its Hermite conjugate

$$[S]^{-1} = [S]^{*t}$$

If we use the tables in subsection 3.2.5 we can relate the S-matrix to the hybrid and cascade matrices. The S-matrix can be obtained from measurements.

3.7.3 Matching of load impedances

When a load is impedance matched to a transmission line, the transmission of the signal is optimized in some sense, often w.r.t. the transmitted power. If the signal is a single frequency then there are many ways to obtain an exact match. If the signal contains a band of frequencies and if the load is frequency dependent it is mostly not possible to match the load exactly in the whole frequency band. Instead we try to find an optimal matching.

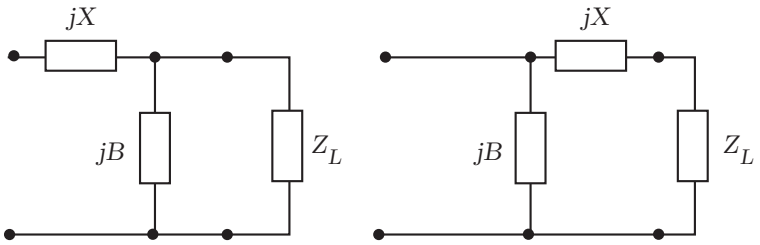


Figure 3.20: Two-ports for matching of Z_L.

Conjugate matching

In circuit theory we know that we maximize the power from a source with inner impedance Z_i by choosing the load impedance equal to the complex conjugate of the inner impedance, i.e., $Z_L = Z_i^*$. In many cases we have a given load impedance Z_L. In order to get maximum power to the load we have to use a lossless matching circuit.

Matching a load to a transmission line

A load impedance Z_L is impedance matched to a transmission line with characteristic impedance Z_0 only if $Z_L = Z_0$. Otherwise power is reflected at the load. We can match a load to a line by adding reactive components. It is quite easy to match a load at a single frequency. We have already seen that the resistive load impedance can be matched to a transmission line with a characteristic impedance Z_0 by using quarter wave transmission line with characteristic impedance, c.f., (3.22)

$$Z_1 = \frac{Z_0^2}{Z_L}$$

If the load has a reactance X we can get rid of the reactance by first adding a reactance $-X$ in series, or a susceptance $B = -X/\sqrt{R^2 + X^2}$ in parallel with Z_L, before we add the quarter wave transformer. Below we present two other methods that can be used. All three methods described here manage to match the impedance at one frequency. It is more difficult to get an impedance match for a band of frequencies and we do not discuss that in detail here.

Two-port matching

We can match the impedance Z_L to a characteristic impedance Z_0 by adding a capacitance and an inductance. The two possible matching two-ports are depicted in figure 3.20

For the two-port to the left we get

$$Z_0 = \mathrm{j}X + \frac{1}{\mathrm{j}B + (G_L + \mathrm{j}B_L)} \tag{3.40}$$

and for the two-port to the right we get

$$\frac{1}{Z_0} = \mathrm{j}B + \frac{1}{\mathrm{j}X + R_L + \mathrm{j}X_L} \tag{3.41}$$

where $Z_L = R_L + \mathrm{j}X_L$ and $Y_L = G_L + \mathrm{j}B_L$. We assume that the transmission line is lossless, i.e., Z_0 is real. By identifying the real and imaginary parts in the two expressions we can solve for X and B. The upper equation gives

$$X = \pm Z_0 \sqrt{\frac{1 - Z_0 G_L}{Z_0 G_L}}$$
$$B = \pm \sqrt{\frac{G_L(1 - Z_0 G_L)}{Z_0}} - B_L$$

We see that we must have $Z_0 G_L \leq 1$, i.e., $R_L + X_L^2/R_L \geq Z_0$ to satisfy the relations. This is satisfied if $R_L \geq Z_0$. Equation (3.41) gives

$$X = \pm \sqrt{R_L(Z_0 - R_L)} - X_L$$
$$B = \pm \frac{\sqrt{(Z_0 - R_L)/R_L}}{Z_0}$$

This requires that $R_L \leq Z_0$. Evidently we should use the two-port to the left when $R_L \geq Z_0$ and the two-port to the right when $R_L \leq Z_0$.

Example 3.11
Apparently we have two choices when $R_L = Z_0$. If we use the left circuit we see that $X = 0$ and $B = -B_L$ and if we choose the right circuit we get $B = 0$ and $X = -X_L$. ∎

3.7.4 Matching with stub

A short transmission line is sometimes called a stub. We can match the impedance Z_L to a characteristic impedance Z_0 by adding a shortened stub with length ℓ and characteristic impedance Z_0 at a distance d from the load. We can either connect the stub in series or in parallel. We have the two parameters d and ℓ to play with and that is enough to achieve impedance match.

The input impedance of the stub is purely reactive and is given by

$$Z_{st} = \mathrm{j}Z_0 \tan \beta\ell$$

The input impedance of the transmission line at the distance d from the load is

$$Z_{\text{in}} = Z_0 \frac{Z_L \cos(\beta d) + \mathrm{j}Z_0 \sin(\beta d)}{Z_0 \cos(\beta d) + \mathrm{j}Z_L \sin(\beta d)}$$

We first connect the stub in parallel with the load. We get impedance match when

$$\frac{1}{Z_0} = \frac{1}{Z_{\text{stub}}} + \frac{1}{Z_{\text{in}}}$$

This gives

$$1 = -\frac{\mathrm{j}}{\tan \beta\ell} + \frac{1 + \mathrm{j}z_L \tan \beta d}{z_L + \mathrm{j} \tan \beta d} \tag{3.42}$$

The real part of the equation gives the expression for d

$$\tan \beta d = \frac{x_L \pm \sqrt{r_L((r_L - 1)^2 + x_L^2)}}{r_L - 1}$$

We pick the smallest d that satisfies this equation. We then get ℓ from the imaginary part of (3.42)

$$\tan\beta\ell = \frac{r_L^2 + (x_L + \tan\beta d)^2}{r_L^2 \tan\beta d - (1 - x_L \tan\beta d)(x_L + \tan\beta d)}$$

This equation has infinitely many solutions and we pick the smallest positive solution. If Z_L is purely resistive the expression is simplified.

The other alternative is to connect the stub in series which gives

$$Z_0 = Z_{\text{stub}} + Z_{\text{in}}$$

from which ℓ and d are solved.

3.8 Smith chart

The Smith chart was first constructed in 1939 by P. H. Smith, a researcher at Bell Telephone Laboratories. It is still a useful tool in microwave technique. Many measurement instruments for microwaves, eg., the network analyzer, can display the frequency dependence of the input impedance of a device in a Smith chart. The Smith chart graph contains almost all information about the device. The Smith chart is also an important tool for matching loads to a transmission lines with known characteristic impedance. In this subsection we give a short introduction to the Smith chart and give some examples on how it can be used.

We assume a lossless transmission line with characteristic impedance Z_0 and load impedance Z. The reflection coefficient at the load is given by

$$\Gamma = \frac{V^-(\ell)}{V^+(\ell)} = \frac{V_n}{V_p} e^{2\mathrm{j}\beta\ell} = \frac{Z - Z_0}{Z + Z_0} = \frac{z - 1}{z + 1} \tag{3.43}$$

where we introduced the normalized impedance $z = Z/Z_0$. We now express z as a function of Γ

$$z = \frac{1 + \Gamma}{1 - \Gamma} \tag{3.44}$$

The relation in (3.43) is a so called Möbius transformation from the complex plane $z = r + \mathrm{j}x$ to the complex plane $\Gamma = \Gamma_r + \mathrm{j}\Gamma_i$. The inverse transformation is given by (3.44) and is a Möbius transformation from the Γ-plane to the $z-$plane. One can show that all Möbius transformations are conformal, i.e., they preserve angles. That means that if two lines cross each other at an angle α in the $z-$plane, then the images of the lines in the γ-plane cross each other at the same angle. The Möbius transformation maps a generalized circle in the $z-$plane to a generalized circle in the γ-plane, and vice versa. A generalized circle is a circle or a straight line. The mapping from the $z-$plane to the $\Gamma-$plane implies:

- A passive load, i.e., a load impedance $z = r + \mathrm{j}x$ with $r \geq 0$, is mapped on a point inside the circle $|\Gamma| \leq 1$. Load impedances with $|z| = \infty$ are mapped to the point $\Gamma = 1$.
- The imaginary axis, $z = \mathrm{j}x$, is mapped to a circle $|\Gamma| = |\mathrm{j}x - 1|/|\mathrm{j}x + 1| = 1$. Loads $z = \pm\mathrm{j}\infty$ are mapped on $\Gamma = 1$ and $z = 0$ is mapped on $\Gamma = -1$.
- The line $z = r + \mathrm{j}x$, $-\infty < x < \infty$ is mapped on a circle that goes through the points $\Gamma = 1$ and $\Gamma = (r - 1)/(r + 1)$ and has its centre on the real axis.

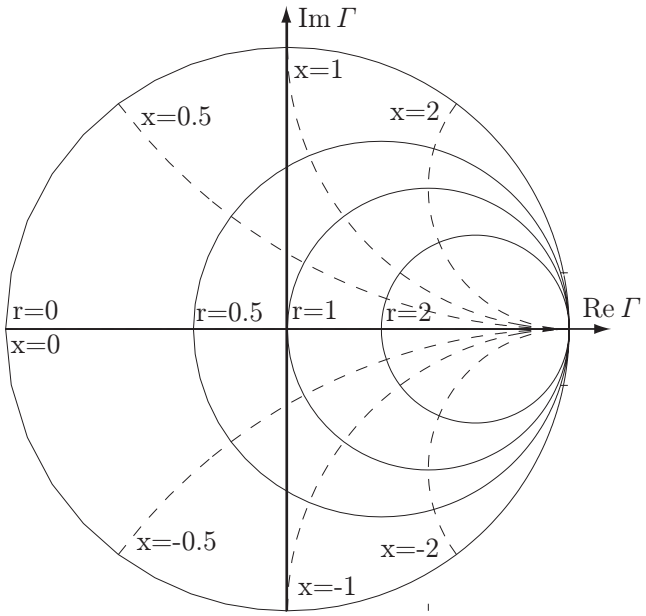


Figure 3.21: The Smith chart for the impedance z. When the admittance y is displayed x is exchanged for b and r for g.

- The positive real axis $z = r$, $r \geq 0$ is mapped on the line $\Gamma = \Gamma_r$, $-1 \leq \Gamma_r \leq 1$.
- The line $z = r + \mathrm{j}x$, $0 \leq r < \infty$ is mapped on the circle segment from the point $\Gamma = (\mathrm{j}x-1)/(\mathrm{j}x+1) = \exp(\mathrm{j}(\pi - 2\arctan x))$ to the point $\Gamma = 1$. The circle segment is perpendicular to the circle $\Gamma = 1$.
- An inductive load, i.e., a load impedance z with $x > 0$, is mapped on a point with $\Gamma_i > 0$.
- A capacitive load, i.e., a load impedance z with $x < 0$, is mapped on a point with $\Gamma_i < 0$.
- If an impedance z is mapped on the point Γ then z^* is mapped on Γ^*.

When we draw the mapping of the straight lines $z = r + \mathrm{j}x$, $-\infty < x < \infty$ and $z = r + \mathrm{j}x$, $0 \leq r < \infty$ in the $\Gamma-$plane we get the Smith-chart. Some important properties of the Smith chart are:

1. A matched impedance load has $Z = Z_0$, i.e., $z = 1$, and is mapped on $\Gamma = 0$.
2. If the reflection coefficient at the load, i.e., at $z = \ell$, is $\Gamma(\ell) = \Gamma_0$ then at a point $z_0 < \ell$ it is
$$\Gamma(z_0) = \Gamma_0 e^{-2\mathrm{j}\beta(\ell - z_0)}$$
Hence we move clock wise an angle $2\beta(\ell - z_0) = 4\pi(\ell - z_0)/\lambda$ along the circle with radius $|\Gamma_0|$ in the Smith chart when the distance to the load increases by a distance $(\ell - z_0)$.
3. When we move quarter of a wavelength, $(\ell - z_0) = \lambda/4$, towards the generator we rotate an angle $2\beta\lambda/4 = \pi$ clock wise. Hence the reflection coefficient changes sign, $\Gamma(\ell - \lambda/4) = -\Gamma_0$.

4. The admittance $y = 1/z$ is on the same circle with radius $|\Gamma|$ as z but at the point rotated 180° around $\Gamma = 0$. Then the values x and r in the Smith chart should be read as the normalized susceptance b and the normalized conductance g.

3.8.1 Matching a load by using the Smith chart

If we only are interested to match an impedance $Z_L = R_L + \mathrm{j}X_L$ to a transmission line with characteristic impedance Z_0 for a specified frequency, the Smith chart is an efficient and fast tool. We can do it in a large number of ways. The matching can be seen as a game where the goal is to obtain an input impedance $Z_{\text{in}} = Z_0$. We start at the position $z_L = Z_L/Z_0$ in the Smith chart and try to reach $z_{\text{in}} = 1$. The rules for this game are as follows:

1. We are allowed to rotate clockwise on a circle with center at $\Gamma = 0$. An angle ϕ of clockwise rotation is related to a translation Δz towards the generator as $\phi = 2\beta\Delta z = 4\pi\Delta z/\lambda$.
2. We may move along curves with r constant in the $z-$diagram and curves with g constant in the $y-$diagram. A move along a curve where r is constant corresponds to adding a reactance x in series with the current input impedance and a move along a curve with g constant corresponds to adding a susceptance b in parallel with the current input impedance.

We cannot add resistors or conductors since they consume power.

3.8.2 Frequency sweep in the Smith chart

Assume an input impedance $Z_{\text{in}}(f)$ that is a complex function of the frequency f. When $Z_{\text{in}}(f)$ is plotted as a function of frequency in the Smith chart the graph contains all information that is needed.

3.9 $z-$dependent parameters

We now consider a transmission line where the line parameters depend on z. We still let the line occupy the distance $0 < z < \ell$. By solving a system of first order ordinary differential equations we can determine the cascade matrix $[K]$. We start with the line equations

$$\frac{d}{dz}\begin{pmatrix} V(z) \\ I(z) \end{pmatrix} = -\begin{pmatrix} 0 & R(z) + \mathrm{j}\omega L(z) \\ G(z) + \mathrm{j}\omega C(z) & 0 \end{pmatrix}\begin{pmatrix} V(z) \\ I(z) \end{pmatrix} = -[D(z)]\begin{pmatrix} V(z) \\ I(z) \end{pmatrix} \tag{3.45}$$

The cascade matrix from 0 to z is given by

$$\begin{pmatrix} V(0) \\ I(0) \end{pmatrix} = [K(z)]\begin{pmatrix} V(z) \\ I(z) \end{pmatrix} \tag{3.46}$$

We notice that this is in accordance with (3.4), since the current $I(z)$ is directed outwards at the positive pole. We know that the boundary condition for the cascade matrix is $[K(0)] = [U]$. If we differentiate (3.46) wrt z we get

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = [K'(z)]\begin{pmatrix} V(z) \\ I(z) \end{pmatrix} + [K(z)]\frac{d}{dz}\begin{pmatrix} V(z) \\ I(z) \end{pmatrix}$$

We insert (3.45) to get the equation for the cascade matrix

$$[K'(z)] = [K(z)][D(z)] \tag{3.47}$$

If the matrix $[D(z)]$ commute with $[D(z']$, i.e., if $[D(z)][D(z')] = [D(z')][D(z)]$, then the formal solution reads

$$[K(z)] = e^{\int_0^z [D(z')]dz'}$$

The exponential function of a matrix is defined by the Maclaurin expansion

$$e^{[A]} = \sum_{n=0}^{\infty} \frac{1}{n!}[A]^n$$

where $[A]^0 = [U]$ is the identity matrix. There are many ways to calculate the exponent of a matrix. One way is to use the Caley-Hamilton theorem, see example.

Example 3.12
For a line $0 \leq z \leq \ell$ with constant material parameters the matrix $[D]$ is constant. The cascade matrix is then given by

$$[K(z)] = e^{[D]z}$$

∎

Example 3.13
Caley-Hamilton theorem: The Caley-Hamilton theorem says that every square matrix satisfies its own characteristic equation. We first prove the theorem and then use it to determine the exponential function of a square matrix.

The characteristic equations shows up when we determine the eigenvalues of a matrix. If $[A]$ is a $n \times n$ matrix the eigenvalues λ and eigenvectors $\boldsymbol{b}$ are determined by

$$[A]\boldsymbol{b} = \lambda\boldsymbol{b} \Rightarrow ([A] - \lambda[U])\boldsymbol{b} = \boldsymbol{0}$$

In order to have non-trivial solutions the determinant of $[A] - \lambda[U]$ must be zero

$$\det([A] - \lambda[U]) = 0$$

That leads to the polynomial equation of grade n for λ

$$a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots a_0 = 0$$

If all solutions λ_k, $k = 1, \ldots n$, to this equation are distinct we can prove the Caley-Hamilton theorem. Let $\boldsymbol{b}_k$ be the eigenvector corresponding to the eigenvalue λ_k. We can write the characteristic equation as

$$Char(\lambda) = \prod_{k=1}^{n}(\lambda - \lambda_k) = 0$$

Now form

$$Char([A]) = \prod_{k=1}^{n}([A] - \lambda_k[U])$$

Then for $1 \leq p \leq n$

$$Char([A])\boldsymbol{b}_p = \prod_{k=1}^{n}([A] - \lambda_k[U])\boldsymbol{b}_p = [0]$$

since the factor number p in the product gives zero. Hence $Char([A])\boldsymbol{b} = \mathbf{0}$ for every $\boldsymbol{b}$ that is a linear combination of eigenvectors. Since the eigenvectors are linearly independent it follows that $Char([A]) = [0]$ and the theorem is proven.

We use the theorem to write $e^{[A]}$ as a polynomial of finite grade of $[A]$. Every polynomial $f(\lambda)$ can be decomposed as:

$$f(\lambda) = Char(\lambda)q(\lambda) + p(\lambda)$$

where $q(\lambda)$ and $p(\lambda)$ are polynomials and where $p(\lambda)$ is a polynomial with grade less than n. If we make this decomposition for the exponential function we see that

$$e^{[A]} = Char([A])q([A]) + p([A]) = p([A])$$

It is now sufficient to determine the polynomial p of grade less than n to determine $e^{[A]}$. The polynomial p is determined by

$$e^{\lambda_k} = p(\lambda_k) = p_0 + p_1\lambda_k + \ldots p_{n-1}\lambda_k^{n-1}\,, k = 1, 2 \ldots n$$

This is a system of equations with n equations and n unknowns. ∎

3.9.1 Solution based on propagators

We again consider an inhomogeneous transmission line on the section $0 < z < \ell$. At $z = 0$ our line is connected to a homogeneous line $z < 0$ with characteristic impedance Z_0. At $z = \ell$ we have connected a load impedance Z_b. The inhomogeneous line is excited by an incident voltage wave $V^+(0)$. For $0 < z < \ell$ we define $V^{\pm}(z)$ by

$$\begin{pmatrix} V^+(z) \\ V^-(z) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & Z(z) \\ 1 & -Z(z) \end{pmatrix}\begin{pmatrix} V(z) \\ I(z) \end{pmatrix} = [A]\begin{pmatrix} V(z) \\ I(z) \end{pmatrix}$$

We introduce the propagators $g^{\pm}(z)$ such that

$$\begin{pmatrix} V^+(z) \\ V^-(z) \end{pmatrix} = \begin{pmatrix} g^+(z) \\ g^-(z) \end{pmatrix} V^+(0) \tag{3.48}$$

This relation is differentiated w.r.t. z

$$\frac{d}{dz}\begin{pmatrix} V^+(z) \\ V^-(z) \end{pmatrix} = \frac{d}{dz}\begin{pmatrix} g^+(z) \\ g^-(z) \end{pmatrix} V^+(0)$$

and then we utilize (3.45) and (3.48) and obtain

$$-[P(z)]\begin{pmatrix} g^+(z) \\ g^-(z) \end{pmatrix} V^+(0) = \frac{d}{dz}\begin{pmatrix} g^+(z) \\ g^-(z) \end{pmatrix} V^+(0)$$

where $[P(z)] = [A(z)][D(z)][A(z)]^{-1} - \dfrac{d[A(z)]}{dz}[A(z)]^{-1}$. Thus the propagators satisfy the same equations as $V^{\pm}(z)$

$$\frac{d}{dz}\begin{pmatrix} g^+(z) \\ g^-(z) \end{pmatrix} = -[P(z)]\begin{pmatrix} g^+(z) \\ g^-(z) \end{pmatrix} \tag{3.49}$$

The boundary conditions for $g^{\pm}(z)$ are $g^+(0) = 1$ and $g^-(\ell) = \Gamma(\ell)g^+(\ell) = \frac{Z_b - Z(\ell)}{Z_b + Z(\ell)}g^+(\ell)$.

Problems in Chapter 3

3.1 a) The time-harmonic voltage is $v(t) = V_0 \cos(\omega t + \pi/4)$. Determine the corresponding complex voltage V.

b)The complex voltage is $V = V_0(1 - \mathrm{j})/\sqrt{2}$ where V_0 is real. Determine $v(t)$.

3.2 Draw the equivalent circuit for a two-port based on the

a) impedance matrix

b) admittance matrix

c) hybrid matrix

b) the inverse hybrid matrix

3.3 Derive the relation between $[Z]$ and $[H]$ for a general two-port.

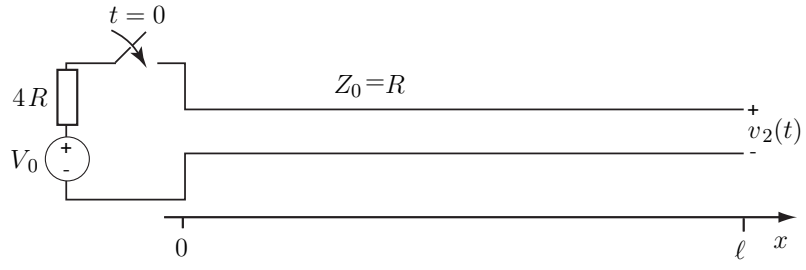
3.4 Show that Z and Y are symmetric matrices for a reciprocal N-port.

3.5 Show that the impedance matrix for a lossless N-port is purely imaginary.

3.6 A lossy transmission line has the following data at 100 MHz:

$$Z_0 = 50\,\Omega \text{ (real)} \quad \alpha = 10^{-3}\,\mathrm{m}^{-1} \quad \beta = 0{,}95\pi\,\mathrm{m}^{-1}$$

Determine L, C, R and G at 100 MHz.



3.7

A lossless transmission line has the length ℓ. The line has the characteristic impedance R and the phase speed v_p. In the left end the line is connected to an ideal voltage source in series with a resistance $4R$, see figure. The other end of the line is open. That means that $i(\ell, t) = 0$. The voltage source gives the voltage

$$\begin{aligned} v(t) &= 0 \qquad & t < 0 \\ v(t) &= V_0 \qquad & t \geq 0 \end{aligned}$$

where V_0 is constant voltage. We are interested in the voltage $v_2(t)$ in the right end of the line. The source and the resistor are connected to the line via short wires.

a) Determine $v_2(t)$ when $t \to +\infty$.

b) Determine $v_2(\ell/2v_\mathrm{p})$.

c) Determine $v_2(3\ell/2v_\mathrm{p})$.

d) Determine $v_2(5\ell/2v_\mathrm{p})$.

e) Determine $v_2(7\ell/2v_\mathrm{p})$.

3.8 A lossy transmission line is terminated by a matched load. The voltage along the line is measured at two positions 20 m apart.The result was 2,8 V and 2,1 V.

a) Determine the attenuation constant α.

b) Determine the attenuation in dB/km.

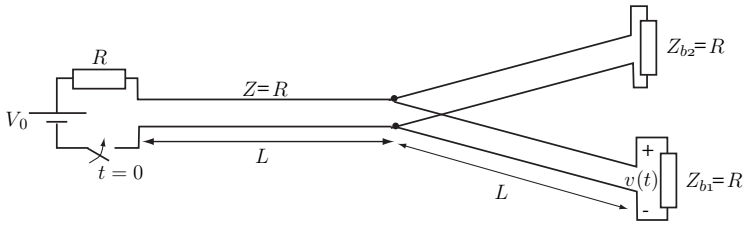
3.9 A lossless transmission line has the length 30 m. The input impedance of the line was measured when the other end was shortened, and when it was open. When it was open the input impedance was j360 V/A and when it was shortened the impedance was j10 V/A. The wavelength along the line was larger than 1 km. Determine the characteristic impedance Z_0 and the phase coefficient β.

3.10 An antenna with the purely resistive impedance 300 Ω is to be matched to a coaxial cable with the characteristic impedance 60 Ω. For this purpose a quarter wave transmission line is used. The quarter wave line consists of a coaxial cable with the relative permittivity $\epsilon = 2$ between the conductors. Determine the length ℓ and the characteristic impedance Z_0 of the quarter wave line when the frequency is 200 MHz.

3.11 A lossless transmission line has the characteristic impedance $Z_0 = 60\,\Omega$. One end is connected to a load resistance $R_b = 180\,\Omega$. Determine the reflection coefficient at the load Γ and the standing wave ratio SWR.

3.12 A lossless line with the characteristic impedance $Z_0 = 50\,\Omega$ has a resistive load R_b. The standing wave ratio is SWR= 3. Determine R_b.

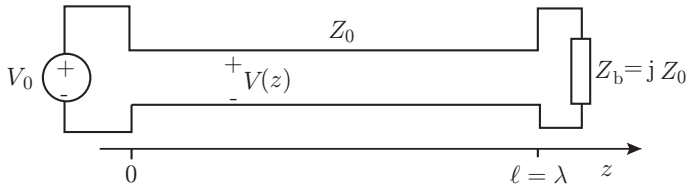
3.13 A lossless transmission line with characteristic impedance $Z_0 = 60\,\Omega$ is terminated by a load impedance $Z_b = (60 + \mathrm{j}60)\,\Omega$. The length of the line is $\lambda/8$, where λ is the wavelength along the line. Determine the input impedance of the line.



3.14

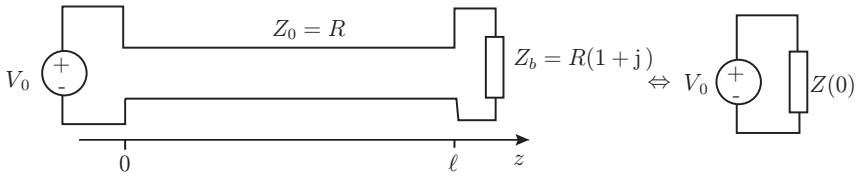
a) A voltage source with output resistance R has the open circuit voltage V_0. At time $t = 0$ the source is connected to a circuit that consists of three lossless transmission lines, see figure. The three transmission lines are identical and has the length L, the characteristic impedance $Z = R$ and the phase speed v_p. Determine the voltage $v(t)$ over the load Z_{b1} for all times.

b) Assume that the three transmission lines are connected to a time harmonic voltage source. The three lines can then be replaced by an equivalent impedance Z_in. Determine Z_in if the frequency is chosen such that each of the lines is a quarter of a wave length long.



3.15

The figure depicts a lossless transmission line with an ideal voltage source and a purely reactive load. The line is one wavelength long. The reactance of the load equals the characteristic impedance of the line. Determine the $z-$values in the interval $[0,\, 2\pi/\beta]$ for which the amplitude of $v(z,t)$ has its maximum.



3.16

The figure shows a lossless line with the load impedance $Z_b = R(1 + j)$. The characteristic impedance of the line is R. It is possible to chose the length ℓ such that the input impedance $Z(0)$ is real.

a) Determine the values of $\beta\ell$ in the interval $0 < \beta\ell < \pi$ for which $Z(0)$ is real.

b) Determine for each of the $\beta\ell$-value in a) the corresponding value of $Z(0)$.

Summary of chapter 3

Transmission lines

Time domain line equations

$$\boxed{\begin{aligned} -\frac{\partial v(z,t)}{\partial z} &= Ri(z,t) + L\frac{\partial i(z,t)}{\partial t} \\ -\frac{\partial i(z,t)}{\partial z} &= Gv(z,t) + C\frac{\partial v(z,t)}{\partial t} \end{aligned}}$$

Frequency domain line equations

$$\boxed{\begin{aligned} -\frac{dV(z)}{dz} &= (R + \mathrm{j}\omega L)I(z) \\ -\frac{dI(z)}{dz} &= (G + \mathrm{j}\omega C)V(z) \end{aligned}}$$

Characteristic impedance

$$\boxed{Z_0 = \begin{cases} \sqrt{\dfrac{L}{C}} & \text{lossless line} \\ \sqrt{\dfrac{R + \mathrm{j}\omega L}{G + \mathrm{j}\omega C}} & \text{lossy line} \end{cases}}$$

Reflection coefficient at the load

$$\boxed{\Gamma = \frac{v_r(\ell, t)}{v_i(\ell, t)} = \frac{R_L - Z_0}{R_L + Z_0}}$$

Input impedance

$$\boxed{Z(0) = \begin{cases} Z_0 \dfrac{Z_L \cos(\beta\ell) + \mathrm{j}Z_0 \sin(\beta\ell)}{Z_0 \cos(\beta\ell) + \mathrm{j}Z_L \sin(\beta\ell)}, & \text{lossless line} \\ Z_0 \dfrac{Z_L \cosh\gamma\ell + Z_0 \sinh\gamma\ell}{Z_L \sinh\gamma\ell + Z_0 \cosh\gamma\ell} & \text{lossy line} \end{cases}}$$

Chapter 4

Electromagnetic fields with a preferred direction

In this chapter we decompose an arbitrary vector field in the longitudinal component along the $z-$direction and the transverse vector in the x-y-plane. We apply this decomposition to the Maxwell equations and analyze the solutions to these equations for a geometry that is constant in the $z-$direction. The equations that are derived in this chapter form the basis for the following chapters.

From now on we use the time dependence $e^{-\mathrm{i}\omega t}$ in contrast to the transmission line chapter where we used $e^{\mathrm{j}\omega t}$. The reason is that most literature on waveguides uses this convention.

4.1 Decomposition of vector fields

An arbitrary vector field $\boldsymbol{F}(\boldsymbol{r})$ can always be decomposed in two perpendicular vectors[1]. We let one component be along the $z-$axis and the other in the xy-plane. A similar decomposition is used for the ∇-operator. We introduce the notations:

$$\begin{cases} \nabla = \nabla_T + \hat{\boldsymbol{z}}\dfrac{\partial}{\partial z} \\ \boldsymbol{F}(\boldsymbol{r}) = \boldsymbol{F}_T(\boldsymbol{r}) + \hat{\boldsymbol{z}}F_z(\boldsymbol{r}) \end{cases}$$

where $\boldsymbol{F}_T$ denotes the vector in the xy-plane. The two components of the vector $\boldsymbol{F}$ are uniquely determined and are obtained as

$$\begin{cases} F_z(\boldsymbol{r}) = \hat{\boldsymbol{z}} \cdot \boldsymbol{F}(\boldsymbol{r}) \\ \boldsymbol{F}_T(\boldsymbol{r}) = \boldsymbol{F}(\boldsymbol{r}) - \hat{\boldsymbol{z}}\left(\hat{\boldsymbol{z}} \cdot \boldsymbol{F}(\boldsymbol{r})\right) = \boldsymbol{F}(\boldsymbol{r}) - \hat{\boldsymbol{z}}F_z(\boldsymbol{r}) = \hat{\boldsymbol{z}} \times \left(\boldsymbol{F}(\boldsymbol{r}) \times \hat{\boldsymbol{z}}\right) \end{cases}$$

In the last equality we used the BAC-CAB-rule, $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B}(\boldsymbol{A} \cdot \boldsymbol{C}) - \boldsymbol{C}(\boldsymbol{A} \cdot \boldsymbol{B})$.

The $z-$component of a vector is called the *longitudinal* component and the xy-component the *transverse* component. We use a related decomposition of the position vector $\boldsymbol{r}$

$$\boldsymbol{r} = \hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y + \hat{\boldsymbol{z}}z = \boldsymbol{\rho} + \hat{\boldsymbol{z}}z$$

[1]The dependence of ω or t is not written explicitly in the argument of the fields in this chapter

Since ∇_T and $\hat{\boldsymbol{z}}$ are perpendicular we decompose the rotation of a vector as

$$\begin{aligned}\hat{\boldsymbol{z}}\cdot(\nabla\times\boldsymbol{F}(\boldsymbol{r})) &= \hat{\boldsymbol{z}}\cdot\left[\left(\nabla_T+\hat{\boldsymbol{z}}\frac{\partial}{\partial z}\right)\times(\boldsymbol{F}_T(\boldsymbol{r})+\hat{\boldsymbol{z}}F_z(\boldsymbol{r}))\right]\\ &= \hat{\boldsymbol{z}}\cdot(\nabla_T\times\boldsymbol{F}_T(\boldsymbol{r})) = -\nabla_T\cdot(\hat{\boldsymbol{z}}\times\boldsymbol{F}_T(\boldsymbol{r}))\end{aligned} \tag{4.1}$$

In the last equality we applied a cyclic permutation. The transverse component of the rotation is

$$\begin{aligned}&\nabla\times\boldsymbol{F}(\boldsymbol{r})-\hat{\boldsymbol{z}}\,(\hat{\boldsymbol{z}}\cdot(\nabla\times\boldsymbol{F}(\boldsymbol{r})))\\ &\quad=\left[\left(\nabla_T+\hat{\boldsymbol{z}}\frac{\partial}{\partial z}\right)\times(\boldsymbol{F}_T(\boldsymbol{r})+\hat{\boldsymbol{z}}F_z(\boldsymbol{r}))\right]-\hat{\boldsymbol{z}}\,(\hat{\boldsymbol{z}}\cdot(\nabla_T\times\boldsymbol{F}_T(\boldsymbol{r})))\\ &\quad=\nabla_T\times\hat{\boldsymbol{z}}F_z(\boldsymbol{r})+\hat{\boldsymbol{z}}\frac{\partial}{\partial z}\times\boldsymbol{F}_T(\boldsymbol{r})=\hat{\boldsymbol{z}}\times\frac{\partial}{\partial z}\boldsymbol{F}_T(\boldsymbol{r})-\hat{\boldsymbol{z}}\times\nabla_T F_z(\boldsymbol{r})\end{aligned} \tag{4.2}$$

since $\nabla_T\times\boldsymbol{F}_T(\boldsymbol{r})=\hat{\boldsymbol{z}}\,(\hat{\boldsymbol{z}}\cdot(\nabla_T\times\boldsymbol{F}_T(\boldsymbol{r})))$.

The decompositions we have described so far are valid for all vector fields. We now apply them to the fields in the Maxwell equations.

4.2 Decomposition of the Maxwell field equations

The Maxwell field equations for a dispersive material are

$$\begin{cases}\nabla\times\boldsymbol{E}(\boldsymbol{r},\omega)=\mathrm{i}\omega\mu_0\mu(\boldsymbol{r},\omega)\boldsymbol{H}(\boldsymbol{r},\omega)\\ \nabla\times\boldsymbol{H}(\boldsymbol{r},\omega)=-\mathrm{i}\omega\epsilon_0\epsilon(\boldsymbol{r},\omega)\boldsymbol{E}(\boldsymbol{r},\omega)\end{cases}$$

We use the decomposition of a vector field that was introduced in section 4.1. The longitudinal components of the equations become, see (4.1),

$$\begin{cases}\hat{\boldsymbol{z}}\cdot(\nabla_T\times\boldsymbol{E}_T(\boldsymbol{r},\omega))=\mathrm{i}\omega\mu_0\mu(\boldsymbol{r},\omega)H_z(\boldsymbol{r},\omega)\\ \hat{\boldsymbol{z}}\cdot(\nabla_T\times\boldsymbol{H}_T(\boldsymbol{r},\omega))=-\mathrm{i}\omega\epsilon_0\epsilon(\boldsymbol{r},\omega)E_z(\boldsymbol{r},\omega)\end{cases} \tag{4.3}$$

The transverse parts are, see (4.2),

$$\begin{cases}\hat{\boldsymbol{z}}\times\dfrac{\partial}{\partial z}\boldsymbol{E}_T(\boldsymbol{r},\omega)-\hat{\boldsymbol{z}}\times\nabla_T E_z(\boldsymbol{r},\omega)=\mathrm{i}\omega\mu_0\mu(\boldsymbol{r},\omega)\boldsymbol{H}_T(\boldsymbol{r},\omega)\\ \hat{\boldsymbol{z}}\times\dfrac{\partial}{\partial z}\boldsymbol{H}_T(\boldsymbol{r},\omega)-\hat{\boldsymbol{z}}\times\nabla_T H_z(\boldsymbol{r},\omega)=-\mathrm{i}\omega\epsilon_0\epsilon(\boldsymbol{r},\omega)\boldsymbol{E}_T(\boldsymbol{r},\omega)\end{cases}$$

We can solve for the $z-$derivatives of the transverse fields by acting with $\hat{\boldsymbol{z}}\times$ on these equations

$$\begin{cases}\dfrac{\partial}{\partial z}\boldsymbol{E}_T(\boldsymbol{r},\omega)=\nabla_T E_z(\boldsymbol{r},\omega)-\mathrm{i}\omega\mu_0\mu(\boldsymbol{r},\omega)\hat{\boldsymbol{z}}\times\boldsymbol{H}_T(\boldsymbol{r},\omega)\\ \dfrac{\partial}{\partial z}\boldsymbol{H}_T(\boldsymbol{r},\omega)=\nabla_T H_z(\boldsymbol{r},\omega)+\mathrm{i}\omega\epsilon_0\epsilon(\boldsymbol{r},\omega)\hat{\boldsymbol{z}}\times\boldsymbol{E}_T(\boldsymbol{r},\omega)\end{cases} \tag{4.4}$$

These equations can be written as a 4×4 system of equations

$$\begin{aligned}&\frac{\partial}{\partial z}\begin{pmatrix}\boldsymbol{E}_T(\boldsymbol{r},\omega)\\ \eta_0\boldsymbol{H}_T(\boldsymbol{r},\omega)\end{pmatrix}+\mathrm{i}\frac{\omega}{c_0}\begin{pmatrix}\boldsymbol{0} & \mu(\boldsymbol{r},\omega)\hat{\boldsymbol{z}}\times\\ -\epsilon(\boldsymbol{r},\omega)\hat{\boldsymbol{z}}\times & \boldsymbol{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{E}_T(\boldsymbol{r},\omega)\\ \eta_0\boldsymbol{H}_T(\boldsymbol{r},\omega)\end{pmatrix}\\ &\qquad=\begin{pmatrix}\nabla_T E_z(\boldsymbol{r},\omega)\\ \eta_0\nabla_T H_z(\boldsymbol{r},\omega)\end{pmatrix}\end{aligned}$$

where $\eta_0 = \sqrt{\mu_0/\epsilon_0}$ is the wave impedance of vacuum. The operator $\hat{\boldsymbol{z}}\times$ on the right hand side gives a rotation by an angle $\pi/2$ in the x-y-plane of the vector it acts on. We see that the longitudinal components E_z and H_z act as sources to the transverse components.

The decomposition in longitudinal and transverse fields wrt the $z-$axis is a general decomposition of the electromagnetic fields in isotropic materials. In the problems that are treated in this book the decomposition is very useful since the $z-$axis is along the guiding structure, see chapters 5 and 8.

4.3 Specific z-dependence of the fields

From now on we let the $z-$axis be parallel to the guiding structures that are treated in this book. We treat fields with the $z-$dependence $\exp(ik_z z)$. If E_z and H_z have this z-dependence it follows from (4.3) that the transverse components of the fields have the same $z-$dependence and thus all fields can be written as

$$\boldsymbol{E}(\boldsymbol{r},\omega) = \boldsymbol{E}(\boldsymbol{\rho},k_z,\omega)e^{\mathrm{i}k_z z}$$

The coefficient k_z is referred to as the *longitudinal wave number*. The Maxwell equations now simplifies to partial differential equations in the transverse coordinates x and y

$$\begin{cases} \hat{\boldsymbol{z}}\cdot(\nabla_T \times \boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega)) = \mathrm{i}\omega\mu_0\mu(\omega)H_z(\boldsymbol{\rho},k_z,\omega) \\ \hat{\boldsymbol{z}}\cdot(\nabla_T \times \boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega)) = -\mathrm{i}\omega\epsilon_0\epsilon(\omega)E_z(\boldsymbol{\rho},k_z,\omega) \end{cases} \tag{4.5}$$

and

$$\begin{cases} \mathrm{i}k_z\hat{\boldsymbol{z}}\times\boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega) - \mathrm{i}\omega\mu_0\mu(\omega)\boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega) = \hat{\boldsymbol{z}}\times\nabla_T E_z(\boldsymbol{\rho},k_z,\omega) \\ \mathrm{i}k_z\hat{\boldsymbol{z}}\times\boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega) + \mathrm{i}\omega\epsilon_0\epsilon(\omega)\boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega) = \hat{\boldsymbol{z}}\times\nabla_T H_z(\boldsymbol{\rho},k_z,\omega) \end{cases} \tag{4.6}$$

We observe that the transverse components of the vectors $\boldsymbol{E}$ and $\boldsymbol{H}$, i.e. $\boldsymbol{E}_T$ and $\boldsymbol{H}_T$ can be expressed in terms of the longitudinal components E_z and H_z. We see this by operating with $\hat{\boldsymbol{z}}\times$ on the first of the transverse equations in (4.6), utilizing $\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C}) = \boldsymbol{B}(\boldsymbol{A}\cdot\boldsymbol{C}) - \boldsymbol{C}(\boldsymbol{A}\cdot\boldsymbol{B})$ (BAC-CAB-rule), and by eliminating $\hat{\boldsymbol{z}}\times\boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega)$ using the second equation in (4.6). If we treat the second equation in (4.6) in the same manner we get

$$\begin{cases} -\mathrm{i}k_z\boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega) - \dfrac{\omega\mu_0\mu(\omega)}{k_z}\Big[\hat{\boldsymbol{z}}\times\nabla_T H_z(\boldsymbol{\rho},k_z,\omega) \\ \qquad\qquad -\mathrm{i}\omega\epsilon_0\epsilon(\omega)\boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega)\Big] = -\nabla_T E_z(\boldsymbol{\rho},k_z,\omega) \\ -\mathrm{i}k_z\boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega) + \dfrac{\omega\epsilon_0\epsilon(\omega)}{k_z}\Big[\hat{\boldsymbol{z}}\times\nabla_T E_z(\boldsymbol{\rho},k_z,\omega) \\ \qquad\qquad +\mathrm{i}\omega\mu_0\mu(\omega)\boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega)\Big] = -\nabla_T H_z(\boldsymbol{\rho},k_z,\omega) \end{cases}$$

or

$$\begin{cases} \boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega) = \mathrm{i}\dfrac{k_z\nabla_T E_z(\boldsymbol{\rho},k_z,\omega) - \omega\mu_0\mu(\omega)\hat{\boldsymbol{z}}\times\nabla_T H_z(\boldsymbol{\rho},k_z,\omega)}{\frac{\omega^2}{c_0^2}\epsilon(\omega)\mu(\omega) - k_z^2} \\ \boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega) = \mathrm{i}\dfrac{k_z\nabla_T H_z(\boldsymbol{\rho},k_z,\omega) + \omega\epsilon_0\epsilon(\omega)\hat{\boldsymbol{z}}\times\nabla_T E_z(\boldsymbol{\rho},k_z,\omega)}{\frac{\omega^2}{c_0^2}\epsilon(\omega)\mu(\omega) - k_z^2} \end{cases}$$

We rewrite these equations

$$\begin{cases} \boldsymbol{E}_T(\boldsymbol{\rho}, k_z, \omega) = \dfrac{\mathrm{i}}{k_t^2} \left\{ k_z \nabla_T E_z(\boldsymbol{\rho}, k_z, \omega) - \omega \mu_0 \mu(\omega) \hat{\boldsymbol{z}} \times \nabla_T H_z(\boldsymbol{\rho}, k_z, \omega) \right\} \\ \boldsymbol{H}_T(\boldsymbol{\rho}, k_z, \omega) = \dfrac{\mathrm{i}}{k_t^2} \left\{ k_z \nabla_T H_z(\boldsymbol{\rho}, k_z, \omega) + \omega \epsilon_0 \epsilon(\omega) \hat{\boldsymbol{z}} \times \nabla_T E_z(\boldsymbol{\rho}, k_z, \omega) \right\} \end{cases} \tag{4.7}$$

where we introduced the *transverse wave number* k_t

$$k_t^2 = \frac{\omega^2}{c_0^2} \epsilon(\omega) \mu(\omega) - k_z^2 \tag{4.8}$$

The relation between the wave numbers $k(\omega)$, k_z, and k_t is

$$k^2 = k_t^2 + k_z^2$$

Since the transverse components of the the electric and magnetic fields are determined by the $z-$components of the fields, it is sufficient to determine the z-components of the electric and magnetic fields in order to construct the transverse parts. Thus the full vector problem that we started with has been reduced to much simpler scalar problems.

Each of the longitudinal components, $E_z(\boldsymbol{\rho}, k_z, \omega)$ och $H_z(\boldsymbol{\rho}, k_z, \omega)$, satisfies a partial differential equation in the variables x och y. We easily get these equations from (2.8) and (2.9) on page 15. The result is

$$\begin{cases} \nabla_T^2 E_z(\boldsymbol{\rho}, k_z, \omega) + k_t^2 E_z(\boldsymbol{\rho}, k_z, \omega) = 0 \\ \nabla_T^2 H_z(\boldsymbol{\rho}, k_z, \omega) + k_t^2 H_z(\boldsymbol{\rho}, k_z, \omega) = 0 \end{cases} \tag{4.9}$$

where the transverse wave number k_t is defined in (4.8). The transverse components also satisfy a system of partial differential equations. They are less useful since they are vectorial.

Problems in Chapter 4

*__4.1__ Let $\boldsymbol{A} = \nabla \times (\nabla \times \boldsymbol{F})$. Determine $\boldsymbol{A}_T$ och A_z expressed in $\boldsymbol{F}_T$, F_z, ∇_T och $\frac{\partial}{\partial z}$.

__4.2__ Show that one can relate $\boldsymbol{E}_T$ and $\boldsymbol{H}_T$ in the following way:

$$\text{If } E_z = 0 \text{ then } \boldsymbol{E}_T = -\frac{\omega \mu_0 \mu(\omega)}{k_z} \hat{\boldsymbol{z}} \times \boldsymbol{H}_T$$

and

$$\text{If } H_z = 0 \text{ then } \boldsymbol{H}_T = \frac{\omega \epsilon_0 \epsilon(\omega)}{k_z} \hat{\boldsymbol{z}} \times \boldsymbol{E}_T$$

We have here assumed that all fields have the specific z-dependence $\exp(\mathrm{i} k_z z)$.

Summary of chapter 4

Decomposition $\boldsymbol{E}(\boldsymbol{r},\omega) = \boldsymbol{E}(\boldsymbol{\rho},k_z,\omega)e^{ik_z z}$

Maxwell equtions

$$\begin{cases} \hat{\boldsymbol{z}} \cdot (\nabla_T \times \boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega)) = \mathrm{i}\omega\mu_0\mu(\omega)H_z(\boldsymbol{\rho},k_z,\omega) \\ \hat{\boldsymbol{z}} \cdot (\nabla_T \times \boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega)) = -\mathrm{i}\omega\epsilon_0\epsilon(\omega)E_z(\boldsymbol{\rho},k_z,\omega) \end{cases}$$

$$\begin{cases} \mathrm{i}k_z\hat{\boldsymbol{z}} \times \boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega) - \mathrm{i}\omega\mu_0\mu(\omega)\boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega) = \hat{\boldsymbol{z}} \times \nabla_T E_z(\boldsymbol{\rho},k_z,\omega) \\ \mathrm{i}k_z\hat{\boldsymbol{z}} \times \boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega) + \mathrm{i}\omega\epsilon_0\epsilon(\omega)\boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega) = \hat{\boldsymbol{z}} \times \nabla_T H_z(\boldsymbol{\rho},k_z,\omega) \end{cases}$$

Transverse components

$$\begin{cases} \boldsymbol{E}_T(\boldsymbol{\rho},k_z,\omega) = \dfrac{\mathrm{i}}{k_t^2}\left\{k_z\nabla_T E_z(\boldsymbol{\rho},k_z,\omega) - \omega\mu_0\mu(\omega)\hat{\boldsymbol{z}} \times \nabla_T H_z(\boldsymbol{\rho},k_z,\omega)\right\} \\ \boldsymbol{H}_T(\boldsymbol{\rho},k_z,\omega) = \dfrac{\mathrm{i}}{k_t^2}\left\{k_z\nabla_T H_z(\boldsymbol{\rho},k_z,\omega) + \omega\epsilon_0\epsilon(\omega)\hat{\boldsymbol{z}} \times \nabla_T E_z(\boldsymbol{\rho},k_z,\omega)\right\} \end{cases}$$

$$k_t^2 = \frac{\omega^2}{c_0^2}\epsilon(\omega)\mu(\omega) - k_z^2$$

Equations for the longitudinal components

$$\begin{aligned} \nabla_T^2 E_z(\boldsymbol{\rho},k_z,\omega) + k_t^2 E_z(\boldsymbol{\rho},k_z,\omega) = 0 \\ \nabla_T^2 H_z(\boldsymbol{\rho},k_z,\omega) + k_t^2 H_z(\boldsymbol{\rho},k_z,\omega) = 0 \end{aligned}$$

Chapter 5

Waveguides at fix frequency

Waveguides are structures that guide waves along a given direction. Figure 5.1 gives an example of geometry for a waveguide. The surface of the waveguide is denoted S and the normal to the surface $\hat{\boldsymbol{n}}$. Note that the normal $\hat{\boldsymbol{n}}$ is a function of the coordinates x and y, but not of the coordinate z. The cross section of the waveguide is denoted Ω and it is circumscribed by the curve Γ, c.f., figure 5.4. Figure 5.4a shows a waveguide with a simply connected cross section Ω, while the figure 5.4b shows a waveguide with an inner surface (the curve Γ consists of two non-connected parts). The analysis in this chapter is valid for waveguides with general cross section.

Two types of waveguides that are studied in this book. The first type is referred to as closed waveguide, or hollow waveguide and has metallic walls that enclose the region. The other type is the open waveguide, for which parts of the enclosing surface is not metallic. Resonance cavities and dielectric resonators are related to hollow waveguides and dielectric waveguides, respectively, and these are also analyzed in this book.

Figure 5.2 shows two different hollow waveguides and an optical fiber, which is an open waveguide. The figure also shows reflector antennas that are fed by circular and rectangular horn antennas. In such antenna systems hollow waveguides are crucial. Figure 5.3 shows another application for hollow waveguides and hollow cavities. It is a klystron that generates electromagnetic fields with high power and a waveguide that leads this power to a cavity in a linear accelerator.

This chapter treats the hollow waveguides. A special type of open waveguides referred to as dielectric waveguides are treated in chapter 8. To analyse the hollow waveguides mathematically we need boundary conditions and wave equations. The boundary conditions for the metallic walls are treated in section 5.1. The derivation of the wave equation from the Maxwell equations in a source free region is given in section 5.2 and 5.3. In the same sections the solutions to the equations are discussed. In section 5.4 the solutions are expressed in terms of expansions in orthogonal and complete sets of basis functions. Some very important examples of waveguides are presented in 5.5. The normalizations for the sets of basis functions are given in section 5.7. Based upon these normalizations we derive expressions for the power flow and the losses in the walls and present them in sections 5.8 and 5.9.

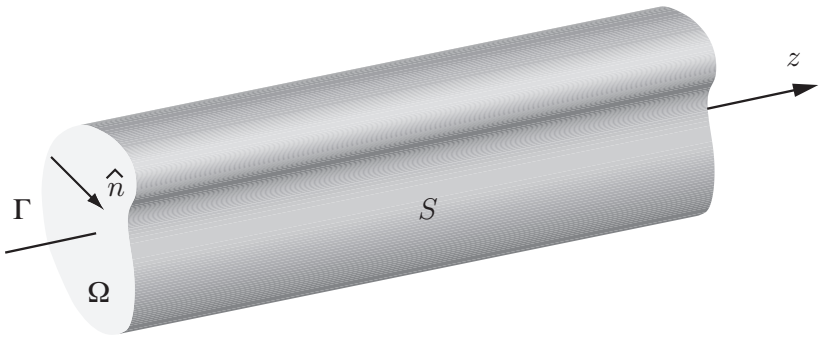


Figure 5.1: Geometry for waveguide.

5.1 Boundary conditions

We now analyze the boundary conditions for the electric and magnetic fields on the metallic surface S of a hollow waveguide. We assume isotropic material in the interior, i.e., the constitutive relations are given by

$$\begin{cases} \boldsymbol{D}(\boldsymbol{r},\omega) = \epsilon_0\epsilon(\boldsymbol{r},\omega)\boldsymbol{E}(\boldsymbol{r},\omega) \\ \boldsymbol{B}(\boldsymbol{r},\omega) = \mu_0\mu(\boldsymbol{r},\omega)\boldsymbol{H}(\boldsymbol{r},\omega) \end{cases}$$

The sufficient boundary conditions on a perfectly conducting surface are, c.f., (1.17) on page 7,

$$\begin{cases} \hat{\boldsymbol{n}} \times \boldsymbol{E}(\boldsymbol{r},\omega) = \boldsymbol{0} \\ \hat{\boldsymbol{n}} \cdot \boldsymbol{H}(\boldsymbol{r},\omega) = 0 \end{cases} \qquad \boldsymbol{r} \text{ on } S$$

since $\boldsymbol{B} = \mu_0\mu\boldsymbol{H}$ for an isotropic material.

We express the boundary conditions in terms of the decomposed fields in section 4.1

$$\begin{cases} \hat{\boldsymbol{n}} \times (\boldsymbol{E}_T(\boldsymbol{r},\omega) + \hat{\boldsymbol{z}}E_z(\boldsymbol{r},\omega)) = \boldsymbol{0} \\ \hat{\boldsymbol{n}} \cdot (\boldsymbol{H}_T(\boldsymbol{r},\omega) + \hat{\boldsymbol{z}}H_z(\boldsymbol{r},\omega)) = 0 \end{cases} \qquad \boldsymbol{r} \text{ on } S$$

The unit normal vector $\hat{\boldsymbol{n}}$ to the surface S has no z-component, i.e., $\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{z}} = 0$. Since $\hat{\boldsymbol{n}} \times \boldsymbol{E}_T$ only has a component along the z-axis, while $\hat{\boldsymbol{n}} \times \hat{\boldsymbol{z}}$ is perpendicular to the $z-$axis (directed tangential to Γ), each term in the first equation has to be zero. The second term in the second equation is zero since $\hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{z}}$ are perpendicular. The conditions are equivalent to

$$\begin{cases} E_z(\boldsymbol{r},\omega) = 0 \\ \hat{\boldsymbol{z}} \cdot (\hat{\boldsymbol{n}} \times \boldsymbol{E}_T(\boldsymbol{r},\omega)) = \hat{\boldsymbol{n}} \cdot (\boldsymbol{E}_T \times \hat{\boldsymbol{z}}) = 0 \qquad \boldsymbol{r} \text{ on } S \\ \hat{\boldsymbol{n}} \cdot \boldsymbol{H}_T(\boldsymbol{r},\omega) = 0 \end{cases} \tag{5.1}$$

These equations are valid on the entire surface S, which implies that also the following z-derivative is zero (remember that $\hat{\boldsymbol{n}}$ is independent of z):

$$\hat{\boldsymbol{n}} \cdot \frac{\partial}{\partial z}\boldsymbol{H}_T(\boldsymbol{r},\omega) = 0 \qquad \boldsymbol{r} \text{ on } S$$

We utilize (4.4)

$$\frac{\partial}{\partial z}\boldsymbol{H}_T(\boldsymbol{r},\omega) = \nabla_T H_z(\boldsymbol{r},\omega) + \mathrm{i}\omega\epsilon_0\epsilon(\boldsymbol{r},\omega)\hat{\boldsymbol{z}} \times \boldsymbol{E}_T(\boldsymbol{r},\omega)$$

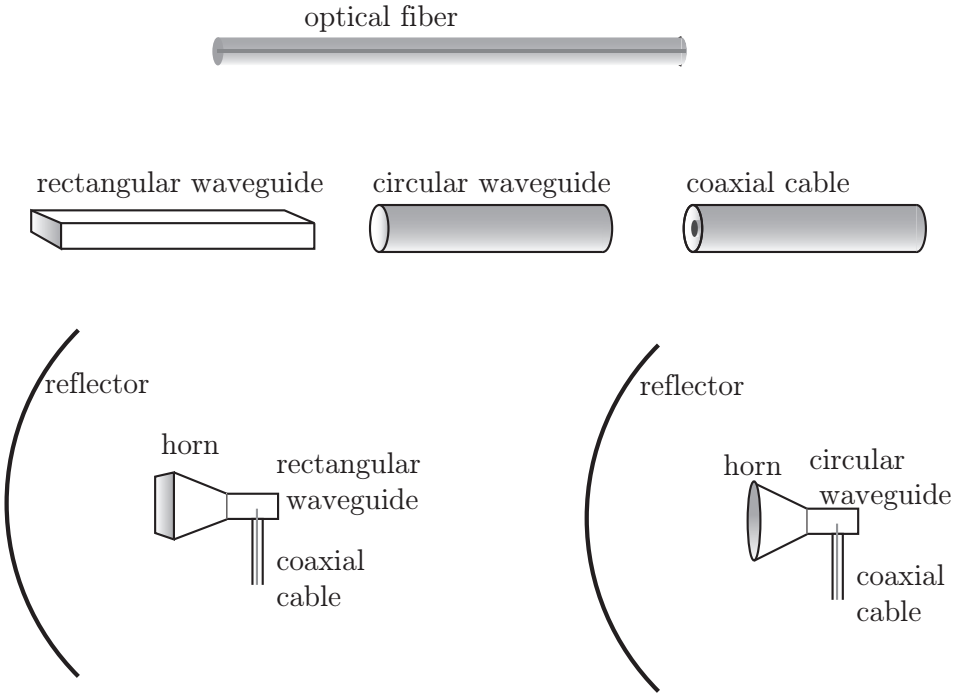


Figure 5.2: Examples of waveguides. Reflector antenna with a rectangular and circular feed horn.

in order to eliminate the z-derivative, and get

$$\hat{\boldsymbol{n}} \cdot (\nabla_T H_z(\boldsymbol{r}, \omega) + \mathrm{i}\omega\epsilon_0\epsilon(\boldsymbol{r}, \omega)\hat{\boldsymbol{z}} \times \boldsymbol{E}_T(\boldsymbol{r}, \omega)) = 0 \qquad \boldsymbol{r} \text{ on } S$$

By utilizing the original boundary condition (5.1), the boundary conditions on the surface S are reduced to

$$\begin{cases} E_z(\boldsymbol{r}, \omega) = 0 \\ \dfrac{\partial}{\partial n} H_z(\boldsymbol{r}, \omega) = 0 \end{cases} \qquad \boldsymbol{r} \text{ on } S \tag{5.2}$$

where $\frac{\partial}{\partial n} H_z(\boldsymbol{r}, \omega) = \hat{\boldsymbol{n}} \cdot \nabla_T H_z(\boldsymbol{r}, \omega)$. These boundary conditions only contain the $z-$components of the fields and they are sufficient for determining the waves that can exist in a hollow waveguide.

5.2 TM- and TE-modes

In this section we solve the Maxwell equations in a waveguide with general cross-section Ω and perfectly conducting walls S. The conditions in (5.2) separates the z-component of the electric field, E_z from the $z-$component of the magnetic field, H_z. We look for solutions where either E_z *or* H_z is zero, i.e.,

$$\begin{cases} H_z(\boldsymbol{r}, \omega) = 0 & \text{(TM-case)} \\ E_z(\boldsymbol{r}, \omega) = 0 & \text{(TE-case)} \end{cases}$$

The first case is the transverse magnetic case (TM-case), where the magnetic field lacks $z-$component. The other case is the transverse electric case (TE-case). The solutions to

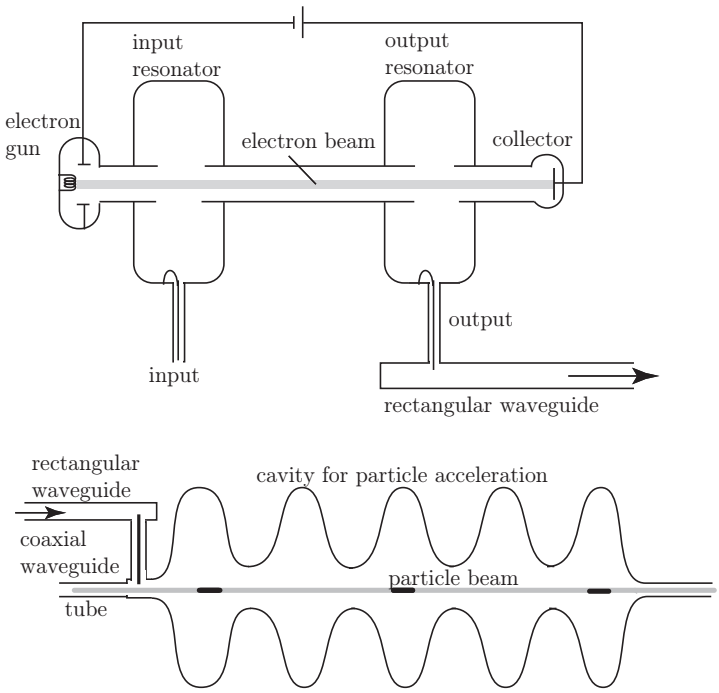


Figure 5.3: Upper: Klystron that generates the field for a cavity in an accelerator. Lower: Feed of a cavity in an accelerator.

the two cases do not couple since there is no coupling via the differential equations or the boundary conditions. We will later also discuss the conditions that have to be satisfied in order to obtain solutions with *both* E_z and H_z zero.

We let the region $z_1 < z < z_2$ be source free, i.e., $\boldsymbol{J} = \boldsymbol{0}$, c.f., figur 5.5 and determine the waves that can exist in this region. Regions with sources are treated in section 5.10.

We first describe our strategy for finding general solutions. The waveguide is assumed to be filled with an isotropic, homogeneous material with material parameters $\epsilon(\omega)$ and $\mu(\omega)$. The $z-$components of the equations (2.8) and (2.9) on page 15, and the boundary conditions for $E_z(\boldsymbol{r},\omega)$ and $H_z(\boldsymbol{r},\omega)$ are summarized as

$$\begin{cases} \nabla^2 E_z(\boldsymbol{r},\omega) + k^2(\omega) E_z = 0 \\ E_z(\boldsymbol{r},\omega) = 0 \qquad \boldsymbol{r} \text{ on } S \end{cases} \qquad z \in [z_1, z_2], \boldsymbol{\rho} \in \Omega \text{ (TM-case)}$$

$$\begin{cases} \nabla^2 H_z(\boldsymbol{r},\omega) + k^2(\omega) H_z = 0 \\ \dfrac{\partial H_z}{\partial n}(\boldsymbol{r},\omega) = 0 \qquad \boldsymbol{r} \text{ on } S \end{cases} \qquad z \in [z_1, z_2], \boldsymbol{\rho} \in \Omega \text{ (TE-case)} \tag{5.3}$$

where the wave number in the material is

$$k^2(\omega) = \frac{\omega^2}{c_0^2}\epsilon(\omega)\mu(\omega)$$

From the solutions to these equations we can determine the entire vector field $\boldsymbol{E}$ and $\boldsymbol{H}$ from (4.7) on page 68. The first case in (5.3) is the transverse magnetic case (TM-case) where $H_z = 0$. The other case is the transverse electric case (TE-case) where $E_z = 0$.

We use the method of separation of variables to solve the two boundary value problems in (5.3). The method is frequently used in mathematical physics and in our case it leads

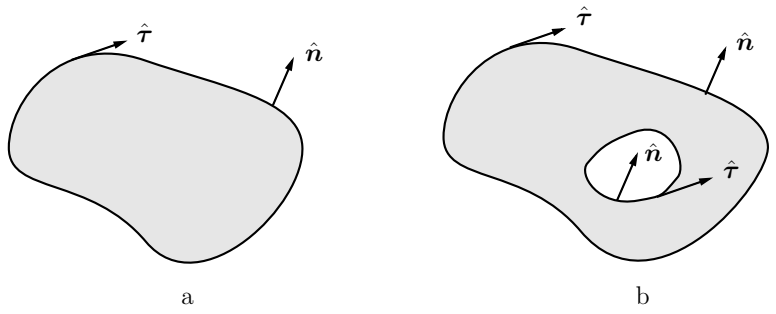


Figure 5.4: Cross sections of waveguides. The z-axis is directed perpendicular to the plane of the paper and $\hat{\boldsymbol{\tau}} = \hat{\boldsymbol{z}} \times \hat{\boldsymbol{n}}$.

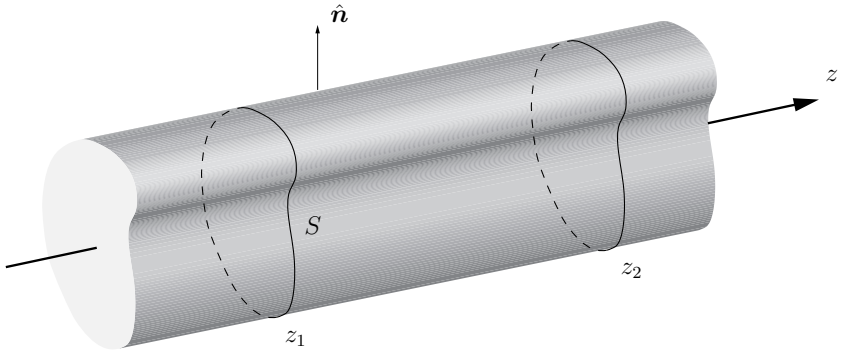


Figure 5.5: The source free region in the waveguide.

to a complete set of functions in the transverse coordinates x and y. The $z-$component of the electric (TM-case) or magnetic field (TE-case) is expanded in this system. The other components follow from the relations between the transverse and longitudinal components.

5.2.1 The longitudinal components of the fields

We make the following ansatz

$$\begin{cases} E_z(\boldsymbol{r}) = v(\boldsymbol{\rho})a(z), & \text{(TM-case)} \\ H_z(\boldsymbol{r}) = w(\boldsymbol{\rho})b(z), & \text{(TE-case)} \end{cases}$$

where $\boldsymbol{\rho} = \hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y$ and η_0 is the wave impedance for vacuum. We insert this into (5.3)

$$\begin{cases} a(z)\nabla_T^2 v(\boldsymbol{\rho}) + v(\boldsymbol{\rho})\dfrac{\partial^2 a}{\partial z^2}(z) + k^2 v(\boldsymbol{\rho})a(z) = 0 \\ v(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases}$$

and

$$\begin{cases} b(z)\nabla_T^2 w(\boldsymbol{\rho}) + w(\boldsymbol{\rho})\dfrac{\partial^2 b}{\partial z^2}(z) + k^2 w(\boldsymbol{\rho})b(z) = 0 \\ \dfrac{\partial w}{\partial n}(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases}$$

After division with $v(\boldsymbol{\rho})a(z)$ and $w(\boldsymbol{\rho})b(z)$, respectively, we get

$$\begin{cases} \dfrac{\nabla_T^2 v(\boldsymbol{\rho})}{v(\boldsymbol{\rho})} = -\dfrac{1}{a(z)}\dfrac{\partial^2 a}{\partial z^2}(z) - k^2 \\ v(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \qquad \begin{cases} \dfrac{\nabla_T^2 w(\boldsymbol{\rho})}{w(\boldsymbol{\rho})} = -\dfrac{1}{b(z)}\dfrac{\partial^2 b}{\partial z^2}(z) - k^2 \\ \dfrac{\partial w}{\partial n}(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases}$$

In these differential equations the left hand side only depends on the variables x and y, while the right hand side only depends on z. This can only be satisfied if both sides are equal to a constant and we denote this constant $-k_t^2$ for reasons that will soon be obvious. We identify the following two eigenvalue problems for the hollow waveguide

$$\begin{cases} \nabla_T^2 v(\boldsymbol{\rho}) + k_t^2 v(\boldsymbol{\rho}) = \dfrac{\partial^2 v(\boldsymbol{\rho})}{\partial x^2} + \dfrac{\partial^2 v(\boldsymbol{\rho})}{\partial y^2} + k_t^2 v(\boldsymbol{\rho}) = 0 \\ v(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \qquad \text{(TM-case)}$$

and

$$\begin{cases} \nabla_T^2 w(\boldsymbol{\rho}) + k_t^2 w(\boldsymbol{\rho}) = \dfrac{\partial^2 w(\boldsymbol{\rho})}{\partial x^2} + \dfrac{\partial^2 w(\boldsymbol{\rho})}{\partial y^2} + k_t^2 w(\boldsymbol{\rho}) = 0 \\ \dfrac{\partial w}{\partial n}(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \qquad \text{(TE-case)}$$

Later in this section we give explicit examples on geometries and their corresponding sets of systems of eigenfunction, but here we continue with the general analysis. The eigenvalue problems for the TM- and TE-case are expressed in the transverse coordinates x and y. There is only a countable set of values of k_t^2 (k_t is the transverse wave number) for which there exist non-trivial solutions. These values of k_t^2 are called the eigenvalues of the problem and can be numbered in their order of size. For most geometries the eigenvalues of the TE- and TM-cases are different, but we use the same notation for the two cases for practical reasons. One can prove that the eigenvalues are positive, $k_t^2 > 0$, c.f., example 5.1, and number the eigenvalues according to:

$$0 < k_{t1}^2 \leq k_{t2}^2 \leq k_{t3}^2 \leq \dots$$

There is only a finite number of eigenvalues that have the same values. The eigenfunction corresponding to eigenvalue number n is denoted $v_n(\boldsymbol{\rho})$ for the TM-case and $w_n(\boldsymbol{\rho})$ for the TE-case, i.e., they are solutions to [1]

$$\begin{cases} \nabla_T^2 v_n(\boldsymbol{\rho}) + k_{tn}^2 v_n(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \in \Omega \\ v_n(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \qquad \text{(TM-case)} \qquad (5.4)$$

and

$$\begin{cases} \nabla_T^2 w_n(\boldsymbol{\rho}) + k_{tn}^2 w_n(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \in \Omega \\ \dfrac{\partial w_n(\boldsymbol{\rho})}{\partial n} = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \qquad \text{(TE-case)} \qquad (5.5)$$

Note that these eigenfunctions are determined by the geometry of the cross section which is defined by Ω. They are independent of the angular frequency ω and of the material

[1] In the two-dimensional case it is often practical to count the eigenvalues in a sequence with two indices mn, see examples presented later in this section.

in the waveguide, i.e., independent of $\epsilon(\omega)$ and $\mu(\omega)$. We always let the eigenfunctions $v_n(\boldsymbol{\rho})$ and $w_n(\boldsymbol{\rho})$ be *real valued*.

Each of the sets of eigenfunctions, $\{v_n(\boldsymbol{\rho})\}_{n=1}^{\infty}$ and $\{w_n(\boldsymbol{\rho})\}_{n=1}^{\infty}$, constitutes a complete set of functions in the plane. We can expand an arbitrary function, defined in the region Ω, in this set of functions. The functions $E_z(\boldsymbol{r},\omega)$ and $H_z(\boldsymbol{r},\omega)$ are expanded as

$$\begin{cases} E_z(\boldsymbol{r},\omega) = \displaystyle\sum_{n=1}^{\infty} a_n(z,\omega) v_n(\boldsymbol{\rho}) \\ \\ H_z(\boldsymbol{r},\omega) = \displaystyle\sum_{n=1}^{\infty} b_n(z,\omega) w_n(\boldsymbol{\rho}) \end{cases} \tag{5.6}$$

Example 5.1
All eigenvalues to the TM- and TE-cases are non-negative numbers. This can be shown by integrating the vector rule $\nabla \cdot (f\nabla f) = \nabla f \cdot \nabla f + f\nabla^2 f$, over the cross section Ω. Gauss' theorem in the plane gives

$$\oint_{\Gamma} f(\boldsymbol{\rho}) \frac{\partial}{\partial n} f(\boldsymbol{\rho}) \, \mathrm{d}l = \iint_{\Omega} \left(\nabla_T f(\boldsymbol{\rho})\right)^2 \, \mathrm{d}x\mathrm{d}y + \iint_{\Omega} f(\boldsymbol{\rho}) \nabla_T^2 f(\boldsymbol{\rho}) \, \mathrm{d}x\mathrm{d}y$$

Notice that if the function is independent of the z-coordinate, then $\nabla f = \nabla_T f$.

We first consider the TM-case and let $f(\boldsymbol{\rho}) = v_n(\boldsymbol{\rho})$. Due to the boundary condition on the boundary curve Γ, $v_n = 0$, the line integral vanishes. We then use the differential equation for the eigenvalue problem (5.4) to get the equality

$$\iint_{\Omega} \left(\nabla_T v_n(\boldsymbol{\rho})\right)^2 \, \mathrm{d}x\mathrm{d}y = k_{tn}^2 \iint_{\Omega} \left(v_n(\boldsymbol{\rho})\right)^2 \, \mathrm{d}x\mathrm{d}y \tag{5.7}$$

and the inequality

$$k_{tn}^2 \iint_{\Omega} \left(v_n(\boldsymbol{\rho})\right)^2 \, \mathrm{d}x\mathrm{d}y \geq 0$$

If v_n is not identically zero, the inequality implies that the eigenvalue for the TM-case is non-negative, $k_{tn}^2 \geq 0$.

By letting $f(\boldsymbol{\rho}) = w_n(\boldsymbol{\rho})$, and by using (5.5) and the boundary condition $\frac{\partial}{\partial n} w_n(\boldsymbol{\rho}) = 0$, the corresponding relation for the TE-case

$$\iint_{\Omega} \left(\nabla_T w_n(\boldsymbol{\rho})\right)^2 \, \mathrm{d}x\mathrm{d}y = k_{tn}^2 \iint_{\Omega} \left(w_n(\boldsymbol{\rho})\right)^2 \, \mathrm{d}x\mathrm{d}y \tag{5.8}$$

and the inequality

$$k_{tn}^2 \iint_{\Omega} \left(w_n(\boldsymbol{\rho})\right)^2 \, \mathrm{d}x\mathrm{d}y \geq 0$$

are obtained. Unless w_n is not identically zero, the eigenvalue for the TE-case is also non-negative, $k_{tn}^2 \geq 0$.

We can prove an even stronger result, namely that the eigenvalue $k_{tn}^2 = 0$ leads to a contradiction and that the eigenvalues are positive for the TM- and TE-cases[2]. From

[2]The eigenvalue $k_{tn}^2 = 0$ is possible for the TEM-case.

equations (5.7) and (5.8) we find that $k_{t_n}^2 = 0$ implies that $\nabla_T v_n = \nabla_T w_n = 0$ in Ω, i.e., $v_n =$ constant and $w_n =$ constant in Ω. In the TM-case this implies that $v_n = 0$ in Ω since the boundary condition $v_n = 0$ on Γ implies that the constant has to be zero. Thus there is a contradiction and the eigenvalues $k_{t_n}^2$ in the TM-case are all positive. To prove the same result for the TE-case we use a result from page 66 and (4.3).

$$H_z(\boldsymbol{r},\omega) = \frac{1}{i\omega\mu_0\mu(\omega)}\hat{\boldsymbol{z}}\cdot(\nabla_T \times \boldsymbol{E}_T(\boldsymbol{r},\omega)) = \frac{1}{i\omega\mu_0\mu(\omega)}\hat{\boldsymbol{z}}\cdot(\nabla \times \boldsymbol{E}_T(\boldsymbol{r},\omega))$$

Since $w_m =$constant in Ω, H_z in the left hand side cannot depend on $\boldsymbol{\rho}$. Stoke's theorem on the cross section Ω gives

$$\iint\limits_{\Omega} H_z \,\mathrm{d}x\mathrm{d}y = \frac{1}{i\omega\mu_0\mu(\omega)} \iint\limits_{\Omega} \hat{\boldsymbol{z}}\cdot(\nabla \times \boldsymbol{E}_T(\boldsymbol{r},\omega))\,\mathrm{d}x\mathrm{d}y = \frac{1}{i\omega\mu_0\mu(\omega)} \oint_{\Gamma} \boldsymbol{E}_T(\boldsymbol{r},\omega)\cdot \mathrm{d}\boldsymbol{r} = 0$$

due to the boundary condition $\hat{\boldsymbol{n}} \times \boldsymbol{E} = \boldsymbol{0}$. We get

$$H_z \iint\limits_{\Omega} \mathrm{d}x\mathrm{d}y = 0$$

which implies that $H_z = 0$ or $w_n = 0$, which is a contradiction, and as above it follows that all eigenvalues $k_{t_n}^2$ in the TE-case are positive. ∎

Example 5.2

We now prove that the eigenfunctions v_n and v_m, or w_n and w_m, that belong to different eigenvalues $k_{t_n}^2$ and $k_{t_m}^2$ in the TM- and TE-cases are orthogonal. We start with Gauss' theorem in the plane ($f_n = v_n$ in the TM-case and $f_n = w_n$ in the TE-case)

$$\begin{aligned} 0 = \oint_{\Gamma} (f_n \nabla_T f_m - f_m \nabla_T f_n)\cdot \hat{\boldsymbol{n}}\, dl &= \iint\limits_{\Omega} \nabla_T \cdot (f_n \nabla_T f_m - f_m \nabla_T f_n)\,\mathrm{d}x\mathrm{d}y \\ &= \iint\limits_{\Omega} (\nabla_T f_n \cdot \nabla_T f_m - \nabla_T f_m \cdot \nabla_T f_n + f_n \nabla_T^2 f_m - f_m \nabla_T^2 f_n)\,\mathrm{d}x\mathrm{d}y \\ &= (k_{t_n}^2 - k_{t_m}^2) \iint\limits_{\Omega} f_n f_m \,\mathrm{d}x\mathrm{d}y \end{aligned}$$

where we have used the eigenvalue equation (5.4) or (5.5). If the eigenvalues are different

$$\iint\limits_{\Omega} f_n f_m \,\mathrm{d}x\mathrm{d}y = 0$$

i.e., the eigenfunctions v_n and v_m or w_n and w_m that belongs to the eigenvalues $k_{t_n}^2$ and $k_{t_m}^2$ are orthogonal. ∎

We insert the expansions in equation (5.6) into the original equation (5.3). By shifting differentiation and summation and utilizing the properties of the eigenfunctions $\{v_n(\boldsymbol{\rho})\}_{n=1}^{\infty}$ and $\{w_n(\boldsymbol{\rho})\}_{n=1}^{\infty}$, we get the following ordinary differential equations for the Fourier coefficients a_n and b_n:

$$\begin{cases} \dfrac{\partial^2 a_n}{\partial z^2}(z,\omega) + \left(\dfrac{\omega^2}{c_0^2}\epsilon(\omega)\mu(\omega) - k_{t_n}^2\right) a_n(z,\omega) = 0 \\ \dfrac{\partial^2 b_n}{\partial z^2}(z,\omega) + \left(\dfrac{\omega^2}{c_0^2}\epsilon(\omega)\mu(\omega) - k_{t_n}^2\right) b_n(z,\omega) = 0 \end{cases}$$

The general solutions to these equations are

$$\begin{cases} a_n(z,\omega) = a_n^{\pm} e^{\pm \mathrm{i} k_{zn}(\omega) z} \\ b_n(z,\omega) = b_n^{\pm} e^{\pm \mathrm{i} k_{zn}(\omega) z} \end{cases}$$

where the *longitudinal wavenumber* k_{zn} is

$$k_{zn}(\omega) = \left(\frac{\omega^2}{c_0^2} \epsilon(\omega)\mu(\omega) - k_{tn}^2 \right)^{\frac{1}{2}} \tag{5.9}$$

The longitudinal wavenumber is a complex number that depends on frequency. This is in contrast to the transverse wavenumber that is real and independent of frequency. The branch of the complex square root for the longitudinal wavenumber is in this book chosen such that the real and imaginary part of k_{zn} are both non-negative. The relation between the wavenumber $k(\omega)$, the longitudinal wavenumber k_{zn} and the transverse wavenumber k_{tn} is

$$k^2(\omega) = k_{tn}^2 + k_{zn}^2(\omega)$$

c.f., also the analysis in section 4.3. It should be emphasized that in most cases the waveguides are filled with air, or vacuum, and unless the frequency is very large ($f > 50$ GHz) the material can be considered to be lossless with $\epsilon = \mu = 1$. Note that the real part of k_{zn} is less than the wavenumber, $k(\omega)$, of the material in the waveguide, which means that the phase velocity in the $z-$direction, $v_p = \omega/k_z$, is larger than the speed of light in the material. Since no information is transported with the phase velocity this does not violate the theory of special relativity.

We conclude that the longitudinal components of the electric and magnetic fields have the following general series expansions

$$\begin{cases} E_z(\boldsymbol{r},\omega) = \sum_{n=1}^{\infty} v_n(\boldsymbol{\rho}) \left(a_n^+(\omega) e^{\mathrm{i} k_{zn}(\omega) z} + a_n^-(\omega) e^{-\mathrm{i} k_{zn}(\omega) z} \right) \\ H_z(\boldsymbol{r},\omega) = \sum_{n=1}^{\infty} w_n(\boldsymbol{\rho}) \left(b_n^+(\omega) e^{\mathrm{i} k_{zn}(\omega) z} + b_n^-(\omega) e^{-\mathrm{i} k_{zn}(\omega) z} \right) \end{cases} \tag{5.10}$$

Each term in these sums corresponds to a *waveguide mode*. The coefficients $a_n\pm$ and $b_n^{\pm}$ are determined by the excitation of the waves in the waveguide. Note that the longitudinal wavenumbers k_{zn} are different in the sums, and that k_{zn} in general is a complex number. The plus sign in the exponent corresponds to a wave traveling in the positive $z-$direction, while the minus sign corresponds to a wave traveling in the negative z-direction. A general expression for a field propagating in the positive z-direction is

$$\begin{cases} E_z(\boldsymbol{r},\omega) = \sum_{n=1}^{\infty} v_n(\boldsymbol{\rho}) a_n^+(\omega) e^{\mathrm{i} k_{zn}(\omega) z} \\ H_z(\boldsymbol{r},\omega) = \sum_{n=1}^{\infty} w_n(\boldsymbol{\rho}) b_n^+(\omega) e^{\mathrm{i} k_{zn}(\omega) z} \end{cases} \tag{5.11}$$

In the lossless case we say that the mode number n is a propagating mode if $k_{tn} < k = \omega/c$, since then k_{zn} is a real number and hence the wave is not attenuated in the $z-$direction. If $k_{tn} > k$ then the mode is a non-propagating mode since $\mathrm{Re}\, k_{zn} = 0$ and

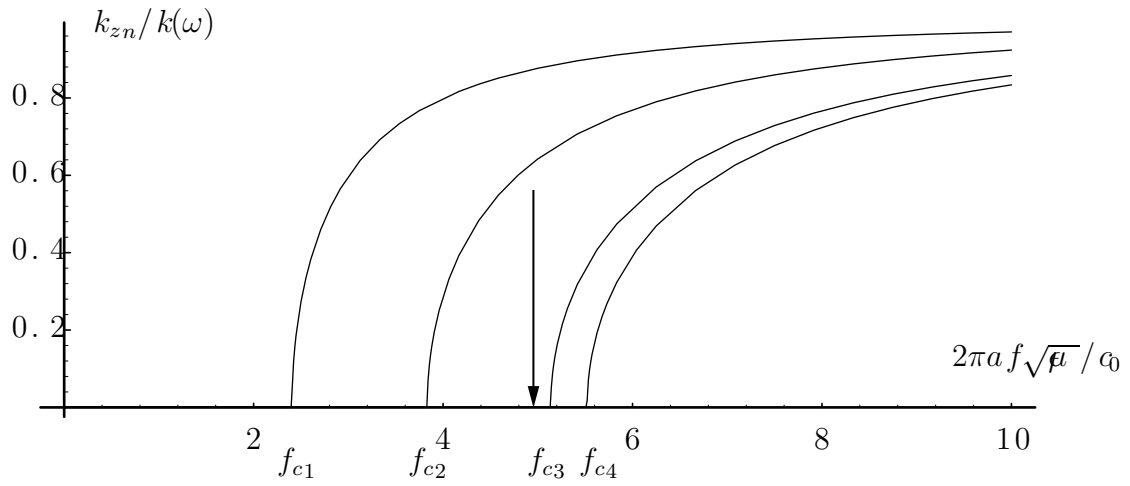


Figure 5.6: Waveguide dispersion for a circular waveguide (radius a) as a function of frequency f. The dispersion relations for the first four modes are shown. The arrow shows a frequency where only two of the modes are propagating. The longitudinal wavenumber k_{zn} is normalized with the wavenumber $k(\omega)$and the frequency f is nornalized with $c_0/(2\pi a\sqrt{\epsilon\mu})$, c.f., the explicit expressions of eigenvalues in table 5.4.

$\mathrm{Im}\, k_{zn} < 0$ which means that the wave is attenuated. If $k_{tn} = k$ we say that the mode is at its cut-off frequency. Then $k_{zn} = 0$ and the wave is a standing wave in the $xy-$plane. In the sum (5.11) only a finite number of modes are propagating at a fixed frequency. The reason is that k_{tn} is a non-decreasing sequence of real numbers. The cut-off frequency $f_{cn} = \omega_{cn}/(2\pi)$ for mode n is given by $k_{zn} = 0$, i.e.,

$$f_{cn} = \frac{k_{tn}c_0}{2\pi\sqrt{\epsilon\mu}} \tag{5.12}$$

The relation between the frequency f, the longitudinal wavenumber k_{zn} and the cut-off frequency f_{cn} is

$$k_{zn} = \frac{2\pi}{c_0}\sqrt{\epsilon\mu}\left(f^2 - f_{cn}^2\right)^{\frac{1}{2}}$$

The relation between the longitudinal wavenumber and the frequency is called the dispersion relation. The dispersion relation for the lowest modes in a circular waveguide is depicted in figure 5.6. The waveguide dispersion is different from the material dispersion that is given by the frequency dependence of the wavenumber $k(\omega)$.

When there are losses in the waveguide, i.e., if at least one of the the material parameters $\epsilon(\omega)$ or $\mu(\omega)$ is complex, we define the cut-off frequency to be the frequency where $k_{tn} = \mathrm{Re}\, k(\omega)$. In the lossy case, propagating modes have a small positive imaginary part of k_{zn} and then even propagating modes are attenuated in the direction of propagation. Only for frequencies where the material is lossless, i.e., where $\epsilon(\omega)$ and $\mu(\omega)$ are real, modes can propagate without attenuation. Later in this chapter we will take into account the fact that the walls of a waveguide are not perfectly conducting. Then the losses due to currents in the walls give rise to attenuation of the modes.

5.2.2 Transverse components of the fields

To make the picture complete we have to determine the transverse components of the fields for each mode. We observe that the z-components of the electric and magnetic fields have the explicit $z-$dependence $\exp\{\pm \mathrm{i}k_z z\}$. To generate the transverse components of the fields we use the results from section 4.3. Equation (4.7) on page 68 provides us with the relations we need

$$\begin{cases} \boldsymbol{E}_T(\boldsymbol{\rho}, k_z, \omega) = \dfrac{\mathrm{i}}{k_t^2}\left\{\pm k_z \nabla_T E_z(\boldsymbol{\rho}, k_z, \omega) - \omega\mu_0\mu(\omega)\hat{\boldsymbol{z}} \times \nabla_T H_z(\boldsymbol{\rho}, k_z, \omega)\right\} \\ \boldsymbol{H}_T(\boldsymbol{\rho}, k_z, \omega) = \dfrac{\mathrm{i}}{k_t^2}\left\{\pm k_z \nabla_T H_z(\boldsymbol{\rho}, k_z, \omega) + \omega\epsilon_0\epsilon(\omega)\hat{\boldsymbol{z}} \times \nabla_T E_z(\boldsymbol{\rho}, k_z, \omega)\right\} \end{cases}$$

where the relation (5.9) has been used. The result is

$$\begin{cases} \boldsymbol{E}_T(\boldsymbol{r}, \omega) = \displaystyle\sum_{n=1}^{\infty} \frac{\mathrm{i}}{k_{tn}^2}\Big\{a_n^+ k_{zn}\nabla_T v_n(\boldsymbol{\rho}) - b_n^+\omega\mu_0\mu(\omega)\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})\Big\}e^{\mathrm{i}k_{z_n}z} \\ \qquad\qquad - \displaystyle\sum_{n=1}^{\infty} \frac{\mathrm{i}}{k_{tn}^2}\Big\{a_n^- k_{zn}\nabla_T v_n(\boldsymbol{\rho}) + b_n^-\omega\mu_0\mu(\omega)\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})\Big\}e^{-\mathrm{i}k_{z_n}z} \\ \boldsymbol{H}_T(\boldsymbol{r}, \omega) = \displaystyle\sum_{n=1}^{\infty} \frac{i}{k_{tn}^2}\Big\{b_n^+ k_{zn}\nabla_T w_n(\boldsymbol{\rho}) + a_n^+\omega\epsilon_0\epsilon(\omega)\hat{\boldsymbol{z}} \times \nabla_T v_n(\boldsymbol{\rho})\Big\}e^{\mathrm{i}k_{z_n}z} \\ \qquad\qquad - \displaystyle\sum_{n=1}^{\infty} \frac{i}{k_{tn}^2}\Big\{b_n^- k_{zn}\nabla_T w_n(\boldsymbol{\rho}) - a_n^-\omega\epsilon_0\epsilon(\omega)\hat{\boldsymbol{z}} \times \nabla_T v_n(\boldsymbol{\rho})\Big\}e^{-\mathrm{i}k_{z_n}z} \end{cases} \tag{5.13}$$

The transverse components are well defined quantities since $k_t^2 > 0$. In particular a mode propagating in the positive z-direction has the transverse fields

$$\begin{cases} \boldsymbol{E}_T(\boldsymbol{r}, \omega) = \displaystyle\sum_{n=1}^{\infty} \frac{\mathrm{i}}{k_{tn}^2}\Big\{a_n^+ k_{zn}\nabla_T v_n(\boldsymbol{\rho}) - b_n^+\omega\mu_0\mu(\omega)\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})\Big\}e^{\mathrm{i}k_{z_n}z} \\ \boldsymbol{H}_T(\boldsymbol{r}, \omega) = \displaystyle\sum_{n=1}^{\infty} \frac{\mathrm{i}}{k_{tn}^2}\Big\{b_n^+ k_{zn}\nabla_T w_n(\boldsymbol{\rho}) + a_n^+\omega\epsilon_0\epsilon(\omega)\hat{\boldsymbol{z}} \times \nabla_T v_n(\boldsymbol{\rho})\Big\}e^{\mathrm{i}k_{z_n}z} \end{cases}$$

At this stage we know how to obtain the entire electric and magnetic fields for the TE and TM waveguide modes in a hollow waveguide.

5.3 TEM-modes

In the previous sections we have assumed that either the electric or the magnetic field has a $z-$component that is non-zero. Under this assumption we found complete sets of functions, c.f., section 5.2. In particular we found that the transverse wavenumber is always positive and that no TE- or TM-modes with eigenvalue $k_{tn} = 0$ exist.

In this section we analyze if there are any solutions to the Maxwell field equations in the waveguide for which both $E_z = H_z = 0$. Such solutions have only transverse components of the fields and are called TEM-modes. In transmission lines the fields are mostly TEM-waves (this is at least the case when the material is independent of the transverse coordinates $\boldsymbol{\rho}$ and the transmission line is lossless). The transmission lines

have been analyzed in chapter 3 by solving the equations for the voltage and current along the line. In this section we do the analysis in terms of the electric and magnetic fields. We start from the equations in chapter 4.

If the z-components of the fields are zero we obtain from (4.3) and (4.4)

$$\begin{cases} \hat{\boldsymbol{z}} \cdot (\nabla_T \times \boldsymbol{E}_T(\boldsymbol{r}, \omega)) = 0 \\ \hat{\boldsymbol{z}} \cdot (\nabla_T \times \boldsymbol{H}_T(\boldsymbol{r}, \omega)) = 0 \end{cases}$$

and

$$\begin{cases} \dfrac{\partial}{\partial z} \boldsymbol{E}_T(\boldsymbol{r}, \omega) = -\mathrm{i}\omega\mu_0\mu(\omega)\hat{\boldsymbol{z}} \times \boldsymbol{H}_T(\boldsymbol{r}, \omega) \\ \dfrac{\partial}{\partial z} \boldsymbol{H}_T(\boldsymbol{r}, \omega) = \mathrm{i}\omega\epsilon_0\epsilon(\omega)\hat{\boldsymbol{z}} \times \boldsymbol{E}_T(\boldsymbol{r}, \omega) \end{cases} \tag{5.14}$$

Since $\nabla_T \times \boldsymbol{E}_T(\boldsymbol{r}, \omega)$ only has a z-component we see that

$$\nabla_T \times \boldsymbol{E}_T(\boldsymbol{r}, \omega) = \boldsymbol{0}$$

everywhere. Then there exists a scalar potential $\psi(\boldsymbol{r}, \omega)$ such that

$$\boldsymbol{E}_T(\boldsymbol{r}, \omega) = -\nabla_T \Psi(\boldsymbol{r}, \omega) \tag{5.15}$$

The potential Ψ is not uniquely defined, since every potential that differs by a function of z from Ψ gives the same field $\boldsymbol{E}_T$. Equations (2.7) and (2.8) on page 15 give

$$\begin{cases} \nabla^2 \boldsymbol{E}_T(\boldsymbol{r}, \omega) + k^2(\omega)\boldsymbol{E}_T(\boldsymbol{r}, \omega) = \boldsymbol{0} \\ \nabla_T \cdot \boldsymbol{E}_T(\boldsymbol{r}, \omega) = 0 \end{cases}$$

where the wavenumber, as usual, is $k^2(\omega) = \omega^2\epsilon(\omega)\mu(\omega)/c_0^2$. When we insert $\boldsymbol{E}_T(\boldsymbol{r}, \omega) = -\nabla_T \Psi(\boldsymbol{r}, \omega)$ into this equation we get

$$\begin{cases} \nabla_T \left(\nabla^2 \Psi(\boldsymbol{r}, \omega) + k^2(\omega)\Psi(\boldsymbol{r}, \omega) \right) = \boldsymbol{0} \\ \nabla_T^2 \Psi(\boldsymbol{r}, \omega) = 0 \end{cases}$$

We eliminate the derivatives w.r.t. the transverse coordinates by using the lower equation. The result is

$$\nabla_T \left\{ \frac{d^2}{dz^2} \Psi(\boldsymbol{r}, \omega) + k^2(\omega)\Psi(\boldsymbol{r}, \omega) \right\} = \boldsymbol{0}$$

which gives

$$\frac{d^2}{dz^2} \Psi(\boldsymbol{r}, \omega) + k^2(\omega)\Psi(\boldsymbol{r}, \omega) = C(z)$$

The fields are independent of the value of $C(z)$ and we can utilize that the potential is not uniquely determined and let $C(z)$ be zero. We have now shown that the scalar potential Ψ satisfies

$$\frac{d^2}{dz^2} \Psi(\boldsymbol{r}, \omega) + k^2(\omega)\Psi(\boldsymbol{r}, \omega) = 0$$

The solutions to this equation are

$$\Psi(\boldsymbol{r}, \omega) = \psi^{\pm}(\boldsymbol{\rho}) e^{\pm \mathrm{i}k(\omega)z}$$

We see that for the TEM-mode $k_z = k(\omega)$, i.e., the wavenumber for propagation in the $\pm z$-direction is the same as the wavenumber for the material. That means that the

phase velocity of the TEM-mode equals the phase velocity of a plane wave in the material in an unbounded space. There is no cut-off frequency for a TEM-mode, which means that if the material is lossless, all frequencies propagate without attenuation. The functions $\psi^{\pm}(\boldsymbol{\rho})$ do not depend on ω, as seen later in the analysis. Based on this result we construct $\boldsymbol{E}_T(\boldsymbol{r},\omega)$ and $\boldsymbol{H}_T(\boldsymbol{r},\omega)$ using equations (5.14) and (5.15). The result is

$$\begin{cases} \boldsymbol{E}_T(\boldsymbol{r},\omega) = -\nabla_T\psi^+(\boldsymbol{\rho})e^{\mathrm{i}k(\omega)z} - \nabla_T\psi^-(\boldsymbol{\rho})e^{-\mathrm{i}k(\omega)z} \\ \boldsymbol{H}_T(\boldsymbol{r},\omega) = (\eta_0\eta(\omega))^{-1}\hat{\boldsymbol{z}} \times \left(\nabla_T\psi^+(\boldsymbol{\rho})e^{\mathrm{i}k(\omega)z} - \nabla_T\psi^-(\boldsymbol{\rho})e^{-\mathrm{i}k(\omega)z}\right) \end{cases} \tag{5.16}$$

where we have introduced the relative wave impedance for the material in the waveguide $\eta(\omega)$ as

$$\eta(\omega) = \sqrt{\frac{\mu(\omega)}{\epsilon(\omega)}} \tag{5.17}$$

Thus the relation between the $\boldsymbol{E}$ and $\boldsymbol{H}$ fields for the TEM-mode is the same as for a plane wave.

A necessary condition for TEM-modes to exist in a waveguide with a homogenous material is that a non-constant scalar potential $\psi(\boldsymbol{\rho})$ exists in the waveguide. This potential is the solution to the two-dimensional Laplace equation. We obtain the two-dimensional Laplace equation from the condition $\nabla \cdot \boldsymbol{E} = 0$ and (5.15)

$$\nabla_T^2\psi(\boldsymbol{\rho}) = 0 \tag{5.18}$$

The boundary condition is $\psi =$ constant on every simply connected part of the curve Γ (this is obtained from $\hat{\boldsymbol{n}} \times \boldsymbol{E} = \boldsymbol{0}$ i.e., $\frac{\partial\psi}{\partial\tau} = 0$).

The problem has a non-trivial solution unless the region Ω is simply connected. An example of a region where a solution exists is shown in figure 5.4b. We also see that the solution ψ is independent of ω. The potential is normalized such that

$$\iint_{\Omega} \nabla_T\psi(\boldsymbol{\rho}) \cdot \nabla_T\psi(\boldsymbol{\rho})\,\mathrm{d}x\mathrm{d}y = 1 \tag{5.19}$$

If we compare the longitudinal wavenumber k_z for the TEM-mode with the corresponding wavenumbers for the TM- and TE-cases, we see that the TEM-case corresponds to the forbidden case $k_t = 0$.

In chapter 3 we used $V^+(z)$ and $V^-(z)$ as the voltages traveling in the positive and negative $z-$directions, respectively. These voltages are related to the voltages $\psi^{\pm}(\boldsymbol{\rho})e^{\pm\mathrm{i}k(\omega)z}$ by

$$V^{\pm}(z) = \psi^{\pm}(\boldsymbol{\rho})e^{\pm\mathrm{i}k(\omega)z}|_{\boldsymbol{\rho}\,\text{on conductor one}} - \psi^{\pm}(\boldsymbol{\rho})e^{\pm\mathrm{i}k(\omega)z}|_{\boldsymbol{\rho}\,\text{on conductor two}}$$

5.3.1 Waveguides with several conductors

Consider a waveguide with N separated simply connected conductors. We denote the border of conductor number n by Γ_n and consider the problems

$$\begin{aligned} &\nabla_T^2\psi_n(\boldsymbol{\rho}) = 0 \\ &\psi_n(\boldsymbol{\rho}) = \begin{cases} V_n & \text{on}\,\Gamma_n \\ V_0 & \text{on}\,\Gamma_m,\ m = 1\ldots N, m \neq n \end{cases} \end{aligned} \tag{5.20}$$

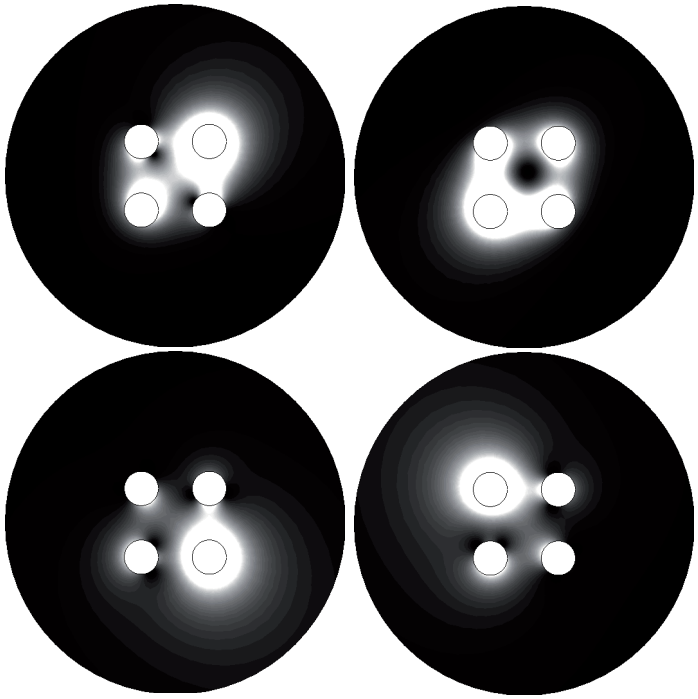


Figure 5.7: The power flow density of the four TEM-modes in a waveguide consisting of four metallic circular conductors in a hollow circular metallic cylinder. The light areas have large power flow density.

where $n = 1 \ldots N-1$ and where $V_n \neq V_0$. It is seen that these $N-1$ problems are independent since we cannot form any of these potentials by superposition of the other ones. The potential for the case when $n = N$ can be formed from the other $N-1$ potentials by superposition. Thus there are $N-1$ independent TEM-modes in a waveguide with N simply connected conductors.

Figure 5.7 depicts the power flow density for the four TEM-modes in a waveguide consisting of four metallic circular conductors in a hollow circular metallic cylinder. The four modes have been obtained from COMSOL. Each of the modes is a superposition of the modes obtained from the problems in (5.20) with $N = 5$.

5.4 Vector basis functions in hollow waveguides

Based on (5.10) and (5.13) we define a set of vector valued basis functions $\boldsymbol{E}^{\pm}_{n\nu}(\boldsymbol{r},\omega)$, that are suitable for the expansion of the electric field in a hollow waveguide.

The vector basis functions for TM-modes that propagate in $\pm z$-direction are defined as [3]

$$\begin{cases} \boldsymbol{E}^{\pm}_{n\nu}(\boldsymbol{r},\omega) = \{\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega) \pm v_n(\boldsymbol{\rho})\hat{\boldsymbol{z}}\}\, e^{\pm \mathrm{i}k_{zn}z} \\ \boldsymbol{H}^{\pm}_{n\nu}(\boldsymbol{r},\omega) = \pm \boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho},\omega) e^{\pm \mathrm{i}k_{zn}z} \end{cases} \qquad \nu = \mathrm{TM} \tag{5.21}$$

[3]Note that the functions are normalized such that $E_z = v_n$ and $H_z = w_n$ are dimensionless. The $\boldsymbol{E}$ and $\boldsymbol{H}$ are to be multiplied with an amplitude with dimension A/m for TE-modes and V/m for TM-modes.

The TE vector basis functions are defined by

$$\begin{cases} \boldsymbol{E}^{\pm}_{n\nu}(\boldsymbol{r},\omega) = \boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega)e^{\pm \mathrm{i}k_{z_n}z} \\ \boldsymbol{H}^{\pm}_{n\nu}(\boldsymbol{r},\omega) = \{\pm\boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho},\omega) + w_n(\boldsymbol{\rho})\hat{\boldsymbol{z}}\}\, e^{\pm \mathrm{i}k_{z_n}z} \end{cases} \qquad \nu = \mathrm{TE} \tag{5.22}$$

We define the vector basis for TEM-modes as

$$\begin{cases} \boldsymbol{E}^{\pm}_{n\nu}(\boldsymbol{r},\omega) = \boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega)e^{\pm \mathrm{i}kz} \\ \boldsymbol{H}^{\pm}_{n\nu}(\boldsymbol{r},\omega) = \pm\boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho},\omega)e^{\pm \mathrm{i}kz} \end{cases} \qquad \nu = \mathrm{TEM} \tag{5.23}$$

where the $n-$index is needed since there exist more than one TEM-mode if there are more than two conductors. In the expressions n is the mode index, while the index ν has the three values, $\nu =$ TM,TE,TEM.

The connections between w_n, v_n, ψ and the transverse components $\boldsymbol{E}_{Tn\nu}$ and $\boldsymbol{H}_{Tn\nu}$ are obtained from (4.7) and (5.16):

$$\begin{cases} \boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega) = \begin{cases} \frac{\mathrm{i}}{k_{t_n}^2}k_{zn}(\omega)\nabla_T v_n(\boldsymbol{\rho}), & \nu = \mathrm{TM} \\ -\frac{\mathrm{i}\omega}{k_{t_n}^2}\mu_0\mu(\omega)\hat{\boldsymbol{z}}\times\nabla_T w_n(\boldsymbol{\rho}), & \nu = \mathrm{TE} \\ -\nabla_T\psi_N(\boldsymbol{\rho},\omega), & \nu = \mathrm{TEM} \end{cases} \\ \boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho},\omega) = Z_{n\nu}^{-1}\hat{\boldsymbol{z}}\times\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega) \end{cases} \tag{5.24}$$

where $Z_{n\nu}$ is the mode impedance defined by

$$Z_{n\nu} = \begin{cases} \frac{k_{zn}}{\omega\epsilon_0\epsilon}, & \nu = \mathrm{TM} \\ \frac{\omega\mu_0\mu}{k_{zn}}, & \nu = \mathrm{TE} \\ \eta_0\eta, & \nu = \mathrm{TEM} \end{cases} \tag{5.25}$$

As before, $v_n(\boldsymbol{\rho})$ and $w_n(\boldsymbol{\rho})$ are the real eigenfunctions to the eigenvalue problems in (5.4) and (5.5). We normalize them such that

$$\begin{cases} \iint\limits_{\Omega} v_n(\boldsymbol{\rho})v_{n'}(\boldsymbol{\rho})\,\mathrm{d}x\mathrm{d}y = \delta_{n,n'} \\ \iint\limits_{\Omega} w_n(\boldsymbol{\rho})w_{n'}(\boldsymbol{\rho})\,\mathrm{d}x\mathrm{d}y = \delta_{n,n'} \end{cases} \tag{5.26}$$

where δ_{ij} is the Kronecker delta (c.f., Appendix D).

The vector basis functions in (5.24) constitute a complete set of vector basis functions for the Maxwell field equations in a source free region of the waveguide. With this notation a general solution can be written in a very condensed form

$$\begin{cases} \boldsymbol{E}(\boldsymbol{r},\omega) = \sum\limits_{\substack{n \\ \nu=\mathrm{TM,TE,TEM}}} \left(a^{+}_{n\nu}\boldsymbol{E}^{+}_{n\nu}(\boldsymbol{r},\omega) + a^{-}_{n\nu}\boldsymbol{E}^{-}_{n\nu}(\boldsymbol{r},\omega)\right) \\ \boldsymbol{H}(\boldsymbol{r},\omega) = \sum\limits_{\substack{n \\ \nu=\mathrm{TM,TE,TEM}}} \left(a^{+}_{n\nu}\boldsymbol{H}^{+}_{n\nu}(\boldsymbol{r},\omega) + a^{-}_{n\nu}\boldsymbol{H}^{-}_{n\nu}(\boldsymbol{r},\omega)\right) \end{cases} \qquad \begin{cases} z \in [z_1, z_2] \\ \boldsymbol{\rho} \in \Omega \end{cases} \tag{5.27}$$

where summation is over the mode indices n and $\nu =$ TM,TE,TEM. An advantage with this definition of the vector basis functions is that the expansion coefficients $a^{\pm}_{n\nu}$ are the same for the expansions of the electric and magnetic fields. The expansion coefficients $a^{\pm}_{n\nu}$ are determined by the sources of the fields c.f., section 5.10. In general the coefficients $a^{+}_{n\nu}$ are determined by sources to the left of the source free region and $a^{-}_{n\nu}$ are determined by sources to the right of the source free region.

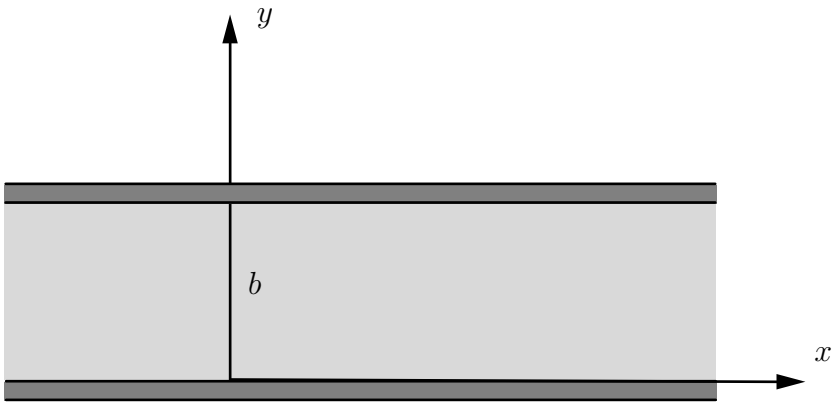


Figure 5.8: Geometry for a planar waveguide.

5.4.1 The fundamental mode

The mode with the lowest cut-off frequency is called the fundamental mode. The bandwidth, BW, of a waveguide is the width of the frequency band where only the fundamental mode can propagate. Let f_0 be the cut-off frequency of the fundamental mode and f_1 the cut-off frequency for the next mode. The fractional bandwidth is defined by

$$b_{\mathrm{f}} = \frac{\mathrm{BW}}{f_{\mathrm{center}}} = 2\frac{f_1 - f_0}{f_1 + f_0} \tag{5.28}$$

where $\mathrm{BW} = f_1 - f_0$ is the bandwidth and $f_{\mathrm{center}} = 0.5(f_0 + f_1)$ is the center frequency. It is very common to use the fundamental mode for transportation of power or information. If the frequency is in the frequency band where only the fundamental mode can propagate then there is only one phase speed involved and that is crucial in order to control the phase of the waves.

5.5 Examples

We now give examples of very important, cross-sections for which we can derive explicit expressions of the vector basis functions.

5.5.1 Planar waveguide

We start with the simplest case, which is the planar waveguide. The geometry of this waveguide is depicted in figure 5.8.

The solution is obtained from the following one-dimensional eigenvalue problems:

$$\begin{cases} \dfrac{d^2Y(y)}{dy^2} + \gamma Y(y) = 0, \quad 0 \le y \le b \\ Y(y) = 0, \quad y = 0, b \end{cases}$$

and

$$\begin{cases} \dfrac{d^2Y(y)}{dy^2} + \gamma Y(y) = 0, \quad 0 \le y \le b \\ \dfrac{dY}{dy}(y) = 0, \quad y = 0, b \end{cases}$$

	Eigenfunctions v_n, w_n, ψ	**Eigenvalues** $k_{t\,n}^2$
TM_n	$v_n = \sqrt{\frac{2}{b}} \sin\left(\frac{n\pi y}{b}\right)$	$\pi^2 \frac{n^2}{b^2}$
TE_n	$w_n = \sqrt{\frac{2}{b}} \cos\left(\frac{n\pi y}{b}\right)$	$\pi^2 \frac{n^2}{b^2}$
TEM	$\nabla_T \psi = \sqrt{\frac{1}{b}} \hat{\boldsymbol{y}}$	0

Table 5.1: A table of the normalized eigenfunctions to equations (5.4), (5.5) and (5.18) for the planar waveguide, c.f., figure 5.8 for the definition of the geometry. The integer n has the values $n = 1, 2, 3, \ldots$.

The solutions to these problems are given by

$$Y_n(y) = \sin\left(\frac{n\pi y}{b}\right), \quad n = 1, 2, 3, \ldots$$

and

$$Y_n(y) = \cos\left(\frac{n\pi y}{b}\right), \quad n = 0, 1, 2, 3, \ldots$$

respectively. The eigenvalues are in both cases $\gamma = n^2\pi^2/b^2$. These sets of functions are complete. From the functions we construct the functions v_n and w_n, that in this case only depend on the coordinate y. The normalized TM-case basis functions are

$$v_n(y) = \sqrt{\frac{2}{b}} \sin\left(\frac{n\pi y}{b}\right), \quad n = 1, 2, 3, \ldots$$

and the TE-case basis functions are

$$w_n(y) = \sqrt{\frac{2}{b}} \cos\left(\frac{n\pi y}{b}\right), \quad n = 1, 2, 3, \ldots$$

The index $n = 0$ does nor correspond to a TM-mode, as seen from page 78.

The planar waveguide has two separated surfaces which means that a TEM-mode exists. We use (5.18) to determine the TEM-mode basis functions

$$\begin{cases} \dfrac{d^2\psi(y)}{dy^2} = 0, \quad 0 \le y \le b \\ \psi(y) = \begin{cases} C_1, & y = 0 \\ C_2, & y = b \end{cases} \end{cases}$$

The solution is

$$\psi(y) = C_1 + \frac{C_2 - C_1}{b} y$$

and the TEM-mode has the normalized basis function (see (5.16))

$$\nabla_T \psi^{\pm}(y) = \sqrt{\frac{1}{b}} \hat{\boldsymbol{y}}$$

The results are collected in table 5.1.

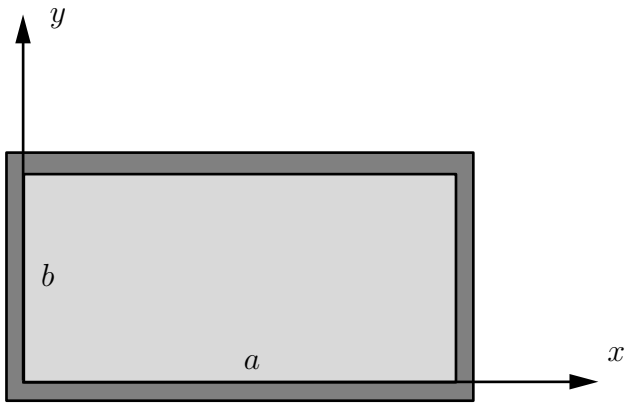


Figure 5.9: The geometry for a waveguide with rectangular cross-section.

5.5.2 Waveguide with rectangular cross-section

We continue and determine the eigenfunctions for the rectangular waveguide. This is the most common type of hollow waveguide. The geometry is depicted in figure 5.9. The surface is simply connected and hence no TEM-mode exists. The convention is to let the longest side of the rectangle be along the x-axis.

The eigenvalues that are to be solved are

$$\begin{cases} \dfrac{\partial^2 v(\boldsymbol{\rho})}{\partial x^2} + \dfrac{\partial^2 v(\boldsymbol{\rho})}{\partial y^2} + k_t^2 v(\boldsymbol{\rho}) = 0 \\ v(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \qquad \text{(TM-case)}$$

and

$$\begin{cases} \dfrac{\partial^2 w(\boldsymbol{\rho})}{\partial x^2} + \dfrac{\partial^2 w(\boldsymbol{\rho})}{\partial y^2} + k_t^2 w(\boldsymbol{\rho}) = 0 \\ \dfrac{\partial w}{\partial n}(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \qquad \text{(TE-case)}$$

The solution is based on the following one-dimensional eigenvalue problems:

$$\begin{cases} \dfrac{d^2 X(x)}{dx^2} + \gamma X(x) = 0, \quad 0 \le x \le a \\ X(x) = 0, \quad x = 0, a \end{cases}$$

and

$$\begin{cases} \dfrac{d^2 \tilde{X}(x)}{dx^2} + \gamma \tilde{X}(x) = 0, \quad 0 \le x \le a \\ \dfrac{d\tilde{X}}{dx}(x) = 0, \quad x = 0, a \end{cases}$$

The solutions to these two problems are

$$X_m(x) = \sin\left(\frac{m\pi x}{a}\right), \quad m = 1, 2, 3, \ldots$$

and

$$\tilde{X}_m(x) = \cos\left(\frac{m\pi x}{a}\right), \quad m = 0, 1, 2, 3, \ldots,$$

	Eigenfunctions v_{mn}, w_{mn}	**Eigenvalues** $k_{t\,mn}^2$
TM_{mn}	$v_{mn} = \dfrac{2}{\sqrt{ab}} \sin\left(\dfrac{m\pi x}{a}\right) \sin\left(\dfrac{n\pi y}{b}\right)$	$\pi^2 \left(\dfrac{m^2}{a^2} + \dfrac{n^2}{b^2}\right)$
TE_{mn}	$w_{mn} = \sqrt{\dfrac{\varepsilon_m \varepsilon_n}{ab}} \cos\left(\dfrac{m\pi x}{a}\right) \cos\left(\dfrac{n\pi y}{b}\right)$	$\pi^2 \left(\dfrac{m^2}{a^2} + \dfrac{n^2}{b^2}\right)$

Table 5.2: Table of normalized eigenfunctions to equations (5.4) and (5.5) for rectangular waveguides, see figure 5.9. The integers m and n can have values $m, n = 0, 1, 2, 3, \ldots$, with the exception that m and n are not zero for TM-modes, and m and n cannot **both** be zero for the TE-modes ($\varepsilon_m = 2 - \delta_{m,0}$), see page 78. The convention in this book is always to have the long side of the rectangle along the x-axis, i.e., $a > b$. The mode with the lowest cut-off frequency is then the TE_{10} mode. This mode is called the fundamental mode and is very important.

respectively. These sets of functions are orthogonal and complete on the interval $x \in [0, a]$. The solution to the two-dimensional eigenvalue problems for the rectangular waveguide are obtained as a product of these sets of one-dimensional eigenfunctions [4], i.e.,

$$\begin{cases} \sin\left(\dfrac{m\pi x}{a}\right) \sin\left(\dfrac{n\pi y}{b}\right), & \text{TM-case} \\ \cos\left(\dfrac{m\pi x}{a}\right) \cos\left(\dfrac{n\pi y}{b}\right), & \text{TE-case} \end{cases}$$

The eigenvalues in the two cases are the same $k_t^2 = \pi^2 \left(m^2/a^2 + n^2/b^2\right)$. The normalized functions are

$$\begin{cases} v_{mn} = \dfrac{2}{\sqrt{ab}} \sin\left(\dfrac{m\pi x}{a}\right) \sin\left(\dfrac{n\pi y}{b}\right), & \text{TM-case} \\ w_{mn} = \sqrt{\dfrac{\varepsilon_m \varepsilon_n}{ab}} \cos\left(\dfrac{m\pi x}{a}\right) \cos\left(\dfrac{n\pi y}{b}\right), & \text{TE-case} \end{cases}$$

where the Neumann-factor is $\varepsilon_m = 2 - \delta_{m,0}$. The results are collected in table 5.2.

Example 5.3
The fundamental mode of a rectangular waveguide with $a > b$ is the TE_{10} mode. It has the cut-off frequency $f_{c10} = \dfrac{c_0}{2a}$ and $w_{10} = \sqrt{\dfrac{2}{ab}} \cos\left(\dfrac{\pi x}{a}\right)$. The normalized electric field is

$$\boldsymbol{E}_{10\mathrm{TE}}(x, \omega) = \hat{\boldsymbol{y}} \frac{\mathrm{i}\omega\mu_0}{\pi} \sqrt{\frac{2a}{b}} \sin\left(\frac{\pi x}{a}\right) \tag{5.29}$$

[4] A common method to create complete sets of functions in two dimensions is to take the product of one-dimensional systems, i.e., if $\{f_m(x)\}_{m=1}^{\infty}$ and $\{g_n(y)\}_{n=1}^{\infty}$ are complete systems on the intervals $x \in [a, b]$ and $y \in [c, d]$, respectively, then

$$\{f_m(x) g_n(y)\}_{m,n=1}^{\infty}$$

is a complete set of functions in the rectangle $[a, b] \times [c, d]$.

m	1	2	0	1	2	3	3	4	0	1
n	0	0	1	1	1	0	1	0	2	2
f_{cmn} (GHz)	3.19	6.38	6.81	7.52	9.33	9.57	11.7	12.8	13.6	14.0
k_{zmn} $(\mathrm{m}^{-1})^a$	43.3	$107i$	$119i$	$136i$	$179i$	$184i$	$233i$	$255i$	$274i$	$282i$
k_{zmn} $(\mathrm{m}^{-1})^b$	144.6	86.6	70.6	22.6	$114i$	$122i$	$188i$	$215i$	$237i$	$246i$

[a]The frequency is $f = 3.8$ GHz.
[b]The frequency is $f = 7.6$ GHz.

Table 5.3: Table of the lowest cut-off frequencies f_{cmn} and the longitudinal wavenumber k_{zmn} for a rectangular waveguide with dimensions 4.7 cm × 2.2 cm. Only TE-modes can have m- or n-values that are zero. For frequencies below the cut-off frequency the longitudinal wavenumber k_{zmn} is imaginary and the corresponding mode is non-propagating. The attenuation of that mode is $\exp(-\mathrm{Im}\{k_{zmn}\}z)$.

If $a > 2b$ then the second mode is TE_{20} that has cut-off frequency $f_{c20} = \dfrac{c_0}{a}$. If $b < a < 2b$ then TE_{01} is the second mode with cut-off frequency $f_{c01} = \dfrac{c_0}{2b}$. In order to maximize the bandwidth it is common to have rectangular waveguides with $a > 2b$. Then the bandwidth is $\mathrm{BW} = \dfrac{c_0}{2a}$ and the fractional bandwidth is $b_\mathrm{f} = 2(c_0/2a)/(3c_0/2a) = 2/3 = 0.67$. ∎

Example 5.4
A rectangular waveguide has dimensions 4.7 cm × 2.2 cm. The cut-off frequencies f_{cmn} for the different modes are easy to calculate from (5.12) and table 5.2. The longitudinal wavenumbers k_{zmn}, given by (5.9), are related to the frequency f and the cut-off frequency f_{cmn} in the following way

$$k_{zmn} = \frac{2\pi}{c_0}\sqrt{\epsilon\mu}\sqrt{f^2 - f^2{}_{cmn}}$$

The results are given in table 5.3. The bandwidth is BW = 3.19 Ghz and the fractional bandwidth is $b_\mathrm{f} = 1$. ∎

5.5.3 Waveguide with circular cross-section

The geometry of the circular waveguide with radius a is depicted in figure 5.10. The geometry has only one simply connected surface and hence there is no TEM-mode. It is best to solve the eigenvalue problem in cylindrical-(polar)coordinates. The eigenvalue problems are given by

$$\begin{cases} \nabla_T^2 v(\boldsymbol{\rho}) + k_t^2 v(\boldsymbol{\rho}) = \dfrac{1}{\rho}\dfrac{\partial}{\partial\rho}\left(\rho\dfrac{\partial v(\boldsymbol{\rho})}{\partial\rho}\right) + \dfrac{1}{\rho^2}\dfrac{\partial^2 v(\boldsymbol{\rho})}{\partial\phi^2} + k_t^2 v(\boldsymbol{\rho}) = 0 \\ v(a,\phi) = 0 \end{cases} \qquad \text{(TM-case)}$$

and

$$\begin{cases} \nabla_T^2 w(\boldsymbol{\rho}) + k_t^2 w(\boldsymbol{\rho}) = \dfrac{1}{\rho}\dfrac{\partial}{\partial\rho}\left(\rho\dfrac{\partial w(\boldsymbol{\rho})}{\partial\rho}\right) + \dfrac{1}{\rho^2}\dfrac{\partial^2 w(\boldsymbol{\rho})}{\partial\phi^2} + k_t^2 w(\boldsymbol{\rho}) = 0 \\ \dfrac{\partial w}{\partial n}(a,\phi) = 0 \end{cases} \qquad \text{(TE-case)}$$

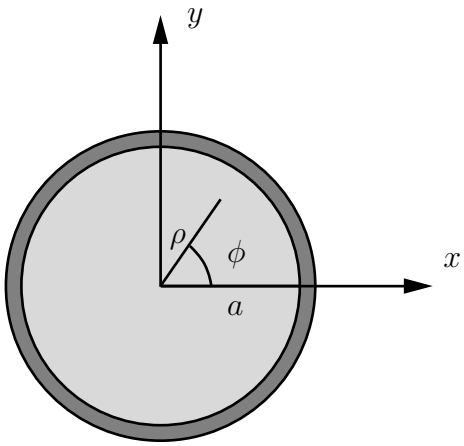


Figure 5.10: Geometry for waveguide with circular cross-section.

We solve these eigenvalue problems by the method of separation of variables. We make the ansatz $v(\rho, \phi) = f(\rho)g(\phi)$ and insert this into the differential equation. After division with $f(\rho)g(\phi)/\rho^2$ we get

$$\frac{\rho}{f(\rho)}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial f(\rho)}{\partial\rho}\right) + k_t^2\rho^2 = -\frac{1}{g(\phi)}\frac{\partial^2 g(\phi)}{\partial\phi^2}$$

The right hand side depends only on ϕ and the left hand side depends only on ρ. That means that they both have to be equal to a constant and we denote this constant γ. We get

$$\begin{cases} \rho\dfrac{\partial}{\partial\rho}\left(\rho\dfrac{\partial f(\rho)}{\partial\rho}\right) + \left(k_t^2\rho^2 - \gamma\right) f(\rho) = 0 \\ \dfrac{\partial^2 g(\phi)}{\partial\phi^2} + \gamma g(\phi) = 0 \end{cases}$$

The solution to the eigenvalue problem in the variable ϕ is

$$g(\phi) = \begin{pmatrix}\cos m\phi \\ \sin m\phi\end{pmatrix}, \qquad m = 0, 1, 2, 3, \ldots$$

Only integer values of m are allowed since the function must be periodic in ϕ with period 2π, i.e., only $\gamma = m^2$, $m = 0, 1, 2, 3, \ldots$ are possible values. The corresponding set of functions is complete on the interval $\phi \in [0, 2\pi)$. The solution to the equation in the ρ-variable is a Bessel function, see appendix A. Only solutions that are regular in $\rho = 0$ are valid, i.e.,

$$f(\rho) = J_m(k_t\rho)$$

The boundary conditions $v_m(a, \phi) = 0$ and $\frac{dw_m}{d\rho}(a, \phi) = 0$ for the TM- and TE-cases, respectively, add extra conditions. For these boundary conditions to be satisfied, the transverse wavenumber has to satisfy

$$k_t a = \begin{cases} \xi_{mn}, & \text{(TM-case)} \\ \eta_{mn}, & \text{(TE-case)} \end{cases}$$

	Eigenfunctions v_{mn}, w_{mn}	**Eigenvalues** $k_{t\,mn}^2$
TM_{mn}	$v_{mn} = \dfrac{\sqrt{\varepsilon_m} J_m(\xi_{mn}\rho/a)}{\sqrt{\pi} a J'_m(\xi_{mn})} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}$	$\dfrac{\xi_{mn}^2}{a^2}$
TE_{mn}	$w_{mn} = \dfrac{\sqrt{\varepsilon_m} \eta_{mn} J_m(\eta_{mn}\rho/a)}{\sqrt{\pi\,(\eta_{mn}^2 - m^2)} a J_m(\eta_{mn})} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}$	$\dfrac{\eta_{mn}^2}{a^2}$

Table 5.4: Table of the normalized eigenfunctions to equations (5.4) and (5.5) for waveguides with circular cross-section, see figure 5.10 for definition of geometry. ($\varepsilon_m = 2 - \delta_{m,0}$). The first values of the positive zeros ξ_{mn} to $J_m(x)$ and the positive zeros η_{mn} to $J'_m(x)$, i.e., $J_m(\xi_{mn}) = 0$ and $J'_m(\eta_{mn}) = 0$, $m = 0, 1, 2, 3, \ldots$, $n = 1, 2, 3, \ldots$ are listed in tables A.1 and A.2 in appendix A. The mode with the lowest cut-off frequency is the TE_{11} mode.

where ξ_{mn} and η_{mn}, $n = 1, 2, 3, \ldots$, are zeros to the Bessel function $J_m(x)$ and to the derivative of the Bessel function, respectively, i.e., $J_m(\xi_{mn}) = 0$ and $J'_m(\eta_{mn}) = 0$. Numerical values of the first of these zeros are given in appendix A.

The sets of functions $\{J_m(\xi_{mn}\rho/a)\}_{n=1}^{\infty}$, $\{J'_m(\eta_{mn}\rho/a)\}_{n=1}^{\infty}$ are both complete on the interval $\rho \in [0, a]$ for every value of m. The complete set of functions in the circle is, in analogy with the rectangular waveguide, given by the product of the sets of basis functions. The results of the normalized functions (the normalization integrals are given in appendix A) are

$$\begin{cases} v_{mn} = \dfrac{\sqrt{\varepsilon_m} J_m(\xi_{mn}\rho/a)}{\sqrt{\pi} a J'_m(\xi_{mn})} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}, & \text{TM-case} \\ w_{mn} = \dfrac{\sqrt{\varepsilon_m} \eta_{mn} J_m(\eta_{mn}\rho/a)}{\sqrt{\pi\,(\eta_{mn}^2 - m^2)} a J_m(\eta_{mn})} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}, & \text{TE-case} \end{cases}$$

where $\varepsilon_m = 2 - \delta_{m,0}$. The results are collected in table 5.4.

Example 5.5
The fundamental mode is the TE_{11} mode. The cut-off frequency is given by $f_{c_{11}} = \dfrac{c_0 \eta_{11}}{2\pi a}$ where a is the radius of the cylinder and $\eta_{11} = 1.841$ is the first zero of $J'_1(x)$. The second mode is the TM_{01} mode with cut-off frequency $f_{c_{01}} = \dfrac{c_0 \xi_{01}}{2\pi a}$ where $\xi_{01} = 2.405$ is the first zero of $J_0(x)$. The bandwidth is $\mathrm{BW} = f_{c_{01}} - f_{c_{11}} = \dfrac{c_0}{2\pi a}(2.405 - 1.841) = 0.564\dfrac{c_0}{2\pi a}$. The fractional bandwidth is $b_{\mathrm{f}} = 2\dfrac{0.564}{1.841 + 2.405} = 0.265$. ∎

5.6 Analyzing waveguides with FEM

Waveguides with arbitrary cross-sections can be analyzed with numerical methods and in this book we use the finite element method. The specific calculations are done with the commercial software package COMSOL Multiphysics. We use COMSOL to find the cut-off frequencies for the TE- and TM-modes in a hollow waveguide filled with a homogenous non-conducting material with permittivity ϵ. We also let COMSOL determine the electric

and magnetic fields and the power flow density for the lowest modes. In COMSOL we do the following steps:

- We choose **2D> Radio frequency> Electromagnetic waves> Eigenfrequency study**.
- We draw the cross section of the waveguide.
- In **Study>Eigenfrequency** we define how many modes that are to be determined and the cut-off frequency where COMSOL starts to look for eigenfrequencies.
- We let COMSOL solve the eigenvalue problem. It then shows the electric field in the cross section of the waveguide for the different modes. It also gives the cut-off frequencies f_c for the modes. From the cut-off frequencies we get the corresponding k_t from $k_t = \omega/c_0 = 2\pi f_c/c_0$. There are spurious solutions with very low frequencies, or complex frequencies that COMSOL presents. These can be recognized in the field plot since they have a fuzzy field plot.
- To distinguish TE- from TM-modes we plot the $z-$component of the electric field. If the plot is fuzzy with very small field values then the mode is a TE-mode, otherwise it is a TM-mode.
- The fields that COMSOL presents are not normalized. We use a normalization such that $\iint_\Omega |E_z(\boldsymbol{\rho})|^2 \,\mathrm{d}x\mathrm{d}y = 1$ for the TM-modes and $\iint_\Omega |H_z(\boldsymbol{\rho})|^2 \,\mathrm{d}x\mathrm{d}y = 1$ for the TE-modes. To obtain this normalization we divide all field values with $\iint_\Omega |E_z(\boldsymbol{\rho})|^2 \,\mathrm{d}x\mathrm{d}y$ for the TM-modes and $\iint_\Omega |H_z(\boldsymbol{\rho})|^2 \,\mathrm{d}x\mathrm{d}y$ for the TE-modes. To integrate we right clock on **Derived values** and choose integration and surface integral.
- Notice that there are many options of surface graphs to choose from.

Example 5.6

We analyze the ridge waveguide. This is a waveguide with a large bandwidth since the fundamental mode has a very low cut-off frequency. The cross-section is depicted in figure 5.11. We use the scheme for COMSOL to obtain the modes. The TM modes have cut-off frequencies $f_c = 3.88$ GHz, 3.88 GHz, 5.83 GHz, 5.83 GHz, 6.43 GHz, 6.43 GHz and the TE modes have cut-off frequencies $f_c = 0.663$ GHz, 2.51 GHz, 2.51 GHz, 2.86 GHz, 3.14 GHz. In figure 5.12 we see the active power flow density for one of the TM-mode with the cut-off frequency 5.83 GHz. The other mode with the same cut-off frequency has its power flow in the right part of the waveguide. This mode is very close to the TM_{12} mode in a 5 cm $\times$ 6 cm rectangular waveguide. The TM_{12} mode has cut-off frequency $f_c = 5.831$ GHz which is very close to the cut-off frequency 5.827 GHz obtained for the mode in figure 5.11. The fractional bandwidth of this waveguide is $b_f = 2(f_{\text{upper}} - f_{\text{lower}})/(f_{\text{upper}} + f_{\text{lower}})$. The ridge waveguide has the fractional bandwidth of 1.16. This is to be compared with the value 0.67 for a rectangular waveguide with $a > 2b$.

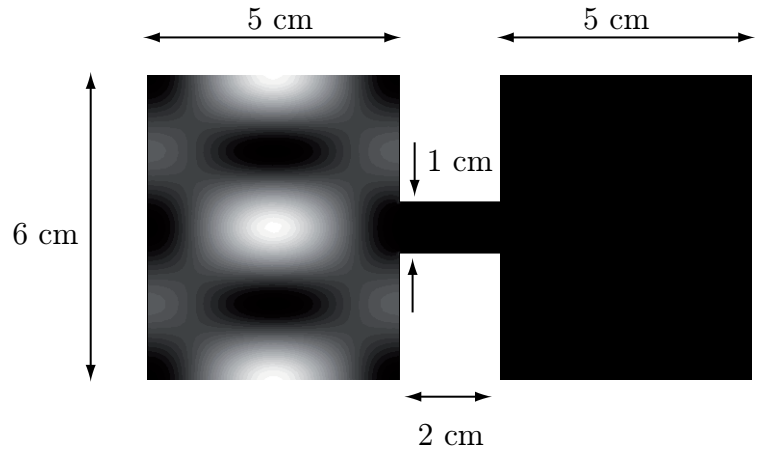


Figure 5.11: Geometry for the ridge waveguide and the power flow density for the TM-mode with cut-off frequency 5.83 GHz.

5.7 Normalization integrals

In this section we solve the following integrals

$$\begin{aligned} &\iint_{\Omega} \hat{\boldsymbol{z}} \cdot \{\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega) \times \boldsymbol{H}^{*}_{Tn'\nu'}(\boldsymbol{\rho},\omega)\}\, \mathrm{d}x\mathrm{d}y \\ &\iint_{\Omega} \hat{\boldsymbol{z}} \cdot \{\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega) \times \boldsymbol{H}_{Tn'\nu'}(\boldsymbol{\rho},\omega)\}\, \mathrm{d}x\mathrm{d}y \end{aligned} \tag{5.30}$$

The real part of the first integral is the power that the mode $n\nu$ transports in the waveguide, see section avs:Effekt. The second integral is needed when we treat sources in waveguides in section 5.10. The values of the integrals are given in equations (5.37) and (5.39).

The eigenfunctions $v_n(\boldsymbol{\rho})$ and $w_n(\boldsymbol{\rho})$ are real functions that are normalized according to (5.26), i.e.,

$$\begin{cases} \displaystyle\iint_{\Omega} v_n(\boldsymbol{\rho}) v_{n'}(\boldsymbol{\rho})\, \mathrm{d}x\mathrm{d}y = \delta_{n,n'} \\ \displaystyle\iint_{\Omega} w_n(\boldsymbol{\rho}) w_{n'}(\boldsymbol{\rho})\, \mathrm{d}x\mathrm{d}y = \delta_{n,n'} \end{cases}$$

From the relation $\nabla \cdot (g\nabla f) = g(\nabla^2 f) + \nabla g \cdot \nabla f$ and Gauss' theorem (in two dimensions) we obtain

$$\begin{aligned} \iint_{\Omega} \nabla_T g(\boldsymbol{\rho}) \cdot \nabla_T f(\boldsymbol{\rho})\, \mathrm{d}x\mathrm{d}y = &- \iint_{\Omega} g(\boldsymbol{\rho}) \nabla_T^2 f(\boldsymbol{\rho})\, \mathrm{d}x\mathrm{d}y \\ &+ \int_{\Gamma} g(\boldsymbol{\rho}) \hat{\boldsymbol{n}} \cdot \nabla_T f(\boldsymbol{\rho})\, \mathrm{d}l \end{aligned} \tag{5.31}$$

where $\mathrm{d}l$ is the line element along the border curve Γ.

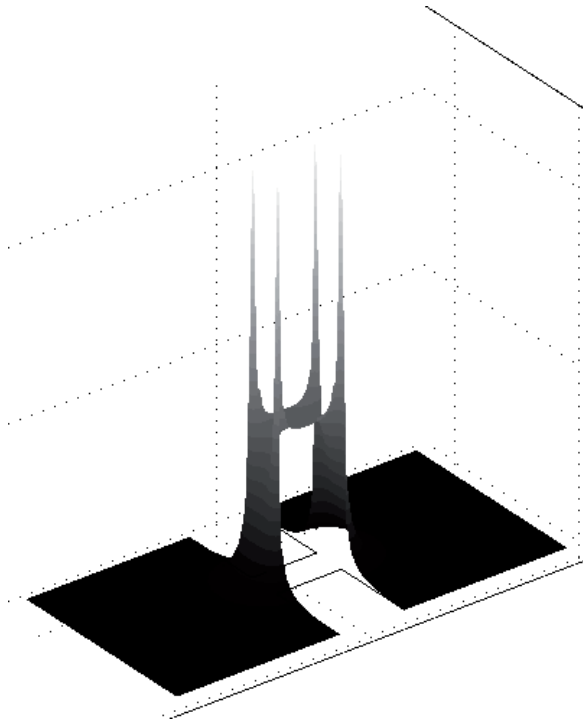


Figure 5.12: The power flow density for the fundamental mode of the ridge waveguide, i.e., the TE-mode with cut-off frequency 0.663 GHz. The power flow is concentrated to the narrow section. Notice that the electric field is very strong at edges.

If we use the eigenfunctions for the TM-case such that $f(\boldsymbol{\rho}) = v_n(\boldsymbol{\rho})$ and $g(\boldsymbol{\rho}) = v_{n'}(\boldsymbol{\rho})$, the curve integral is zero since $v_n(\boldsymbol{\rho}) = 0$ on the boundary curve Γ. We get

$$\begin{aligned} \iint_\Omega \nabla_T v_n(\boldsymbol{\rho}) \cdot \nabla_T v_{n'}(\boldsymbol{\rho}) \,\mathrm{d}x\mathrm{d}y &= -\iint_\Omega v_n(\boldsymbol{\rho}) \nabla_T^2 v_{n'}(\boldsymbol{\rho}) \,\mathrm{d}x\mathrm{d}y \\ &= k_{tn'}^2 \iint_\Omega v_n(\boldsymbol{\rho}) v_{n'}(\boldsymbol{\rho}) \,\mathrm{d}x\mathrm{d}y = k_{tn}^2 \delta_{n,n'} \end{aligned} \tag{5.32}$$

where we have used equations (5.4) and (5.26).

If we use the eigenfunctions for the TE-case, i.e., $f(\boldsymbol{\rho}) = w_n(\boldsymbol{\rho})$ and $g(\boldsymbol{\rho}) = w_{n'}(\boldsymbol{\rho})$ in (5.31), the curve integral is zero since $\frac{\partial}{\partial n} w_n(\boldsymbol{\rho}) = 0$ on the boundary curve Γ, and hence

$$\begin{aligned} \iint_\Omega \nabla_T w_n(\boldsymbol{\rho}) \cdot \nabla_T w_{n'}(\boldsymbol{\rho}) \,\mathrm{d}x\mathrm{d}y &= -\iint_\Omega w_n(\boldsymbol{\rho}) \nabla_T^2 w_{n'}(\boldsymbol{\rho}) \,\mathrm{d}x\mathrm{d}y \\ &= k_{tn'}^2 \iint_\Omega w_n(\boldsymbol{\rho}) w_{n'}(\boldsymbol{\rho}) \,\mathrm{d}x\mathrm{d}y = k_{tn}^2 \delta_{n,n'} \end{aligned} \tag{5.33}$$

where we used equations (5.5) and (5.26). We see that the gradient of the eigenfunctions are othogonal, but not normalized.

Four more combinations of integrals are of interest in later sections:

$$\begin{cases} \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\nabla_T v_n(\boldsymbol{\rho}) \times \nabla_T w_{n'}(\boldsymbol{\rho})\} \, \mathrm{d}x\mathrm{d}y \\ \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{(\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times (\hat{\boldsymbol{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho}))\} \, \mathrm{d}x\mathrm{d}y \end{cases} \tag{5.34}$$

and

$$\begin{cases} \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\nabla_T v_n(\boldsymbol{\rho}) \times (\hat{\boldsymbol{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho}))\} \, \mathrm{d}x\mathrm{d}y \\ \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{(\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times \nabla_T w_{n'}(\boldsymbol{\rho})\} \, \mathrm{d}x\mathrm{d}y \end{cases} \tag{5.35}$$

We rewrite the first integral in (5.34) by using formulas for the nabla-operator and Stoke's theorem (the unit normal vector to Ω is $\hat{\boldsymbol{z}}$).

$$\begin{aligned} \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\nabla_T v_n(\boldsymbol{\rho}) \times \nabla_T w_{n'}(\boldsymbol{\rho})\} \, \mathrm{d}x\mathrm{d}y &= \hat{\boldsymbol{z}} \cdot \iint\limits_{\Omega} \nabla_T \times (v_n(\boldsymbol{\rho}) \nabla_T w_{n'}(\boldsymbol{\rho})) \, \mathrm{d}x\mathrm{d}y \\ = \iint\limits_{\Omega} \nabla \times (v_n(\boldsymbol{\rho}) \nabla_T w_{n'}(\boldsymbol{\rho})) \cdot \hat{\boldsymbol{n}} \, \mathrm{d}x\mathrm{d}y &= \int\limits_{\Gamma} (v_n(\boldsymbol{\rho}) \nabla_T w_{n'}(\boldsymbol{\rho})) \cdot \mathrm{d}\boldsymbol{r} = 0 \end{aligned}$$

since $v_n(\boldsymbol{\rho}) = 0$ on the boundary curve Γ. The linear element is $d\boldsymbol{r} = \hat{\boldsymbol{z}} \times \hat{\boldsymbol{n}} dl = \hat{\boldsymbol{\tau}} dl$. In the same manner we can show that the second integral in (5.34) is zero by using $(\boldsymbol{c} \times \boldsymbol{a}) \times (\boldsymbol{c} \times \boldsymbol{b}) = \boldsymbol{c}\,((\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c})$ and utilizing

$$\begin{aligned} \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{(\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times (\hat{\boldsymbol{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho}))\} \, \mathrm{d}x\mathrm{d}y \\ = \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\nabla_T w_n(\boldsymbol{\rho}) \times \nabla_T v_{n'}(\boldsymbol{\rho})\} \, \mathrm{d}x\mathrm{d}y = 0 \end{aligned}$$

The integrals in (5.35) are simplified by using the BAC-CAB-rule.

$$\begin{aligned} \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\nabla_T v_n(\boldsymbol{\rho}) \times (\hat{\boldsymbol{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho}))\} \, \mathrm{d}x\mathrm{d}y \\ = \iint\limits_{\Omega} \nabla_T v_n(\boldsymbol{\rho}) \cdot \nabla_T v_{n'}(\boldsymbol{\rho}) \, \mathrm{d}x\mathrm{d}y \\ \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{(\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times \nabla_T w_{n'}(\boldsymbol{\rho})\} \, \mathrm{d}x\mathrm{d}y \\ = - \iint\limits_{\Omega} \nabla_T w_n(\boldsymbol{\rho}) \cdot \nabla_T w_{n'}(\boldsymbol{\rho}) \, \mathrm{d}x\mathrm{d}y \end{aligned}$$

By using the orthogonality integrals (5.32) and (5.33) we can summarize the integrals in

(5.34) and (5.35) as

$$\begin{cases} \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\nabla_T v_n(\boldsymbol{\rho}) \times \nabla_T w_{n'}(\boldsymbol{\rho})\} \, \mathrm{d}x\mathrm{d}y = 0 \\ \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{(\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times (\hat{\boldsymbol{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho}))\} \, \mathrm{d}x\mathrm{d}y = 0 \\ \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\nabla_T v_n(\boldsymbol{\rho}) \times (\hat{\boldsymbol{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho}))\} \, \mathrm{d}x\mathrm{d}y = k_{t_n}^2 \delta_{n,n'} \\ \iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{(\hat{\boldsymbol{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times \nabla_T w_{n'}(\boldsymbol{\rho})\} \, \mathrm{d}x\mathrm{d}y = -k_{t_n}^2 \delta_{n,n'} \end{cases} \tag{5.36}$$

We use the integrals in (5.36) to calculate the following normal surface integral:

$$\iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\boldsymbol{E}_{T n\nu}(\boldsymbol{\rho}, \omega) \times \boldsymbol{H}^*_{T n'\nu'}(\boldsymbol{\rho}, \omega)\} \, \mathrm{d}x\mathrm{d}y$$

This integral is used in section 5.8. By using the definitions of $\boldsymbol{E}_{T n\nu}(\boldsymbol{\rho}, \omega)$ and $\boldsymbol{H}_{T n\nu}(\boldsymbol{\rho}, \omega)$ in (5.24) we find by using (5.36) that

$$\iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\boldsymbol{E}_{T n\nu}(\boldsymbol{\rho}, \omega) \times \boldsymbol{H}^*_{T n'\nu'}(\boldsymbol{\rho}, \omega)\} \, \mathrm{d}x\mathrm{d}y = 2P^E_{n\nu} \delta_{n,n'} \delta_{\nu,\nu'} \tag{5.37}$$

We have here introduced the mode power $P^E_{n\nu}$ as

$$P^E_{n\nu} = \begin{cases} \dfrac{\omega}{2k_{t_n}^2} \begin{cases} k_{z_n}(\omega)\epsilon_0\epsilon^*(\omega), & \nu = \mathrm{TM} \\ k^*_{z_n}(\omega)\mu_0\mu(\omega), & \nu = \mathrm{TE} \end{cases} \\ \dfrac{1}{2\eta_0\eta^*}, \quad \nu = \mathrm{TEM} \end{cases} \tag{5.38}$$

In lossless materials the mode powers are real, and thus purely active, for frequencies above the cut-off frequency, and imaginary, and thus purely reactive, for frequencies below the cut-off frequency.

In section 5.10 we need the following integral

$$\iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot \{\boldsymbol{E}_{T n\nu}(\boldsymbol{\rho}, \omega) \times \boldsymbol{H}_{T n'\nu'}(\boldsymbol{\rho}, \omega)\} \, \mathrm{d}x\mathrm{d}y = -2U_{n\nu} \delta_{n,n'} \delta_{\nu,\nu'} \tag{5.39}$$

where

$$U_{n\nu} = \begin{cases} \dfrac{\omega k_{z_n}(\omega)}{2k_{t_n}^2} \begin{cases} \epsilon_0\epsilon(\omega), & \nu = \mathrm{TM} \\ \mu_0\mu(\omega), & \nu = \mathrm{TE} \end{cases} \\ \dfrac{1}{2\eta_0\eta}, \quad \nu = \mathrm{TEM} \end{cases} \tag{5.40}$$

5.8 Power flow density

The active power flow density of the electromagnetic field is given by the Poynting vector(see (1.20) on page 9)

$$<\boldsymbol{S}(t)> (\boldsymbol{r}, \omega) = \frac{1}{2} \mathrm{Re} \{\boldsymbol{E}(\boldsymbol{r}, \omega) \times \boldsymbol{H}^*(\boldsymbol{r}, \omega)\}$$

We are in particular interested of the z-component of this vector, since it gives the time average of the power per unit area transported along the waveguide

$$\hat{\boldsymbol{z}}\cdot <\boldsymbol{S}(t)> (\boldsymbol{r},\omega) = \frac{1}{2}\hat{\boldsymbol{z}}\cdot \mathrm{Re}\left\{\boldsymbol{E}(\boldsymbol{r},\omega)\times \boldsymbol{H}^*(\boldsymbol{r},\omega)\right\}$$

We introduce the decomposition of the vector fields in their longitudinal and transverse parts according to section 4.1.

$$\begin{aligned}\hat{\boldsymbol{z}}\cdot <\boldsymbol{S}(t)> (\boldsymbol{r},\omega) &= \frac{1}{2}\hat{\boldsymbol{z}}\cdot \mathrm{Re}\left\{(\boldsymbol{E}_T(\boldsymbol{r},\omega)+\hat{\boldsymbol{z}}E_z(\boldsymbol{r},\omega))\times(\boldsymbol{H}_T^*(\boldsymbol{r},\omega)+\hat{\boldsymbol{z}}H_z^*(\boldsymbol{r},\omega))\right\}\\ &= \frac{1}{2}\hat{\boldsymbol{z}}\cdot \mathrm{Re}\left\{\boldsymbol{E}_T(\boldsymbol{r},\omega)\times\boldsymbol{H}_T^*(\boldsymbol{r},\omega)\right\}\end{aligned}$$

We use the general waveguide solution given by equations (5.27), (5.21) and (5.22), in order to rewrite the expression for the power flow in the z-direction. We get

$$\begin{aligned}\hat{\boldsymbol{z}}\cdot <\boldsymbol{S}(t)> (\boldsymbol{r},\omega) = \sum_{\substack{n,n'\\ \nu,\nu'=\mathrm{TM,TE,TEM}}} \frac{1}{2}\hat{\boldsymbol{z}}\cdot \mathrm{Re}\Bigg\{ &\left(a_{n\nu}^{+}\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega)e^{\mathrm{i}k_{zn}z}+a_{n\nu}^{-}\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega)e^{-\mathrm{i}k_{zn}z}\right)\\ &\times\left(a_{n'\nu'}^{+*}\boldsymbol{H}_{Tn'\nu'}^{*}(\boldsymbol{\rho},\omega)e^{-\mathrm{i}k_{zn'}^{*}z}-a_{n'\nu'}^{-*}\boldsymbol{H}_{Tn'\nu'}^{*}(\boldsymbol{\rho},\omega)e^{\mathrm{i}k_{zn'}^{*}z}\right)\Bigg\}\end{aligned}$$

The time average of the total power flow at z is

$$\iint_{\Omega}\hat{\boldsymbol{z}}\cdot <\boldsymbol{S}(t)> (\boldsymbol{r},\omega)\,\mathrm{d}x\mathrm{d}y$$

In this surface integral over the cross section the integral in (5.37) appears. The integration over the cross section Ω of the Poynting vector implies that all terms with products of v_n and w_n vanish, and the power flow in the waveguide is

$$\begin{aligned}\iint_{\Omega}\hat{\boldsymbol{z}}\cdot <\boldsymbol{S}(t)> (\boldsymbol{r},\omega)\,\mathrm{d}x\mathrm{d}y &= \mathrm{Re}\sum_{\substack{n\\ \nu=\mathrm{TM,TE,TEM}}} P_{n\nu}^{E}\Big\{\left|a_{n\nu}^{+}\right|^2 e^{\mathrm{i}(k_{zn}-k_{zn}^{*})z}-\left|a_{n\nu}^{-}\right|^2 e^{-\mathrm{i}(k_{zn}-k_{zn}^{*})z}\\ &\qquad + a_{n\nu}^{-}a_{n\nu}^{+*}e^{-\mathrm{i}(k_{zn}+k_{zn}^{*})z}-a_{n\nu}^{+}a_{n\nu}^{-*}e^{\mathrm{i}(k_{zn}+k_{zn}^{*})z}\Big\}\\ &= \mathrm{Re}\sum_{\substack{n\\ \nu=\mathrm{TM,TE,TEM}}} P_{n\nu}^{E}\left\{\left|a_{n\nu}^{+}\right|^2 e^{-2\,\mathrm{Im}\,k_{zn}z}-\left|a_{n\nu}^{-}\right|^2 e^{2\,\mathrm{Im}\,k_{zn}z}+2\mathrm{i}\,\mathrm{Im}\left(a_{n\nu}^{-}a_{n\nu}^{+*}e^{-2\mathrm{i}\,\mathrm{Re}\,k_{zn}z}\right)\right\}\end{aligned}$$

where the mode power, $P_{n\nu}^{E}$, is given by (5.38), which is a general expression. In addition to the terms that express power flow in the $+z$- and $-z$-directions, $|a_{n\nu}^{+}|^2$ and $|a_{n\nu}^{-}|^2$, respectively, there are terms that express the interaction of waves propagating in opposite directions (standing wave phenomena).

We now specialize to the lossless case, i.e., with $\epsilon(\omega)$ and $\mu(\omega)$ real, and wave propagation either in the positive, or the negative, $z-$direction. The expression simplifies to

$$\iint_{\Omega}\hat{\boldsymbol{z}}\cdot <\boldsymbol{S}(t)> (\boldsymbol{r},\omega)\,\mathrm{d}x\mathrm{d}y = \mathrm{Re}\sum_{\substack{n\\ \nu=\mathrm{TM,TE}}} P_{n\nu}^{E}\left|a_{n\nu}^{+}\right|^2 e^{-2\,\mathrm{Im}\,k_{zn}z} \tag{5.41}$$

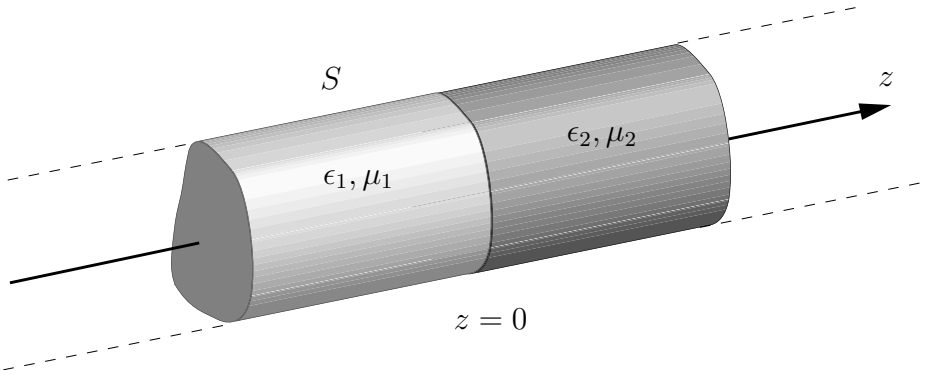


Figure 5.13: Reflection and transmission in a waveguide filled with two isotropic materials with material parameters ϵ_1, μ_1 and ϵ_2, μ_2.

for the waves propagating in the $+z$-direction and

$$\iint_{\Omega} \hat{\boldsymbol{z}} \cdot <\boldsymbol{S}(t)> (\boldsymbol{r}, \omega) \, \mathrm{d}x\mathrm{d}y = -\operatorname{Re} \sum_{\substack{n \\ \nu=\mathrm{TM,TE}}} P^E_{n\nu} \left| a^-_{n\nu} \right|^2 e^{2 \operatorname{Im} k_{zn} z}$$

for the waves propagating in the $-z$-direction. These sums are finite since k_{zn}, and hence $P^E_{n\nu}$, is purely imaginary when $f < f_{cn}$, or equivalently $k_{tn} > k(\omega)$.

Example 5.7
In a microwave oven there is a door with a metallic screen that is perforated with circular holes. The holes are filled with air. The screen is a high pass filter since it lets visible light pass but blocks microwaves efficiently. If we assume that the holes in the screen have a radius of 0.5 mm, the lowest cut-off frequency is $f_{c\mathrm{TE}_{11}} = 1.841 c_0 / 2\pi a = 1.76 \cdot 10^{11}$ Hz, see table 5.4. The frequencies for visible light are far above this cut-off frequency and thus light can pass through the holes. The frequency used by microwave ovens is $f = 2.45$ GHz, i.e., well below the cut-off frequency for the fundamental mode. The attenuation of the power at this frequency is $\alpha_{\mathrm{TE}_{11}} = \frac{4\pi}{c_0} \sqrt{f_{cn}^2 - f^2} = 7363 \text{ m}^{-1}$. With a thickness of 0.5 mm this corresponds to an attenuation of 16 dB of the microwaves. ∎

Example 5.8
Sources in the region $z < 0$ generates a waveguide mode. The region $z < 0$ is assumed to be filled with a lossless material with parameters ϵ_1 and μ_1, while the region $z > 0$ is filled with a lossless material with ϵ_2 and μ_2, see figure 5.13. Determine the reflection and transmission coefficients for the waves.

We denote the electric and magnetic fields generated by the sources $\boldsymbol{E}^+_{n\nu}(\boldsymbol{r}, \omega)$ and $\boldsymbol{H}^+_{n\nu}(\boldsymbol{r}, \omega)$. Since the material parameters are discontinuous in the plane $z = 0$ there is a reflected wave in the region $z < 0$. The electric and magnetic fields in the region $z < 0$ is a sum of the incident and reflected fields, see (5.27) (normalization is $a^+_{n\nu} = 1$)

$$\begin{cases} \boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{E}^+_{n\nu}(\boldsymbol{r}, \omega) + \sum\limits_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} r_{n'\nu'} \boldsymbol{E}^-_{n'\nu'}(\boldsymbol{r}, \omega) \\ \boldsymbol{H}(\boldsymbol{r}, \omega) = \boldsymbol{H}^+_{n\nu}(\boldsymbol{r}, \omega) + \sum\limits_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} r_{n'\nu'} \boldsymbol{H}^-_{n'\nu'}(\boldsymbol{r}, \omega) \end{cases} \quad z \leq 0$$

where ϵ and μ have values ϵ_1 and μ_1.

A general ansatz for the region $z > 0$ is

$$\begin{cases} \boldsymbol{E}(\boldsymbol{r},\omega) = \sum\limits_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} t_{n'\nu'} \boldsymbol{E}^{+}_{n'\nu'}(\boldsymbol{r},\omega) \\ \boldsymbol{H}(\boldsymbol{r},\omega) = \sum\limits_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} t_{n'\nu'} \boldsymbol{H}^{+}_{n'\nu'}(\boldsymbol{r},\omega) \end{cases} \qquad z \geq 0$$

where ϵ and μ have values ϵ_2 and μ_2. Due to the boundary conditions (continuity of the transverse components of $\boldsymbol{E}$ and $\boldsymbol{H}$) and the orthogonality relations (5.36) each mode, $n\nu$, generated from the source couples only to the same mode at the interface $z = 0$. The fields in the region $z < 0$ then reduce to

$$\begin{cases} \boldsymbol{E}(\boldsymbol{r},\omega) = \boldsymbol{E}^{+}_{n\nu}(\boldsymbol{r},\omega) + r_{n\nu}\boldsymbol{E}^{-}_{n\nu}(\boldsymbol{r},\omega) \\ \boldsymbol{H}(\boldsymbol{r},\omega) = \boldsymbol{H}^{+}_{n\nu}(\boldsymbol{r},\omega) + r_{n\nu}\boldsymbol{H}^{-}_{n\nu}(\boldsymbol{r},\omega) \end{cases} \qquad z \leq 0$$

and the fields in the region $z > 0$ reduce to

$$\begin{cases} \boldsymbol{E}(\boldsymbol{r},\omega) = t_{n\nu}\boldsymbol{E}^{+}_{n\nu}(\boldsymbol{r},\omega) \\ \boldsymbol{H}(\boldsymbol{r},\omega) = t_{n\nu}\boldsymbol{H}^{+}_{n\nu}(\boldsymbol{r},\omega) \end{cases} \qquad z \geq 0$$

The boundary condition on the interface $z = 0$ give (see equations (5.24), (5.21) and (5.22))

$$\begin{cases} \left(k_1^2 - k_{tn}^2\right)^{1/2}(1 + r_{n\nu}) = t_{n\nu}\left(k_2^2 - k_{tn}^2\right)^{1/2} \\ (1 - r_{n\nu})\,\epsilon_1 = t_{n\nu}\epsilon_2 \end{cases} \qquad \nu = \mathrm{TM}$$

and

$$\begin{cases} (1 + r_{n\nu})\,\mu_1 = t_{n\nu}\mu_2 \\ \left(k_1^2 - k_{tn}^2\right)^{1/2}(1 - r_{n\nu}) = t_{n\nu}\left(k_2^2 - k_{tn}^2\right)^{1/2} \end{cases} \qquad \nu = \mathrm{TE}$$

where $k_1 = \frac{\omega}{c_0}\sqrt{\epsilon_1\mu_1}$ and $k_2 = \frac{\omega}{c_0}\sqrt{\epsilon_2\mu_2}$ are the wave numbers in region 1 and 2, respectively. Note that the transverse wavenumber k_{tn}, and the basis functions $v_n(\boldsymbol{\rho})$ and $w_n(\boldsymbol{\rho})$, are identical in the two regions $z < 0$ and $z > 0$. The solutions to these equations are

$$\begin{cases} r_{n\nu} = \dfrac{\epsilon_1\left(k_2^2 - k_{tn}^2\right)^{1/2} - \epsilon_2\left(k_1^2 - k_{tn}^2\right)^{1/2}}{\epsilon_1\left(k_2^2 - k_{tn}^2\right)^{1/2} + \epsilon_2\left(k_1^2 - k_{tn}^2\right)^{1/2}} \\ t_{n\nu} = \dfrac{2\epsilon_1\left(k_1^2 - k_{tn}^2\right)^{1/2}}{\epsilon_1\left(k_2^2 - k_{tn}^2\right)^{1/2} + \epsilon_2\left(k_1^2 - k_{tn}^2\right)^{1/2}} \end{cases} \qquad \nu = \mathrm{TM}$$

and

$$\begin{cases} r_{n\nu} = \dfrac{\mu_2\left(k_1^2 - k_{tn}^2\right)^{1/2} - \mu_1\left(k_2^2 - k_{tn}^2\right)^{1/2}}{\mu_2\left(k_1^2 - k_{tn}^2\right)^{1/2} + \mu_1\left(k_2^2 - k_{tn}^2\right)^{1/2}} \\ t_{n\nu} = \dfrac{2\mu_1\left(k_1^2 - k_{tn}^2\right)^{1/2}}{\mu_2\left(k_1^2 - k_{tn}^2\right)^{1/2} + \mu_1\left(k_2^2 - k_{tn}^2\right)^{1/2}} \end{cases} \qquad \nu = \mathrm{TE}$$

■

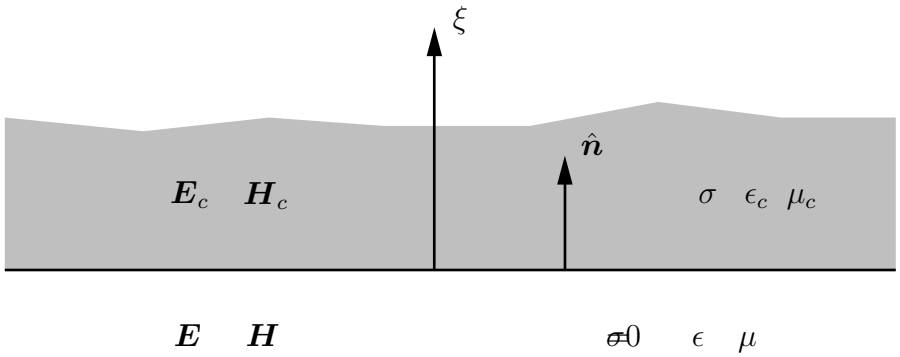


Figure 5.14: The interface between the non-perfectly conducting wall and the inner region of the waveguide.

5.9 Losses in walls

So far the walls of the waveguides have been perfect conductors. This means that a waveguide mode with cut-off frequency below the frequency of the field is not attenuated if the the material in the waveguide is lossless. In this section a more realistic model is presented where the conductivity of the walls is assumed to be very large, but finite. By use of appropriate approximations the theory for waveguides with perfectly conducting walls can be generalized and expressions for the attenuation of propagating modes can be derived. The approximations are valid, as will be seen, for frequencies that are not close to the cut-off frequency of the mode.

We start by analyzing reflection of plane waves impinging on a plane metallic surface. The metal has permittivity ϵ_c, permeability μ_c and conductivity σ. The condition for a good conductor

$$\sigma \gg \omega\epsilon_0\epsilon_c$$

is assumed to be fullfilled. The fields in the metal are denoted $\boldsymbol{E}_c$, $\boldsymbol{D}_c$, $\boldsymbol{B}_c$ and $\boldsymbol{H}_c$ while the fields outside the metal do not have any index. We let $\hat{\boldsymbol{n}}_c$ be the normal to the surface, directed into the metal, and introduce the coordinate ξ along the $\hat{\boldsymbol{n}}_c$-direction, see figure 5.14. The boundary conditions at the surface are given by (see (1.16) on page 7)

$$\hat{\boldsymbol{n}}_c \times \boldsymbol{H} = \hat{\boldsymbol{n}}_c \times \boldsymbol{H}_c$$
$$\hat{\boldsymbol{n}}_c \times \boldsymbol{E} = \hat{\boldsymbol{n}}_c \times \boldsymbol{E}_c$$

i.e., the tangential components of $\boldsymbol{E}$ and $\boldsymbol{H}$ are continuous. There are no surface currents since σ is finite.

The wavenumber $k(\omega)$ for the metal is

$$k(\omega) = \frac{\omega\sqrt{\mu_c}}{c_0}\left(\epsilon_c + i\frac{\sigma}{\omega\epsilon_0}\right)^{1/2}$$

The material parameters ϵ_c and μ_c are assumed to be real. We extract a factor ϵ_c and use the approximation $\sigma \gg \omega\epsilon_0\epsilon_c$. We get

$$k(\omega) = \frac{\omega\sqrt{\epsilon_c\mu_c}}{c_0}\left(1 + i\frac{\sigma}{\omega\epsilon_c\epsilon_0}\right)^{1/2} \approx \frac{\omega\sqrt{\epsilon_c\mu_c}}{c_0}\left(i\frac{\sigma}{\omega\epsilon_c\epsilon_0}\right)^{1/2} = \frac{1+i}{\sqrt{2}}\sqrt{\sigma\mu_0\mu_c\omega}$$

Material	σ (S/m)	$f = 50$ (Hz)	$f = 1$ (MHz)	$f = 1$ (GHz)
Silver	$6.30 \cdot 10^7$	8.97 (mm)	0.063 (mm)	0.0020 (mm)
Copper	$5.96 \cdot 10^7$	9.22	0.065	0.0021
Gold	$4.52 \cdot 10^7$	10.6	0.075	0.0024
Aluminium	$3.78 \cdot 10^7$	11.6	0.082	0.0026
Iron ($\mu = 10^3$)	$1.04 \cdot 10^7$	0.70	0.005	0.00016
Fresh water	0.001	2250 (m)	†[a]	†[a]
Saltwater	4	35.6	0.25 (m)	†[a]

[a] At this frequency the approximation $\sigma \gg \omega\epsilon_0\epsilon$ is not valid and $\epsilon \approx 80$.

Table 5.5: Table of the skin depth δ in different materials at different frequencies. The conductivities of the metals are at temperature 20° C. The values for fresh and saltwater are approximative.

since $(i)^{1/2} = (1+i)/\sqrt{2}$. A plane wave with dependence $\exp(\mathrm{i}k\xi)$ is decomposed as

$$e^{\mathrm{i}k\xi} = e^{\mathrm{i}\xi/\delta} e^{-\xi/\delta}$$

where

$$\delta = \sqrt{\frac{2}{\omega\mu_0\mu_c\sigma}} \tag{5.42}$$

The quantity δ is the skin depth of the material and is the characteristic depth where the electric field has been attenuated a factor e^{-1}.

At microwave frequencies the skin depth is much smaller than the dimensions of the waveguide, see table 5.5. That implies that in the walls the ξ-derivatives of the fields are much greater than the derivatives in the tangential directions and for this reason we neglect the tangential derivatives of $\boldsymbol{E}_c$ and $\boldsymbol{H}_c$ as

$$\nabla \simeq \hat{\boldsymbol{n}}_c \frac{\partial}{\partial\xi} \tag{5.43}$$

Since the normal component of $\boldsymbol{H}$ at a perfectly conducting surface is zero, we can assume that the normal component of $\boldsymbol{H}_c$ is negligible compared to the tangential components. Thus the magnetic field in the metal has a component in a direction $\hat{\boldsymbol{\tau}}$ tangential to the boundary curve Γ and a component along the $\hat{\boldsymbol{z}}$-direction, but no component along the $\xi-$direction. We will soon see that this is in accordance with the approximation in (5.43).

When we utilize the approximation (5.43), the Maxwell field equations are simplified to

$$\begin{cases} \boldsymbol{H}_c = -\dfrac{\mathrm{i}}{\omega\mu_0\mu_c}\hat{\boldsymbol{n}}_c \times \dfrac{\partial \boldsymbol{E}_c}{\partial\xi} \\ \boldsymbol{E}_c = \dfrac{1}{\sigma}\hat{\boldsymbol{n}}_c \times \dfrac{\partial \boldsymbol{H}_c}{\partial\xi} \end{cases} \tag{5.44}$$

Note that the displacement current $-i\omega\epsilon_0\epsilon_c\boldsymbol{E}_c$ is negligible compared to the current density $\sigma\boldsymbol{E}_c$ in Ampère's law, since $\sigma \gg \omega\epsilon_0\epsilon_c$. We eliminate the electric field from the equations

by operating with $\hat{\boldsymbol{n}}_c \times \frac{\partial}{\partial \xi}$ on the lower of the two equations in (5.44)

$$\hat{\boldsymbol{n}}_c \times \frac{\partial \boldsymbol{E}_c}{\partial \xi} = \frac{1}{\sigma}\hat{\boldsymbol{n}}_c \times \left(\hat{\boldsymbol{n}}_c \times \frac{\partial^2 \boldsymbol{H}_c}{\partial \xi^2}\right) = \frac{1}{\sigma}\left(\hat{\boldsymbol{n}}_c \left(\hat{\boldsymbol{n}}_c \cdot \frac{\partial^2 \boldsymbol{H}_c}{\partial \xi^2}\right) - \frac{\partial^2 \boldsymbol{H}_c}{\partial \xi^2}\right)$$

When we insert this into Faraday's law we get

$$\boldsymbol{H}_c = -\frac{\mathrm{i}}{\omega\sigma\mu_0\mu_c}\left(\hat{\boldsymbol{n}}_c \left(\hat{\boldsymbol{n}}_c \cdot \frac{\partial^2 \boldsymbol{H}_c}{\partial \xi^2}\right) - \frac{\partial^2 \boldsymbol{H}_c}{\partial \xi^2}\right)$$

The normal component of this equation gives $\hat{\boldsymbol{n}}_c \cdot \boldsymbol{H}_c = 0$, which we anticipated earlier. Since $\hat{\boldsymbol{n}}_c \cdot \frac{\partial^2 \boldsymbol{H}_c}{\partial \xi^2} = \frac{\partial^2}{\partial \xi^2}(\hat{\boldsymbol{n}}_c \cdot \boldsymbol{H}_c) = 0$ we obtain the equation

$$\frac{\partial^2 \boldsymbol{H}_c}{\partial \xi^2} + \mathrm{i}\omega\mu_0\mu_c\sigma\boldsymbol{H}_c = 0$$

with solution

$$\boldsymbol{H}_c = \boldsymbol{H}_\parallel e^{-\xi/\delta}e^{\mathrm{i}\xi/\delta} \tag{5.45}$$

where δ is the skin depth of the metal, c.f., equation (5.42), and $\boldsymbol{H}_\parallel$ is the tangential component of the magnetic field at the surface, which can be decomposed in components along the directions $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{z}}$. Based on the approximations in this section we conclude that the amplitude of the tangential components are the same as for a perfectly conducting surface. The corresponding electric field is obtained by inserting (5.45) into Ampère's law in equation (5.44).

$$\boldsymbol{E}_c \simeq \frac{\mathrm{i}-1}{\sigma\delta}\left(\hat{\boldsymbol{n}}_c \times \boldsymbol{H}_\parallel\right) e^{-\xi/\delta}e^{\mathrm{i}\xi/\delta}$$

The tangential components of the electric and magnetic fields are continuous and if we let $\xi = 0$ we get the impedance boundary condition that relates the tangential components of the magnetic field to the tangential components of the magnetic field at the surface of a conducting surface, c.f., equation (1.18) in section 1.1

$$\boldsymbol{E} - \hat{\boldsymbol{n}}(\boldsymbol{E} \cdot \hat{\boldsymbol{n}}) = -\eta_s \hat{\boldsymbol{n}} \times \boldsymbol{H} \tag{5.46}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electric and magnetic fields at the surface, $\eta_s = \frac{1-\mathrm{i}}{\sigma\delta}$ is the impedance of the metal and $\hat{\boldsymbol{n}} = -\hat{\boldsymbol{n}}_c$ is the normal unit vector, directed out from the metal.

We are now in position to determine the ohmic losses in the metal. The time average of the power loss density in the metal is given by

$$p = \frac{1}{2}\,\mathrm{Re}\left\{\boldsymbol{J} \cdot \boldsymbol{E}_c^*\right\} = \frac{\sigma}{2}\boldsymbol{E}_c \cdot \boldsymbol{E}_c^* = \frac{1}{\sigma\delta^2}|\boldsymbol{H}_\parallel|^2 e^{-2\xi/\delta}$$

We get the power loss per unit area by integrating this expression in the ξ-direction.

$$\frac{dP_c}{da} = \int_0^\infty p(\xi)\,\mathrm{d}\xi = \frac{1}{2\sigma\delta}|\boldsymbol{H}_\parallel|^2 = \frac{1}{2}R_s|\boldsymbol{H}_\parallel|^2 \tag{5.47}$$

where we introduced

$$R_s = \frac{1}{\sigma\delta} = \text{surface resistance} \tag{5.48}$$

c.f., (3.31) in chapter 3. We have chosen to integrate all the way to infinity for practical reasons. We know that the metallic wall has finite thickness but since the skin depth is very

small and the integrand falls off exponentially only the very first part of the integration interval contributes to the integrand. Since $\boldsymbol{H}_{\parallel}$ in our approximation is assumed to be the same as for a perfectly conducting surface we can relate $\boldsymbol{H}_{\parallel}$ to the surface current density of a perfectly conducting surface

$$\boldsymbol{H}_{\parallel} = -\hat{\boldsymbol{n}}_c \times \boldsymbol{J}_S$$

and hence

$$\frac{dP_c}{da} = \frac{1}{2} R_s |\boldsymbol{J}_S|^2 \tag{5.49}$$

We now turn to the objective of this section, namely the attenuation of waves in a waveguide. The walls of the waveguide are very good, but not perfect, conductors, and for simplicity we assume that the waveguide is filled with a lossless material with real material parameters ϵ and μ. In a first order approximation we obtain an attenuation of a mode due to the losses in the waveguide walls. The frequency of the wave exceeds the cut-off frequency. We study one mode at a time and only propagation in the positive z-direction. The E_z and H_z components of the mode are given by

$$\begin{cases} E_z(\boldsymbol{r}) = a_n^+ v_n(\boldsymbol{\rho}) e^{\mathrm{i}k_{z_n} z} & \text{TM-mod} \\ H_z(\boldsymbol{r}) = b_n^+ w_n(\boldsymbol{\rho}) e^{\mathrm{i}k_{z_n} z} & \text{TE-mod} \end{cases}$$

where the eigenfunctions v_n and w_n are normalized according to (5.26). The corresponding tangential components of the magnetic fields are given by equations (5.13)

$$|\boldsymbol{H}_{\parallel_T}|^2 = \begin{cases} |a_n^+|^2 \left(\dfrac{\omega \epsilon_0 \epsilon}{k_{t_n}^2} \right)^2 |\hat{\boldsymbol{z}} \times \nabla_T v_n|^2, & \text{TM-mod} \\ |b_n^+|^2 \dfrac{k_{z_n}^2}{k_{t_n}^4} |\nabla_T w_n|^2, & \text{TE-mode} \end{cases}$$

Thus the total magnetic field on the metal surface is given by

$$|\boldsymbol{H}_{\parallel}|^2 = |\boldsymbol{H}_{\parallel_T}|^2 + |H_z|^2 = \begin{cases} |a_n^+|^2 \dfrac{(\omega \epsilon_0 \epsilon)^2}{k_{t_n}^4} |\hat{\boldsymbol{z}} \times \nabla_T v_n|^2, & \text{TM-mod} \\ \\ |b_n^+|^2 \left(\dfrac{k_{z_n}^2}{k_{t_n}^4} |\nabla_T w_n|^2 + |w_n|^2 \right), & \text{TE-mode} \end{cases} \tag{5.50}$$

We have seen that, see (5.41), the time average of the power flow, P, in a lossless waveguide with perfectly conducting walls is

$$P = \iint_{\Omega} \boldsymbol{S} \cdot \hat{\boldsymbol{z}} \, \mathrm{d}x\mathrm{d}y = \begin{cases} |a_n^+|^2 \dfrac{\epsilon_0 \omega k_{z_n}}{2 k_{t_n}^2} \epsilon, & \text{TM-mod} \\ |b_n^+|^2 \dfrac{\mu_0 \omega k_{z_n}}{2 k_{t_n}^2} \mu, & \text{TE-mod} \end{cases} \tag{5.51}$$

In the case of perfectly conducting walls the power flow is independent of the $z-$coordinate and the mode is not attenuated. When the walls are not perfectly conducting the mode dissipates power when it propagates due to the ohmic losses in the walls. Hence the power flow P is an exponentially decreasing function of z. On a distance dz of the waveguide the mode dissipates the power

$$dP = -dz \oint_{\Gamma} \frac{dP_c}{da} \, \mathrm{d}l \tag{5.52}$$

We can express the power loss per unit area, dP_c/da, in the coefficients a_n^+ and b_n^+ by utilizing equations (5.47) and (5.50). We express a_n^+ and b_n^+ in terms of P by using equation (5.51) and get an equation for the power flow

$$\frac{dP}{dz} = -2\alpha P \tag{5.53}$$

with solution

$$P(z) = P(0)e^{-2\alpha z}$$

where

$$\alpha = \begin{cases} \dfrac{R_S \omega \epsilon \epsilon_0}{2k_{t_n}^2 k_{z_n}} \oint_\Gamma |\hat{\boldsymbol{z}} \times \nabla_T v_n|^2 \, \mathrm{d}l, & \text{TM-mod} \\ \\ \dfrac{R_S k_{t_n}^2}{2\omega\mu_0\mu k_{z_n}} \oint_\Gamma \left(\frac{k_{z_n}^2}{k_{t_n}^4} |\nabla_T w_n|^2 + |w_n|^2 \right) \mathrm{d}l, & \text{TE-mod} \end{cases}$$

These expressions give the attenuation in a waveguide with walls that are not perfectly conducting. It should be noted that for frequencies very close to the cut-off frequency the expression has to be modified since α then goes to infinity. We refer to the book by Collin [5] for this analysis.

5.9.1 Losses in waveguides with FEM: method 1

It is quite straightforward to determine the attenuation in a waveguide with FEM. We then use the relations in equations (5.49), (5.52) and (5.53)

$$\alpha = -\frac{1}{2P(z)}\frac{dP(z)}{dz} = \frac{1}{2P(z)} \oint_\Gamma \frac{dP_c(z)}{da} \, \mathrm{d}a = \frac{R_s}{4P(z)} \oint_\Gamma |J_s|^2 \, \mathrm{d}a \tag{5.54}$$

With COMSOL the following steps give α:

- We choose the, **2D>Electromagnetic waves>Eigenfrequency study**.
- We draw the cross section of the waveguide.
- We define the material in the waveguide. We either define the material ourselves, or pick it from the list. Usually it is air, which is in the list.
- In **Study>Eigenfrequency** we define how many modes that are to be determined and the cut-off frequency where COMSOL starts to look for eigenfrequencies.
- We let COMSOL solve the eigenvalue problem. It then shows the electric field in the cross section of the waveguide for the different modes. It also gives the cut-off frequencies f_c for the modes. From the cut-off frequencies we get the corresponding k_t from $k_t = \omega/c_0 = 2\pi f_c/c_0$.
- We now choose the mode that we like to study. We open up a new **Study>Mode analysis**. From the eigenfrequency of the mode and the frequency of the field we can calculate the value of k_z. We choose **Study>Mode analysis>Out-of-plane wavenumber** and enter the frequency and the value for k_z in box for **search for modes around**. One can also use **effective mode index** which is defined by $k_z = n_{\text{eff}} k_0$ where $k_0 = \omega/c_0$ is the wavenumber in vacuum.

- We let COMSOL solve the problem and we check that we get the correct value of k_z.
- We use **Results>Derived values>Line integration** and integrate $|J_s|^2$ by choosing the square of the predefined surface current density **emw.normJs**^{2}. We then calculate $P(z)$ by **Results>Derived values>Surface integration>emw.Poavz** (power flow, time average, z-component). From these two integrals we obtain α_p from (5.54).

5.9.2 Losses in waveguides with FEM: method 2

It is somewhat easier to use the impedance boundary condition to determine the attenuation. We then do the following steps in Comsol:

- We choose the, **2D>Electromagnetic waves>Mode analyzis**.
- We draw the cross section of the waveguide.
- Define the materials. One should be air (or vacuum) and is for the interior of the waveguide. The other one is the material in the walls. In **Geometric entity level:** we mark **Boundary** for the metal material and add the boundaries of the waveguide.
- Right click on Electromagnetic waves and let all boundaries have **Impedance Boundary Condition**.
- In **Study>Mode Analysis** we define how many modes that are to be analyzed and the frequency that we are interested in. We also add the effective mode index where COMSOL starts to look for eigenfrequencies.
- We let COMSOL solve the eigenvalue problem. It then shows the electric field in the cross section of the waveguide for the different modes. It also gives the complex effective mode index n. From this we get α as $\alpha = \omega \mathrm{Im}\{n\}/c$.

Example 5.9
Using impedance boundary conditions we get that a rectangular copper waveguide with dimension $a = 0.3$ m and $b = 0.15$ m has $n = 0.7045 + \mathrm{i}1.74 \cdot 10^{-5}$ for the TE_{10}-mode at 704 MHz. This gives $\alpha = 0.0002565$ Np/m. In ESS the distance between the klystrons and the cavities is approximately 20 meters. It means that the attenuation of the power is $e^{-2\alpha 20}$=0.9898. Approximately one percent of the power is lost in the waveguides. The average power fed to the cavities is 5 MW which means that the loss is on the order of 50 kW. By increasing the size of the waveguide to 0.35×0.175 m^2 the effective mode index is changed to $n = 0.7937 + \mathrm{i}1.208 \cdot 10^{-5}$. Then $e^{-2\alpha 20}$=0.9925 and the average loss is 38 kW. ∎

5.10 Sources in waveguides

In this section we determine the fields that are generated by known current densities in a waveguide. We assume that the waveguide has constant cross-section and perfectly

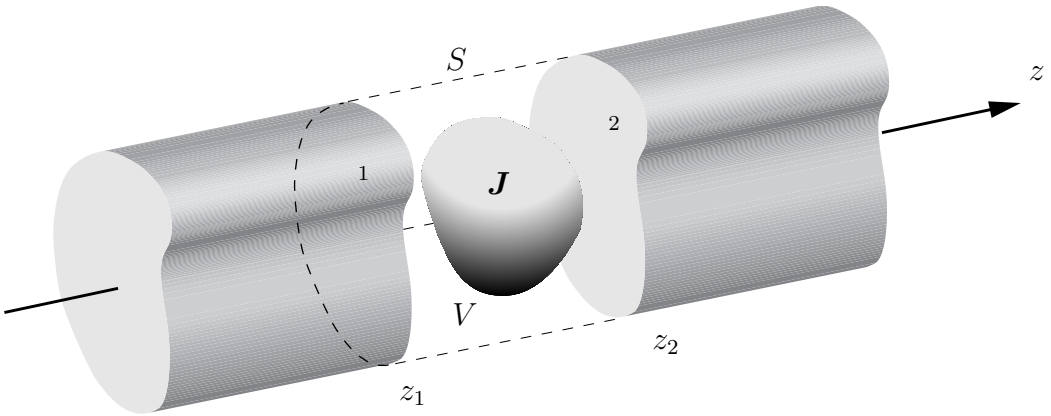


Figure 5.15: Waveguide with a source region V, confined by the surfaces Ω_1, Ω_2 and the lateral surface S.

conducting walls. The boundary to the cross section is assumed to be simply connected and hence there are no TEM-modes. It is easy to generalize to the case with non-perfectly conducting walls by using the theory in the previous section, and also to waveguides with TEM-modes. We assume a finite region in the waveguide confined between the cross-sections Ω_1 and Ω_2, see figure 5.15, where there is a known time-harmonic current density

$$\boldsymbol{J}(\boldsymbol{r},t) = \mathrm{Re}\{\boldsymbol{J}(\boldsymbol{r},\omega)e^{-\mathrm{i}\omega t}\}$$

Since the mode solutions that we developed and analyzed in section 5.4, constitute a complete set of functions in a source free region, the fields generated by the source is a superposition of propagating and non-propagating modes. The suitable vector valued expansion functions are given in (5.27), with the corresponding definitions in equations (5.24), (5.21) and (5.22).

As in the previous sections we use two indices for the modes, an index $\nu = \mathrm{TM,TE}$, and an n-index for the modes for each ν. As before the scalar functions v_n and w_n are the solutions to the eigenvalue problems in equations (5.4) and (5.5), normalized according to (5.26). The normalized vector functions satisfy the orthogonality relations in (5.37).

We let V be the volme between the surfaces Ω_1 and Ω_2 at $z = z_1$ and $z = z_2$, where $z_1 < z_2$, see figure 5.15. In the region $z \geq z_2$ the fields are propagating in the positive z-direction, since the sources of the fields are in the region $z < z_2$. When $z \geq z_2$ the expansion of the fields reads

$$\begin{cases} \boldsymbol{E}^+(\boldsymbol{r},\omega) = \displaystyle\sum_{\nu=\mathrm{TM,TE}}^{n} a^+_{n\nu}\boldsymbol{E}^+_{n\nu}(\boldsymbol{r},\omega) \\ \boldsymbol{H}^+(\boldsymbol{r},\omega) = \displaystyle\sum_{\nu=\mathrm{TM,TE}}^{n} a^+_{n\nu}\boldsymbol{H}^+_{n\nu}(\boldsymbol{r},\omega) \end{cases} \qquad z \geq z_2 \tag{5.55}$$

Analogously the fields are propagating in the negative z-direction when $z \leq z_1$ and then

$$\begin{cases} \boldsymbol{E}^-(\boldsymbol{r},\omega) = \displaystyle\sum_{\nu=\mathrm{TM,TE}}^{n} a^-_{n\nu}\boldsymbol{E}^-_{n\nu}(\boldsymbol{r},\omega) \\ \boldsymbol{H}^-(\boldsymbol{r},\omega) = \displaystyle\sum_{\nu=\mathrm{TM,TE}}^{n} a^-_{n\nu}\boldsymbol{H}^-_{n\nu}(\boldsymbol{r},\omega) \end{cases} \qquad z \leq z_1 \tag{5.56}$$

To determine the coefficients $a_{n\nu}^{\pm}$ we use the Lorentz' reciprocity theorem, which we first derive. Let $\boldsymbol{E}$ and $\boldsymbol{H}$ be the fields generated by the current density $\boldsymbol{J}$, i.e., $\boldsymbol{E}$ and $\boldsymbol{H}$ satisfy

$$\begin{cases} \nabla \times \boldsymbol{E} = \mathrm{i}\omega\mu_0\mu\boldsymbol{H} \\ \nabla \times \boldsymbol{H} = \boldsymbol{J} - \mathrm{i}\omega\epsilon_0\epsilon\boldsymbol{E} \end{cases}$$

where the electric field $\boldsymbol{E}$ satisfies the boundary condition $\hat{\boldsymbol{n}} \times \boldsymbol{E} = \boldsymbol{0}$ on the perfectly conducting surface of the waveguide. By definition the vector fields $\boldsymbol{E}_{n\nu}^{\pm}$, $\boldsymbol{H}_{n\nu}^{\pm}$ for the modes satisfy the homogenous Maxwell equations (no current densities), i.e.,

$$\begin{cases} \nabla \times \boldsymbol{E}_{n\nu}^{\pm} = \mathrm{i}\omega\mu_0\mu\boldsymbol{H}_{n\nu}^{\pm} \\ \nabla \times \boldsymbol{H}_{n\nu}^{\pm} = -\mathrm{i}\omega\epsilon_0\epsilon\boldsymbol{E}_{n\nu}^{\pm} \end{cases}$$

and $\hat{\boldsymbol{n}} \times \boldsymbol{E}_{n\nu}^{\pm} = 0$ on the perfectly conducting surfaces.

The following identity is obtained from the Maxwell field equations and vector formulas for the nabla-operator

$$\begin{aligned} &\nabla \cdot \left(\boldsymbol{E}_{n\nu}^{\pm} \times \boldsymbol{H} - \boldsymbol{E} \times \boldsymbol{H}_{n\nu}^{\pm}\right) = \boldsymbol{H} \cdot \left(\nabla \times \boldsymbol{E}_{n\nu}^{\pm}\right) - \boldsymbol{E}_{n\nu}^{\pm} \cdot (\nabla \times \boldsymbol{H}) \\ &- \boldsymbol{H}_{n\nu}^{\pm} \cdot (\nabla \times \boldsymbol{E}) + \boldsymbol{E} \cdot \left(\nabla \times \boldsymbol{H}_{n\nu}^{\pm}\right) = -\boldsymbol{J} \cdot \boldsymbol{E}_{n\nu}^{\pm} \end{aligned}$$

We integrate this relation over the volume V and utilize the divergence theorem

$$\iint_{S_0} (\boldsymbol{E}_{n\nu}^{\pm} \times \boldsymbol{H} - \boldsymbol{E} \times \boldsymbol{H}_{n\nu}^{\pm}) \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S = -\iiint_{V} \boldsymbol{J} \cdot \boldsymbol{E}_{n\nu}^{\pm} \, \mathrm{d}v \tag{5.57}$$

The surface S_0 consists of the two cross-sections Ω_1 and Ω_2, and the part of the environmental surface S between the surfaces $z = z_1$ and $z = z_2$, see figure 5.15, and $\hat{\boldsymbol{n}}$ is the *outward directed* normal to S_0. Since the environmental surface is the perfectly conducting wall, where $\hat{\boldsymbol{n}} \times \boldsymbol{E}_{n\nu}^{\pm} = \hat{\boldsymbol{n}} \times \boldsymbol{E} = \boldsymbol{0}$, the corresponding surface integral does not contribute. The integrals over Ω_1 and Ω_2 give contributions that can be determined from the expansions in equations (5.55) and (5.56) and the orthogonality relation (5.37).

For the field $\boldsymbol{E}_{n\nu}^{+}$ on the surface Ω_1 we use the expansions in equations (5.56), (5.21) and (5.22). The orthogonality relation (5.39) gives

$$\begin{aligned} &\iint_{\Omega_1} (\boldsymbol{E}_{n\nu}^{+} \times \boldsymbol{H} - \boldsymbol{E} \times \boldsymbol{H}_{n\nu}^{+}) \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S \\ &\quad = -\sum_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} a_{n'\nu'}^{-} \iint_{\Omega_1} \left(\boldsymbol{E}_{n\nu}^{+} \times \boldsymbol{H}_{n'\nu'}^{-} - \boldsymbol{E}_{n'\nu'}^{-} \times \boldsymbol{H}_{n\nu}^{+}\right) \cdot \hat{\boldsymbol{z}} \, \mathrm{d}S \\ &\quad = \sum_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} a_{n'\nu'}^{-} e^{\mathrm{i}(k_{z_n} - k_{z_{n'}})z_1} \iint_{\Omega_1} (\boldsymbol{E}_{Tn\nu} \times \boldsymbol{H}_{Tn'\nu'} + \boldsymbol{E}_{Tn'\nu'} \times \boldsymbol{H}_{Tn\nu}) \cdot \hat{\boldsymbol{z}} \, \mathrm{d}S \\ &\quad = -4a_{n\nu}^{-} U_{n\nu} \end{aligned}$$

Note that on the surface Ω_1, the outward directed normal is $-\hat{\boldsymbol{z}}$. On the surface Ω_2 we

use the expansion (5.55) and get

$$\begin{aligned}
&\iint\limits_{\Omega_2} (\boldsymbol{E}^+_{n\nu} \times \boldsymbol{H} - \boldsymbol{E} \times \boldsymbol{H}^+_{n\nu}) \cdot \hat{\boldsymbol{n}} \,\mathrm{d}S \\
&\quad = \sum_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} a^+_{n'\nu'} \iint\limits_{\Omega_2} \left(\boldsymbol{E}^+_{n\nu} \times \boldsymbol{H}^+_{n'\nu'} - \boldsymbol{E}^+_{n'\nu'} \times \boldsymbol{H}^+_{n\nu}\right) \cdot \hat{\boldsymbol{z}} \,\mathrm{d}S \\
&\quad = \sum_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} a^+_{n'\nu'} e^{\mathrm{i}(k_{zn}+k_{zn'})z_2} \iint\limits_{\Omega_2} (\boldsymbol{E}_{Tn\nu} \times \boldsymbol{H}_{Tn'\nu'} - \boldsymbol{E}_{Tn'\nu'} \times \boldsymbol{H}_{Tn\nu}) \cdot \hat{\boldsymbol{z}} \,\mathrm{d}S = 0
\end{aligned}$$

By using (5.57) we obtain the expression

$$a^-_{n\nu} = \frac{1}{4U_{n\nu}} \iiint\limits_V \boldsymbol{J} \cdot \boldsymbol{E}^+_{n\nu} \,\mathrm{d}v$$

where $U_{n\nu}$ is given by (5.40). In the same manner we get the coefficients for $\boldsymbol{E}^-_{n\nu}$

$$\begin{aligned}
&\iint\limits_{\Omega_1} (\boldsymbol{E}^-_{n\nu} \times \boldsymbol{H} - \boldsymbol{E} \times \boldsymbol{H}^-_{n\nu}) \cdot \hat{\boldsymbol{n}} \,\mathrm{d}S \\
&\quad = -\sum_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} a^-_{n'\nu'} \iint\limits_{\Omega_1} \left(\boldsymbol{E}^-_{n\nu} \times \boldsymbol{H}^-_{n'\nu'} - \boldsymbol{E}^-_{n'\nu'} \times \boldsymbol{H}^-_{n\nu}\right) \cdot \hat{\boldsymbol{z}} \,\mathrm{d}S \\
&\quad = \sum_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} a^-_{n'\nu'} e^{-\mathrm{i}(k_{zn}+k_{zn'})z_1} \iint\limits_{\Omega_1} (\boldsymbol{E}_{Tn\nu} \times \boldsymbol{H}_{Tn'\nu'} - \boldsymbol{E}_{Tn'\nu'} \times \boldsymbol{H}_{Tn\nu}) \cdot \hat{\boldsymbol{z}} \,\mathrm{d}S = 0
\end{aligned}$$

and

$$\begin{aligned}
&\iint\limits_{\Omega_2} (\boldsymbol{E}^-_{n\nu} \times \boldsymbol{H} - \boldsymbol{E} \times \boldsymbol{H}^-_{n\nu}) \cdot \hat{\boldsymbol{n}} \,\mathrm{d}S \\
&\quad = \sum_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} a^+_{n'\nu'} \iint\limits_{\Omega_2} \left(\boldsymbol{E}^-_{n\nu} \times \boldsymbol{H}^+_{n'\nu'} - \boldsymbol{E}^+_{n'\nu'} \times \boldsymbol{H}^-_{n\nu}\right) \cdot \hat{\boldsymbol{z}} \,\mathrm{d}S \\
&\quad = \sum_{\substack{n' \\ \nu'=\mathrm{TM,TE}}} a^+_{n'\nu'} e^{-\mathrm{i}(k_{zn}-k_{zn'})z_2} \iint\limits_{\Omega_2} (\boldsymbol{E}_{Tn\nu} \times \boldsymbol{H}_{Tn'\nu'} + \boldsymbol{E}_{Tn'\nu'} \times \boldsymbol{H}_{Tn\nu}) \cdot \hat{\boldsymbol{z}} \,\mathrm{d}S \\
&\quad = -4a^+_{n\nu} U_{n\nu}
\end{aligned}$$

This leads to

$$a^+_{n\nu} = \frac{1}{4U_{n\nu}} \iiint\limits_V \boldsymbol{J} \cdot \boldsymbol{E}^-_{n\nu} \,\mathrm{d}v$$

The final result is:

$$a^\pm_{n\nu} = \frac{1}{4U_{n\nu}} \iiint\limits_V \boldsymbol{J} \cdot \boldsymbol{E}^\mp_{n\nu} \,\mathrm{d}v \tag{5.58}$$

Example 5.10
Consider a current density $\boldsymbol{J}$ that is confined to a thin wire along the curve C with tangential unit vector $\hat{\boldsymbol{\tau}}$. The volume integrals in (5.58) are reduced to line integrals.

$$a_{n\nu}^{\pm} = \frac{1}{4U_{n\nu}} \int_C I(\boldsymbol{\rho}) \boldsymbol{E}_{n\nu}^{\mp} \cdot \mathrm{d}\boldsymbol{r} \tag{5.59}$$

where $\mathrm{d}\boldsymbol{r} = \hat{\boldsymbol{\tau}}\mathrm{d}l$. ∎

Example 5.11
If the curve C is in a transverse plane $z = z_0$ we see from equations (5.59), (5.21), (5.22) and (5.23) that $a_{n\nu}^{+} = a_{n\nu}^{-} \exp(-2\mathrm{i}k_z z_0)$. ∎

Example 5.12
A small, planar, closed wire with $\boldsymbol{r}$-independent time-harmonic current I has its center in $\boldsymbol{r}_0$. The wire is represented by a magnetic elementary dipole $\boldsymbol{m} = IA\hat{\boldsymbol{n}}$, where A is the area of the planar surface spanned by the wire and $\hat{\boldsymbol{n}}$ is the unit normal to A directed according to the right hand rule. Since the loop is small $\boldsymbol{H}_{n\nu}^{\pm}(\boldsymbol{r}) \simeq \boldsymbol{H}_{n\nu}^{\pm}(\boldsymbol{r}_0)$ in A. From (5.59), Stoke's theorem, and the induction law we get

$$\begin{aligned} a_{n\nu}^{\pm} &= \frac{I}{4U_{n\nu}} \oint_C \boldsymbol{E}_{n\nu}^{\mp} \cdot \mathrm{d}\boldsymbol{r} = \frac{I}{4U_{n\nu}} \iint_\Omega (\nabla \times \boldsymbol{E}_{n\nu}^{\mp}) \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S \\ &\simeq \frac{I}{4U_{n\nu}} \mathrm{i}\omega\mu_0\mu \boldsymbol{H}_{n\nu}^{\mp}(\boldsymbol{r}_0) \cdot \iint_\Omega \hat{\boldsymbol{n}} \, \mathrm{d}S = \frac{\mathrm{i}\omega\mu_0\mu}{4U_{n\nu}} \boldsymbol{H}_{n\nu}^{\mp}(\boldsymbol{r}_0) \cdot \boldsymbol{m} \end{aligned}$$

∎

Example 5.13
Consider an electric elementary dipole, $\boldsymbol{p} = \mathrm{i}I\,\mathrm{d}\boldsymbol{r}/\omega$ where the vector $\mathrm{d}\boldsymbol{r}$ is much smaller than the wavelength and the dimensions of the waveguide. When this dipole is placed in a waveguide we see from (5.59) that

$$a_{n\nu}^{\pm} = -\frac{\mathrm{i}\omega}{4U_{n\nu}} \boldsymbol{p} \cdot \boldsymbol{E}_{n\nu}^{\mp}$$

Assume that the dipole is placed in the point $\boldsymbol{r}_0$. The field from the dipole is given by

$$\boldsymbol{E}(\boldsymbol{r}) = \begin{cases} -\dfrac{\mathrm{i}\omega}{4} \displaystyle\sum_{\substack{n \\ \nu=\mathrm{TM,TE}}} \frac{1}{U_{n\nu}} (\boldsymbol{p} \cdot \boldsymbol{E}_{n\nu}^{-}(\boldsymbol{r}_0)) \boldsymbol{E}_{n\nu}^{+}(\boldsymbol{r}), & z > z_0 \\ -\dfrac{\mathrm{i}\omega}{4} \displaystyle\sum_{\substack{n \\ \nu=\mathrm{TM,TE}}} \frac{1}{U_{n\nu}} (\boldsymbol{p} \cdot \boldsymbol{E}_{n\nu}^{+}(\boldsymbol{r}_0)) \boldsymbol{E}_{n\nu}^{-}(\boldsymbol{r}), & z < z_0 \end{cases}$$

This is rewritten as

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{\mathrm{i}\omega}{4} \boldsymbol{p} \cdot \mathcal{G}(\boldsymbol{r}, \boldsymbol{r}_0)$$

The function $\mathcal{G}$ is given by

$$\mathcal{G}(\boldsymbol{r}, \boldsymbol{r}_0) = \begin{cases} -\displaystyle\sum_{\substack{n \\ \nu=\mathrm{TM,TE}}} \frac{1}{U_{n\nu}} \boldsymbol{E}_{n\nu}^{-}(\boldsymbol{r}_0) \boldsymbol{E}_{n\nu}^{+}(\boldsymbol{r}), & z > z_0 \\ -\displaystyle\sum_{\substack{n \\ \nu=\mathrm{TM,TE}}} \frac{1}{U_{n\nu}} \boldsymbol{E}_{n\nu}^{+}(\boldsymbol{r}_0) \boldsymbol{E}_{n\nu}^{-}(\boldsymbol{r}), & z < z_0 \end{cases}$$

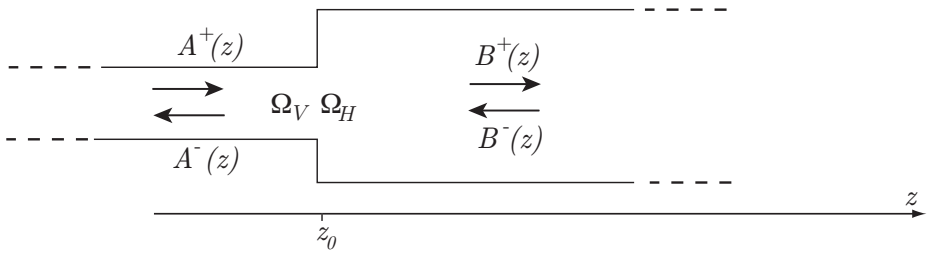


Figure 5.16: The geometry for the mode matching method.

and is called the Green dyadic, which is a vectorial analogue of the Green function for scalar fields. ∎

5.11 Mode matching method

The mode matching method is an efficient method to determine the reflected and transmitted fields at the transition between waveguides with different dimensions.

We first consider the case with two lossless waveguides that are joined at $z = z_0$. The waveguide in $z > z_0$ has a cross section surface Ω_H and the waveguide in $z < z_0$ a cross section surface Ω_V. The surface Ω_V is a subsurface to Ω_H, as indicated in figure 5.16. We assume modes with angular frequency ω propagating in the positive and negative directions in both waveguides. The expansions of the fields are, according to (5.27) on page 85,

$$\begin{cases} \boldsymbol{E}(\boldsymbol{r}) = \sum\limits_{n} \sum\limits_{\nu=\mathrm{TM,TE,TEM}} \left(a^+_{n\nu} \boldsymbol{E}^+_{n\nu}(\boldsymbol{r}) + a^-_{n\nu} \boldsymbol{E}^-_{n\nu}(\boldsymbol{r}) \right) \\ \boldsymbol{H}(\boldsymbol{r}) = \sum\limits_{n} \sum\limits_{\nu=\mathrm{TM,TE,TEM}} \left(a^+_{n\nu} \boldsymbol{H}^+_{n\nu}(\boldsymbol{r}) + a^-_{n\nu} \boldsymbol{H}^-_{n\nu}(\boldsymbol{r}) \right) \end{cases} \quad z < z_0$$

$$\begin{cases} \boldsymbol{E}(\boldsymbol{r}) = \sum\limits_{n} \sum\limits_{\nu=\mathrm{TM,TE,TEM}} \left(b^+_{n\nu} \boldsymbol{E}^+_{n\nu}(\boldsymbol{r}) + b^-_{n\nu} \boldsymbol{E}^-_{n\nu}(\boldsymbol{r}) \right) \\ \boldsymbol{H}(\boldsymbol{r}) = \sum\limits_{n} \sum\limits_{\nu=\mathrm{TM,TE,TEM}} \left(b^+_{n\nu} \boldsymbol{H}^+_{n\nu}(\boldsymbol{r}) + b^-_{n\nu} \boldsymbol{H}^-_{n\nu}(\boldsymbol{r}) \right) \end{cases} \quad z > z_0$$

where the amplitudes of the incident modes, $a^+_{n\nu}$ and $b^-_{n\nu}$, are assumed to be known. We

extract the transverse components of the fields from these sums

$$\begin{cases} \boldsymbol{E}_T^V(\boldsymbol{r}) = \sum\limits_{\substack{n \\ \nu=\mathrm{TM,TE,TEM}}} \left(a_{n\nu}^+ e^{\mathrm{i}k_{z_n}^V z} + a_{n\nu}^- e^{-\mathrm{i}k_{z_n}^V z} \right) \boldsymbol{E}_{T\,n\nu}^V(\boldsymbol{\rho}) \\ \boldsymbol{H}_T^V(\boldsymbol{r}) = \sum\limits_{\substack{n \\ \nu=\mathrm{TM,TE,TEM}}} \left(a_{n\nu}^+ e^{\mathrm{i}k_{z_n}^V z} - a_{n\nu}^- e^{-\mathrm{i}k_{z_n}^V z} \right) \boldsymbol{H}_{T\,n\nu}^V(\boldsymbol{\rho}) \end{cases} \quad z < z_0$$

$$\begin{cases} \boldsymbol{E}_T^H(\boldsymbol{r}) = \sum\limits_{\substack{n \\ \nu=\mathrm{TM,TE,TEM}}} \left(b_{n\nu}^+ e^{\mathrm{i}k_{z_n}^H z} + b_{n\nu}^- e^{-\mathrm{i}k_{z_n}^H z} \right) \boldsymbol{E}_{T\,n\nu}^H(\boldsymbol{\rho}) \\ \boldsymbol{H}_T^H(\boldsymbol{r}) = \sum\limits_{\substack{n \\ \nu=\mathrm{TM,TE,TEM}}} \left(b_{n\nu}^+ e^{\mathrm{i}k_{z_n}^H z} - b_{n\nu}^- e^{-\mathrm{i}k_{z_n}^H z} \right) \boldsymbol{H}_{T\,n\nu}^H(\boldsymbol{\rho}) \end{cases} \quad z > z_0$$

Note that transverse mode functions are different in the two waveguides since the transverse wavenumber k_{t_n} depends on the cross section of the waveguide. We denote the mode functions in the waveguides $\boldsymbol{E}_{T\,n\nu}^V$ and $\boldsymbol{E}_{T\,n\nu}^H$. The same indices are used for the mode functions for the magnetic field and for the longitudinal wavenumbers, i.e., $k_{z_n}^{H,V}$. At $z = z_0$ the transverse components of the electric and magnetic fields are continuous on Ω_V, while the transverse electric field is zero over the remaining part of Ω_H, i.e., in the part $\boldsymbol{\rho} \in \Omega_H$ but $\boldsymbol{\rho} \notin \Omega_V$. This leads to

$$\begin{aligned} \boldsymbol{E}_T^H(\boldsymbol{\rho}, z_0) &= \begin{cases} \boldsymbol{0}, & \boldsymbol{\rho} \in \Omega_H \text{ and } \boldsymbol{\rho} \notin \Omega_V \\ \boldsymbol{E}_T^V(\boldsymbol{\rho}, z_0), & \boldsymbol{\rho} \in \Omega_V \end{cases} \\ \boldsymbol{H}_T^H(\boldsymbol{\rho}, z_0) &= \boldsymbol{H}_T^V(\boldsymbol{\rho}, z_0), \qquad \boldsymbol{\rho} \in \Omega_V \end{aligned} \tag{5.60}$$

We now evaluate the normal surface integral over Ω_H of the vector product between $\boldsymbol{H}_{T\,n'\nu'}^{H*}$ and the upper equation in (5.60). We utilize the normalization integral (5.37) on page 97

$$\iint\limits_{\Omega_H} \hat{\boldsymbol{z}} \cdot \left(\boldsymbol{E}_{T\,n\nu}^H(\boldsymbol{\rho}) \times \boldsymbol{H}_{T\,n'\nu'}^{H*}(\boldsymbol{\rho}) \right) \mathrm{d}S = 2P_{n\nu}^H \delta_{nn'} \delta_{\nu\nu'}$$

where the mode powers in the lossless waveguide are given by, see (5.38),

$$P_{n\nu}^{H,V} = \begin{cases} \dfrac{\epsilon_0 \epsilon \omega k_{z_n}^{H,V}}{2\left(k_{t_n}^{H,V}\right)^2}, & \nu = \mathrm{TM} \\ \dfrac{\mu_0 \mu \omega k_{z_n}^{H,V*}}{2\left(k_{t_n}^{H,V}\right)^2}, & \nu = \mathrm{TE} \\ \dfrac{1}{2\eta_0 \eta}, & \nu = \mathrm{TEM} \end{cases} \tag{5.61}$$

The following relation is obtained

$$P(B^+(z_0) + B^-(z_0)) = Q^t(A^+(z_0) + A^-(z_0)) \tag{5.62}$$

where $A^\pm(z_0)$ and $B^\pm(z_0)$ are column vectors $A_{n\nu}^\pm(z) = a_{n\nu}^\pm e^{\pm \mathrm{i}k_{z_n}^V z}$ and $B_{n\nu}^\pm(z) = b_{n\nu}^\pm e^{\pm \mathrm{i}k_{z_n}^H z}$ where ^{t} denotes transpose. The matrix Q is given by

$$Q_{n\nu,n'\nu'} = \frac{1}{2} \iint\limits_{\Omega_V} \hat{\boldsymbol{z}} \cdot \left(\boldsymbol{E}_{T\,n\nu}^V(\boldsymbol{\rho}) \times \boldsymbol{H}_{T\,n'\nu'}^{H*}(\boldsymbol{\rho}) \right) \mathrm{d}S$$

while P is the diagonal matrix

$$P_{n\nu,n'\nu'} = P^H_{n\nu}\delta_{nn'}\delta_{\nu\nu'}$$

Next we evaluate the normal surface integral over Ω_V of the vector product of $\boldsymbol{E}^{V*}_{T\,n'\nu'}$ and the lower of the equations in (5.60) and get the relation

$$Q^*(B^+(z_0) - B^-(z_0)) = R^*(A^+(z_0) - A^-(z_0)) \tag{5.63}$$

where R is the diagonal matrix

$$R_{n\nu,n'\nu'} = P^V_{n\nu}\delta_{nn'}\delta_{\nu\nu'}$$

The mode power $P^V_{n\nu}$ is given by (5.61).

The system of equations (5.62) and (5.63) gives the relations between the amplitudes of the incident modes, $A^+(z_0)$, $B^-(z_0)$, and the amplitudes of the outgoing modes $A^-(z_0)$, $B^+(z_0)$:

$$\begin{pmatrix} A^-(z_0) \\ B^+(z_0) \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A^+(z_0) \\ B^-(z_0) \end{pmatrix} = \mathbf{S} \begin{pmatrix} A^+(z_0) \\ B^-(z_0) \end{pmatrix}$$

where $\mathbf{S}$ is the scattering matrix with elements given by

$$\begin{cases} S_{11} = (Q^*P^{-1}Q^t + R^*)^{-1}(R^* - Q^*P^{-1}Q^t) \\ S_{12} = 2(Q^*P^{-1}Q^t + R^*)^{-1}Q^* \\ S_{21} = 2(Q^tR^{*-1}Q^* + P)^{-1}Q^t \\ S_{22} = (Q^tR^{*-1}Q^* + P)^{-1}(Q^tR^{*-1}Q^* - P) \end{cases}$$

We have so far considered waveguides for which $\Omega_V < \Omega_H$. The case when $\Omega_V > \Omega_H$ leads to similar expressions for the relations between the amplitudes

$$\begin{pmatrix} A^-(z_0) \\ B^+(z_0) \end{pmatrix} = \begin{pmatrix} \widetilde{S}_{11} & \widetilde{S}_{12} \\ \widetilde{S}_{21} & \widetilde{S}_{22} \end{pmatrix} \begin{pmatrix} A^+(z_0) \\ B^-(z_0) \end{pmatrix}$$

where the scattering matrix is

$$\begin{cases} \widetilde{S}_{11} = (R + \widetilde{Q}^tP^{*-1}\widetilde{Q}^*)^{-1}(\widetilde{Q}^tP^{*-1}\widetilde{Q}^* - R) \\ \widetilde{S}_{12} = 2(R + \widetilde{Q}^tP^{*-1}\widetilde{Q}^*)^{-1}\widetilde{Q}^t \\ \widetilde{S}_{21} = 2(\widetilde{Q}^*R^{-1}\widetilde{Q}^t + P^*)^{-1}\widetilde{Q}^* \\ \widetilde{S}_{22} = (\widetilde{Q}^*R^{-1}\widetilde{Q}^t + P^*)^{-1}(P^* - \widetilde{Q}^*R^{-1}\widetilde{Q}^t) \end{cases}$$

and where

$$\widetilde{Q}_{n\nu,n'\nu'} = \frac{1}{2}\iint\limits_{\Omega_H} \hat{\boldsymbol{z}} \cdot \left(\boldsymbol{E}^H_{T\,n\nu}(\boldsymbol{\rho}) \times \boldsymbol{H}^{V*}_{T\,n'\nu'}(\boldsymbol{\rho})\right) \mathrm{d}S$$

Example 5.14

Consider a transition between the two planar waveguides depicted in figure 5.17. An incident TEM wave propagates in the positive z-direction and gives rise to a reflected and transmittet TEM-wave at the transition $z = 0$. The frequency is low enough such

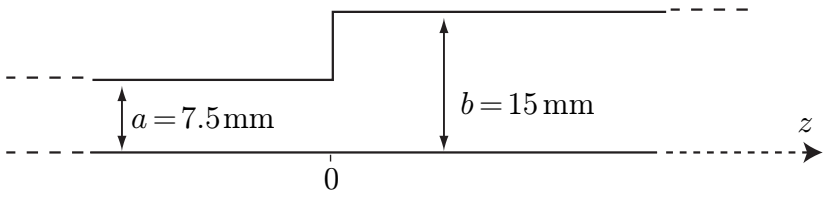


Figure 5.17: Transition between two planar waveguides.

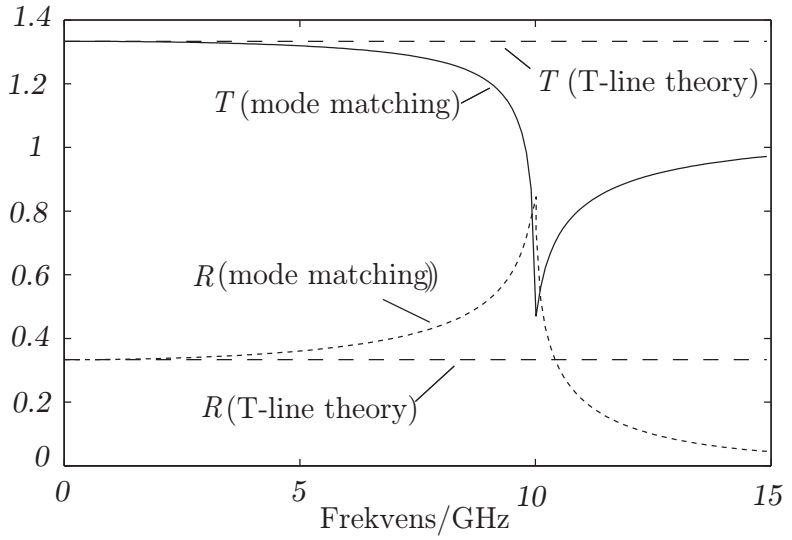


Figure 5.18: Reflection and transmission coefficients for the TEM-mode.

that no other modes than the TEM-mode can propagate. For low enough frequencies the transition can be treated by transmission line theory. This gives the reflection coefficient $\Gamma = (Z_2 - Z_1)/(Z_2 + Z_1)$ and the transmission coefficient $T = 2Z_2/(Z_2 + Z_1)$. For a planar waveguide the characteristic impedance is given by $Z = d\eta/w$, where d is the distance between the plates and w is the width of the two conductors. The reflection and transmission coefficients are then given by

$$\begin{aligned} R &= \frac{b-a}{b+a} \\ T &= \frac{2b}{b+a} \end{aligned} \tag{5.64}$$

These coefficients can be compared with the corresponding coefficients obtained from the mode matching method. In this case $B^-(z_0) = 0$. The scattering matrix element S_{11} acts as a reflection matrix and the matrix element $S_{21}\sqrt{\frac{Z_2}{Z_1}}$ as a transmission matrix, where $\sqrt{\frac{Z_2}{Z_1}} = \sqrt{\frac{b}{a}}$ is the quotient of the characteristic impedances of the transmission lines at port 1 and 2, c.f., (3.39). In figure 5.18 the reflection coefficient for the TEM wave calculated from the mode matching method is compared with the corresponding coefficients calculated from transmission line theory. In this case $a = 7.5$ mm and $b = 15$ mm. We see that transmission line theory is only accurate up to approximately 1 GHz.

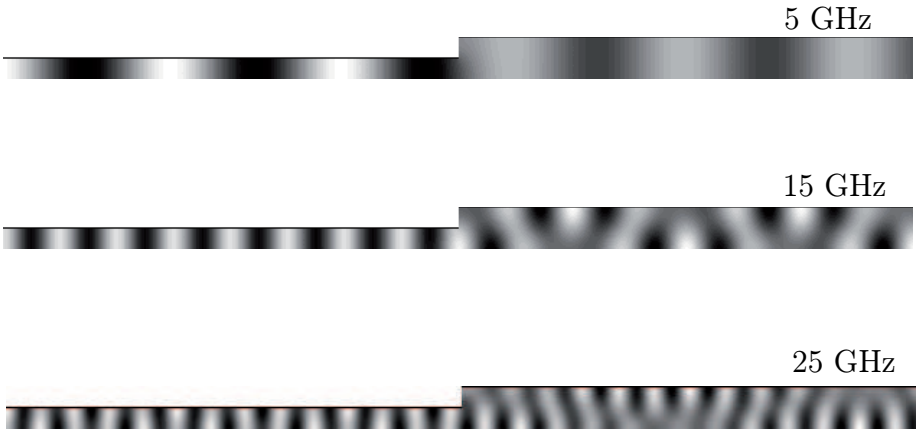


Figure 5.19: The magnetic field distribution for the junction in figure 5.17 generated by COMSOL at 5, 15, and 25 GHz. The incident TEM-wave enters from the left. The total length of the waveguide is 32 cm. At 5 GHz only the TEM wave propagates. At 15 GHz only the TEM wave propagates in the left waveguide and both the TEM and TM_1 modes propagate in the right part. At 25 GHz TEM and TM_1 propagate in the left part and TEM, TM_1, TM_2 propagate in the right part.

This despite that the cut-off frequency for the next propagating modes TE_1 and TM_1 is 10 GHz. ∎

Example 5.15
To analyze the transition in figure 5.17 with COMSOL we choose **2D>Electromagnetic waves>Frequency domain**. We draw the geometry and define the boundary conditions. All of the surfaces are perfect conductors except the vertical surface to the left, which is the input port, and the vertical surface to the right, which is the output port. In boundary conditions we specify the vertical surface to the left to be the input port and the vertical surface to the right to be the output port. Now we can specify the frequency in **Study>Frequency domain** and let COMSOL calculate the field in the waveguide. Figure 5.19 shows the magnetic field at 5 GHz, 15 GHz and 25 GHz. The cut-off frequencies for the TM_1 mode is 10 GHz in the right part of the waveguide and 20 GHz in the left part. This is in accordance with the three figures.
∎

It seems that we can handle junctions with FEM. FEM is more flexible than the mode matching technique since it does not rely on analytical expressions for the waveguide modes. We might then get the impression that the mode matching method is redundant, but this is not all true. There are a number of cases where the mode-matching technique is superior to FEM. If there are long distances between junctions then the mode matching technique is very efficient. The mode matching technique decomposes the waves in their mode sums and that is not as straightforward with FEM.

5.11.1 Cascading

A waveguide with several discontinuities can be treated by a cascading method. Assume the geometry depicted in figure 5.20. The relation between $A^{\pm}(z_0)$ and $B^{\pm}(z_0)$ can be

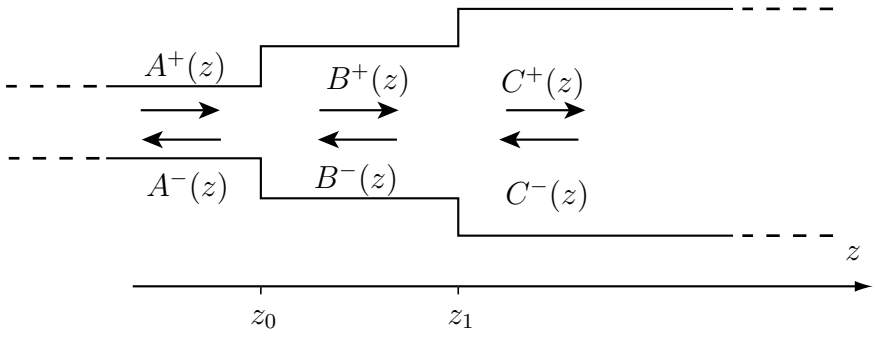


Figure 5.20: Three cascaded waveguides.

written, see equations (5.62) and (5.63)

$$\begin{pmatrix} P_0 & P_0 \\ Q_0^* & -Q_0^* \end{pmatrix} \begin{pmatrix} B^+(z_0) \\ B^-(z_0) \end{pmatrix} = \begin{pmatrix} Q_0^t & Q_0^t \\ R_0^* & -R_0^* \end{pmatrix} \begin{pmatrix} A^+(z_0) \\ A^-(z_0) \end{pmatrix}$$

The relation between $B^{\pm}(z_1)$ and $C^{\pm}(z_1)$ is given by

$$\begin{pmatrix} P_1 & P_1 \\ Q_1^* & -Q_1^* \end{pmatrix} \begin{pmatrix} C^+(z_1) \\ C^-(z_1) \end{pmatrix} = \begin{pmatrix} Q_1^t & Q_1^t \\ R_1^* & -R_1^* \end{pmatrix} \begin{pmatrix} B^+(z_1) \\ B^-(z_1) \end{pmatrix}$$

The relation between $B^{\pm}(z_1)$ and $B^{\pm}(z_0)$ is

$$\begin{pmatrix} B^+(z_1) \\ B^-(z_1) \end{pmatrix} = \begin{pmatrix} E^+(z_1 - z_0) & 0 \\ 0 & E^-(z_1 - z_0) \end{pmatrix} \begin{pmatrix} B^+(z_0) \\ B^-(z_0) \end{pmatrix}$$

where $E^{\pm}_{n\nu,n'\nu'}(z) = \delta_{nn'}\delta_{\nu\nu'} \exp(\pm \mathrm{i} k_{zn} z)$. By matrix multiplication we obtain the relations between $A^{\pm}(z_0)$ and $C^{\pm}(z_1)$. This is straightforward to generalize to a waveguide with an arbitrary number of transitions.

A continuous (tapered) transition from one waveguide to another can be treated by cascading a large number of waveguides with constant cross sections.

5.11.2 Waveguides with bends

With FEM we can analyze waveguides that are not straight. In the example in figure 5.21 a TE_{10} mode enters the left port and exits at the upper port. It is straightforward to draw the geometry in COMSOL. All of the surfaces are perfect conductors except the ports. At the ports we specify that the mode is the first TE-mode, i.e., the TE_{10} mode. One can let COMSOL calculate the reflection and transmission coefficient as a function of frequency.

5.12 Transmission lines in inhomogeneous media by FEM

On page 48 in Chapter 3 we presented one method for the determination of the line parameters for a transmission line where the conductors are in a medium with x and y dependent ϵ and σ. We can now analyze these transmission lines by regarding them

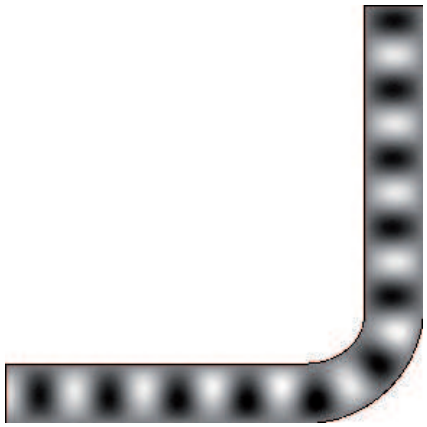


Figure 5.21: The electric field for a TE_{10} mode traveling in a waveguide with a bend. The field is calculated by COMSOL. The frequency is 3 GHz. The cross section of the waveguide in the straight sections is 1cm×0.5 cm. The figure shows the xz-plane and the electric field is perpendicular to the plane.

as waveguides. We first analyze the microstrip on page 51 in Chapter 3 and compare the results with the results obtained from the approximate method. We then consider a microstrip with thickness on the order of the skin depth.

Example 5.16
Consider first the the microstrip in figure 3.17. We do the following steps:

- Choose **2D>Electromagnetic waves>Mode analysis**.
- Draw a rectangle of size 60×50 mm where the bottom is the ground plane. Draw the substrate as a 60×0.5 mm rectangle on top of the ground plane. Draw the microstrip as a 3×0.1 mm rectangle and put it on top of the substrate. We create the region between the outer rectangle and the microstrip as our computational domain.
- In order to calculate the voltage between the microstrip and the ground plane we draw a straight vertical line from the ground plane to the point in the middle of the lower line of the microstrip.
- We define the permittivity and conductivity of the substrate in **+Material** and the material of the microstrip (and if we like alos the ground plane).
- We use impedance boundary condition for the boundaries of the microstrip. The skin depth is $\delta = \sqrt{2/(\omega\sigma\mu_0)} = 6.5$ μm, which is much smaller than the thickness of the strip, and then the impedance boundary condition is a good approximation.
- Let the frequency be 30 GHz in the **Study>Mode analysis**. We add 2 in **Study>Mode analysis>Search for effective mode index at**. We do this since we know that the wave will almost be a TEM-mode in a medium with $\epsilon = 4$ which means that the effekctive mode index k_z/k_0 is close to $\sqrt{\epsilon}$.
- Solve. The solution we look for has the power flow density confined to the region between the microstrip and the ground plane. In this calculation we get the effective

mode index $n_{\text{eff}} = 1.8711 + \text{j}0.3806$ which corresponds to $\gamma = \text{i}n_{\text{eff}}k_0 = 0.7977 + \text{j}3.922$. This is close to the value $\gamma = 0.7789 + \text{j}3.905\,\text{m}^{-1}$ that we got with the more approximate method presented in Chapter 3. The value obtained with the method presented here is accurate.

We can obtain the characteristic impedance Z_0 and the line parameters in two different ways. The first method is to first determine Z_0 by $Z_0 = V/I$ where V is the voltage between the two conductors and I is the current in one of the conductors. The voltage V is obtained by a line integral of the electric field from the ground plane to the microstrip and I is obtained by a line integral of the surface current density in the z-direction around the outer rectangle. The surface current of the outer rectangle gives better accuracy than that of the microstrip since the current density is rapidly varying close to the edges of the microstrip. Once V and I are determined we obtain Z_0 from $Z_0 = V/I$. In this example we get $Z_0 = 21.756 + \text{j}4.604$ V/A. We then notice that, c.f., equations (3.24) and (3.25),

$$\begin{aligned} Z_0\gamma &= R + \text{j}\omega L \\ \frac{\gamma}{Z_0} &= G + \text{j}\omega C \end{aligned}$$

Thus

$$\begin{aligned} R &= \text{Real}\{Z_0\gamma\} \\ L &= \frac{\text{Imag}\{Z_0\gamma\}}{\omega} \\ G &= \text{Real}\{\gamma Z_0^{-1}\} \\ C &= \frac{\text{Imag}\{\gamma Z_0^{-1}\}}{\omega} \end{aligned} \tag{5.65}$$

We get the values $R = 0.006985\ \Omega$/m, $L = 141.6$ nH/m, $G = 0.0716$ S/m and $C = 262.8$ pF/m. These are quite close to the ones we got in Chapter 3.

The other way to obtain the line parameters is slightly more complicated. First the inductance per unit length is obtained by first calculating the time average of the magnetic energy W_m and then the current I in the $z-$direction as the boundary integral of the $z-$component of the surface current density of the outer rectangle. The magnetic energy is a predefined quantity in COMSOL and is straightforward to obtain by integration over the computational domain. Then

$$L = \frac{4W_m}{|I|^2}$$

This gives the value $L = 139.8$ nH/m, which is in good agreement with the value $L = 139.5$ nH/m obtained from the method in Chapter 3. We can obtain the current by integrating over the microstrip, but the rapid variation of the current density on the microstrip deteriorates the accuracy and it is better to integrate the current over the ground plane.

To obtain the capacitance we use the energy relation

$$\frac{1}{4}C|V|^2 = W_E = \text{Time average of electric energy}$$

where the voltage V is obtained by integrating the electric field from the ground plane to the microstrip. Here we utilize the line that is the left side of the extra rectangle. By

Figure 5.22: The power flow density for the microstrip.

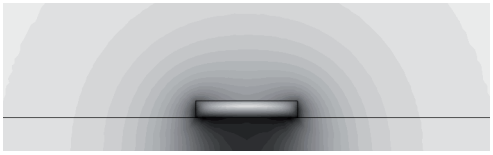


Figure 5.23: The absolute value of the magnetic field for the microstrip in example 17. Notice that the magnetic field penetrates the microstrip.

integrating the $y-$component of the electric field along this side we get the voltage. The electric energy is obtained by integrating the electric energy density over all subdomains. In this case the value is $1.92 \cdot 10^{-8}$ N. The value of the capacitance is $C = 265.9$ pF/m, which is to be compared with 265.5 pF/m with the approximate method.

The conductance we get from

$$\frac{1}{2}|V|^2 G = P_s = \text{time average of dissipated power per unit length in the substrate}$$

The dissipated power density is a predefined quantity. This gives $G = 0.0695$ S/m which is to be compared with 0.0694 S/m with the method in Chapter 3. We see that the simpler method in Chapter 3 gives very small errors in the line parameters. ∎

Example 5.17

Consider the microstrip in figure 5.23 The dimensions of the microstrip is $3\,\mu\text{m} \times 0.5\,\mu\text{m}$ and it consists of a metal with conductivity $4 \cdot 10^7$ S/m. The substrate has relative permittivity $\epsilon = 4$, conductivity $\sigma_s = 0.01$ S/m, and thickness $1\,\mu$m. The ground plane is assumed to have infinite conductivity. The frequency is 30 GHz. We determine the propagation constant γ and the line parameters. We do the following steps:

- Choose **2D>Electromagnetic waves>Mode analysis**.
- Draw a rectangle of size 60×50 μm where the bottom is the ground plane. Add the substrate and the microstrip.
- In order to calculate the voltage between the microstrip and the ground plane we draw a rectangle with dimension 1×0.5 μm (the width is not important) and let the

left side of the rectangle go from the ground plane to the point in the middle of the lower line of the microstrip. Thus the rectangle is a part of the substrate.

- We define the permittivity and conductivity of the substrate in **+Material** and the microstrip and let the frequency be 100 MHz in the **Study>Mode analysis**.
- We add 2 in the **Study>Mode analysis>Search for effective mode index at**. We do this since we know that the wave will almost be a TEM-mode in a medium with $\epsilon = 4$ which means that the effektive mode index k_z/k_0 is close to $\sqrt{\epsilon}$.
- Solve. We look for the mode with the power flow density confined to the region below the microrstrip. The effective mode index for that mode is $n_{\text{eff}} = 1.91 - j0.36$, which corresponds to the propagation constant $\gamma = -226 + \text{i}1202\ \text{m}^{-1}$.
- It is possible to calculate the inductance, capacitance and resistance. We can either use $Z_0 = V/I$ and (5.65) or use

$$L = \frac{4W_m}{|I|^2}$$
$$C = \frac{4W_e}{|V|^2}$$
$$R = \frac{2P_c}{|I|^2}$$
$$G = \frac{2P_s}{|V|^2}$$

where the electric and magnetic energies W_m and W_e are obtained by integration of the corresponding predefined densities over all subdomains. The dissipated power in the microstrip, P_c, and in the substrate, P_s, are obtained by integrating the resistive heating over the microstrip and the substrate, respectively. The current I is obtained by integration of the current density over the microstrip, or, over the groundplane. The voltage V is obtained from the boundary integral along the left side of the extra rectangle. In this example the two methods give the same values of the line parameters

$$L = 232.5\,\text{nH/m}$$
$$C = 159\,\text{pF/m}$$
$$R = 17.1\,\text{k}\Omega\text{/m}$$
$$G = 0.0395\,\text{S/m}$$

The values have been checked by comparing with different meshes and sizes of the computational domain. If we use the expression $\gamma = \sqrt{(R - \text{i}\omega L)(G - \text{i}\omega C)}$ we get $\gamma = -220 + \text{i}1168\ \text{m}^{-1}$, which differs by approximately two percent from the value obtained from the effective mode index. This indicates that the transmission line model with line parameters is a quite good approximation also in this case.

■

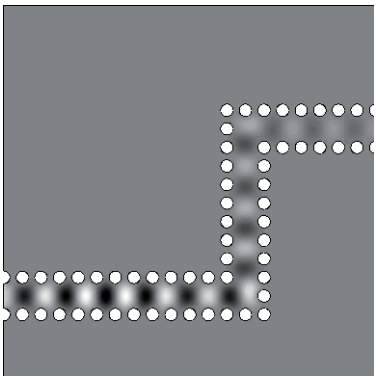


Figure 5.24: The SIW at 50 GHz. The circles are the cross sections of the circular metallic vias connecting two parallel conducting planes that are separated by a dielectric substrate. Only the fundamental mode is propagating.

Figure 5.25: The SIW at 30 GHz. No modes are propagating since the frequency is below the cut-off frequency for the fundamental mode.

5.13 Substrate integrated waveguides

Many applications today use frequencies above 10 GHz and then even a printed circuit might not be discrete. One way to transmit signals at such frequencies is to use substrate integrated waveguides (SIW). These are formed by a pattern of vias between parallel ground planes. Vias are vertical thin circular conductors that serve as connections between layers in a stratified substrate. In figure 5.24 a SIW in a substrate with $\epsilon = 5$ is depicted. The frequency is 50 GHz and the the input port is at the lower left corner and the output port is at the upper right corner. The radius of the circular vias is 500 μm and the distance between the centers of two consecutive vias is 1500 μm. The distance between the centers of two parallel vias is 3000 μm. The waveguide is similar to a planar waveguide with width 2000 μm. Such a waveguide has the cut-off frequency 33 GHz. At frequencies below cut-off frequency for the fundamental mode, as in figure 5.25 , no waves propagate. In figure 5.27 we have made the corners smoother in order to avoid reflections. We see that the transmission from the input port to the output port is much better when the smooth corners are used.

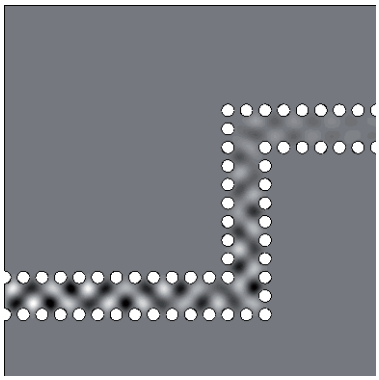


Figure 5.26: The SIW at 70 GHz. At this frequency both the fundamental and the second mode are propagating.

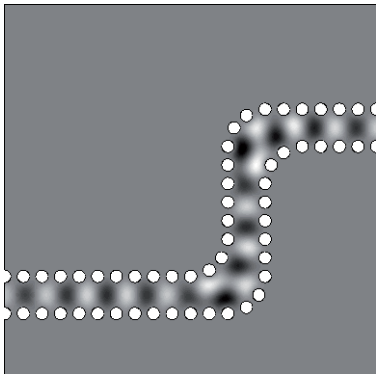


Figure 5.27: The SIW with smooth bends, at 50 GHz. The reflections at the corners are reduced as compared to the SIW in figure 5.24.

Problems in Chapter 5

5.1 Show that the fields for the propagating TE- and TM-modes in a planar waveguide can be written as a superposition of two linearly polarized, plane waves. Determine in what directions the two waves propagate. How do they propagate at a frequency close to the cut-off frequency and at a very high frequency?

5.2 The time average of the power flow in a lossless waveguide is always zero for non-propagating modes. Show that the time average of the power flow in a waveguide filled with a conductive material is larger than zero for all frequencies and all modes. Where does the power go?

*__5.3__ Determine the expressions for the TE- and TM-modes in a coaxial cable, with radius a of the inner conductor and b of the inner surface of the outer conductor, see figure 5.28.

5.4 In a circular waveguide the fundamental mode is the TE_{11}-modes. Its cut-off frequency is given by $f_{c\mathrm{TE}_{11}} = 1.841c_0/2\pi a$, see table 5.4. By inserting a planar strip in the waveguide the frequency of the fundamental mode can be reduced. Determine

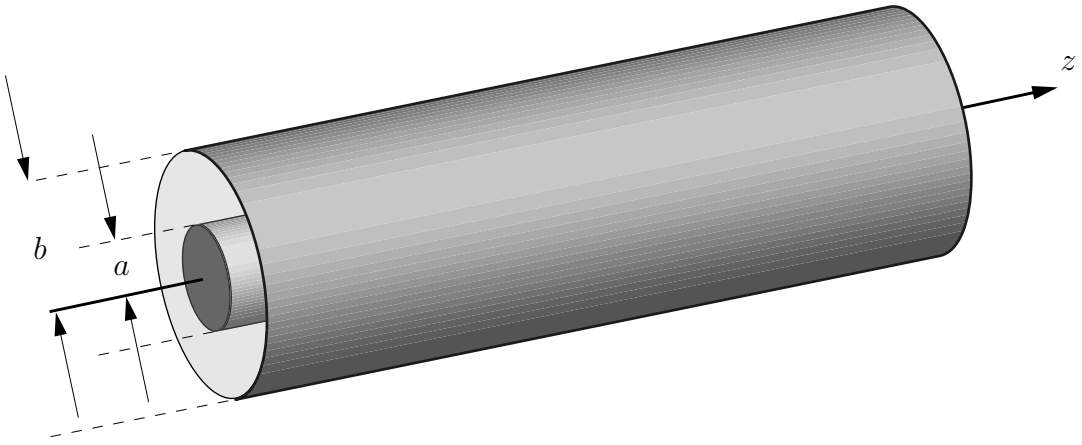


Figure 5.28: The geometry for the coaxial cable in problem 5.3.

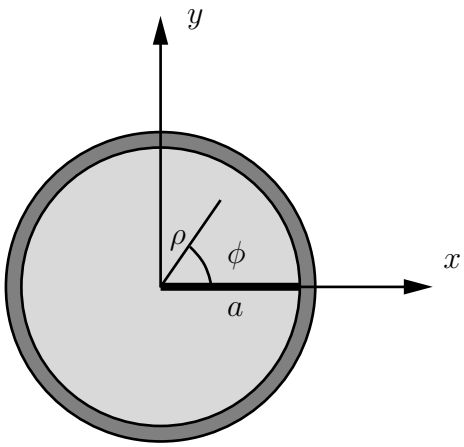


Figure 5.29: Geometry for problem 5.4. The circular waveguide with the metal strip.

the modes for the geometry in figure 5.29, and determine the reduction of the lowest cut-off frequency.

Hint: The Bessel functions with half integer indices can be expressed in trigonometric functions

$$J_{m+1/2}(z) = \sqrt{\frac{2}{\pi}} z^{m+1/2} \left(-\frac{1}{z}\frac{d}{dz} \right)^m \frac{\sin z}{z} = \sqrt{\frac{2z}{\pi}} j_m(z)$$

where $j_m(z)$ is a spherical Bessel function. In particular

$$J_{1/2}(z) = \sqrt{\frac{2}{\pi z}} \sin z, \qquad J_{3/2}(z) = \sqrt{\frac{2z}{\pi}} (\sin z - z \cos z)$$

5.5 A rectangular waveguide with walls $a = 4$ cm and $b = 3$ cm is in the region $z < 0$ filled with air ($\epsilon = 1$) and for $z > 0$ filled with a dielectric material with $\epsilon = 2$. the TM-mode with the lowest cut-off frequency propagates in the positive z-direction for $z < 0$. The frequency is chosen such that it is the same as the cut-off frequency for the second lowest TM-mode in the region. Let P_i and P_r be the time averages

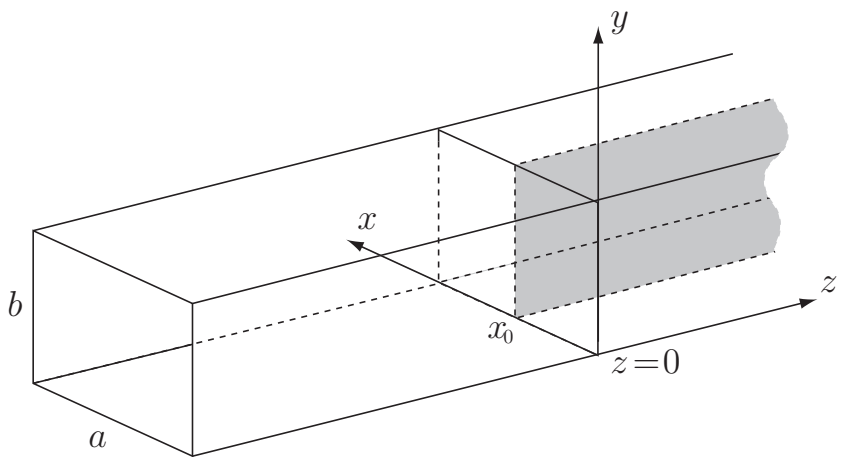


Figure 5.30: Geometry for problem 5.7.

of the total power flow in the waveguide for $z < 0$, for the incident and reflected modes, respectively

a) Determine the quotient P_r/P_i.

b) Is there any frequency for which $P_r = 0$?

5.6 A circular hollow waveguide has the radius $a = 3$ cm.

a) Determine the modes that can propagate at the frequency $f = 5$ GHz when the waveguide is filled with air.

b) Assume that the waveguide is filled with a plastic material with relative permittivity $\epsilon = 3$ and conductivity $\sigma = 10^{-11}$ S/m. Determine the attenuation of the dominant mode in dB/km as a function of frequency.

5.7 Consider a rectangular waveguide with size $0 < x < a$, $0 < y < b$, $a = 6$ cm, $b = 4$ cm. In the region $z > 0$, a metallic plate is inserted, see figure 5.30. The plate i parallel with the y-z-plane and is placed at $x = x_0$. The walls of the waveguide and the plate are perfect conductors. For a certain value of x_0 we measure P_i and P_r, where P_i and P_r are the time averages of the power flow in positive and negative z-direction, respectively, in the region $z < 0$. If only the fundamental mode TE_{10} propagates in positive z-direction for $z < 0$ we have $P_r/P_i = 1$ for all frequencies below 3.75 GHz and $P_r/P_i < 1$ for frequencies above 3.75 GHz.

a) Determine x_0.

b) What is the quotient P_r/P_i when a TE_{03} mode propagates in the positive $z-$direction for $z < 0$ at the frequency 20 GHz?

c) What is P_r/P_i when a TE_{30} mode propagates in positive $z-$direction for $z < 0$ at the frequency 10 GHz?

5.8 A waveguide has a cross section in the shape of a quarter circle with radius R. Determine all TE- and TM-modes for the waveguide.

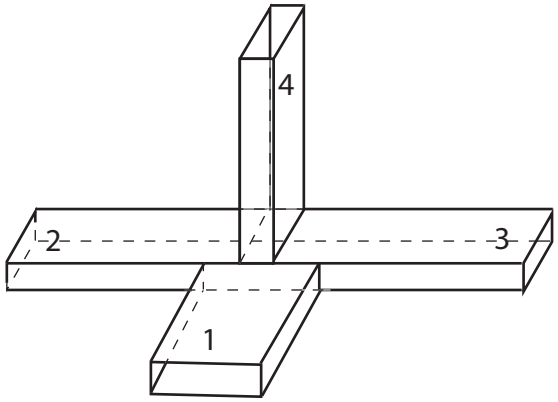


Figure 5.31: Geometry for problem 5.10.

5.9 A hollow waveguide is terminated by a perfectly conducting surface at $z = 0$. Show that the surface current that is induced by an incident mode gives rise to the reflected mode and a "transmitted" mode that cancels the incident mode.

5.10 The figure shows a magic T. All of the waveguides have the same dimensions $a \times b$ where $a > b$. We send in the fundamental mode TE_{10} mode in different ports. The frequency is in the frequency band where only the TE_{10} is propagating.

a) Assume that a TE_{10} mode is sent into 1 and that we can neglect the field that is reflected back in port 1. How much of the incident power is transmitted through ports 2, 3 and 4, respectively? Assume that the electric field in port 2 is directed upwards at a certain time and distance d from the symmetry plane. How is the electric field then directed in port 3 at the same time and distance from the symmetry plane?

b) Assume that the mode TE_{10} is sent into port 4 and that we can neglect the reflected field in port 4. How much of the incident power is transmitted through ports 1, 2 and 3, respectively? Assume that the electric field in port 2 is directed upwards at a certain time and distance d from the symmetry plane. How is the electric field then directed in port 3 at the same time and distance from the symmetry plane?

Summary of chapter 5

Mode solutions—TM- and TE-modes

Eigenvalue problem

$$\begin{cases} \nabla_T^2 v_n(\boldsymbol{\rho}) + k_{t_n}^2 v_n(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \in \Omega \\ v_n(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \quad \text{(TM-case)}$$

$$\begin{cases} \nabla_T^2 w_n(\boldsymbol{\rho}) + k_{t_n}^2 w_n(\boldsymbol{\rho}) = 0 \qquad \boldsymbol{\rho} \in \Omega \\ \dfrac{\partial w_n(\boldsymbol{\rho})}{\partial n} = 0 \qquad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \quad \text{(TE-case)}$$

Normalization

$$\iint_\Omega v_n(\boldsymbol{\rho}) v_{n'}(\boldsymbol{\rho})\,\mathrm{d}x\mathrm{d}y = \delta_{n,n'}$$

$$\iint_\Omega w_n(\boldsymbol{\rho}) w_{n'}(\boldsymbol{\rho})\,\mathrm{d}x\mathrm{d}y = \delta_{n,n'}$$

Eigenmodes

$$\begin{cases} \boldsymbol{E}_{n\nu}^{\pm}(\boldsymbol{r},\omega) = \{\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega) \pm v_n(\boldsymbol{\rho})\hat{\boldsymbol{z}}\}\, e^{\pm \mathrm{i} k_{z_n} z} \\ \boldsymbol{H}_{n\nu}^{\pm}(\boldsymbol{r},\omega) = \pm \boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho},\omega) e^{\pm \mathrm{i} k_{z_n} z} \end{cases} \quad \nu = \mathrm{TM}$$

$$\begin{cases} \boldsymbol{E}_{n_\nu}^{\pm}(\boldsymbol{r},\omega) = \boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega) e^{\pm \mathrm{i} k_{z_n} z} \\ \boldsymbol{H}_{n\nu}^{\pm}(\boldsymbol{r},\omega) = \{\pm \boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho},\omega) + (\omega)_n(\boldsymbol{\rho})\hat{\boldsymbol{z}}\}\, e^{\pm \mathrm{i} k_{z_n} z} \end{cases} \quad \nu = \mathrm{TE}$$

Transverse mode components

$$\boxed{\begin{aligned}\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho},\omega) &= \frac{\mathrm{i}}{k_{tn}^2(\omega)}\begin{cases} k_{zn}(\omega)\nabla_T v_n(\boldsymbol{\rho}), & \nu = \mathrm{TM} \\ -\omega\mu_0\mu(\omega)\hat{\boldsymbol{z}}\times\nabla_T w_n(\boldsymbol{\rho}), & \nu = \mathrm{TE}\end{cases} \\ \boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho},\omega) &= \frac{\mathrm{i}}{k_{tn}^2(\omega)}\begin{cases} \omega\epsilon_0\epsilon(\omega)\hat{\boldsymbol{z}}\times\nabla_T v_n(\boldsymbol{\rho}), & \nu = \mathrm{TM} \\ k_{zn}(\omega)\nabla_T w_n(\boldsymbol{\rho}), & \nu = \mathrm{TE}\end{cases}\end{aligned}}$$

Mode solutions—TEM-mode

Potential problem

$$\boxed{\begin{aligned}&\nabla_T^2\psi(\boldsymbol{\rho}) = 0 \\ &\psi(\boldsymbol{\rho}) = \text{ konstant} \qquad \boldsymbol{\rho} \text{ on } \Gamma\end{aligned}}$$

Normalization

$$\boxed{\iint_\Omega \nabla_T\psi(\boldsymbol{\rho})\cdot\nabla_T\psi(\boldsymbol{\rho})\,\mathrm{d}x\mathrm{d}y = 1}$$

Eigenmodes

$$\boxed{\begin{cases}\boldsymbol{E}_\nu^\pm(\boldsymbol{r},\omega) = \boldsymbol{E}_{T\nu}(\boldsymbol{\rho},\omega)e^{\pm\mathrm{i}kz} \\ \boldsymbol{H}_\nu^\pm(\boldsymbol{r},\omega) = \pm\boldsymbol{H}_{T\nu}(\boldsymbol{\rho},\omega)e^{\pm\mathrm{i}kz}\end{cases} \quad \nu = \mathrm{TEM}}$$

Transverse components

$$\boxed{\begin{aligned}&\boldsymbol{E}_{T\,\mathrm{TEM}}(\boldsymbol{\rho},\omega) = -\nabla_T\psi(\boldsymbol{\rho},\omega)\\&\eta_0\boldsymbol{H}_{T\,\mathrm{TEM}}(\boldsymbol{\rho},\omega) = \frac{k(\omega)c_0}{\omega\mu(\omega)}\hat{\boldsymbol{z}}\times\boldsymbol{E}_{T\,\mathrm{TEM}}(\boldsymbol{\rho},\omega)\end{aligned}}$$

General expansions in modes

$$\boxed{\begin{cases}\boldsymbol{E}(\boldsymbol{r},\omega) = \displaystyle\sum_{\substack{n\\ \nu=\mathrm{TM,TE,TEM}}}\left(a^+_{n\nu}\boldsymbol{E}^+_{n\nu}(\boldsymbol{r},\omega) + a^-_{n\nu}\boldsymbol{E}^-_{n\nu}(\boldsymbol{r},\omega)\right)\\ \boldsymbol{H}(\boldsymbol{r},\omega) = \displaystyle\sum_{\substack{n\\ \nu=\mathrm{TM,TE,TEM}}}\left(a^+_{n\nu}\boldsymbol{H}^+_{n\nu}(\boldsymbol{r},\omega) + a^-_{n\nu}\boldsymbol{H}^-_{n\nu}(\boldsymbol{r},\omega)\right)\end{cases}\quad\begin{cases}z\in[z_1,z_2]\\ \boldsymbol{\rho}\in\Omega\end{cases}}$$

Orthogonality relation

$$\boxed{\iint_\Omega \hat{\boldsymbol{z}}\cdot\{\boldsymbol{E}_{T\,n\nu}(\boldsymbol{\rho},\omega)\times\boldsymbol{H}^*_{T\,n'\nu'}(\boldsymbol{\rho},\omega)\}\,\mathrm{d}S = 2P^E_{n\nu}\delta_{n,n'}\delta_{\nu,\nu'}}$$

Mode power

$$\boxed{P^E_{n\nu} = \begin{cases}\dfrac{\omega}{2k_{tn}^2}\begin{cases}k_{zn}(\omega)\epsilon_0\epsilon^*(\omega), & \nu=\mathrm{TM}\\ k^*_{zn}(\omega)\mu_0\mu(\omega), & \nu=\mathrm{TE}\end{cases}\\ \dfrac{1}{2\eta_0\eta^*}, \quad \nu=\mathrm{TEM}\end{cases}}$$

Power flow

General expressions

$$\iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot <\boldsymbol{S}(t)> (\boldsymbol{r},\omega)\,\mathrm{d}x\mathrm{d}y$$
$$= \sum_{\substack{n \\ \nu=\mathrm{TM,TE}}} \mathrm{Re}\, P^E_{n\nu} \left\{ \left|a^+_{n\nu}\right|^2 e^{-2\,\mathrm{Im}\,k_{zn}z} - \left|a^-_{n\nu}\right|^2 e^{2\,\mathrm{Im}\,k_{zn}z} + 2\mathrm{i}\,\mathrm{Im}\left(a^-_{n\nu} a^{+*}_{n\nu} e^{-2\mathrm{i}\,\mathrm{Re}\,k_{zn}z}\right)\right\}$$

Propagating modes

$$\iint\limits_{\Omega} \hat{\boldsymbol{z}} \cdot <\boldsymbol{S}(t)> (\boldsymbol{r},\omega)\,\mathrm{d}x\mathrm{d}y = \pm \sum_{k_{tn}<k(\omega)} P^E_{n\nu} \left|a^\pm_{n\nu}\right|^2$$

Losses in walls

$$P(z) = P(0)e^{-\alpha z}$$
$$\alpha = \begin{cases} \dfrac{\omega\epsilon\epsilon_0}{\sigma\delta k_{tn}^2 k_{zn}} \oint_\Gamma |\hat{\boldsymbol{z}} \times \nabla_T v_n|^2\,\mathrm{d}l, & \text{TM-mod} \\ \dfrac{k_{tn}^2}{\omega\sigma\delta\mu_0\mu k_{zn}} \oint_\Gamma \left(\dfrac{k_{zn}^2}{k_{tn}^4}|\nabla_T w_n|^2 + |w_n|^2\right)\mathrm{d}l, & \text{TE-mod} \end{cases}$$

Sources and determination of expansion coefficients

$$a^\pm_{n\nu} = \frac{1}{4U_{n\nu}} \iiint\limits_V \boldsymbol{J} \cdot \boldsymbol{E}^\mp_{n\nu}\,\mathrm{d}v$$

Chapter 6

Resonance cavities

A finite volume of air or vacuum that is enclosed by a metallic surface constitutes a resonance cavity. Only electromagnetic fields with certain frequencies can exist in the cavity. These fields are called cavity modes, or eigenmodes, and the corresponding frequencies are called eigenfrequencies, or resonance frequencies. In this chapter we describe how cavity modes and resonance frequencies can be obtained by analytical and numerical methods. Resonance cavities are frequently used as bandpass and bandstop filters in microwave systems. The losses are much smaller than in traditional bandpass filters based on circuit components and that makes the filters based on cavities very narrow banded. In modern particle accelerators the particles are accelerated by the electric fields in microwave cavities. Another important application is klystrons and magnetrons, which are generators for time-harmonic electromagnetic waves. Magnetrons are used in radars and also in microwave ovens. Klystrons are used as sources in, eg., particle accelerators and high power communication systems.

6.1 General cavities

In this section we give the general theory for electromagnetic fields in cavities. We proceed and analyze how fields are generated in cavities by a current density. Finally this theory is applied to the generation of fields in cavities that are to be used at ESS, c.f., figures 6.1 and 6.2.

6.1.1 The resonances in a lossless cavity with sources

Assume a volume V with a homogeneous lossless material bounded by a perfectly conducting surface S with inward pointing normal $\hat{\boldsymbol{n}}$. The electromagnetic fields in this hollow cavity have to satisfy the Maxwell equations and the boundary conditions at the walls of the cavity. We have seen that in a source free region the Maxwell equations imply the vector Helmholtz equation. The boundary condition for the electric field at the walls is that $\hat{\boldsymbol{n}} \times \boldsymbol{E}(\omega, \boldsymbol{r}) = \boldsymbol{0}$. The eigenvalue problem for the electric field is

$$\begin{cases} \nabla^2 \boldsymbol{E}(\boldsymbol{r}) + k^2 \boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{0} & \boldsymbol{r} \in V \\ \hat{\boldsymbol{n}} \times \boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{0} & \boldsymbol{r} \in S \end{cases}$$

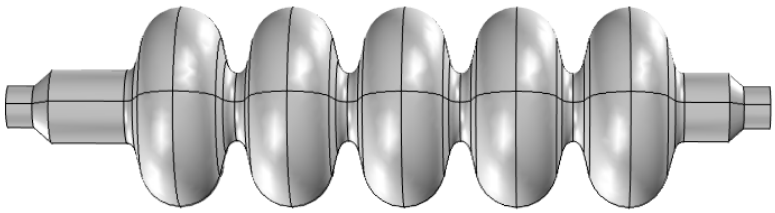


Figure 6.1: The elliptic cavity to be used in the LINAC of ESS

A general vector field $\boldsymbol{E}$ can according to Helmholtz theorem be written as a sum of a solenoidal function $\nabla \times \boldsymbol{N}$ and an irrotational function $-\nabla\Phi$ such that

$$\boldsymbol{E}(\boldsymbol{r}) = \nabla \times \boldsymbol{N}(\boldsymbol{r}) - \nabla\Phi(\boldsymbol{r})$$

In order to express the electric field from a source in a cavity both solenoidal and irrotational vector functions are needed. We will later show that only the solenoidal fields can exist in an empty cavity. The solenoidal functions satisfy the eigenvalue problem

$$\begin{cases} \nabla^2 \boldsymbol{E}_n + k_n^2 \boldsymbol{E}_n = \boldsymbol{0} & \text{in } V \\ \hat{\boldsymbol{n}} \times \boldsymbol{E}_n = \boldsymbol{0} & \text{on } S \\ \nabla \cdot \boldsymbol{E}_n = 0 & \text{in } V \text{ and on } S \end{cases}$$

where k_n^2 is the eigenvalue for eigenfunction number n. The irrotational part of the field only exist when there are sources inside the cavity. In most applications they are less important than the solenoidal fields and for this reason we do not discuss them here.

The eigenfunctions can be proven to be orthogonal and they can be chosen to be real. We let the eigenfunctions be real and normalize them such that

$$\iiint_V \boldsymbol{E}_m \cdot \boldsymbol{E}_n \, \mathrm{d}V == \delta_{n,m} \tag{6.1}$$

for all m and n. From the identity $\nabla\cdot(\boldsymbol{E}_n\times(\nabla\times\boldsymbol{E}_m)) = (\nabla\times\boldsymbol{E}_m)\cdot(\nabla\times\boldsymbol{E}_n) - k_m^2\boldsymbol{E}_m\cdot\boldsymbol{E}_n$ and Gauss theorem it follows that

$$\iiint_V (\nabla \times \boldsymbol{E}_m) \cdot (\nabla \times \boldsymbol{E}_n) \, \mathrm{d}V = k_m^2 \iiint_V \boldsymbol{E}_m \cdot \boldsymbol{E}_n \, \mathrm{d}V = k_m^2 \delta_{n,m} \tag{6.2}$$

The set of solenoidal vector eigenfunctions constitute a complete set of vector functions for the volume V. It means that any solenoidal vector field $\boldsymbol{A}(\boldsymbol{r},t)$ that satisfies $\hat{\boldsymbol{n}} \times \boldsymbol{A}(\boldsymbol{r},t) = \boldsymbol{0}$ on S can be expanded in a series of the solenoidal vector eigenfunctions. In a source free cavity $\nabla \cdot \boldsymbol{E} = 0$ and hence the irrotational part of $\boldsymbol{E}$ is zero and we expand the electric field in a series of solenoidal eigenfunctions.

$$\boldsymbol{E}(\boldsymbol{r},t) = \sum_{n=0}^{\infty} e_n(t) \boldsymbol{E}_n(\boldsymbol{r}) \tag{6.3}$$

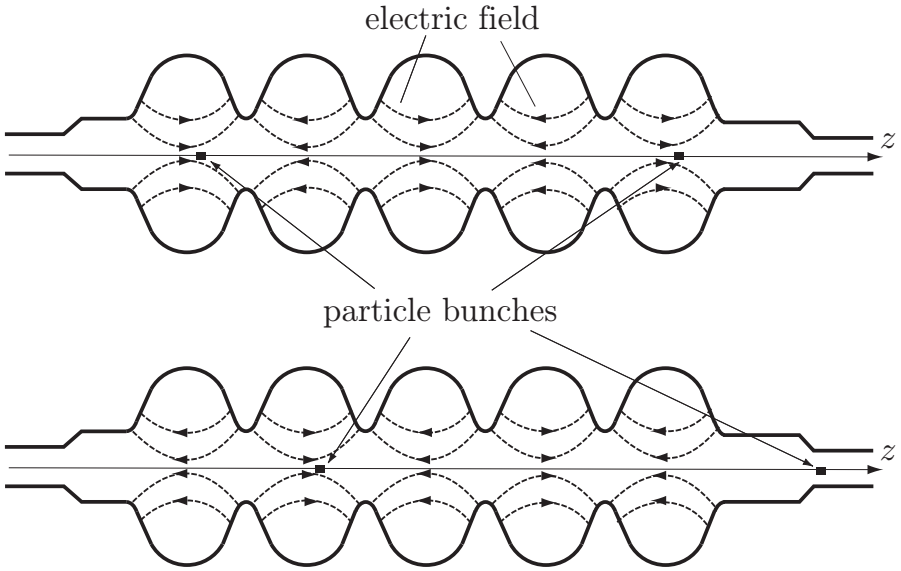


Figure 6.2: The cross section of the elliptic cavity and the electric field. The bunches are traveling to the right. The frequency of the fundamental mode is 704 MHz and the bunches come with a frequency of 352MHz. The lower figure is at half a period later than the upper one.

From the Maxwell equations it is seen that the electric field satisfies the source free wave equation

$$\nabla^2 \boldsymbol{E} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = \boldsymbol{0}$$

We insert the series expansion into this equation

$$\sum_{m=0}^{\infty} k_m^2 e_m(t)\boldsymbol{E}_m(\boldsymbol{r}) + \sum_{m=0}^{\infty} \frac{1}{c^2}\frac{\partial^2 e_m(t)}{\partial t^2}\boldsymbol{E}_m(\boldsymbol{r}) = \boldsymbol{0}$$

We take the scalar product of the wave equation with $\boldsymbol{E}_n$, integrate over the volume of the cavity and use the orthogonality of the eigenfunctions, see (6.1). This leads to

$$\frac{\partial^2 e_n(t)}{\partial t^2} + \omega_n^2 e_n(t) = 0 \tag{6.4}$$

Thus in an empty cavity $e_n(t) = A\cos(\omega_n t + \phi)$, where $\omega_n = ck_n$ is the resonance angular frequency and A is a constant. We let the phase ϕ be zero, for convenience. The conclusion is that only fields with the resonance frequencies $f_n = \omega_n/(2\pi)$ can exist in the empty cavity.

6.1.2 Q-factor for a cavity

At room temperature, the walls of resonance cavities are not perfectly conducting which leads to power losses, c.f., section 5.9, and non-zero bandwidths. The losses add a dissipative term in (6.4)

$$\frac{d^2 e_n(t)}{dt^2} + 2\alpha_n \frac{de_n(t)}{dt} + \omega_n^2 e_n(t) = 0 \tag{6.5}$$

The general solution is $e_n(t) = A_n \cos(\omega t + \phi)e^{-\alpha_n t}$, where $\omega = \sqrt{\omega_n^2 - \alpha_n^2}$. We assume that the losses are very small, i.e., $\alpha_n \ll \omega_n$ such that $\omega \approx \omega_n$.

The Q-factor (Quality factor) is a measure of the losses in the cavity. We now derive an expression for the Q-factor and its relation to the bandwidth of the cavity.

We define the Q-value as

$$Q = 2\pi \frac{\text{the stored energy in the cavity time averaged over one period}}{\text{the dissipated energy during one period at resonance}} \tag{6.6}$$

The stored and dissipated energies are evaluated during the same period. The time average of the energy in the system for mode n is due to the orthogonality

$$U_n(t) = \frac{\epsilon_0 \epsilon}{2T} \int_t^{t+T} |A_n|^2 \cos^2(\omega_n t' + \phi) e^{-2\alpha_n t'} \, \mathrm{d}t' \int_V |\boldsymbol{E}_n(\boldsymbol{r})|^2 \, \mathrm{d}V \approx \frac{\epsilon_0 \epsilon}{4} |A_n|^2 e^{-2\alpha_n t}$$

where we have used $\alpha_n T \ll 1$. It means that the time average of the energy is given by $U_n(t) = U_n(0)e^{-2\alpha_n t}$, where $U_n(0)$ is the time average of the energy at $t = 0$. The dissipated power during the period $[0, T]$ is $P_n(0) = U_n(0) - U_n(T) = U_n(0)(1 - e^{-2\alpha_n T})$ where $T = 2\pi/\omega_0$ is the period. Since $\alpha_n \ll \omega_n$ it must be that $2\alpha_n T \ll 1$ and $e^{-2\alpha_n T} \approx 1 - 2\alpha_n T$. That gives

$$Q = \frac{2\pi}{2\alpha_n T} = \frac{\omega_n}{2\alpha_n}$$

From circuit theory, see example below, we know that the Q-value is related to the bandwidth as

$$\text{BW} = f_+ - f_- = \frac{\alpha_n}{\pi} = \frac{f_n}{Q} = \text{bandwidth}$$

The dissipated energy during one period is given by

$$U_d(t) = R_S \int_t^{t+T} \iint_S |\boldsymbol{J}_{Sn}(\boldsymbol{r}, t)|^2 \, \mathrm{d}S = R_S \int_t^{t+T} \iint_S |\hat{\boldsymbol{n}} \times \boldsymbol{H}_n(\boldsymbol{r}, t)|^2 \, \mathrm{d}S$$

where $R_S = \dfrac{1}{\sigma_c \delta} = \sqrt{\dfrac{\omega \mu_c \mu_0}{2\sigma_c}}$ is the surface resistance, c.f., (3.31) and (5.48). The magnetic field is related to the electric field by the induction law, i.e.,

$$\mu_0 \frac{\partial \boldsymbol{H}_n(\boldsymbol{r}, t)}{\partial t} = -e_n(t) \nabla \times \boldsymbol{E}_n(\boldsymbol{r})$$

Example 6.1

Each mode in the cavity can be modeled by the parallel circuit in figure 6.3. From Kirchhoff's current law we get the equation for the voltage

$$\frac{d^2 v(t)}{dt^2} + 2\alpha \frac{dv(t)}{dt} + \omega_0^2 v(t) = \frac{1}{C} \frac{di(t)}{dt} \tag{6.7}$$

where $\alpha = (2RC)^{-1}$ is the attenuation constant and $\omega_0 = 1/\sqrt{LC}$ is the resonance frequency. This is the same equation we get for the amplitude $e_n(t)$ of mode number n in

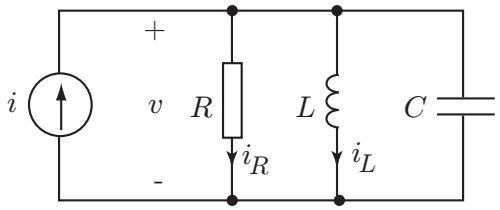


Figure 6.3: The circuit analogues of a cavity mode. The values of R, L, and C are different for different modes.

a cavity, i.e., with the electric field $\boldsymbol{E}(\boldsymbol{r},t) = e_n(t)\boldsymbol{E}_n(\boldsymbol{r})$. If $i(t) = 0$ then the solution is the damped sinusoidal signal

$$v(t) = A\cos(\omega t + \phi)e^{-\alpha t},$$

where $\omega = \sqrt{\omega_0^2 - \alpha^2}$. The analysis of the parallell resonance circuit can be found in most textbooks on basic circuit theory. It is then seen that the Q-value is given by

$$Q = \frac{2\pi}{2\alpha T} = \frac{\omega_0}{2\alpha}$$

and that the bandwidth is related to Q

$$\text{BW} = f_+ - f_- = \frac{\alpha}{\pi} = \frac{f_0}{Q} = \text{bandwidth}$$

where $f_\pm$ are the frequencies where the time averaged stored energy is half of the stored energy at the resonance frequency. The stored electric energy $W_e(t)$, stored magnetic energy $W_m(t)$ and the dissipated power $P_d(t)$ in the circuit are related to the corresponding quantities in the cavity as

$$\begin{aligned} W_e(t) &= \frac{\epsilon_0\epsilon}{2}\iiint_V |\boldsymbol{E}_n(\boldsymbol{r},t)|^2\,\mathrm{d}V = \frac{1}{2}C(v(t))^2 \\ W_m(t) &= \frac{\mu_0}{2}\iiint_V |\boldsymbol{H}_n(\boldsymbol{r},t)|^2\,\mathrm{d}V = \frac{1}{2}L(i_L(t))^2 \\ P_d(t) &= R_S\iint_S |\boldsymbol{J}_{Sn}(\boldsymbol{r},t)|^2\,\mathrm{d}S = R(i_R(t))^2 = \frac{(v(t))^2}{R} \end{aligned} \tag{6.8}$$

∎

Example 6.2
Assume that there are several modes present in a source free cavity with metallic, but not perfectly conducting, walls. In the time domain the electromagnetic field is given by

$$\boldsymbol{E}(\boldsymbol{r},t) = \sum_n e_n(t)\boldsymbol{E}_n(\boldsymbol{r}) \tag{6.9}$$

The dissipated power is given by

$$P_d(t) = R_S\iint_S |\sum_n \boldsymbol{J}_{Sn}(\boldsymbol{r},t)|^2\,\mathrm{d}S \tag{6.10}$$

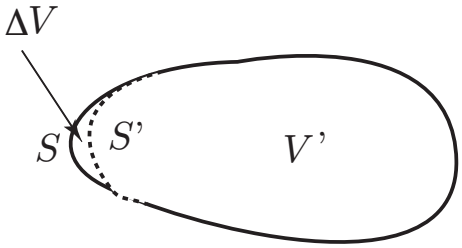


Figure 6.4: The volume V eclipsed by the surface S and the perturbed surface S'. The volume between S and S' is denoted ΔV.

Is this power equal to the sum of the dissipated powers for the modes?

The answer is yes. To prove it we use Poynting's theorem in section 1.2

$$\begin{aligned}\iint_S \boldsymbol{E}\times\boldsymbol{H}\cdot\hat{\boldsymbol{n}}\,\mathrm{d}S &= \iiint_V \nabla\cdot\boldsymbol{S}\,\mathrm{d}v \\ &= -\iiint_V \left[\boldsymbol{H}\cdot\frac{\partial\boldsymbol{B}}{\partial t}+\boldsymbol{E}\cdot\frac{\partial\boldsymbol{D}}{\partial t}\right]\mathrm{d}v - \iiint_V \boldsymbol{E}\cdot\boldsymbol{J}\,\mathrm{d}v\end{aligned} \tag{6.11}$$

where the volume V is enclosed by the surface S. Now let S be a surface deep enough into the metal such that the electromagnetic fields are zero. The dissipated power in the metal is the term $P_d(t) = \iiint_V \boldsymbol{E}\cdot\boldsymbol{J}\,\mathrm{d}v$, where only the volume in the metal gives a contribution. Thus

$$P_d(t) = -\iiint_V \left[\boldsymbol{H}\cdot\frac{\partial\boldsymbol{B}}{\partial t}+\boldsymbol{E}\cdot\frac{\partial\boldsymbol{D}}{\partial t}\right]\mathrm{d}v \tag{6.12}$$

Since $e_n(t) = A_n\cos\omega_n t e^{-\alpha_n t}$ we get

$$P_d(t) = -\iiint_V \left[\sum_n \boldsymbol{H}_n(\boldsymbol{r})e_n(t)\cdot\sum_m \boldsymbol{B}_m(\boldsymbol{r})\frac{\partial e_m(t)}{\partial t}+\sum_n \boldsymbol{E}_n(\boldsymbol{r})e_n(t)\cdot\sum_m \boldsymbol{D}_m(\boldsymbol{r})\frac{\partial e_m(t)}{\partial t}\right]\mathrm{d}v \tag{6.13}$$

We use the orthogonality and get

$$P_d(t) = -2\epsilon_0\epsilon\sum_n e_n(t)\frac{\partial e_n(t)}{\partial t} = 2\epsilon_0\epsilon\sum_n |A_n|^2\left(\alpha_n\cos^2(\omega_n t)+\cos(\omega_n t)\sin(\omega_n t)\right)e^{-2\alpha_n t} \tag{6.14}$$

The time average over one period from t to $t+T$ is

$$P_d = \epsilon_0\sum_n |A_n|^2\alpha_n e^{-2\alpha_n t} \tag{6.15}$$

Thus the time average of the total dissipated power is the sum of the dissipated powers for the modes. ∎

6.1.3 Slater's theorem

Slater's theorem is due to J. C. Slater [18] who presented the theorem in 1946. It is of importance in the design of cavities for accelerators. By using the theorem one can

determine how the resonance frequencies of a cavity are affected by small changes of the geometry in a cavity. As will be seen in the next subsection it is also very important for measurements of the electric and magnetic fields of a cavity.

Consider a cavity with a volume V that is enclosed by the bounded surface S. The mode number n has the resonance frequency f_n. We then make a small (infinitesimal) perturbation of the surface and volume such that the surface S is pushed into the volume V. The new surface is called S' and the new volume V'. The frequency of mode n is changed to $f = f_n + \Delta f$. Slater's theorem gives the following expression for f

$$\frac{f^2 - f_n^2}{f_n^2} = \iiint_{\Delta V} \left(\frac{\mu_0}{\epsilon_0} |\boldsymbol{H}_n(\boldsymbol{r})|^2 - |\boldsymbol{E}_n(\boldsymbol{r})|^2 \right) \mathrm{d}V \tag{6.16}$$

The integration is over $V - V'$, see figure 6.4. Here $\boldsymbol{H}_n = (\mathrm{i}\omega_n\mu_0)^{-1}\nabla \times \boldsymbol{E}_n$ is the magnetic field of mode n. Notice that $\frac{1}{4}\epsilon_0|\boldsymbol{E}_n|^2$ is the electric energy density and $\frac{1}{4}\mu_0|\boldsymbol{H}_n|^2$ the magnetic energy density of mode n. If we remove a volume where the electric energy is larger than the magnetic energy the frequency deceases, otherwise it increases.

To prove Slater's theorem we start with the Helmholtz equations for the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{H}$ in the source free volume V'

$$\nabla \times (\nabla \times \boldsymbol{E}) - k^2 \boldsymbol{E} = \boldsymbol{0}$$
$$\nabla \times (\nabla \times \boldsymbol{H}) - k^2 \boldsymbol{H} = \boldsymbol{0}$$

We consider mode number n and let $\boldsymbol{E}$ and $\boldsymbol{H}$ be the perturbed versions of $\boldsymbol{E}_n$ and $\boldsymbol{H}_n$. Since the perturbation is small, the wavenumber of the perturbed cavity, $k = \omega/c$, is close to the wavenumber of the unperturbed cavity, $k_n = \omega_n/c$. We take the scalar product of the upper equation and $\boldsymbol{E}_n$ and use the relation $\nabla \cdot (\boldsymbol{A} \times \boldsymbol{B}) = \boldsymbol{B} \cdot (\nabla \times \boldsymbol{A}) - \boldsymbol{A} \cdot (\nabla \times \boldsymbol{B})$ two times to get

$$(\nabla \times (\nabla \times \boldsymbol{E})) \cdot \boldsymbol{E}_n = \nabla \cdot ((\nabla \times \boldsymbol{E}) \times \boldsymbol{E}_n + \boldsymbol{E} \times (\nabla \times \boldsymbol{E}_n)) + k_n^2 \boldsymbol{E} \cdot \boldsymbol{E}_n$$

We integrate over V' and use Gauss theorem

$$(k^2 - k_n^2) \iiint_{V'} \boldsymbol{E} \cdot \boldsymbol{E}_n \, \mathrm{d}V = \iint_{S'} \hat{\boldsymbol{n}} \cdot ((\nabla \times \boldsymbol{E}) \times \boldsymbol{E}_n + \boldsymbol{E} \times (\nabla \times \boldsymbol{E}_n)) \, \mathrm{d}S$$

The last integral is zero since $\hat{\boldsymbol{n}} \times \boldsymbol{E} = \boldsymbol{0}$ on S' and then

$$(k^2 - k_n^2) \iiint_{V'} \boldsymbol{E} \cdot \boldsymbol{E}_n \, \mathrm{d}V = \iint_{S'} \hat{\boldsymbol{n}} \cdot ((\nabla \times \boldsymbol{E}) \times \boldsymbol{E}_n) \, \mathrm{d}S$$

We make use of the the completeness of the system $\boldsymbol{E}_n$. Since V' is inside V we can expand $\boldsymbol{E}$ in a series of $\boldsymbol{E}_n$ as

$$\boldsymbol{E} = \sum_{n=0}^{\infty} \boldsymbol{E}_n \iiint_{V'} \boldsymbol{E} \cdot \boldsymbol{E}_n \, \mathrm{d}V$$

The eigenfunction $\boldsymbol{E}$ is very close to $\boldsymbol{E}_n$ in V', and we approximate $\boldsymbol{E}$ as

$$\boldsymbol{E} \approx \boldsymbol{E}_n \iiint_{V'} \boldsymbol{E} \cdot \boldsymbol{E}_n \, \mathrm{d}V$$

Figure 6.5: The metallic sphere on the symmetry axis of an axially symmetric cavity. The frequency shift of the mode number n is proportional to the square of the electric field at the position of the sphere of the same mode.

That means that

$$(k^2 - k_n^2) \iiint_{V'} \boldsymbol{E}_n \cdot \boldsymbol{E}_n \, \mathrm{d}V \approx \iint_{S'} \hat{\boldsymbol{n}} \cdot ((\nabla \times \boldsymbol{E}_n) \times \boldsymbol{E}_n) \, \mathrm{d}S$$

The right hand side can now be rewritten by Gauss theorem on the volume ΔV. The volume ΔV is bounded by the surfaces S and S' and we know that $\hat{\boldsymbol{n}} \times \boldsymbol{E}_n = 0$ on S. That gives

$$\begin{aligned}(k^2 - k_n^2) \iiint_{V'} \boldsymbol{E}_n \cdot \boldsymbol{E}_n \, \mathrm{d}V \approx \iint_{S'} \hat{\boldsymbol{n}} \cdot ((\nabla \times \boldsymbol{E}_n) \times \boldsymbol{E}_n) \, \mathrm{d}S &= - \iiint_{\Delta V} \nabla \cdot ((\nabla \times \boldsymbol{E}_n) \times \boldsymbol{E}_n) \, \mathrm{d}V \\ &= - \iiint_{\Delta V} \boldsymbol{E}_n \cdot (\nabla \times (\nabla \times \boldsymbol{E}_n)) - (\nabla \times \boldsymbol{E}_n) \cdot (\nabla \times \boldsymbol{E}_n) \, \mathrm{d}V\end{aligned}$$

We use $\nabla \times (\nabla \times \boldsymbol{E}_n) = k_n^2 \boldsymbol{E}_n$, $\boldsymbol{H}_n = (\mathrm{i}\omega_n \mu_0)^{-1} \nabla \times \boldsymbol{E}_n$, $k = 2\pi f/c$, $k_n = 2\pi f_n/c$, and the normalization of $\boldsymbol{E}_n$ over V.

$$f^2 - f_n^2 = f_n^2 \iiint_{\Delta V} \left(\frac{\mu_0}{\epsilon_0} |\boldsymbol{H}_n|^2 - |\boldsymbol{E}_n|^2 \right) \mathrm{d}V \tag{6.17}$$

which is the same as (6.16).

6.1.4 Measuring electric and magnetic fields in cavities

We can utilize Slater's theorem in order to measure the electric and magnetic fields in a cavity. To do this we use a small metallic sphere of radius R_0. The radius of the sphere is much smaller than the wavelength, c.f., figure 6.5. The sphere gives a shift of the resonance frequencies of the cavity, in concordance with the Slater's theorem, and by measuring this shift, the electric or magnetic fields of the resonance modes can be determined.

When a sphere with radius R is placed in a constant electric field $\boldsymbol{E}_0 = E_0 \hat{\boldsymbol{z}} = E_0 \cos\theta \hat{\boldsymbol{r}} - E_0 \sin\theta \hat{\boldsymbol{\theta}}$ it creates an electric field in the vicinity of the sphere. The space dependence of this field is the same as the electrostatic dipole field. The total electric field close to the sphere is given by

$$\boldsymbol{E}_{\text{sphere}}(r, \theta) = \boldsymbol{E}_0 + \boldsymbol{E}_{\text{dipole}} = E_0 \hat{\boldsymbol{z}} + \frac{p}{4\pi\epsilon_0 r^3} (2\hat{\boldsymbol{r}} \cos\theta + \hat{\boldsymbol{\theta}} \sin\theta)$$

where p is the induced dipole moment. The dipole moment is obtained from the fact that the tangential component $\boldsymbol{E}_{\text{sphere}}$ is zero on the surface of the sphere. The tangential component is the $\theta-$component and since $\hat{\boldsymbol{\theta}} \cdot \hat{\boldsymbol{z}} = -\sin\theta$ we get

$$\hat{\boldsymbol{\theta}} \cdot \boldsymbol{E}_{\text{sphere}}(R_0, \theta) = -E_0 \sin\theta + \frac{p}{4\pi\epsilon_0 R^3} \sin\theta = 0$$

Thus $p = 4\pi\epsilon_0 R^3 E_0$ and

$$\boldsymbol{E}_{\text{sphere}}(r,\theta) = E_0\cos\theta\left(1+\frac{2R^3}{r^3}\right)\hat{\boldsymbol{r}} - E_0\sin\theta\left(1-\frac{R^3}{r^3}\right)\hat{\boldsymbol{\theta}}$$

First consider the case when the magnetic field is zero. We have to be careful when we determine the integral

$$\iiint_{\Delta V}\left(|\boldsymbol{E}_n|^2\right)\,\mathrm{d}V$$

since the field $\boldsymbol{E}_n$ depends on the radius of the sphere. We need to start with a sphere of radius 0 and add a layer with thickness δR, calculate the corresponding change in frequency δf Then add a new layer and get a new δf and so on. Assume that we are at the stage that the radius is R and the frequency f and that we add a layer δR. The frequency shift is then given by

$$\begin{aligned}\frac{(f+\delta f)^2 - f^2}{f^2} &= -2\pi\int_0^\pi\int_R^{R+\delta R}|\boldsymbol{E}_{\text{sphere}}(r,\theta)|^2 r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\\ &\approx -2\pi\int_0^\pi|\boldsymbol{E}_{\text{sphere}}(R,\theta)|^2 R^2\sin\theta\,\mathrm{d}\theta\,\delta R\\ &= -12\pi R^2 E_0\,\delta R\end{aligned}$$

We let $\delta f = \mathrm{d}f$ be infinitesimal and get $((f+\mathrm{d}f)^2 - f^2)/f^2 \approx 2\mathrm{d}f/f$. We now integrate f from f_n to $f_n + \Delta f$ and R from 0 to R_0. This gives

$$\ln\left(\frac{f_n+\Delta f}{f_n}\right) = -2\pi R_0^3 E_0^2$$

We utilize that $\Delta f \ll f_n$ and that $\ln(1+x)\approx x$ for small x. Then

$$\Delta f \approx -2\pi R_0^3 E_0^2 f_n \tag{6.18}$$

Next consider the case when the electric field is zero and the magnetic field is non-zero. We get the perturbation of the magnetic field by placing a perfectly conducting sphere into a constant magnetic field $\boldsymbol{H} = H_0\hat{\boldsymbol{z}}$. The boundary condition on the sphere is that the normal component of the total magnetic field is zero on the surface of the sphere. This gives very similar calculations as we used for the electric field. The magnetic field outside the sphere is

$$\boldsymbol{H}_{\text{sphere}}(r,\theta) = H_0\left(1-\frac{R^3}{r^3}\right)\cos\theta\hat{\boldsymbol{r}} - H_0\left(1+\frac{R^3}{2r^3}\right)\sin\theta\hat{\boldsymbol{\theta}}$$

This leads to

$$\Delta f \approx \frac{\mu_0}{\epsilon_0}\pi H_0^2 R_0^3 f_n$$

If we add the contributions from the electric and magnetic cases we get

$$\Delta f \approx f_n\left(\frac{\mu_0}{\epsilon_0}H_0^2 - 2E_0^2\right)\pi R_0^3$$

where H_0 is the value of $|\boldsymbol{H}_n|$ and E_0 the value of $|\boldsymbol{E}_n|$ at the position of the sphere. Notice that the electric field is normalized, c.f., Eq.(6.1).

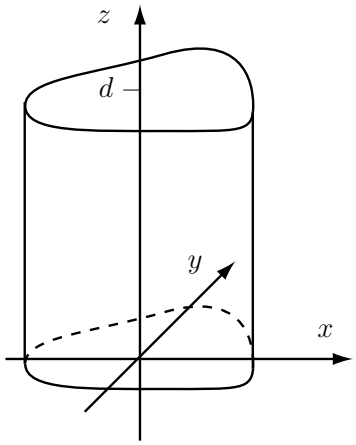


Figure 6.6: Geometry for cylindric resonance cavity.

Cavities that are used in accelerators are often axially symmetric and the mode that is used in the cavity corresponds to the TM_{010} mode of a cylindric cavity. The beam is then propagating along the symmetry axis, where the magnetic field is zero. By measuring the resonance frequency of the TM_{010} mode while a sphere is moved along the symmetry axis, one can determine the variation of the electric field along the symmetry axis for the TM_{010} mode by using Eq. (6.18).

6.2 Example: Cylindrical cavities

We analyze a common type of resonance cavity that consists of a hollow waveguide terminated by metallic plane surfaces at $z = 0$ and $z = d$, see figure 6.6. In order to determine the fields that can exist in such a cavity we need boundary conditions for the z-component of the electric and magnetic fields at $z = 0$ and $z = d$. Since $\boldsymbol{E}_T(\boldsymbol{\rho}, 0) = \boldsymbol{E}_T(\boldsymbol{\rho}, d) = \mathbf{0}$ for all $\boldsymbol{\rho}$, it follows that $\nabla_T \cdot \boldsymbol{E}_T(\boldsymbol{\rho}, 0) = \nabla_T \cdot \boldsymbol{E}_T(\boldsymbol{\rho}, d) = 0$. There are no charges inside the cavity and then $\nabla \cdot \boldsymbol{E}(\boldsymbol{r}) = 0$, i.e., $0 = \nabla_T \boldsymbol{E}_T(\boldsymbol{r}) + \partial E_z(\boldsymbol{r})/\partial z$. It follows that the z-derivative of E_z is zero at the end surfaces. The magnetic field $\boldsymbol{H}$ is zero in the metal and $\boldsymbol{B}$ is always continuous and then it follows that H_z is zero at the end surfaces. We conclude that the boundary conditions at $z = 0$ and $z = d$ are

$$\begin{cases} \dfrac{\partial E_z(\boldsymbol{\rho}, 0)}{\partial z} = \dfrac{\partial E_z(\boldsymbol{\rho}, d)}{\partial z} = 0 \\ H_z(\boldsymbol{\rho}, 0) = H_z(\boldsymbol{\rho}, d) = 0 \end{cases} \tag{6.19}$$

The fields in the cavity is a superposition of waveguide modes propagating in the positive and negative z-directions. From the expansions in (5.21) and (5.22) the z-components of the fields of mode n are expressed as

$$\begin{cases} E_z(\boldsymbol{r}) = (a^+_{n\nu} e^{\mathrm{i}k_z z} - a^-_{n\nu} e^{-\mathrm{i}k_z z}) v_n(\boldsymbol{\rho}) & \nu = \mathrm{TM} \\ H_z(\boldsymbol{r}) = (a^+_{n\nu} e^{\mathrm{i}k_z z} + a^-_{n\nu} e^{-\mathrm{i}k_z z}) w_n(\boldsymbol{\rho}) & \nu = \mathrm{TE} \end{cases}$$

The boundary conditions give $a^+_{n\nu} = -a^-_{n\nu}$ and $\sin k_z d = 0$. Hence k_z can only take the discrete values

$$k_{z\ell} = \frac{\ell\pi}{d} \quad \begin{cases} \ell = 0,\, 1,\, 2 \ldots & \nu = \mathrm{TM} \\ \ell = 1,\, 2 \ldots & \nu = \mathrm{TE} \end{cases}$$

There exists no TE-mode with value $\ell = 0$ since then $H_z = 0$. The frequencies that can exist in the cavity are determined by $k^2 = k_{t_n}^2 + k_{z_\ell}^2$ and thus

$$f_{n\ell} = \frac{c}{2\pi}\sqrt{k_{t_n}^2 + \left(\frac{\ell\pi}{d}\right)^2} \tag{6.20}$$

The fields for the corresponding resonances follow from (5.21) and (5.22)

$$\begin{cases} \boldsymbol{E}_{n\ell}(\boldsymbol{r}) = A_{nl\nu}\sqrt{\frac{\varepsilon_\ell}{d}}\left(\mathrm{i}\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho})\sin\frac{\ell\pi z}{d} + v_n(\boldsymbol{\rho})\cos\frac{\ell\pi z}{d}\hat{\boldsymbol{z}}\right) & \nu = \mathrm{TM} \\ \boldsymbol{H}_{n\ell}(\boldsymbol{r}) = A_{nl\nu}\sqrt{\frac{\varepsilon_\ell}{d}}\boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho})\cos\frac{\ell\pi z}{d} & \ell = 0, 1, 2\ldots \end{cases} \tag{6.21}$$

where $\varepsilon_\ell = 2 - \delta_{\ell,0}$ and

$$\begin{cases} \boldsymbol{E}_{n\ell}(\boldsymbol{r}) = A_{nl\nu}\mathrm{i}\sqrt{\frac{2}{d}}\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho})\sin\frac{\ell\pi z}{d} & \nu = \mathrm{TE} \\ \boldsymbol{H}_{n\ell}(\boldsymbol{r}) = A_{nl\nu}\sqrt{\frac{2}{d}}\left(\boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho})\cos\frac{\ell\pi z}{d} + \mathrm{i}w_n(\boldsymbol{\rho})\sin\frac{\ell\pi z}{d}\hat{\boldsymbol{z}}\right) & \ell = 1, 2\ldots \end{cases} \tag{6.22}$$

Notice that $\boldsymbol{H}_{nl}$ and $\boldsymbol{E}_{nl}$ are related via the induction law $\boldsymbol{H}_{nl}(\boldsymbol{r}) = (\mathrm{i}\omega\mu_0)^{-1}\nabla\times\boldsymbol{E}_{nl}(\boldsymbol{r})$. Both $\boldsymbol{E}_{Tn\nu}(\boldsymbol{\rho})$ and $\boldsymbol{H}_{Tn\nu}(\boldsymbol{\rho})$ are imaginary, as seen from Eq. (5.24). Thus $\boldsymbol{E}_{nl}$ is real and $\boldsymbol{H}_{nl}$ imaginary. The amplitudes $A_{nl\nu}$ are determined from the normalization

$$\iiint_V |\boldsymbol{E}_{nl}|^2\,\mathrm{d}V = \frac{\mu_0}{\epsilon_0\epsilon}\iiint_V |\boldsymbol{H}_{nl}|^2\,\mathrm{d}V = 1$$

Then

$$\begin{aligned} A_{nl\mathrm{TM}} &= \frac{k_{tn}}{k_{nl}} \\ A_{nl\mathrm{TE}} &= \frac{k_{tn}}{\omega_{nl}\mu_0} \end{aligned} \tag{6.23}$$

We are now prepared to determine the Q-value and bandwidth of a cylindrical cavity with metal walls. The conductivity of the walls is σ_c and the relative permeability μ. The electric field of the cavity mode number n is given by

$$\boldsymbol{E}(\boldsymbol{r}, t) = \cos(\omega_n t + \phi)e^{-\alpha t}\boldsymbol{E}_n(\boldsymbol{r})$$

where $\boldsymbol{E}_n(\boldsymbol{r})$ is given by (6.21) for TM-modes and (6.22) for TE-modes. The time average of the losses in the walls gets a contribution $P_e(t)$ from the end surfaces and a contribution $P_m(t)$ from the lateral surface,

$$P_c(t) = P_e(t) + P_m(t)$$

The time averaged stored energy $U(t) = U_e(t) + U_m(t) = 2U_e(t)$, is due to the normalization

$$U(t) = 2U_e(t) = \frac{\epsilon_0\epsilon}{2}e^{-2\alpha_n t} \tag{6.24}$$

The time average of the losses at the end surfaces are

$$P_e(t) = \frac{R_S}{2}\iint_\Omega |\boldsymbol{H}_T(\boldsymbol{\rho}, 0)|^2 + |\boldsymbol{H}_T(\boldsymbol{\rho}, d)|^2\,\mathrm{d}x\mathrm{d}y e^{-2\alpha t} = R_S\iint_\Omega |\boldsymbol{H}_T(\boldsymbol{\rho}, 0)|^2\,\mathrm{d}x\mathrm{d}y e^{-2\alpha t}$$

where $R_S = \dfrac{1}{\sigma_c \delta} = \sqrt{\dfrac{\omega \mu_c \mu_0}{2\sigma_c}}$ is the surface resistance, c.f., (5.47). From (5.24) and the orthogonality relation in (5.33) we get

$$P_e(t) = \begin{cases} R_S \dfrac{\varepsilon_\ell \epsilon_0 \epsilon}{\mu_0 d} e^{-2\alpha_n t} & \text{TM} \\ \\ R_S \dfrac{k_{z\ell}^2 \epsilon_0 \epsilon}{\mu_0^2 k_{tn}^2 d} e^{-2\alpha_n t} & \text{TE} \end{cases}$$

We see that the quotient $P_e(t)/U(t)$ is independent of the cross section Ω and only depends on the length of the cavity

$$\frac{P_e(t)}{U(t)} = \begin{cases} R_S \dfrac{2\varepsilon_\ell}{\mu_0 d} & \text{TM} \\ \\ R_S \dfrac{2k_{z\ell}^2}{\mu_0 k_{n\ell}^2 d} & \text{TE} \end{cases}$$

The time average of the losses in the lateral surface is

$$P_m(t) = \frac{1}{2} R_S \int_0^d \int_\Gamma |\hat{\boldsymbol{n}} \times \boldsymbol{H}(\boldsymbol{r})|^2 d\Gamma dz e^{-2\alpha t}$$

where $\hat{\boldsymbol{n}}$ is the normal to the lateral surface and Γ is the boundary curve. We now utilize equations (5.21)–(5.24) and (6.23) to get

$$P_m(t) = \begin{cases} \dfrac{R_S \epsilon_0 \epsilon}{2\mu_0 k_{tn}^2} e^{-2\alpha_n t} \displaystyle\int_\Gamma |\hat{\boldsymbol{n}} \cdot \nabla_T v_n(\boldsymbol{\rho})|^2 d\Gamma & \text{TM} \\ \\ \dfrac{R_S \epsilon_0 \epsilon}{2\mu_0 k_{nl}^2} e^{-2\alpha_n t} \displaystyle\int_\Gamma \frac{k_{z\ell}^2}{k_{tn}^4} |\hat{\boldsymbol{n}} \times \nabla_T w_n(\boldsymbol{\rho})|^2 + |w_n(\boldsymbol{\rho})|^2 d\Gamma & \text{TE} \end{cases} \tag{6.25}$$

We conclude that the Q-factor is related to the attenuation coefficient α via $Q = \omega_{n\ell}/2\alpha$ and that α is given by

$$\alpha = \frac{P_c(t)}{2U(t)} = \begin{cases} \dfrac{R_S}{\mu_0} \left(\dfrac{\varepsilon_\ell}{d} + \dfrac{1}{2k_{tn}^2} \displaystyle\int_\Gamma |\hat{\boldsymbol{n}} \cdot \nabla_T v_n(\boldsymbol{\rho})|^2 d\Gamma \right) & \text{TM} \\ \\ \dfrac{R_S}{\mu_0 k_{n\ell}^2} \left(\dfrac{k_{z\ell}^2}{d} + \dfrac{1}{2} \displaystyle\int_\Gamma \frac{k_{z\ell}^2}{k_{tn}^2} |\hat{\boldsymbol{n}} \times \nabla_T w_n(\boldsymbol{\rho})|^2 + k_{tn}^2 |w_n(\boldsymbol{\rho})|^2 d\Gamma \right) & \text{TE} \end{cases} \tag{6.26}$$

Example 6.3
We give an example with a circular cylindric cavity. The resonances are denoted TE_{mnl} and TM_{mnl} where the first two indices m and n are related to the variation in the azimuthal and radial directions, and the index l to the variation in the $z-$direction. For a circular waveguide with radius a the eigenfunctions v_n and w_n are given by the table 5.4 on

page 92. By inserting those expressions into (6.25) the power losses in the lateral surface are obtained

$$P_m(t) = R_S e^{-2\alpha t} \begin{cases} (k_{mn\ell}a)^2 \xi_{mn}^{-2} \eta_0^{-2} a^{-1} & \text{TM} \\ \dfrac{\eta_{mn}^2}{(\eta_{mn}^2 - m^2)a} \left(m^2 \dfrac{(k_{z\ell}a)^2}{\eta_{mn}^4} + 1 \right) & \text{TE} \end{cases}$$

Hence the Q-factors are

$$\begin{cases} Q_{TM} = \dfrac{\eta_0 k_{mn\ell} a}{2R_S(1 + \varepsilon_\ell a/d)} \\ Q_{TE} = \dfrac{(k_{mn\ell}a)^3 \eta_0 (1 - m^2 \eta_{mn}^{-2})}{2R_S \eta_{mn}^2 \left(1 + (k_{z\ell}a)^2 m^2 \eta_{mn}^{-4} + 2\left(1 - m^2 \eta_{mn}^{-2}\right) k_{z\ell}^2 a^3 d^{-1} \eta_{mn}^{-2}\right)} \end{cases}$$

Since $k_{mn\ell}a = \sqrt{\eta_{mn}^2 + (\ell\pi a/d)^2}$ and $k_{mn\ell}a = \sqrt{\xi_{mn}^2 + (\ell\pi a/d)^2}$ the $Q-$values Q_{TM} and Q_{TE} can both be written in the form $f(a/d)/R_S$. ∎

6.3 Example: Spherical cavities

Consider a spherical cavity with radius a, filled with air or vacuum. In order to determine the resonances we need general solutions to Maxwells equations in spherical coordinates.

6.3.1 Vector spherical harmonics

Vector spherical harmonics are vector valued functions that constitute an orthogonal set on the unit sphere. They are defined as, cf. [3]

$$\begin{aligned} \boldsymbol{A}_{1\sigma ml}(\theta,\phi) &= \frac{1}{\sqrt{l(l+1)}} \nabla \times (\boldsymbol{r} Y_{\sigma ml}(\theta,\phi)) \\ \boldsymbol{A}_{2\sigma ml}(\theta,\phi) &= \hat{\boldsymbol{r}} \times \boldsymbol{A}_{1\sigma ml}(\theta,\phi) = \frac{1}{\sqrt{l(l+1)}} r \nabla Y_{\sigma ml}(\theta,\phi) \\ \boldsymbol{A}_{3\sigma ml}(\theta,\phi) &= \hat{r} Y_{\sigma ml}(\theta,\phi). \end{aligned} \tag{6.27}$$

where $Y_{\sigma ml}$ are the spherical harmonics:

$$Y_{\sigma ml}(\theta,\phi) = \sqrt{\frac{\varepsilon_m}{2\pi}} \sqrt{\frac{2l+1}{2} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) \begin{pmatrix} \cos m\phi \\ \sin m\phi, \end{pmatrix} \tag{6.28}$$

where $\varepsilon_m = 2 - \delta_{m0}$ and τ, m, l take the values

$$\sigma = \begin{pmatrix} \mathrm{e} \\ \mathrm{o} \end{pmatrix}, \quad m = 0,\, 1,\, 2,\, \ldots,\, l, \quad l = 0,\, 1, \ldots \tag{6.29}$$

When $l = 0$ the vector spherical harmonics are defined to be zero. The functions $P_l^m(\cos\theta)$ are associated Legendre functions, c.f., [3]. The vector spherical harmonics constitute an orthogonal set of vector functions on the unit sphere

$$\int_\Omega \boldsymbol{A}_{\tau n}(\theta,\phi) \cdot \boldsymbol{A}_{\tau' n'}(\theta,\phi)\, \mathrm{d}\Omega = \delta_{\tau\tau'}\delta_{nn'}, \tag{6.30}$$

where the integration is over the unit sphere and where n is the multi index $n = \sigma ml$.

6.3.2 Regular spherical vector waves

The regular divergence-free spherical vector waves are defined by

$$\begin{cases} \boldsymbol{u}_{1n}(\boldsymbol{r}) = j_l(kr)\boldsymbol{A}_{1\sigma ml}(\theta,\phi) \\ \boldsymbol{u}_{2n}(\boldsymbol{r}) = \dfrac{1}{k}\nabla\times(j_l(kr)\boldsymbol{A}_{1\sigma ml}(\theta,\phi)) \\ \qquad\quad = j_l'(kr)\boldsymbol{A}_{2\sigma ml}(\theta,\phi) + \dfrac{1}{kr}j_l(kr)(\boldsymbol{A}_{2\sigma ml}(\theta,\phi) + \sqrt{l(l+1)}\boldsymbol{A}_{3\sigma ml}(\theta,\phi)), \end{cases} \tag{6.31}$$

where $j_l(kr)$ is the spherical Bessel function and $n = \sigma ml$ is a multiindex. They satisfy the vector Helmholtz equation

$$\nabla\times(\nabla\times\boldsymbol{u}_{\tau nn}(\boldsymbol{r})) - k^2\boldsymbol{u}_{\tau nn}(\boldsymbol{r}) = \boldsymbol{0}$$

This is the same equation that the electric and magnetic fields satisfy in a source free region. The waves are called regular since they have finite amplitude everywhere. There are also non-regular waves that have the Bessel functions, j_l replaced by Neumann functions, n_l. The irregular waves will be used later in this chapter when we determine the resonances in the region between two concentric spheres.

6.3.3 Resonance frequencies in a spherical cavity

Consider a spherical cavity with radius a and filled with air, or vacuum. The boundary condition is $\hat{\boldsymbol{r}}\times\boldsymbol{E} = \boldsymbol{0}$. The electric field $\boldsymbol{E}$ can be expanded in the system of regular divergence-free spherical vector waves. The case $\boldsymbol{E} = \boldsymbol{u}_{1n}$ is referred to as the TE-case since the electric field lacks radial component. In the case $\boldsymbol{E} = \boldsymbol{u}_{2n}(\boldsymbol{r})$ the magnetic field is given by $\boldsymbol{H} = \dfrac{1}{\mathrm{i}\omega\mu_0}\nabla\times\boldsymbol{u}_2 = \dfrac{k}{\mathrm{i}\omega\mu_0}\boldsymbol{u}_{1n}$. Since $\boldsymbol{H}$ lacks radial component the case is referred to as the TM-case. It follows that the resonances are obtained from

$$\begin{cases} \hat{\boldsymbol{r}}\times\boldsymbol{u}_{1n}(a\hat{\boldsymbol{r}}) = \boldsymbol{0} & \mathrm{TE-case} \\ \hat{\boldsymbol{r}}\times\boldsymbol{u}_{2n}(a\hat{\boldsymbol{r}}) = \boldsymbol{0} & \mathrm{TM-case} \end{cases} \tag{6.32}$$

That leads to the equations for the wavenumbers

$$j_l(ka) = 0,\ \mathrm{TE} \tag{6.33}$$

and

$$j_l'(ka) + (ka)^{-1}j_l(ka) = 0,\ \mathrm{TM} \tag{6.34}$$

Table (4) gives the ten lowest resonance frequencies, with 15 digits accuracy, for a sphere with radius $a = 1$ m and the corresponding τ and l values. The resonances are degenerated such that there are $2l+1$ different modes for a resonance with polar index l, the modes corresponds to $m = 0$, $\sigma = \mathrm{e}$ and $m = 1, 2, \ldots l$, $\sigma = \begin{pmatrix}\mathrm{e}\\ \mathrm{o}\end{pmatrix}$. The number of resonance modes increases as the cube of the frequency for all resonance cavities. It means that the density of modes is proportional to the square of the frequency. Figure 6.7 gives $N^{1/3}$ as a function of ka, where N is the number of modes with wavenumber less than k. The largest k in the graph is given by $ka = 230.5$ and that corresponds to a radius of $a = 36.7$ wavelengths. There are 1.733 million modes with a wavenumber smaller than the largest

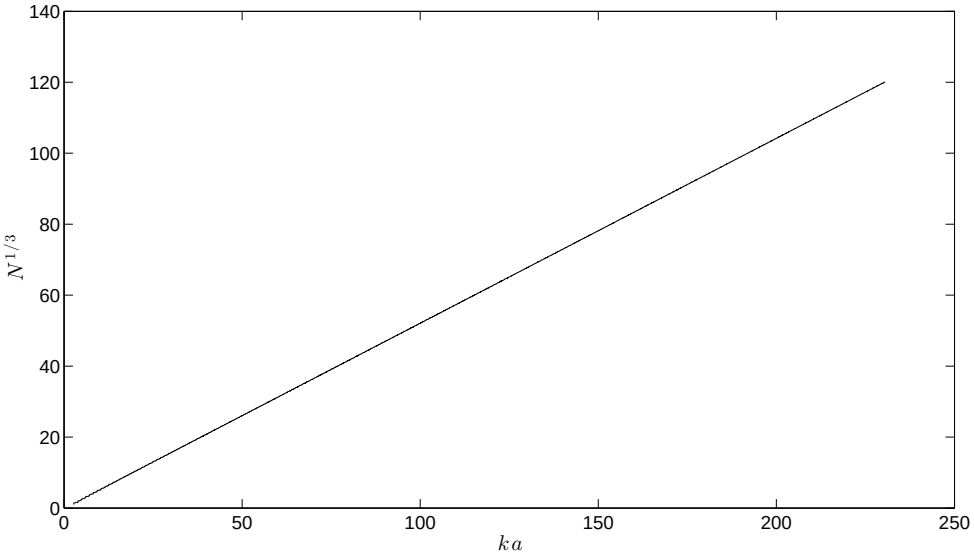


Figure 6.7: The cube root, $N^{1/3}$, of the total number of modes for a sphere with radius a with wavenumber less than k, as a function of ka. The relation $N^{1/3} \sim ka$ is valid for all cavities.

k in the graph. The number of resonances with a wavelength less than 500 nm in a sphere with a radius of 1 meter is $2.8 \cdot 10^{20}$, a huge number.

The time averages of the electric and magnetic energies are equal at resonance. For the TE-case $\boldsymbol{E} = \boldsymbol{u}_{1n}$ and then the time average of the electric field is $U_e = \frac{\epsilon_0}{4} \iiint_V |\boldsymbol{u}_1|^2 \, \mathrm{d}v$, where V is the volume of the sphere. Due to the orthogonality of $\boldsymbol{A}_{\tau n}$ we get

$$\iiint_V |\boldsymbol{u}_1|^2 \, \mathrm{d}v = \int_0^a \left(j_l(kr)\right)^2 r^2 \, \mathrm{d}r \tag{6.35}$$

From appendix A.3 we get $j_l(x) = \sqrt{\frac{2}{\pi x}} J_{l+0.5}(x)$ and the integral

$$\int x(J_\nu(x))^2 \, \mathrm{d}x = \frac{x^2}{2}(J'_\nu(x))^2 + \frac{1}{2}\left(x^2 - \nu^2\right)(J_\nu(x))^2 + \text{constant} \tag{6.36}$$

and get

$$U_e = \frac{\epsilon_0}{4} \iiint_V |\boldsymbol{u}_{1n}|^2 \, \mathrm{d}v = \frac{a^3 \epsilon_0}{8} (j'_l(ka))^2 \tag{6.37}$$

For $\tau = 2$ we have $\boldsymbol{E} = \boldsymbol{u}_2$. Then

$$\boldsymbol{H} = \frac{1}{\mathrm{i}\omega\mu_0} \nabla \times \boldsymbol{u}_2 = \frac{1}{\mathrm{i}\omega\mu_0 k} \nabla \times (\nabla \times \boldsymbol{u}_1) = \frac{k}{\mathrm{i}\omega\mu_0} \boldsymbol{u}_1$$

Table 6.1: The 10 lowest resonance frequencies in a spherical vacuum cavity with radius $a = 1$ m and perfectly conducting surface. n_r is the order number for the zeros of (6.33) and (6.34). All digits in the frequencies are correct.

order	frequency (MHz)	τ	l-value	n_r
1	130.911744010408	TM	1	1
2	184.662441148350	TM	2	1
3	214.396074654639	TE	1	1
4	237.299051157491	TM	3	1
5	274.994531395624	TE	2	1
6	289.236527486054	TM	4	1
7	291.851935634332	TM	1	2
8	333.418355236651	TE	3	1
9	340.684892690868	TM	5	1
10	355.135373846648	TM	2	2

It is seen that

$$\begin{aligned}\int_V |\boldsymbol{u}_1|^2 \,\mathrm{d}v &= \int_0^a \left(j_l(kr)\right)^2 r^2 \,\mathrm{d}r = k^{-3}\frac{ka}{8}\left(1+4(ka)^2-(2l+1)^2\right)\left(j_l(ka)\right)^2 \\ &= k^{-3}\frac{ka}{2}\left((ka)^2-l(l+1)\right)\left(j_l(ka)\right)^2\end{aligned} \tag{6.38}$$

The electric and magnetic energies are

$$U_e = U_m = \frac{\mu_0}{4}\left(\frac{k}{\omega\mu_0}\right)^2 \int_V |\boldsymbol{u}_1|^2 \,\mathrm{d}v = \frac{a^3\epsilon_0}{8(ka)^2}\left((ka)^4-(ka)^2 l(l+1)\right)\left(\frac{j_l(ka)}{ka}\right)^2 \tag{6.39}$$

6.3.4 Q-values

We define the Q-value as

$$Q = 2\pi \frac{\text{the stored energy in the cavity time averaged over one period}}{\text{the dissipated energy during one period at resonance}} \tag{6.40}$$

Consider first $\tau = 1$, i.e., $\boldsymbol{E}_n = \boldsymbol{u}_{1n}$. According to (6.37) the time averaged stored energy is

$$U = 2U_e = \frac{\epsilon_0}{2}\iiint_V |\boldsymbol{u}_{1n}|^2 \,\mathrm{d}v = a^3\frac{\epsilon_0}{4}(j_l'(ka))^2 \tag{6.41}$$

The dissipated power during one period, T, is

$$P = \frac{T}{2}R_S \iint_S |\boldsymbol{J}_S|^2 \,\mathrm{d}S = \frac{T}{2}R_S \iint |\hat{\boldsymbol{n}} \times \boldsymbol{H}|^2 \,\mathrm{d}S$$

where $\boldsymbol{H} = -\dfrac{\mathrm{i}}{\omega\mu_0}\nabla \times \boldsymbol{u}_1 = -\dfrac{\mathrm{i}k}{\omega\mu_0}\boldsymbol{u}_2$. The surface resistance is $R_S = \sqrt{\dfrac{\omega_n\mu_c\mu_0}{2\sigma_c}} = \dfrac{1}{\sigma_c\delta}$ where σ_c is the conductivity of the metal and $\delta = \sqrt{\dfrac{2}{\omega_n\sigma_c\mu_c\mu_0}}$ is the skin depth.

The orthogonality of the vector spherical harmonics, equations (6.48) and (6.33) give

$$P = \frac{T}{2} R_S \left(\frac{k}{\omega\mu_0}\right)^2 \iint_S |\boldsymbol{u}_2|^2 \,\mathrm{d}S = a^2 \frac{T}{2} R_S \left(\frac{k}{\omega\mu_0}\right)^2 \left(j_l'(ka)\right)^2 = a^2 \frac{T}{2} R_S \eta_0^{-2} \left(j_l'(ka)\right)^2$$

where $\eta_0 = \sqrt{\dfrac{\mu_0}{\epsilon_0}}$. The Q-value is

$$Q = 2\pi \frac{U}{P} = a \frac{1}{2R_S} \omega\mu_c\mu_0 = \frac{a}{\delta} \tag{6.42}$$

For $\tau = 2$ the stored energy is given by (6.39)

$$U = \frac{\epsilon_0}{2} \iiint_V |\boldsymbol{u}_{2n}|^2 \,\mathrm{d}v = \frac{a^3\epsilon_0}{4(ka)^2} \left((ka)^4 - (ka)^2 l(l+1)\right) \left(\frac{j_l(ka)}{ka}\right)^2 \tag{6.43}$$

The dissipated power integrated over one period is

$$P = \frac{T}{2} R_S \iint |\boldsymbol{J}_S|^2 \,\mathrm{d}S = \frac{T}{2} R_S \iint |\hat{\boldsymbol{n}} \times \boldsymbol{H}|^2 \,\mathrm{d}S$$

where $\boldsymbol{H} = -\dfrac{\mathrm{i}}{\omega\mu_0} \nabla \times \boldsymbol{u}_2 = -\dfrac{\mathrm{i}k}{\omega\mu_0} \boldsymbol{u}_1 = -\mathrm{i}\eta_0^{-1} \boldsymbol{u}_1$ The orthogonality of the vector spherical harmonics gives

$$P = \frac{T}{2} R_S \left(\eta_0^{-1}\right)^2 \iint_S |\hat{\boldsymbol{n}} \times \boldsymbol{u}_1|^2 \,\mathrm{d}S = a^2 \frac{T}{2\eta_0^2} R_S (j_l(ka))^2 \tag{6.44}$$

Thus

$$Q = 2\pi \frac{U}{P} = a\omega \frac{\mu_0}{2R_S(ka)^4} \left((ka)^4 - (ka)^2 l(l+1)\right) = \frac{a}{\delta} \left(1 - \frac{l(l+1)}{(ka)^2}\right) \tag{6.45}$$

6.3.5 Two concentric spheres

Consider two concentric conducting spheres with radius a and $b = a + h$. The general solutions are

$$\boldsymbol{E}(\boldsymbol{r}) = \alpha \boldsymbol{u}_{1l}(\boldsymbol{r}) + \beta \boldsymbol{w}_{1l}(\boldsymbol{r}) \tag{6.46}$$

for TE-modes and

$$\boldsymbol{E}(\boldsymbol{r}) = \gamma_l \boldsymbol{u}_{2l}(\boldsymbol{r}) + \kappa_l \boldsymbol{w}_{2l}(\boldsymbol{r}) \tag{6.47}$$

for TM-modes. The vector functions $\boldsymbol{w}_{\tau l}(\boldsymbol{r})$ are defined as

$$\begin{cases} \boldsymbol{w}_{1n}(\boldsymbol{r}) = n_l(kr) \boldsymbol{A}_{1\sigma ml}(\theta, \phi) \\ \boldsymbol{w}_{2n}(\boldsymbol{r}) = \dfrac{1}{k} \nabla \times (n_l(kr) \boldsymbol{A}_{1\sigma ml}(\theta, \phi)) \\ \qquad\quad = n_l'(kr) \boldsymbol{A}_{2\sigma ml}(\theta, \phi) + \dfrac{1}{kr} n_l(kr) (\boldsymbol{A}_{2\sigma ml}(\theta, \phi) + \sqrt{l(l+1)} \boldsymbol{A}_{3\sigma ml}(\theta, \phi)), \end{cases} \tag{6.48}$$

where $n_l(kr)$ is the spherical Neumann function.

The boundary conditions are that $\hat{\boldsymbol{n}} \times \boldsymbol{E} = 0$ for $r = a$ and $r = b$. By utilizing the orthogonality of the vector spherical harmonics two systems of equations are obtained

$$\begin{aligned} \alpha j_l(ka) + \beta n_l(ka) &= 0 \\ \alpha j_l(kb) + \beta n_l(kb) &= 0 \end{aligned} \tag{6.49}$$

for TE-modes and

$$\begin{aligned} \gamma \left(j_l'(ka) + \frac{j_l(ka)}{ka} \right) + \beta \left(n_l'(ka) + \frac{n_l(ka)}{ka} \right) &= 0 \\ \gamma \left(j_l'(kb) + \frac{j_l(kb)}{kb} \right) + \beta \left(n_l'(kb) + \frac{n_l(kb)}{kb} \right) &= 0 \end{aligned} \tag{6.50}$$

for TM-modes. The resonance wavenumbers k are determined from the determinants

$$\begin{cases} j_l(ka)n_l(kb) - j_l(kb)n_l(ka) = 0 & \text{TE} \\ \left(j_l'(ka) + \frac{j_l(ka)}{ka} \right) \left(n_l'(kb) + \frac{n_l(kb)}{kb} \right) - \left(j_l'(kb) + \frac{j_l(kb)}{kb} \right) \left(n_l'(ka) + \frac{n_l(ka)}{ka} \right) = 0 & \text{TM} \end{cases} \tag{6.51}$$

Now consider the case when $h \ll a$. Then the lowest resonance frequencies are for TM-modes. The TE-modes have no radial component of the electric field and then the first resonance is when h is on the order of half a wavelength. To find the lowest TM-modes we make Taylor expansions of $n_l'(kb) + \frac{n_l(kb)}{kb}$ and $j_l'(kb) + \frac{j_l(kb)}{kb}$ such that

$$n_l'(kb) + \frac{n_l(kb)}{kb} = n_l'(ka) + \frac{n_l(ka)}{ka} + kh \left(n_l''(ka) + \frac{n_l'(ka)}{ka} - \frac{n_l(ka)}{(ka)^2} \right) + \mathcal{O}((kh)^2) \tag{6.52}$$

By using the differential equation for spherical Bessel and Neumann functions, $j_l''(ka) = -\frac{2}{ka} j_l'(ka) - \left(1 - \frac{l(l+1)}{(ka)^2} \right) j_l(ka)$ and the Wronski relation in (A.15) in appendix A $j_l(ka)n_l'(ka) - j_l'(ka)n_l(ka) = \frac{1}{(ka)^2}$ the TM-equation leads to the equation for the lowest resonance frequencies

$$ka \approx \sqrt{l(l+1)} \tag{6.53}$$

Schumann resonances

The region between the ionosphere and the ground of the earth acts as a resonance cavity. Even though the conductivity of the ionosphere and the ground are small one may still consider them to be perfect conductors. Since the radius of the earth is $a = 6367$ km and the thickness h of the non-conducting atmosphere between the ground and the ionosphere is 80-100 km we have $h \ll a$ and hence the resonances are approximately given by (6.53). The lowest resonances are $f_1 = 10.6$ Hz, $f_2 = 18.4$ Hz, $f_3 = 26.0$ $f_4 = 33.5$ Hz. This is quite close to the measured frequencies $f_1 = 7.83$ Hz, $f_2 = 14.3$ Hz, $f_3 = 20.8$ Hz and $f_4 = 27.3$ Hz.

The conductivity of the ionosphere varies in the range $10^{-7} - 10^{-4}$ S/m and seawater has a conductivity on the order of 10^{-1} S/m. At the lowest Schuman resonance the corresponding skin depths are 15 km for the ionosphere and 500 m for the ground. This is to be compared with the thickness of 500 km for the ionosphere, the radius of the earth

6367 km, and the thickness $h \approx 85$ km of the layer between them. It means that it is relevant to use the model of the ionosphere and the ground as perfect conductors.

The discrepancy between the frequencies from (6.53) and the measured frequencies are due to that the ground and ionosphere are not perfectly conducting. A better analysis is to use the impedance boundary conditions in (1.18) together with the expansions in (6.46). With a conductivity of the ionosphere of $9 \cdot 10^{-6}$ S/m and the thickness $h = 100$ km between the earth and the ground the resonance frequencies are $f_1 = 8.5$ Hz, $f_2 = 15.3$ Hz, $f_3 = 22.1$ Hz and $f_4 = 29.0$ Hz. The frequencies depend highly on the conductivity of the ionosphere, but also on the thickness h. The frequencies increases with increasing conductivity and also with increasing h. This is of no surprise since when σ and h increases a larger part of the wave will travel in the non-conducting atmosphere.

6.4 Analyzing resonance cavities with FEM

The resonance cavities can be analyzed by FEM. There are three different cases that are of interest:

1. If the cavity is axisymmetric we use **2D axisymmetric** in COMSOL. That is a very fast and accurate solver. The axially symmetric geometry makes it possible to expand the electric and magnetic fields in a Fourier series in the system $e^{\mathrm{i}n\phi}$. Then the problem is reduced to a two-dimensional problem in the cylindrical coordinates r_c and z. Each n value is treated separately.

2. If the resonance cavity consists of a hollow waveguide with plane metallic walls at $z = 0$ and $z = h$ it is easy to analyze it with FEM. First the cut-off frequencies for the different modes are determined using the scheme on page **??**. Then the resonance frequencies of the cavity are obtained from (6.20).

3. If neither of the two previous cases are relevant then we have to use the three-dimensional solver.

We now give an example of the first and third cases.

Example 6.4
We determine some of the resonances of a hollow sphere with radius $a = 1$ m. We use the solver **2D axisymmetric** since the sphere is axially symmetric. All of the field components can be expanded in a Fourier series

$$f(r_c, \phi, z) = \sum_{m=-\infty}^{\infty} f_m(r_c, z) e^{\mathrm{i}m\phi}$$

Fields with different m-values do not couple to each other and ten each m value can be treated separately.

1. First choose **2D axisymmetric**>**Radio Frequency**>**Electromagnetic waves** >**Eigenfrequency**.

2. Draw a circle with radius $a = 0.1$ m and put its center at $(0, 0)$.

3. Choose **Sector angle** 180 degrees and **Rotation** -90 degrees. By that you have a half circle in the right half-plane.

4. Choose **Air** as material.
5. Go to **Electromagnetic waves** and choose **perfect conductor** as boundary condition for the circular line. The symmetry axis has the condition **Axial Symmetry** by default.
6. Choose **Electromagnetic waves** and the azimuthal index m.
7. In **Study>Eigenfrequency** we set the frequency to eg., 1 GHz. This is the frequency where COMSOL starts to look for an eigenfrequency. We can also choose the number of resonances that it will determine.
8. The mesh size is **Normal** by default. If we need a better accuracy then we choose a finer mesh.
9. We now let Comsol solve the problem.
10. COMSOL calculates the lowest resonant frequencies and their electric fields. There might be spurious solutions that are unphysical. The resonance frequency for these solutions are either very far from 1 GHz, or even complex, and the corresponding field plots are fuzzy.

If we are interested in the Q-value for a mode we can determine this in at least two different ways. The first one is to use the expressions for the stored energies and the dissipated power in (6.8), and the relation for Q in (6.40). The absorbed power is given by

$$P_c = \frac{1}{2} R_S \iint_S |\boldsymbol{J}_S|^2 \, \mathrm{d}S = \frac{1}{2} R_S \iint_S |\boldsymbol{H}|^2 \, \mathrm{d}S$$

where S is the surface of the cavity. We use **Results>Derived values>Line integral** to evaluate this integral in COMSOL. Remember to mark **Integration Settings>Compute surface integral**. In the axisymmetric mode in COMSOL the absolute value of the magnetic field is a pre-defined quantity. This is enough to calculate the integral. The stored energy U is obtained from the predefined quantity electric energy density. Integrate the electric energy density by **Results>Derived values>Surface integral** and multiply by 2 since $U = 2U_e$. Remember to mark **Integration Settings>Compute volume integral**. The surface resistance is given by $R_S = \sqrt{\dfrac{\omega \mu_c \mu_0}{2\sigma_c}}$ where σ_c is the conductivity and μ_c the relative permeability of the metal.

The other option is to use the impedance boundary condition. We then add the material of the metal in **Material**. We let the boundary of the half-circle have this material. We then choose **Impedance boundary condition** under **Electromagnetic fields**. The problem is now non-linear in the sense that the boundary condition depends on the eigenfrequency. Comsol can handle this if we add a linearization point. We right click on **Eigenvalue solver** and write our guess of frequency, in this case,s 1e9 in the box below transform point. With that value Comsol can linearize the problem and get an eigenvalue. For this application the value of the linearization point is not crucial. We let Comsol solve the problem. The Q-value is a pre-defined quantity that we find if we go to **Results>Derived values>Global evaluation**, choose the frequency that we are interested in, and and choose **Quality factor**.

In figure 6.8 we see the electric and magnetic fields for the three lowest TE- and TM-resonances.

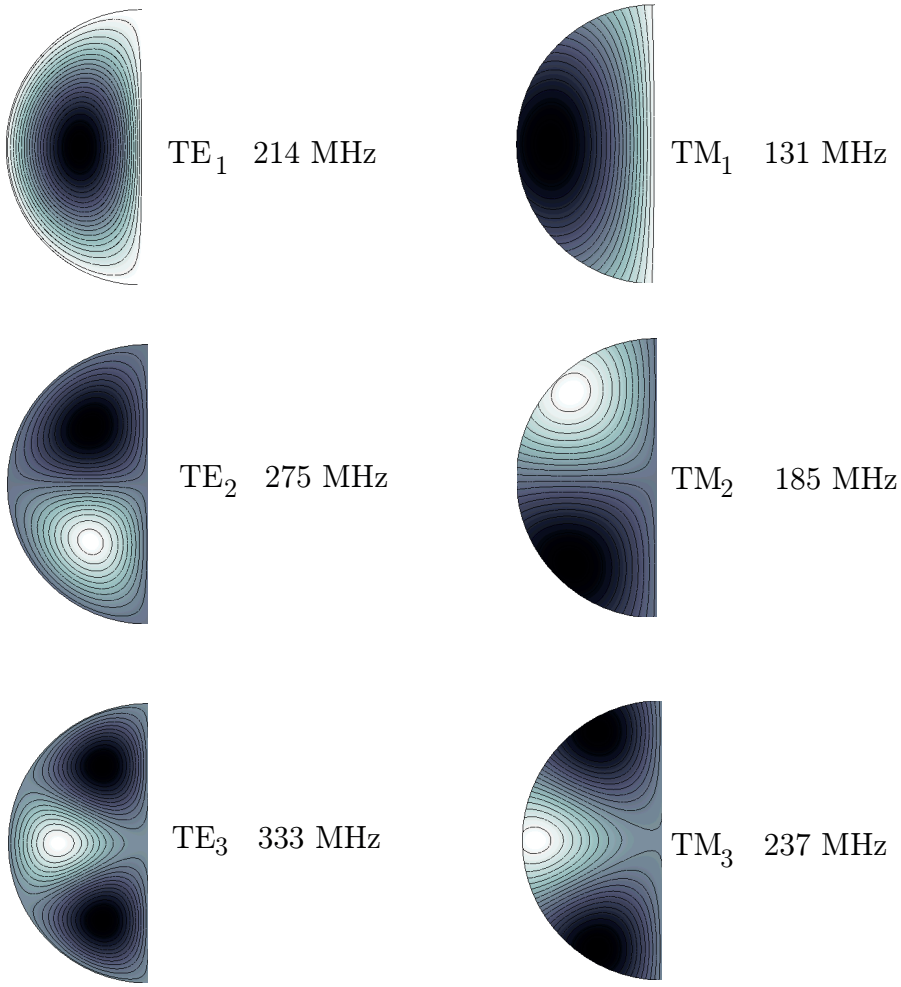


Figure 6.8: The electric field for the three lowest axisymmetric TE-modes and the magnetic field for the three lowest axisymmetric TM-modes in a spherical cavity with radius 1 m. The electric field of the TE-modes and the magnetic field of the TM-modes are directed in the $\hat{\phi}$-direction.

Example 6.5
In order to obtain also the non-axisymmetric modes we choose $m > 0$ in **Electromagnetic waves** .

6.5 Excitation of modes in a cavity

The electromagnetic fields in a cavity are excited by currents. These currents are generally fed via a waveguide or coaxial cable to the cavity, c.f., figure 6.9. The current often has a specified frequency in order to excite only one of the cavity modes. In an accelerator the beam is also a current source and it excites a large number of cavity modes since the

Table 6.2: The 10 lowest resonance frequencies in a spherical vacuum cavity with radius $a = 1$ m and perfectly conducting surface. The table shows the significant figures for three different meshes in Comsol. All values are in MHz. The 2-D axisymmetric solver has been used. The exact frequencies are given in table 6.1

order	normal (267 elements)	fine (859 elements)	extremely fine (11188 elements)
1	130.911	130.9117	130.911744
2	184.66	184.662	184.662441
3	214.39	214.396	214.3961
4	237.30	237.299	237.29905
5	275.00	274.995	274.9945
6	289.24	289.237	289.23653
7	291.8	291.85	291.85194
8	333.4	333.42	333.4184
9	340.6	340.68	340.6849
10	355.1	355.14	355.1354

beam has a wide frequency spectrum. We first examine the excitation problem in the frequency domain, where the complex current density $\boldsymbol{J}(\boldsymbol{r}, \omega)$ is the Fourier transform of the time dependent current density. We proceed by deriving the basic equations in the time domain. The irrotational modes only exist while the current is on. When the source is turned off only the solenoidal mode remain.

Sources in a lossless cavity

The Maxwell equations read

$$\begin{aligned} \nabla \times \boldsymbol{E} &= \mathrm{i}\omega\mu_0 \boldsymbol{H} \\ \nabla \times \boldsymbol{H} &= \boldsymbol{J} - \mathrm{i}\omega\epsilon_0 \boldsymbol{E} \end{aligned}$$

The following identity is used

$$0 = \iiint_V \nabla \cdot (\boldsymbol{E}_n \times \boldsymbol{H}) \,\mathrm{d}V = \iiint_V \boldsymbol{H} \cdot \nabla \times \boldsymbol{E}_n - \boldsymbol{E}_n \cdot \nabla \times \boldsymbol{H} \,\mathrm{d}V \tag{6.54}$$

We utilize the frequency domain version of the expansion in Eq. (6.3) and expand the electric and magnetic fields

$$\begin{aligned} \boldsymbol{E} &= \sum_{n=0}^{\infty} e_n(\omega) \boldsymbol{E}_n(\boldsymbol{r}) \\ \boldsymbol{H} &= (\mathrm{i}\omega\mu_0)^{-1} \nabla \times \boldsymbol{E} = (\mathrm{i}\omega\mu_0)^{-1} \sum_{n=0}^{\infty} e_n(\omega) \nabla \times \boldsymbol{E}_n(\boldsymbol{r}) \end{aligned} \tag{6.55}$$

where $e_n(\omega)$ and $f_n(\omega)$ are the Fourier transforms of $e_n(t)$ and $f_n(t)$, and where we used $\nabla \times \boldsymbol{F}_n = \boldsymbol{0}$. The rotation of the magnetic field is given by

$$\nabla \times \boldsymbol{H} = (\mathrm{i}\omega\mu_0)^{-1} \sum_{n=0}^{\infty} e_n(\omega) \nabla \times (\nabla \times \boldsymbol{E}_n(\boldsymbol{r})) = (\mathrm{i}\omega\mu_0)^{-1} \sum_{n=0}^{\infty} e_n(\omega) k_n^2 \boldsymbol{E}_n(\boldsymbol{r})) \tag{6.56}$$

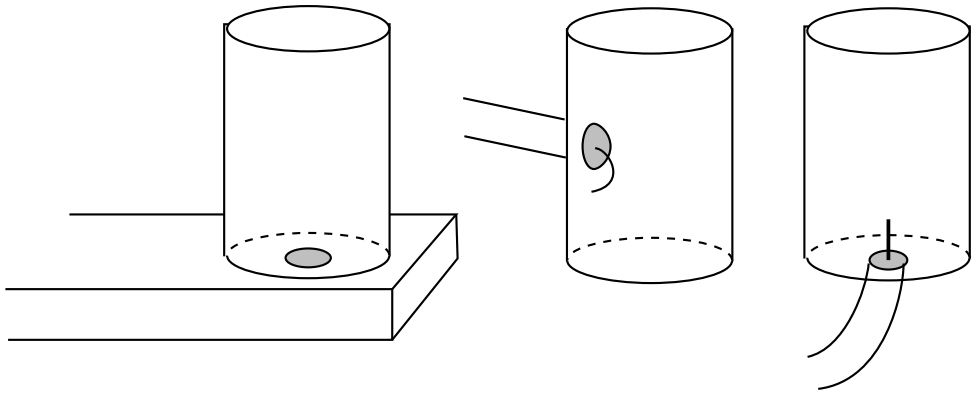


Figure 6.9: Three examples of feeding systems for microwave cavities. The left cavity is fed by a waveguide through a hole in the bottom surface. The waveguide in the middle is fed by a coaxial cable that is attached to the wall of the cavity. The cavity to the right is fed by a coaxial cable where the inner conductor continuous vertically into the cavity.

where we used $\nabla \times (\nabla \times \boldsymbol{E}_n) = -\nabla^2 \boldsymbol{E}_n = k_n^2 \boldsymbol{E}_n$. The Ampére law gives

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} - \mathrm{i}\omega\epsilon_0 \left(\sum_{n=0}^{\infty} e_n \boldsymbol{E}_n(\boldsymbol{r}) + \sum_{n=0}^{\infty} f_n \boldsymbol{F}_n(\boldsymbol{r}) \right) \tag{6.57}$$

We first determine the equation for the coefficients e_n. We insert the expansions in (6.55) and (6.56) into (6.57), take the scalar product with $\boldsymbol{E}_n$, integrate over V, and use the orthogonality of the vector functions (6.1). Thus

$$0 = -\mathrm{i}\frac{k_n^2}{\omega\mu_0} e_n(\omega) - \iiint_V \boldsymbol{J} \cdot \boldsymbol{E}_n \, \mathrm{d}V + \mathrm{i}\omega\epsilon_0 e_n(\omega)$$

and the amplitude of the electric field follows

$$e_n(\omega) = -\mathrm{i}\omega\mu_0 \frac{\iiint_V \boldsymbol{J} \cdot \boldsymbol{E}_n \, \mathrm{d}V}{k^2 - k_n^2} \tag{6.58}$$

We then determine an expression for f_n by taking the scalar product of (6.57) and $\boldsymbol{F}_n$, integrate over V and use the orthogonality (6.1). Then

$$f_n(\omega) = -\frac{\mathrm{i}}{\omega\epsilon_0} \iiint_V \boldsymbol{J} \cdot \boldsymbol{F}_n \, \mathrm{d}V$$

An alternative expression for $f_n(\omega)$ is obtained if we use the equation of continuity $\nabla \cdot \boldsymbol{J} = \mathrm{i}\omega\rho$, where ρ is the space charge density. We then use $\boldsymbol{F}_n = -\nabla\Phi_n$, $\nabla \cdot (\boldsymbol{J}\Phi) = \Phi\nabla \cdot \boldsymbol{J} + \boldsymbol{J} \cdot \nabla\Phi$, Gauss law and the boundary condition that $\Phi = 0$ on S. Eventually this leads to

$$f_n(\omega) = \frac{1}{\epsilon_0} \iiint_V \rho(\boldsymbol{r}, \omega) \Phi_n(\boldsymbol{r}) \, \mathrm{d}V \tag{6.59}$$

Time-domain equations

We obtain the time domain relation for $e_n(t)$ either by Fourier transformation of (6.58), or by an analysis entirely in the time domain. Since multiplication with $-\mathrm{i}\omega$ in the frequency domain corresponds to a time-derivative in the time-domain we obtain the ordinary differential equation for the amplitudes $e_n(t)$ from (6.58)

$$\frac{\partial^2 e_n(t)}{\partial t^2} + \omega_n^2 e_n(t) = -\epsilon_0^{-1} \frac{\partial}{\partial t} \iiint_V \boldsymbol{J}(\boldsymbol{r}, t) \cdot \boldsymbol{E}_n(\boldsymbol{r}) \, \mathrm{d}V \tag{6.60}$$

Sources in a cavity with losses

When there is dissipation of power in the cavity we need to modify the basic expressions above. The coupling of power to the beam introduces dissipation. Power may also escape the cavity through the power coupler of the cavity, through waveguides that are designed to attenuate the higher order modes, as Ohmic power losses in the walls of the cavity, or as power fed back to the beam. We modify the equation for $e_n(t)$ above by adding a dissipation term to (6.60)

$$\frac{\partial^2 e_n(t)}{\partial t^2} + 2\alpha_n \frac{\partial e_n(t)}{\partial t} + \omega_n^2 e_n(t) = -\epsilon_0^{-1} \frac{\partial}{\partial t} \iiint_V \boldsymbol{J}(\boldsymbol{r}, t) \cdot \boldsymbol{E}_n(\boldsymbol{r}) \, \mathrm{d}V \tag{6.61}$$

This leads us to the relation in the frequency domain

$$e_n(\omega) = -\mathrm{i}\omega\mu_0 \frac{\iiint_V \boldsymbol{J}(\boldsymbol{r}, \omega) \cdot \boldsymbol{E}_n(\boldsymbol{r}) \, \mathrm{d}V}{k^2 + 2\mathrm{i}k\alpha_n c^{-1} - k_n^2} \tag{6.62}$$

where $k_n = \omega_n / c$.

6.5.1 Excitation of modes in cavities for accelerators

The electromagnetic fields in a cavity for an accelerator are generated by the feed and by the beam of particles. The feed is usually a coaxial waveguide attached to the wall of the cavity and it feeds the cavity with a field of a given frequency, c.f., figure 5.3. The frequency is close to the mode that accelerates the particles.

The beam generates a wide spectrum of frequencies, and by that excites a large number of cavity modes, as it travels through the cavity . These modes are referred to as higher order modes (HOM). Higher order modes with frequencies close to a multiple of the frequency of the fundamental mode can affect the stability of the beam and give rise to power losses. One often try to prevent this when the cavity is designed.

In this section we analyze the generation of modes in an axisymmetric cavity from a beam that is centered at the axis of symmetry. The axisymmetric cavities are very common in accelerators, and the accelerators at Maxlab and at ESS both use axisymmetric cavities. We will exemplify by considering the generation of higher order modes in the axisymmetric superconducting cavities that are to be used in ESS. The cavities of ESS are referred to as elliptic cavities. The beam of protons in ESS will be 2 ms long and consist of bunches, each having a length of 20-30 ps. The bunches come with a frequency of 352 MHz. It means that the beam consists of 704.000 bunches. Each bunch has a total charge of 0.14 nAs. The beams come with a frequency of 20 Hz, i.e., with a period of 50 ms.

The beam is equivalent to a current density

$$\boldsymbol{J}(\boldsymbol{r},t) \approx \begin{cases} J(r_c,z,t)\hat{\boldsymbol{z}} & r_c < r_0 \\ \boldsymbol{0} & r_c > r_0 \end{cases}$$

and a charge density

$$\rho(\boldsymbol{r},t) \approx \begin{cases} v^{-1}J(r_c,z,t) & r_c < r_0 \\ \boldsymbol{0} & r_c > r_0 \end{cases}$$

We now assume that r_0 is very small such that

$$\boldsymbol{J}(r_c,z,t) \approx \frac{I(z,t)}{2\pi r_c}\delta(r_c)\hat{\boldsymbol{z}}$$
$$\rho(r_c,z,t) \approx \frac{I(z,t)}{2\pi r_c v}\delta(r_c)$$

where $\delta(r_c)$ is the delta function.The current $I(z,t)$ consists of protons traveling with velocity $v\hat{\boldsymbol{z}}$. Then

$$I(z,t) = I(t - z/v)$$

The Fourier transform of $I(t)$ is denoted $I(\omega)$, and hence the Fourier transform of $I(t-z/v)$ is $e^{ik_v z}I(\omega)$, where $k_v = \omega/v$. Notice that $I(t)$ is the current measured at the entrance $z = 0$ of the cavity.

Comment: The speed v is assumed to be constant. This is an approximation since the charges are accelerated.

6.5.2 A single bunch

In the LINAC of ESS the beam consists of a train of very short bunches. The bunch length ~ 20 ps is short enough to approximate the current and the charge density of the bunch with delta pulses

$$\begin{aligned} I(t - z/v) &\approx q\delta(t - z/v) \\ \rho(t - z/v) &\approx \frac{q}{2\pi r_c v}\delta(t - z/v)\delta(r_c) \end{aligned} \tag{6.63}$$

where q is the charge in a bunch. It means that the current density has boiled down to a train of point charges q moving the speed v along the axis of symmetry of the cavity.

Q-factor and losses

When there are losses in the cavity a loss parameter α_n is introduced for each mode. It is related to the $Q-$value of the cavity as, c.f., (6.40)

$$\alpha_n = \frac{\omega_n}{2Q}$$

The $Q-$values of the superconducting cavities are on the order $10^7 - 10^9$, where 10^7 is a typical value for a cavity with beam and 10^9 for an empty cavity. When the excitation of the HOM is to be determined one must use the $Q-$value for an empty cavity. It means that $\alpha < 10$ for the HOM. The beams come with a frequency of 20 Hz, i.e., separated with 50 ms, which means that the attenuation of the amplitude between the beams is no more than $e^{-\alpha 0.05} \approx 0.6$. Hence there will be a large amount of HOM energy in the cavity when a new beam arrives. The $Q-$value for the cavity with the beam is only relevant for the power feed of the cavity.

The equation for the amplitudes in a cavity with losses

When losses are present, eg., by losses in the walls, equation (6.61) for $e_n(t)$ is altered to

$$\frac{\partial^2 e_n(t)}{\partial t^2} + 2\alpha_n \frac{\partial e_n(t)}{\partial t} + \omega_n^2 e_n(t) = -\epsilon_0^{-1} \frac{\partial}{\partial t} \int_0^L I(z,t) E_n(z)\, dz \tag{6.64}$$

where we introduce the notation $E_n(z) = \hat{\boldsymbol{z}} \cdot \boldsymbol{E}_n(0,0,z)$. A very short bunch with charge q, c.f., (6.63) simplifies the right hand side such that

$$\frac{\partial^2 e_n(t)}{\partial t^2} + 2\alpha_n \frac{\partial e_n(t)}{\partial t} + \omega_n^2 e_n(t) = -\frac{qv^2}{\epsilon_0} E_n'(vt)(H(vt) - H(vt - L)) \tag{6.65}$$

where $E_n'(z)$ is the derivative w.r.t. z and $H(vt)$ is the Heaviside step function. The initial conditions to the equation are $e_n(0) = 0$ and $e_n'(0) = 0$. For $t > L/v$ the equation is homogenous with solution $e_n(t) = A_n e^{-\alpha_n t} \cos(\omega_n' t + \phi_n)$ where $\omega_n' = \sqrt{\omega_n^2 + \alpha_n^2}$. Notice that ω_n is $2Q$ times α_n and since Q is on the order of 10^6 or more, the approximation $\omega_n' = \omega_n$ is relevant and will be used hereafter.

A Fourier transformation of the ODE gives the Fourier transform $e_n(\omega)$ as

$$\begin{aligned} e_n(\omega) &= \frac{qv}{\epsilon_0} \frac{\int_0^L e^{\mathrm{i}k_v z} E_n'(z)\, dz}{(\omega^2 + 2\mathrm{i}\omega\alpha_n - \omega_n^2)} \\ &= -i\frac{q\omega}{\epsilon_0} \frac{\int_0^L e^{\mathrm{i}k_v z} E_n(z)\, dz}{(\omega^2 + 2\mathrm{i}\omega\alpha_n - \omega_n^2)} \end{aligned} \tag{6.66}$$

where $k_v = \omega/v$.

When the bunch has left the cavity the amplitude is given by the solution to the homogeneous equation

$$e_n(t) = A_n e^{-\alpha_n t} \cos(\omega_n t + \phi_n) \tag{6.67}$$

We obtain the amplitude A_n and phase ϕ_n from the inverse Fourier transform of (6.66) and residue calculus. This is done in subsection 6.5.4. The function $e_n(\omega)$ has poles at $\omega = -\mathrm{i}\alpha_n \pm \omega_n$ and this implies that, c.f., subsection 6.5.4,

$$A_n = q\frac{\sqrt{\omega_n^2 + \alpha_n^2}}{\omega_n \epsilon_0} \left| \int_0^L e^{\mathrm{i}(\omega_n - \mathrm{i}\alpha_n)z/v} E_n(z)\, dz \right|$$

and

$$\phi_n = -\arg\left\{ q\frac{\omega_n - \mathrm{i}\alpha_n}{\omega_n \epsilon_0} \int_0^L e^{\mathrm{i}(\omega_n - \mathrm{i}\alpha_n)z/v} E_n(z)\, dz \right\}$$

Since $\alpha_n \ll \omega_n$ a relevant approximation is

$$A_n = \frac{q}{\epsilon_0} \left| \int_0^L e^{\mathrm{i}\omega_n z/v} E_n(z)\, dz \right|$$

and

$$\phi_n = -\arg\left\{ \int_0^L e^{\mathrm{i}\omega_n z/v} E_n(z)\, dz \right\}$$

6.5.3 A train of bunches

A pulse in the LINAC of ESS consists of a train of bunches. Typically the pulse lasts for 2 ms and the time between the bunches is $\Delta t = 1/(352 \cdot 10^6)$ s, since the frequency of the cavity that creates the bunches is 352 MHz.

Let time $t = 0$ when the first bunch enters the cavity and assume that there is no field in the cavity for $t < 0$. When M bunches have passed the cavity the total amplitude of mode number n is

$$e_{n\text{tot}}(t) = \sum_{m=0}^{M-1} e_n(t - m\Delta t) \tag{6.68}$$

The Fourier transform is given by

$$\begin{aligned} e_{n\text{tot}}(\omega) &= e_n(\omega) \sum_{m=0}^{M-1} e^{\mathrm{i}m\omega\Delta t} = e_n(\omega) \frac{e^{\mathrm{i}\omega M\Delta t} - 1}{e^{\mathrm{i}\omega\Delta t} - 1} \\ &= e_n(\omega) e^{\mathrm{i}\omega(M-1)\Delta t/2} \frac{\sin(\omega M\Delta t/2)}{\sin(\omega\Delta t/2)} \end{aligned} \tag{6.69}$$

We introduce the bunch train factor

$$F(\omega, M) = e^{\mathrm{i}\omega(M-1)\Delta t/2} \frac{\sin(M\omega\Delta t/2)}{\sin(\omega\Delta t/2)}$$

(6.68) and (6.69) are sufficient in order to obtain the amplitude in the time domain of a cavity mode for all times and its corresponding Fourier transform. For a fixed frequency the Fourier transformation of the amplification is a periodic function of M. The period depends strongly on the product $\omega\Delta t$. Assume an angular frequency

$$\omega = \frac{2\pi}{\Delta t} m + \Delta\omega$$

where m is an integer such that $0 < \Delta\omega < 2\pi$. The absolute value of the bunch train factor is given by

$$|F(\omega, M)| = \frac{|\sin(M\Delta\omega\Delta t/2)|}{|\sin(\Delta\omega\Delta t/2)|}$$

This is a periodic function of M and the period is $M_0 = \dfrac{2\pi}{\Delta\omega\Delta t}$. In the next section we will show that the same periodicity holds in the time domain. We notice that the shortest period is $M_0 = 2$ and is obtained when $\Delta w = 2\pi \cdot 352 \cdot 10^6/2 = \pi/\Delta t$.

6.5.4 Amplitude in time domain

Each pulse of the beam consists of more than 700.000 bunches. Hence the summation in (6.68) gets numerically cumbersome as time increases. There is a way to get around this problem. When the beam has left the cavity the amplitude of the different cavity modes can be obtained by finding the inverse Fourier transform of (6.66) by residue calculus. The function $e_{n\text{tot}}(\omega)$ has poles at $\omega = -\mathrm{i}\alpha_n \pm \omega_n$. Thus

$$e_{n\text{tot}}(\omega) = -\frac{\mathrm{i}\omega q}{2\omega_n\epsilon_0} \int_0^L e^{\mathrm{i}\omega z/v} E_n(z)\, dz \left(\frac{1}{\omega + \mathrm{i}\alpha_n - \omega_n} - \frac{1}{\omega + \mathrm{i}\alpha_n + \omega_n} \right) \frac{e^{\mathrm{i}\omega M\Delta t} - 1}{e^{\mathrm{i}\omega\Delta t} - 1}$$

The Fourier integral contour from $-\infty$ to ∞ can be closed by a half circle in the upper half plane for negative times, which gives $e_{n\mathrm{tot}}(t) = 0$ for $t < 0$ since there are no poles in the upper half plane. For times $t > L/v + (M-1)\Delta t$ it can be closed by a half circle in the lower half plane and then

$$\begin{aligned} e_{n\mathrm{tot}}(t) &= \mathrm{i}\left(\mathrm{Res}\{e_{n\mathrm{tot}}e^{-\mathrm{i}\omega t}, \omega_n - \mathrm{i}\alpha_n\} - \mathrm{Res}\{e_{n\mathrm{tot}}e^{-\mathrm{i}\omega t}, -\omega_n - \mathrm{i}\alpha_n\}\right) \\ &= 2\mathrm{Re}\left\{\mathrm{i}\mathrm{Res}\{e_{n\mathrm{tot}}e^{-\mathrm{i}\omega t}, \omega_n - \mathrm{i}\alpha_n\}\right\} \end{aligned}$$

where Res stands for residue. For $t > L/v + (M-1)\Delta t$, i.e., when the beam has left the cavity, the amplitude of cavity mode number n is given by

$$e_{n\mathrm{tot}}(t) = 2q\mathrm{Re}\left\{\frac{\omega_n - \mathrm{i}\alpha_n}{2\omega_n\epsilon_0}e^{-\alpha_n t}e^{-\mathrm{i}\omega_n t}\int_0^L e^{\mathrm{i}(\omega_n - \mathrm{i}\alpha_n)z/v}E_n(z)\,dz\frac{e^{\mathrm{i}(\omega_n-\mathrm{i}\alpha_n)M\Delta t}-1}{e^{\mathrm{i}(\omega_n-\mathrm{i}\alpha_n)\Delta t}-1}\right\} \tag{6.70}$$

Thus

$$e_{n\mathrm{tot}}(t) = A_{n\mathrm{tot}}\cos(\omega_n t + \phi_n)e^{-\alpha_n t} \tag{6.71}$$

where the amplitude is given by

$$A_{n\mathrm{tot}} = \left|2q\frac{\omega_n - \mathrm{i}\alpha_n}{2\omega_n\epsilon_0}\int_0^L e^{\mathrm{i}(\omega_n-\mathrm{i}\alpha_n)z/v}E_n(z)\,dz\frac{e^{\mathrm{i}(\omega_n-\mathrm{i}\alpha_n)M\Delta t}-1}{e^{\mathrm{i}(\omega_n-\mathrm{i}\alpha_n)\Delta t}-1}\right| \tag{6.72}$$

and the phase is given by

$$\phi_n = -\arg\left\{2q\frac{\omega_n - \mathrm{i}\alpha_n}{2\omega_n\epsilon_0}\int_0^L e^{\mathrm{i}(\omega_n-\mathrm{i}\alpha_n)z/v}E_n(z)\,dz\frac{e^{\mathrm{i}(\omega_n-\mathrm{i}\alpha_n)M\Delta t}-1}{e^{\mathrm{i}(\omega_n-\mathrm{i}\alpha_n)\Delta t}-1}\right\} \tag{6.73}$$

Example 6.6

The fundamental theorem of beam loading: Consider a single bunch with charge q that enters the cavity. First assume that there is no electromagnetic fields in the cavity. When the bunch leaves the cavity it has generated electric fields $e_n(t)\boldsymbol{E}_n(\boldsymbol{r})$, and corresponding magnetic fields, in the different cavity modes. The corresponding energies are $W_{\mathrm{em}}(n,1)$, that are the sum of the electric and magnetic energies in mode number n. If attenuation is neglected these energies are constant until the next bunch arrives. These energies are taken from the kinetic energy of the bunch. Thus the bunch is decelerated by a force in the $z-$direction that can be written as $\boldsymbol{F}_1 = q\kappa e_n(t)E_n(z)\hat{\boldsymbol{z}}$, where κ is a constant.

To determine κ we let the next bunch enter exactly one period of mode number n after the first bunch. When this bunch has left the cavity it has added the same electric field to the cavity as the first bunch. Thus the stored energy of mode n is $4W_{\mathrm{em}}(n,1)$ since the energy is a quadratic quantity in the amplitudes. The second bunch has then delivered the energy $3W_{\mathrm{em}}(n,1)$ to the cavity. When it traveled through the cavity it experienced the field generated by the first bunch and the field generated by itself. The force is then $\boldsymbol{F}_2 = q(1+\kappa)e_n(t)E_n(z)$. The work done by the forces equals the energies that the two

bunches have stored, i.e.,

$$W_{\mathrm{em}}(n,1) = q\kappa \int_0^L e_n(z/v)E_z(z)\,dz$$

$$3W_{\mathrm{em}}(n,1) = q(1+\kappa) \int_0^L e_n(z/v)E_z(z)\,dz$$

Thus $3\kappa = 1 + \kappa$ and $\kappa = 0.5$. The bunch is apparently exposed to half of the field it generates plus the electric field that was present in the cavity when it arrived. This result is called the fundamental theorem of beam loading. ∎

Problems in Chapter 6

6.1 Show that the time averages of the stored electric and magnetic energies in a resonance cavity are equal.

6.2 Assume a resonance cavity with a certain length. Prove that if the length is scaled by a factor K then the Q-factor scales with a factor $\sqrt{K}$ if the conductivity of the walls is constant.

6.3 a) Determine Q for the TE_{101}-mode in a rectangular parallelpiped with walls $a \times b \times d$.

b) determine the resonance frequency and the Q-factor for the TE_{101}-mode when $a = b = 2$ cm, $d = 4$ cm and the cavity is made out of copper ($\sigma_c = 5.8 \cdot 10^7$ S/m).

c) Determine the resonance frequency and the Q-factor for the TE_{101}-mode if $a = b = 20$ cm, $d = 40$ cm and the cavity is made out of copper ($\sigma_c = 5.8 \cdot 10^7$ S/m).

6.4 Estimate the number of resonances with a wavelength larger than 500 nm in a cubic vacuum cavity with volume one cubic meter.

Summary of chapter 6

Resonance cavities

$$\alpha = \begin{cases} \dfrac{R_S}{\mu_0}\left(\dfrac{\varepsilon_\ell}{d} + \dfrac{1}{2k_{t_n}^2}\displaystyle\int_\Gamma |\hat{\boldsymbol{n}}\cdot\nabla_T v_n(\boldsymbol{\rho})|^2 d\Gamma\right) & \text{TM} \\ \dfrac{R_S}{\mu_0 k_{n\ell}^2}\left(\dfrac{2k_{z\ell}^{\,2}}{d} + \dfrac{1}{2}\displaystyle\int_\Gamma \dfrac{k_{z\ell}^{\,2}}{k_{t_n}^{\,2}}|\hat{\boldsymbol{n}}\times\nabla_T w_n(\boldsymbol{\rho})|^2 + k_{t_n}^2|w_n(\boldsymbol{\rho})|^2 d\Gamma\right) & \text{TE} \end{cases}$$

$$Q = \frac{\omega_{n\ell}}{2\alpha} \qquad B = 2\alpha$$

Chapter 7

Transients in waveguides

In chapter 5 we described propagation of time harmonic fields in hollow waveguides. There are applications where the fields in the waveguide are not time harmonic, but consist of a spectrum of frequencies. One way to analyze the propagation of signals with general time dependence is to Fourier transform the input signal and then determine the output signal by an inverse Fourier transform. In this chapter a time-domain method, based upon propagators, is presented. By a convolution of the input signal with the propagator, the signal can be determined at an arbitrary position in the waveguide. The propagator corresponds to the impulse response of the waveguide and is independent of the input signal, the geometry of the waveguide, and the mode number

Wave propagation of a *fixed* mode in a hollow waveguide is at a fixed frequency determined by

$$\begin{aligned} E_z(\boldsymbol{r},\omega) &= v(\boldsymbol{\rho})a(z,\omega) = v(\boldsymbol{\rho})A(\omega)e^{\mathrm{i}k_z(\omega)z} \quad \text{(TM-fallet)} \\ H_z(\boldsymbol{r},\omega) &= w(\boldsymbol{\rho})a(z,\omega) = w(\boldsymbol{\rho})A(\omega)e^{\mathrm{i}k_z(\omega)z} \quad \text{(TE-fallet)} \end{aligned} \tag{7.1}$$

where the longitudinal wavenumber is

$$k_z(\omega) = \left\{k^2(\omega) - k_t^2\right\}^{1/2} = \left\{\frac{\omega^2}{c_0^2} - k_t^2\right\}^{1/2}$$

Note that $v(\boldsymbol{\rho})$, $w(\boldsymbol{\rho})$, k_t^2 and c_0 are independent of the frequency[1]. Even the case with material dispersion can be handled (ϵ frequency dependent). We transform the expressions in equation (7.1) to the time domain in order to study the transient wave phenomena in the waveguide. The inverse Fourier transform of equation (7.1) gives the z-component of the fields in a point z

$$\begin{aligned} E_z(\boldsymbol{r},t) &= v(\boldsymbol{\rho})a(z,t) \\ H_z(\boldsymbol{r},t) &= w(\boldsymbol{\rho})a(z,t) \end{aligned}$$

where

$$a(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} a(z,\omega)e^{-\mathrm{i}\omega t}\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} A(\omega)e^{\mathrm{i}k_z(\omega)z}e^{-\mathrm{i}\omega t}\,d\omega$$

The amplitude $A(\omega)$ is the Fourier transform of $a(0,t)$

$$A(\omega) = \int_{-\infty}^{\infty} a(0,t)e^{\mathrm{i}\omega t}\,\mathrm{d}t$$

[1]With a modification of the analysis in this chapter one can also handle cases where the speed of light c is not equal to speed of light in vacuum.

We add and subtract a term $\exp\{\mathrm{i}\omega z/c_0\}$, that is determined by the asymptotic behavior of the function $k_z(\omega)$ for high frequencies, i.e.,

$$a(z,\omega) = A(\omega)e^{\mathrm{i}\omega z/c_0} + A(\omega)e^{\mathrm{i}\omega z/c_0}\left(e^{\mathrm{i}k_z(\omega)z - i\omega z/c_0} - 1\right)$$

Since a product of two Fourier transforms gives a convolution in the time domain, and multiplication with a factor $\exp(-\mathrm{i}\omega t_0)$ gives a time delay t_0, the $z-$component of the electric field at a point z reads

$$a(z,t) = a(0, t - z/c_0) + \int_{-\infty}^{\infty} P(z, t - z/c_0 - t')a(0,t')\,dt' \tag{7.2}$$

where

$$\begin{cases} P(z, t - z/c_0) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}\left(e^{\mathrm{i}k_z(\omega)z - \mathrm{i}z\omega/c_0} - 1\right)e^{-\mathrm{i}\omega(t-z/c_0)}\,d\omega \\ a(0,t) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} A(\omega)e^{-\mathrm{i}\omega t}\,\mathrm{d}\omega \end{cases}$$

We call the function $P(z,t)$ the propagator kernel, since it maps the field at $z = 0$ on the field at a point z.

We use the following pair of Fourier transforms

$$\begin{cases} \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}\left(1 - e^{-\mathrm{i}b\omega + \mathrm{i}b\left(\omega^2 - a^2\right)^{1/2}}\right)e^{-\mathrm{i}\omega t}\,\mathrm{d}\omega = H(t)ab\dfrac{J_1\left(a\sqrt{t^2+2bt}\right)}{\sqrt{t^2+2bt}} \\ ab\displaystyle\int_0^{\infty}\dfrac{J_1\left(a\sqrt{t^2+2bt}\right)}{\sqrt{t^2+2bt}}e^{\mathrm{i}\omega t}\,dt = 1 - e^{-\mathrm{i}b\omega + \mathrm{i}b\left(\omega^2-a^2\right)^{1/2}} \end{cases}$$

where $H(t)$ is the Heaviside functionen ($H(t) = 1, t \geq 0$, otherwise zero) and furthermore $\mathrm{Im}(\omega^2 - a^2)^{1/2} > 0$, $\arg w \in [0, 2\pi]$. This gives us an explicit expression for $P(z,t)$

$$P(z,t) = -c_0 k_t z\frac{J_1\left(k_t\sqrt{c_0^2t^2 + 2zc_0t}\right)}{\sqrt{c_0^2t^2+2zc_0t}}H(t)$$

and an expression for the amplitude of the waveguide mode

$$a(z, t + z/c_0) = a(0,t) - c_0k_tz\int_{-\infty}^{t}\frac{J_1\left(k_t\sqrt{c_0^2(t-t')^2 + 2c_0z(t-t')}\right)}{\sqrt{c_0^2(t-t')^2+2c_0z(t-t')}}a(0,t')\,dt' \tag{7.3}$$

The parameter t is the time after the wavefront has passed and is called the wave front time (or equivalently retarded time). We notice that the wavefront is moving with the speed c_0.

We see that the coordinates

$$\begin{cases} \zeta = k_t z \\ s = k_t c_0 t \end{cases} \tag{7.4}$$

are suitable dimensionless parameters to describe the wave propagation in the waveguide. With these dimensionless coordinates all modes are propagating as

$$u(\zeta, s+\zeta) = u(0,s) + \int_{-\infty}^{s} P(\zeta, s - s')u(0,s')\,ds'$$

$$P(\zeta, s) = -\zeta\frac{J_1\left(\sqrt{s^2+2\zeta s}\right)}{\sqrt{s^2+2\zeta s}}H(s)$$

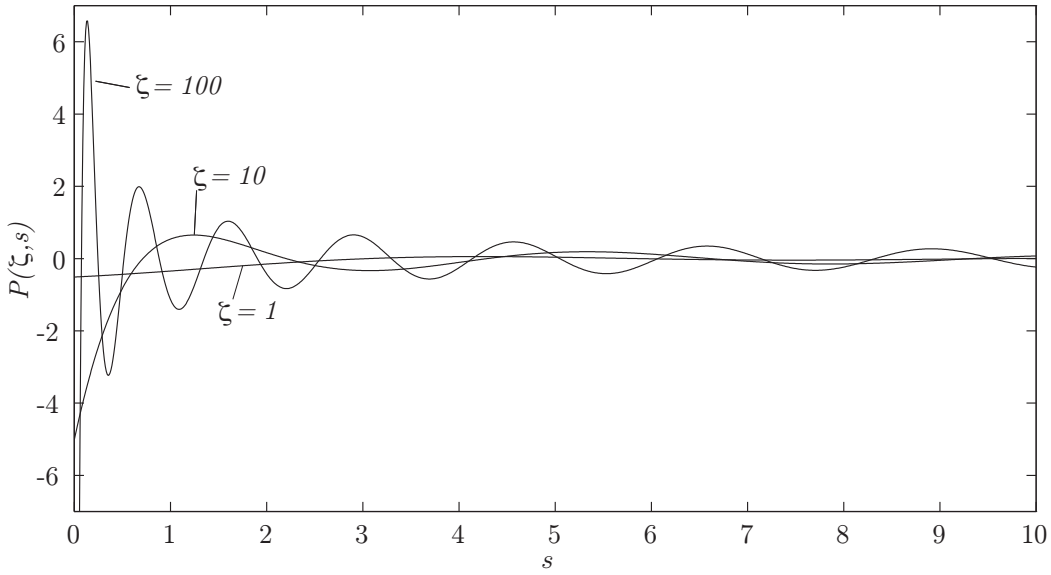


Figure 7.1: The propagator kernel $P(\zeta, s)$ at three different positions ζ. Note that when ζ increases the propagator kernel is compressed and increases in amplitude. One can prove that in the limit $\zeta \to \infty$ it holds that $P(\zeta, s) \to -\delta(s)$.

The different modes only differ by a scaling in time and space given by equation (7.4). In figure 7.1 the propagator kernel $P(\zeta, s)$ is depicted at three different positions ζ.

Example 7.1
If $a(0,t) = \delta(t)$ then

$$a(z, t + z/c_0) = \delta(t) + P(z,t)$$

Hence the wavefront is a delta pulse that moves with the vacuum speed of light and is followed by a tail that is given by $P(z,t)$. In figure 7.1 we see that the oscillations of the tail increases when the wave propagates. ∎

Example 7.2
The expression (7.3) is also valid for negative z values. The propagator kernel $P(z,t)$ can then be used for reconstruction of the pulse at negative z values. One can also use it in order to determine the input signal in order to create a pulse of a given shape, see problem 2. When $t < -2z/c_0$ the square root in the propagator kernel is imaginary and can be written $\sqrt{c_0^2 t^2 + 2c_0 zt} = i\sqrt{-2c_0 zt - c_0^2 t^2}$. One can show that in this case

$$P(z,t) = -c_0 k_t z \frac{I_1(k_t \sqrt{-2c_0 zt - c_0^2 t^2})}{\sqrt{-2c_0 zt - c_0^2 t^2}} H(t)$$

where $I_1(y) = -iJ_1(iy)$ is the modified Bessel functionen. In figure 7.2 we see the graph of $P(\zeta, s)$ for some negative values of ζ. ∎

Problems in Chapter 7

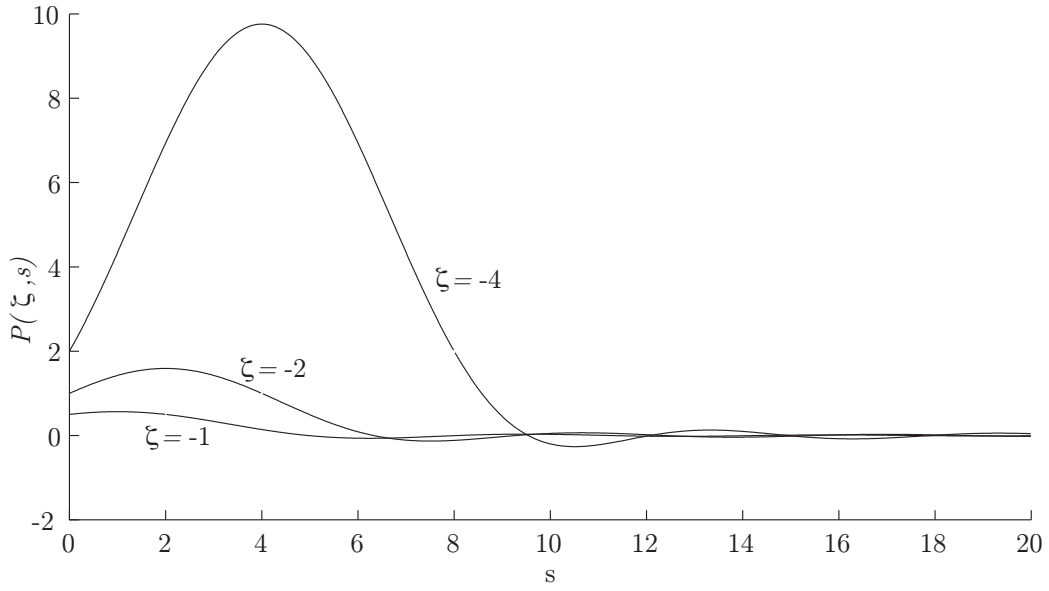


Figure 7.2: The propagator kernel $P(\zeta, s)$ for three values of ζ.

7.1 When a pulse is propagating in a dispersive material of infinite extent a wave phenomena referred to as a precursor is developed. This is the part of the signal that comes directly after the wavefront. The wave equation for a dispersive material is given by

$$\frac{\partial^2 E(z,t)}{\partial z^2} - \frac{1}{c_0^2}\frac{\partial^2 E(z,t)}{\partial t^2} - \frac{1}{c_0^2}\frac{\partial^2}{\partial t^2}\int_{-\infty}^{t} \chi(t-t')E(z,t')\,dt' = 0$$

where $\chi(t)$ is the susceptibilty kernel for the dispersive material and E is a component of the electric field. The susceptibility kernel relates the displacement field (electric density flow) to the electric field through the constitutive relation

$$D(z,t) = \varepsilon_0(E(z,t) + \int_{-\infty}^{t} \chi(t-t')E(z,t')\,dt')$$

For frequencies in the optical regime the Lorentz' model is appropriate for most materials. If the material is lossless this model gives

$$\chi(t) = \omega_p^2 \sin(\omega_0 t)/\omega_0\, H(t)$$

where ω_p is called the plasma frequency, ω_0 is th resonance frequency for the bounded electrons and $H(t)$ is the Heaviside function. If the precursor is defined as the part of the signal that arrives in the time interval $\Delta t << 1/\omega_0$ directly after the wave front, then show that the precursor in a dispersive material satifies the same equation as the transient field in the waveguide.

7.2 A pulse with a mode with transverse wavenumber k_t is sent through a waveguide of length L. We wish that the pulse is a square pulse when it leaves the waveguide,

i.e.,

$$a(L,t) = H(t) - H(t - t_0)$$

Determine the shape of the input signal. The answer may include an integral with known integrand.

7.3 Show that the propagator kernel satisfies the equation

$$P(z_1 + z_2, t) = P(z_1, t) + P(z_2, t) + \int_0^t P(z_1, t - t')P(z_2, t')\, dt'$$

Summary of chapter 7

The propagator

$$E_z(\boldsymbol{r}, t) = v(\boldsymbol{\rho})a(z,t)$$

$$a(z,t) = a(0, t - z/c_0) + \int_{-\infty}^{\infty} P(z, t - z/c_0 - t')a(0,t')\, dt'$$

$$P(z,t) = -c_0 k_t z \frac{J_1\left(k_t\sqrt{c_0^2t^2 + 2zc_0t}\right)}{\sqrt{c_0^2t^2 + 2zc_0t}} H(t)$$

Chapter 8

Dielectric waveguides

In this chapter we analyze dielectric waveguides. In these waveguides the waves are guided in a region with a dielectric material that has lower wave speed, and hence larger index of refraction, than the surrounding regions. From Snell's law we know that a plane wave that impinges on a plane interface between two dielectric materials may be totally reflected. For this to happen two criteria have to be fulfilled. The first is that the index of refraction in the region the wave comes from is larger than in the region it is reflected from. The other is that the angle of incidence is larger than the critical angle $\arcsin(n_2/n_1)$. It is this phenomenon that makes it possible for waves to be trapped in a dielectric plate or cylinder.

A large part of the chapter is devoted to optical fibers. The development of optical fibers started in 1966 when Charles Kao[1] realized that glass could be purified enough to carry light signals with an attenuation of less than 20 dB/km. It took some time before this limit was reached, but in 1970 three researchers, Robert Maurer, Donald Keck and Peter Schultz at Corning Glass Works managed to purify silica such that the attenuation of the waves in a fiber was only 17 dB/km. Five years later they reached 4 dB/km. The optical fiber developed at Corning Glass Works is the same type as is used in todays optical fiber systems. It is a circular dielectric cylinder made out of silica with a thin circular region, referred to as the core, doped with a substance, often germanium dioxide, that increases the index of refraction. The waves are then confined to the core. In the standard optical fiber the attenuation is minimal at the wavelength $\lambda = 1.55\,\mu$m, which is in the infrared region. For this reason most systems use this wavelength. For long distance communication the single mode fiber is used. It means that the radius of the core is so small that only one mode can exist (radius$\sim 10\,\mu$m). The advantage with a single mode fiber is that the dispersion is small. For communication at short distances one can use multimode fibers. These are fibers with a thick core where a large number of modes can propagate. The multi mode fiber systems are cheaper to manufacture than the single mode systems.

The bandwidth of an optical fiber is much larger than for an electric wire. In todays systems one can transmit 40 Gb/s in a single fiber. Optical fibers save space in cable ducts, they are immune to electrical interference and they are very hard to wiretap. Thus optical fibers are superior to electric wires for long distance communication. In 2014 a danish group at the Danish Technical University (DTU) set a new world record by transferring 43 terabit/s using one laser and one fiber.

[1] Kao was awarded the 2009 Nobel prize in physics for his discovery

8.1 Planar dielectric waveguides

We first look at waves that propagate along dielectric plates. Consider a planar dielectric plate that is infinite in the $y-$ and $z-$directions and occupies the region $0 \leq x \leq d$. The plate is lossless and has a relative permittivity ϵ. Outside the plate there is vacuum. We restrict ourselves to time-harmonic wave propagation in the $z-$direction. The waves in the plate can be of TE-type, i.e., with the electric field in the $y-$direction, or of TM-type, with the electric field in the xz-plane. Here we only consider the TM-case and leave the TE-case as a problem.

Inside the plate the electric field is a superposition of an upward and a downward traveling wave. The two waves can always be decomposed into a wave for which the $z-$component of the electric field is an even function of x w.r.t. $x = d/2$ and a wave for which it is an odd function. In the region $x > d$ there are only waves traveling upwards and in the region $x < 0$ there are only waves traveling downwards. We make the following ansatz for the odd modes

$$v(x) = \begin{cases} Ae^{\mathrm{i}k_{0x}(x-d)}, & x > d \\ B\sin(k_{1x}(x-d/2)), & 0 \leq x \leq d \\ -Ae^{-\mathrm{i}k_{0x}x}, & x < 0 \end{cases}$$

where $k_0 = \omega/c_0$ is the wave number in vacuum, $k_1 = k_0\sqrt{\epsilon}$ is the wavenumber in the plate and

$$k_{0x} = \sqrt{k_0^2 - k_z^2}$$
$$k_{1x} = \sqrt{k_1^2 - k_z^2}$$

We get relations between A, B, and k_z by using the boundary conditions at $x = 0$ and $x = d$. Since $v(x)$ is independent of y and antisymmetric it is enough to use the boundary conditions that E_z and D_x are continuous at $x = 0$. The boundary conditions at $x = d$ lead to identical relations.

$$v(0^-) = v(0^+)$$
$$\frac{1}{k_{0x}^2}\frac{\partial v(0^-)}{\partial x} = \frac{\epsilon_1}{k_{1x}^2}\frac{\partial v(0^+)}{\partial x}$$

This gives

$$A = B\sin(k_{1x}d/2)$$
$$\mathrm{i}k_{0x}A = \frac{\epsilon_1 k_{0x}^2}{k_{1x}^2}k_{1x}B\cos(k_{1x}d/2)$$

We combine the two equations and get the equation for non-trivial solutions

$$\tan(k_{1x}d/2) = -\mathrm{i}\frac{\epsilon_1 k_{0x}}{k_{1x}}$$

If k_{0x} has a real part power will be transported in the $x-$direction, i.e., there is a leakage of power, and the wave in the waveguide is attenuated in the $z-$direction. Thus the criteria

for a propagating mode is that k_{0x} is imaginary and k_z is real. Let $k_{0x} = \mathrm{i}q$. Then the equation for q reads

$$\tan(k_{1x}d/2) = \frac{\epsilon_1 q}{k_{1x}}$$
$$k_{1x} = \sqrt{k_1^2 - k_0^2 - q^2}$$

This equation is called a transcendental equation[2]. It has no explicit solutions and has to be solved numerically. For a given d and frequency there are only a finite number of solutions and hence a finite number of propagating modes. The fundamental mode has cut-off frequency zero. To see this we assume that there exists a solution with real q. When $\omega \to 0$ then $k_{1x} \to 0$ and $\tan(k_{1x}d/2) \to k_{1x}d/2$. The transcendental equation simplifies to

$$\frac{k_{1x}^2 d}{2} = \epsilon_1 q$$

or equivalently $k_1^2 - k_0^2 - q^2 = 2q\epsilon_1/d$ with solution

$$q = -\frac{\epsilon_1}{d} \pm \sqrt{\left(\frac{\epsilon_1}{d}\right)^2 + k_1^2 - k_0^2}$$

Only the solution with plus sign in front of the square root gives a propagating mode since the other solution implies an imaginary k_z. As $\omega \to 0$ we use

$$\sqrt{\left(\frac{\epsilon_1}{d}\right)^2 + k_1^2 - k_0^2} = \frac{\epsilon_1}{d}\sqrt{1 + \frac{(k_1^2 - k_0^2)d^2}{\epsilon_1^2}} \approx \frac{\epsilon_1}{d}\left(1 + \frac{(k_1^2 - k_0^2)d^2}{2\epsilon_1^2}\right)$$

Then

$$q \approx \frac{(k_1^2 - k_0^2)d}{2\epsilon_1} = \frac{\omega^2 d(\epsilon_1 - 1)}{2c_0^2\epsilon_1}$$

and

$$k_z = \sqrt{k_0^2 + q^2} \approx k_0\sqrt{1 + k_0^2\left(\frac{d(n_1^2 - 1)}{n_1^2}\right)^2}$$

As $k_0 d \to 0$ then $k_z \to k_0$. The wave is still bounded to the dielectric plate but it propagates with the phase speed of the surrounding medium and since q is small, the field decays slowly in the directions perpendicular to the plate. The even $TM-$mode and the odd and even $TE-$modes are treated in the same manner. The difference is that none of these modes has zero cut-off frequency.

It turns out that all dielectric waveguides have much in common. They all have one or two fundamental modes with zero cut-off frequency. As frequency goes to zero the fundamental mode propagates with the wave speed of the surrounding medium and k_z approaches k_0.

8.2 Cylindrical dielectric waveguides

We assume an infinitely long and straight dielectric cylinder, directed along the $z-$axis, with envelope surface S and cross section surface Ω. The boundary curve to Ω is Γ,

[2]A transcendental function is a function that is not a polynomial

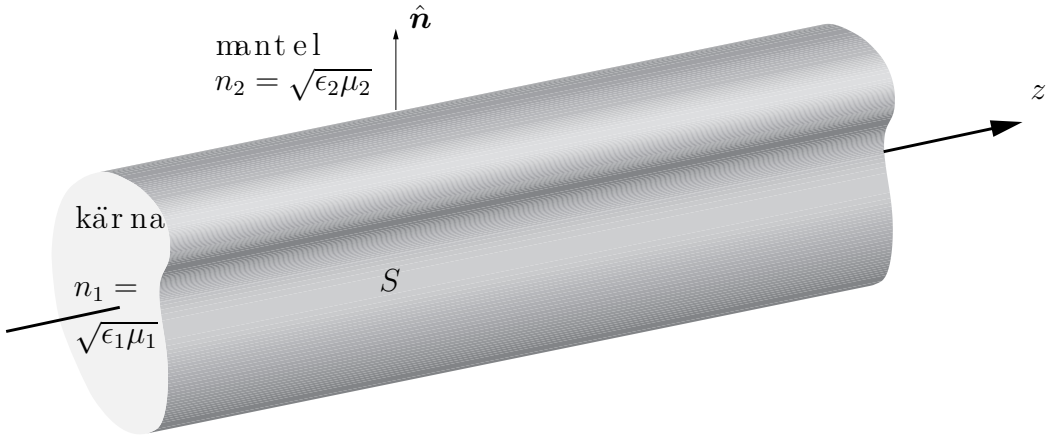


Figure 8.1: The dielectric waveguide with envelope surface S.

see figure 8.1. The cylinder has the relative permittivity ϵ_1 and permeability μ_1 and is surrounded by a region with parameters ϵ_2 and μ_2. The cylinder is denoted *core* and the outer region *cladding*. We assume that the core and cladding are lossless i.e., ϵ and μ are real. The index of refraction in the two regions are given by

$$n_1 = c_0/c_1 = \sqrt{\epsilon_1\mu_1}$$
$$n_2 = c_0/c_2 = \sqrt{\epsilon_2\mu_2}$$

The wavenumber in vacuum is $k_0 = \omega/c_0$, and the wavenumber in the core and cladding are

$$k_1 = \omega/c_1 = k_0 n_1$$
$$k_2 = \omega/c_2 = k_0 n_2$$

8.2.1 The electromagnetic fields

In chapter 4 we introduced a general decomposition of the electromagnetic fields, that is useful for the dielectric waveguides. Since the fields propagate along the waveguide, i.e., in the $z-$direction, we make the following ansatz:

$$\begin{cases} E_z(\boldsymbol{r},\omega) = v(\boldsymbol{\rho},\omega)e^{\mathrm{i}k_z(\omega)z} \\ H_z(\boldsymbol{r},\omega) = w(\boldsymbol{\rho},\omega)e^{\mathrm{i}k_z(\omega)z} \end{cases}$$

Note that we have utilized that k_z has to be the same everywhere in order to satisfy the boundary conditions. In contrast to the hollow waveguide, the functions v and w are frequency dependent. In the rest of the chapter the ω dependence is understood in v, w and k_z. The functions v and w are solutions to the eigenvalue problems

$$\begin{cases} \nabla_T^2 v(\boldsymbol{\rho}) + k_{1t}^2 v(\boldsymbol{\rho}) = 0 \\ \nabla_T^2 w_p(\boldsymbol{\rho}) + k_{1t}^2 w(\boldsymbol{\rho}) = 0 \end{cases} \quad \text{when } \boldsymbol{\rho} \text{ is inside } \Gamma, \tag{8.1}$$

$$\begin{cases} \nabla_T^2 v(\boldsymbol{\rho}) + k_{2t}^2 v(\boldsymbol{\rho}) = 0 \\ \nabla_T^2 w_p(\boldsymbol{\rho}) + k_{2t}^2 w(\boldsymbol{\rho}) = 0 \end{cases} \quad \text{when } \boldsymbol{\rho} \text{ is outside } \Gamma \tag{8.2}$$

The transverse wavenumbers k_{it} are related to the wavenumbers k_i and the longitudinal wavenumber k_z as

$$k_{it}^2 = k_i^2 - k_z^2, \quad i = 1, 2$$

Notice that the transverse wavenumber in the core differs from that in the cladding.

8.2.2 Boundary conditions

We need the boundary conditions at the surface S in order to determine $v(\boldsymbol{\rho})$, $w(\boldsymbol{\rho})$ and k_z. We first determine the boundary conditions for the electric and magnetic fields on the lateral surface S.

There are no surface currents and hence the boundary conditions on S is (c.f., equation (1.16) on page 7)

$$\begin{cases} \hat{\boldsymbol{n}} \times \boldsymbol{E}_1 = \hat{\boldsymbol{n}} \times \boldsymbol{E}_2 \\ \hat{\boldsymbol{n}} \times \boldsymbol{H}_1 = \hat{\boldsymbol{n}} \times \boldsymbol{H}_2 \end{cases} \qquad \boldsymbol{r} \text{ on } S$$

where $\hat{\boldsymbol{n}}$ is the unit normal vector.

In analogy with the analysis of hollow waveguides we introduce the decomposition of $\boldsymbol{E}$ and $\boldsymbol{H}$ fields in a longitudinal and transverse part. In the decomposed fields the boundary conditions are

$$\begin{cases} \hat{\boldsymbol{n}} \times (\boldsymbol{E}_T + \hat{\boldsymbol{z}} E_z) \\ \hat{\boldsymbol{n}} \times (\boldsymbol{H}_T + \hat{\boldsymbol{z}} H_z) \end{cases} \qquad \text{are continuous over } S$$

It follows that

$$\begin{cases} E_z \\ H_z \end{cases} \text{ and } \begin{cases} \hat{\boldsymbol{n}} \times \boldsymbol{E}_T \\ \hat{\boldsymbol{n}} \times \boldsymbol{H}_T \end{cases} \qquad \text{are continuous over } S$$

We use equation (4.7) to rewrite the transverse field components in the longitudinal components, and find that the following quantities are continuous over S:

$$\begin{cases} E_z \\ H_z \end{cases} \quad \text{and} \quad \begin{cases} \dfrac{1}{k_t^2}\left[k_z \hat{\boldsymbol{z}} \cdot (\hat{\boldsymbol{n}} \times \nabla_T E_z) - \omega\mu_0\mu(\omega)\dfrac{\partial H_z}{\partial n}\right] \\ \dfrac{1}{k_t^2}\left[k_z \hat{\boldsymbol{z}} \cdot (\hat{\boldsymbol{n}} \times \nabla_T H_z) + \omega\epsilon_0\epsilon(\omega)\dfrac{\partial E_z}{\partial n}\right] \end{cases}$$

Since $e^{\mathrm{i}k_z z}$ is a common factor for all of the terms, the continuity is also valid for:

$$\begin{cases} v(\boldsymbol{\rho}) \\ w(\boldsymbol{\rho}) \end{cases} \quad \text{and} \quad \begin{cases} \dfrac{1}{k_t^2}\left[k_z \hat{\boldsymbol{z}} \cdot (\hat{\boldsymbol{n}} \times \nabla_T v(\boldsymbol{\rho})) - \omega\mu_0\mu(\omega)\dfrac{\partial w(\boldsymbol{\rho})}{\partial n}\right] \\ \dfrac{1}{k_t^2}\left[k_z \hat{\boldsymbol{z}} \cdot (\hat{\boldsymbol{n}} \times \nabla_T w(\boldsymbol{\rho})) + \omega\epsilon_0\epsilon(\omega)\dfrac{\partial v(\boldsymbol{\rho})}{\partial n}\right] \end{cases} \tag{8.3}$$

In addition to this system of equations we need the conditions that the fields are finite everywhere, in particular at $\rho = \infty$ and inside the core.

8.3 Circular dielectric waveguide

We now specialize to a circular dielectric waveguide. The core is a dielectric cylinder with radius $\rho = a$ and material parameters ϵ_1 and μ_1. In the cladding, $\rho > a$, the material parameters are ϵ_2 and μ_2, see figure 8.2. Both regions are lossless (i.e., ϵ and μ are real).

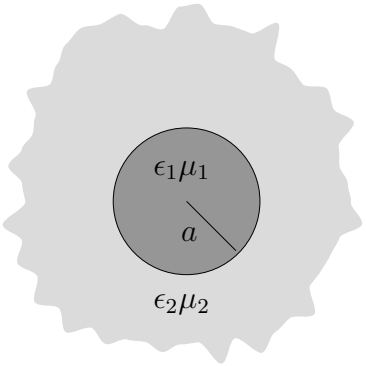


Figure 8.2: Geometry for the circular dielectric waveguide.

We can generalize our results to a waveguide with losses afterwards. We see from equation (8.3) that the $z-$components of the electric and magnetic fields do couple.

As in the case with the hollow waveguide there exist a finite number of propagating modes at each frequency. We will see that there exists one mode that has cut-off frequency zero and hence propagates for all frequencies. If the core is small enough this will be the only propagating mode. We call such a waveguide a single mode waveguide.

8.3.1 Waveguide modes

We now derive the expressions for the fields, the dispersion relations, the cut-off frequencies and the power flow in a circular dielectric waveguide. We have seen that $v(\rho,\phi)$ and $w(\rho,\phi)$ satisfy the Helmholtz equation in the two regions, c.f., equations (8.1) and (8.2). In cylindrical coordinates we have

$$\nabla_T^2 = \frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} \tag{8.4}$$

A separation of Helmholtz equation in cylindrical coordinates gives the eigenfunction system in the azimuthal angle $\{1, \sin m\phi, \cos m\phi\}$, where $m = 1, 2...$, c.f., section 5.5.3 on page 90. The boundary conditions imply that if v has the ϕ-dependence $\sin m\phi$ then w has the ϕ-dependence $\cos m\phi$ and vice versa. We choose E_z to be an odd function in ϕ, and hence H_z an even function [3] If we restrict ourselves to waves propagating in the positive z-direction, we get

$$v(\rho,\phi) = \psi_E(\rho)\sin m\phi$$
$$w(\rho,\phi) = \psi_H(\rho)\cos m\phi$$

where the functions ψ_E and ψ_H satisfy Bessel's differential equation

$$\left(\frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} - \frac{m^2}{\rho^2} + k_t^2\right)\psi(\rho) = 0 \tag{8.5}$$

Bessel's differential equation has two independent solutions, see appendix A, that can be combined in different ways. We use combinations such that the conditions at $\rho = 0$ and $\rho = \infty$ are easy to satisfy.

[3] we can also let E_z be even and H_z odd, i.e., to let $\sin m\phi$ and $\cos m\phi$ shift.

- If k_t is real or complex in the core $\rho < a$ we use the combination
$$\psi(\rho) = AJ_m(k_t\rho) + \tilde{A}N_m(k_t\rho) \tag{8.6}$$
where J_m is the Bessel function and N_m is the Neumann function of order m.

- If k_t is real in the cladding $\rho > a$ we use the combination
$$\psi(\rho) = AH_m^1(k_t\rho) + \tilde{A}H_m^2(k_t\rho) \tag{8.7}$$
where H_m^1 and H_m^2 are the Hankel functions of the first and second kind, respectively, see appendix A.

- If k_t is imaginary in the cladding it is suitable to express $\psi(\rho)$ in modified Bessel functions
$$\psi(\rho) = AK_m(q\rho) + \tilde{A}I_m(q\rho) \tag{8.8}$$
where $q^2 = -k_t^2$.

The different types of Bessel functions are described in appendix A.

We now try to find propagating modes. A necessary condition for a propagating mode is that k_z is real. It will be seen that this implies that $k_{1t}^2 > 0$ and $k_{2t}^2 < 0$. Since the fields have to be finite at $\rho = 0$ and $\rho = \infty$ we make the following ansatz

$$\begin{cases} v_m(\rho,\phi) = AJ_m(h\rho)\sin m\phi \\ w_m(\rho,\phi) = BJ_m(h\rho)\cos m\phi \end{cases} \quad \rho < a$$
$$\begin{cases} v_m(\rho,\phi) = CK_m(q\rho)\sin m\phi \\ w_m(\rho,\phi) = DK_m(q\rho)\cos m\phi \end{cases} \quad \rho > a \tag{8.9}$$

where we have introduced
$$\begin{cases} h = k_{1t} \\ q = -ik_{2t} \end{cases}$$

If we instead let $v_m(\rho,\phi)$ be an even function of ϕ we make the ansatz

$$\begin{cases} v_m(\rho,\phi) = AJ_m(h\rho)\cos m\phi \\ w_m(\rho,\phi) = -BJ_m(h\rho)\sin m\phi \end{cases} \quad \rho < a$$
$$\begin{cases} v_m(\rho,\phi) = CK_m(q\rho)\cos m\phi \\ w_m(\rho,\phi) = -DK_m(q\rho)\sin m\phi \end{cases} \quad \rho > a \tag{8.10}$$

It turns out that the even and odd ansatz above give identical characteristic equations for the relations between A, B, C and D

In order to relate the Fourier coefficients A, B, C and D to each other, and to determine the longitudinal wavenumber k_z (and hence the transverse wavenumbers $k_{1t} = h$ and $k_{2t} = iq$), we utilize that the fields in equation (8.3) are continuous at $\rho = a$. In the circular case $\hat{\boldsymbol{n}} = \hat{\boldsymbol{\rho}}$ in equation (8.3) and hence

$$\begin{cases} v(a^-,\phi) = v(a^+,\phi) \\ w(a^-,\phi) = w(a^+,\phi) \\ \dfrac{1}{h^2}\left[\dfrac{k_z}{a}\dfrac{\partial v(a^-,\phi)}{\partial\phi} - \omega\mu_0\mu_1\dfrac{\partial w(a^-,\phi)}{\partial\rho}\right] = -\dfrac{1}{q^2}\left[\dfrac{k_z}{a}\dfrac{\partial v(a^+,\phi)}{\partial\phi} - \omega\mu_0\mu_2\dfrac{\partial w(a^+,\phi)}{\partial\rho}\right] \\ \dfrac{1}{h^2}\left[\dfrac{k_z}{a}\dfrac{\partial w(a^-,\phi)}{\partial\phi} + \omega\epsilon_0\epsilon_1\dfrac{\partial v(a^-,\phi)}{\partial\rho}\right] = -\dfrac{1}{q^2}\left[\dfrac{k_z}{a}\dfrac{\partial w(a^+,\phi)}{\partial\phi} + \omega\epsilon_0\epsilon_2\dfrac{\partial v(a^+,\phi)}{\partial\rho}\right] \end{cases}$$

The four equations follow

$$\begin{pmatrix} J_m(ha) & 0 & -K_m(qa) & 0 \\ 0 & J_m(ha) & 0 & -K_m(qa) \\ \frac{mk_z}{h^2a}J_m(ha) & -\frac{\omega\mu_0\mu_1}{h}J'_m(ha) & \frac{mk_z}{q^2a}K_m(qa) & -\frac{\omega\mu_0\mu_2}{q}K'_m(qa) \\ \frac{\omega\epsilon_0\epsilon_1}{h}J'_m(ha) & -\frac{mk_z}{h^2a}J_m(ha) & \frac{\omega\epsilon_0\epsilon_2}{q}K'_m(qa) & -\frac{mk_z}{q^2a}K_m(qa) \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{8.11}$$

These equations have non-trivial solutions only if the determinant for the coefficients is zero, i.e.,

$$\begin{vmatrix} J_m(ha) & 0 & -K_m(qa) & 0 \\ 0 & J_m(ha) & 0 & -K_m(qa) \\ \frac{mk_z}{h^2a}J_m(ha) & -\frac{\omega\mu_0\mu_1}{h}J'_m(ha) & \frac{mk_z}{q^2a}K_m(qa) & -\frac{\omega\mu_0\mu_2}{q}K'_m(qa) \\ \frac{\omega\epsilon_0\epsilon_1}{h}J'_m(ha) & -\frac{mk_z}{h^2a}J_m(ha) & \frac{\omega\epsilon_0\epsilon_2}{q}K'_m(qa) & -\frac{mk_z}{q^2a}K_m(qa) \end{vmatrix} = 0$$

This is the characteristic equation that gives the longitudinal wavenumber k_z. We expand the determinant along the upper row and get

$$\begin{aligned} J_m(ha) &\begin{vmatrix} J_m(ha) & 0 & -K_m(qa) \\ -\frac{\omega\mu_1}{c_0h}J'_m(ha) & \frac{mk_z}{q^2a}K_m(qa) & -\frac{\omega\mu_2}{c_0q}K'_m(qa) \\ -\frac{mk_z}{h^2a}J_m(ha) & \frac{\omega\epsilon_2}{c_0q}K'_m(qa) & -\frac{mk_z}{q^2a}K_m(qa) \end{vmatrix} \\ -K_m(qa) &\begin{vmatrix} 0 & J_m(ha) & -K_m(qa) \\ \frac{mk_z}{h^2a}J_m(ha) & -\frac{\omega\mu_1}{c_0h}J'_m(ha) & -\frac{\omega\mu_2}{c_0q}K'_m(qa) \\ \frac{\omega\epsilon_1}{c_0h}J'_m(ha) & -\frac{mk_z}{h^2a}J_m(ha) & -\frac{mk_z}{q^2a}K_m(qa) \end{vmatrix} = 0 \end{aligned}$$

We expand the subdeterminants to get the equation

$$\begin{aligned} m^2\left(\frac{1}{q^2a^2}+\frac{1}{h^2a^2}\right)^2&\left(\frac{k_z}{k_0}\right)^2 \\ &= \left(\frac{\mu_1 J'_m(ha)}{haJ_m(ha)}+\frac{\mu_2 K'_m(qa)}{qaK_m(qa)}\right)\left(\frac{\epsilon_1 J'_m(ha)}{haJ_m(ha)}+\frac{\epsilon_2 K'_m(qa)}{qaK_m(qa)}\right) \end{aligned} \tag{8.12}$$

where we have used $k_0 = \omega/c_0 = \omega\sqrt{\epsilon_0\mu_0}$. This equation is of second degree in the quantity $J'_m(ha)/(haJ_m(ha))$ and has the two solutions

$$\begin{aligned} \frac{J'_m(ha)}{haJ_m(ha)} &= -\frac{K'_m(qa)}{2qaK_m(qa)}\frac{\mu_1\epsilon_2+\mu_2\epsilon_1}{\mu_1\epsilon_1} \\ &\pm\left[\left(\frac{K'_m(qa)}{2qaK_m(qa)}\frac{\mu_1\epsilon_2-\mu_2\epsilon_1}{\mu_1\epsilon_1}\right)^2+\left(\frac{mk_z}{k_1}\right)^2\left(\frac{1}{q^2a^2}+\frac{1}{h^2a^2}\right)^2\right]^{1/2} \end{aligned} \tag{8.13}$$

We have two sets of solutions to equation (8.12), one for the plus sign and one for the minus sign, and that implies that we have two sets of waveguide modes. The solutions that are associated with the equation with the plus sign are referred to as EH-modes and the others as HE-modes. The characteristic equations are the equations for the longitudinal wavenumber k_z. The corresponding transverse wavenumbers in the core and cladding are given by $h = \sqrt{k_1^2 - k_z^2}$ and $k_{2t} = iq$, where $q^2 = k_z^2 - k_2^2 = k_1^2 - k_2^2 - h^2$.

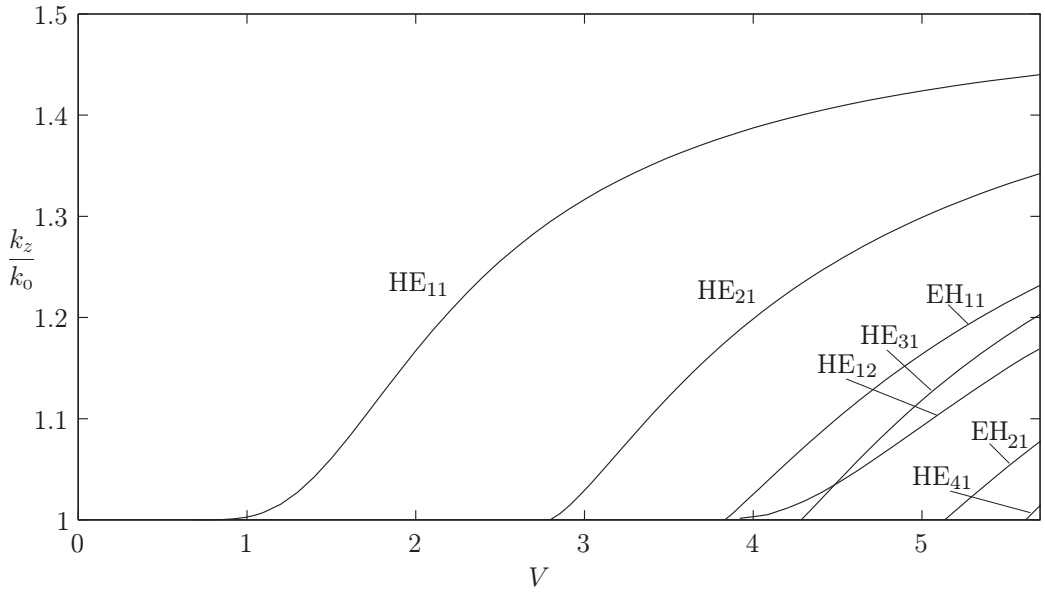


Figure 8.3: The normalized wavenumber k_z/k_0 as a function of the normalized frequency $V = k_0 a\sqrt{(n_1^2 - n_2^2)}$ for the lowest HE and EH modes when $n_1 = 1.5$ and $n_2 = 1$.

We define the cut-off frequency for a mode as the lowest frequency for which the mode propagates without attenuation. If q^2 is real and positive, the waves are exponentially decreasing in the radial direction in the cladding since the argument in the modified Bessel functions in equation (8.9) then is real and positive. No active power is in that case radiated in the radial direction and the wave propagates without attenuation. If q^2 is negative then power is radiated in the radial direction and the wave is attenuated in the z-direction. Mathematically we define the cut-off frequency as the frequency for which $(qa)^2 = 0$, i.e., $qa = 0$. At the cut-off frequency we have $h = \sqrt{k_1^2 - k_2^2} = k_0\sqrt{n_1^2 - n_2^2}$.

When the characteristic equation is solved and k_z is determined we can express three of the coefficients A, B, C and D in the fourth. The fourth one determines the amplitude of the mode and in order to find that we need to know the source of the field. From the equations in (8.11) we get the following relations:

$$\frac{C}{A} = \frac{D}{B} = \frac{J_m(ha)}{K_m(qa)} \tag{8.14}$$

$$\frac{B}{A} = \frac{D}{C} = \frac{mk_z}{k_0}\left(\frac{1}{h^2a^2} + \frac{1}{q^2a^2}\right)\left(\frac{\mu_1 J'_m(ha)}{haJ_m(ha)} + \frac{\mu_2 K'_m(qa)}{qaK_m(qa)}\right)^{-1} \tag{8.15}$$

8.3.2 HE-modes

For $m \geq 1$ only modes with both E_z and H_z non-zero can exist. These modes are called hybrid modes and are not TE- or TM-modes. The characteristic equation that determines

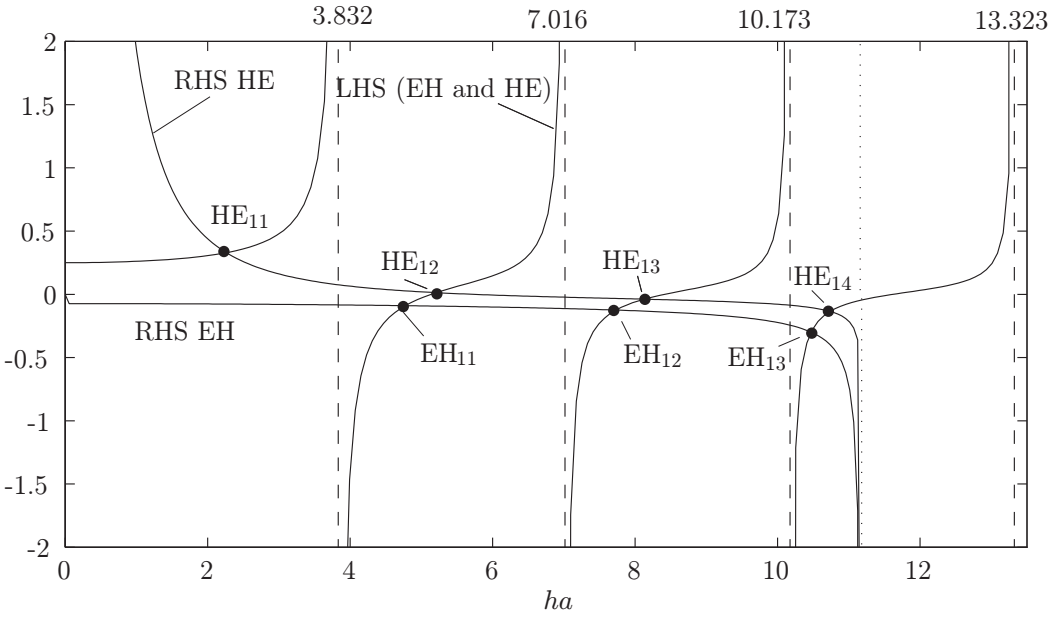


Figure 8.4: The left and right hand sides for the characteristic equation for the EH- and HE-codes with $m = 1$, $\mu_1 = \mu_2 = 1$ $\sqrt{\epsilon_1} = n_1 = 1.5$, $\sqrt{\epsilon_2} = n_2 = 1$ and $k_0 a = 10$. The points of interaction give the transverse wavenumber h for the propagating modes. The vertical dotted line $ha = k_0 a\sqrt{n_1^2 - n_2^2}$ is the limit for the right hand sides of the characteristic equations.

the longitudinal wavenumber for the HE-modes can be written as (see equation (8.13))

$$\begin{aligned}\frac{J_{m+1}(ha)}{haJ_m(ha)} &= \frac{K'_m(qa)}{2qaK_m(qa)}\frac{\mu_1\epsilon_2 + \mu_2\epsilon_1}{\mu_1\epsilon_1} + \frac{m}{(ha)^2} \\ &+ \left[\left(\frac{K'_m(qa)}{2qaK_m(qa)}\frac{\mu_1\epsilon_2 - \mu_2\epsilon_1}{\mu_1\epsilon_1}\right)^2 + \left(\frac{mk_z}{k_1}\right)^2\left(\frac{1}{q^2a^2} + \frac{1}{h^2a^2}\right)^2\right]^{1/2}\end{aligned} \tag{8.16}$$

The left and right hand sides, for the case $m = 1$, are plotted as a function of ha in figure 8.4. The relevant parameters are $m = 1$, $\mu_1 = \mu_2 = 1$ $\sqrt{\epsilon_1} = n_1 = 1.5$, $\sqrt{\epsilon_2} = n_2 = 1$ and $k_0 a = 10$. The points of intersection give the solutions and we see that for this frequency there are four propagating HE-modes with $m = 1$. The vertical dashed lines indicate the values of ha for which the left hand side goes to infinity, while the vertical dotted line indicates the point where $q \to 0$, i.e., where the right hand sides approach $-\infty$. If the frequency decreases the dotted line $ha = k_0 a\sqrt{n_1^2 - n_2^2}$ moves to the left and the number of intersections, and hence propagating modes, are reduced. No matter how much it decreases there is always at least one point of intersection. That means that the lowest HE-mode for $m = 1$ propagates for all frequencies. This is a very important mode and is the one used in single mode fibers.

It is more complicated to determine the cut-off frequencies for the HE-modes than for the EH-modes. For $m = 1$ one can take the first two terms in a power series expansion of $K_1(qa)$ and $K_2(qa)$. One then sees that the right hand side in equation (8.16) goes to

infinity when $qa \to 0$ and hence the cut-off frequencies are given by

$$f_{1n} = \frac{c_0}{2\pi a}\frac{\xi_{1n}}{\sqrt{n_1^2 - n_2^2}} \tag{8.17}$$

where ξ_{1n} is the zero number n of J_1 (the zero $\xi_{11} = 0$ is the first zero in this case), i.e.,

$$J_1(\xi_{1n}) = 0$$

If we take the three first terms in the asymptotic expansion of $K_m(qa)$ when $qa \to 0$ one can prove that the cut-off frequencies for the modes HE_{mn} are given by

$$f_{mn} = \frac{c_0}{2\pi a}\frac{x_{mn}}{\sqrt{n_1^2 - n_2^2}} \tag{8.18}$$

where

$$x_{mn}\frac{J_m(x_{mn})}{J_{m-1}(x_{mn})} = (m-1)\frac{n_1^2 + n_2^2}{n_2^2}, \qquad m > 1$$

The dispersion curves for the lowest HE-modes are depicted in figure 8.3. In the graph the normalized wavenumber, k_z/k_0, is plotted as a function of the normalized frequency $V = k_0 a\sqrt{n_1^2 - n_2^2}$. The normalized cut-off frequencies for the HE_{1n}-modes start where $J_1(V) = 0$, and the cut-off frequencies for the HE_{mn}-modes with $m > 1$ start where $VJ_m(V)/J_{m-1}(V) = (m-1)(n_1^2 + n_2^2)/n_2^2$.

8.3.3 EH-modes

If we utilize that $J_m'(z) = \frac{m}{z}J_m(z) - J_{m+1}$ the characteristic equation that determines the longitudinal wavenumber for the EH-modes, see equation (8.13), can be written as

$$\frac{J_{m+1}(ha)}{haJ_m(ha)} = \frac{K_m'(qa)}{2qaK_m(qa)}\frac{\mu_1\epsilon_2 + \mu_2\epsilon_1}{\mu_1\epsilon_1} + \frac{m}{(ha)^2}$$
$$- \left[\left(\frac{K_m'(qa)}{2qaK_m(qa)}\frac{\mu_1\epsilon_2 - \mu_2\epsilon_1}{\mu_1\epsilon_1}\right)^2 + \left(\frac{mk_z}{k_1}\right)^2\left(\frac{1}{q^2a^2} + \frac{1}{h^2a^2}\right)^2\right]^{1/2}$$

The left hand side is the same as for the HE-modes but the right hand side have changed. The left and right hand sides, for the case $m = 1$, are plotted as a function of ha in figure 8.4. The relevant parameters are $m = 1$, $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.5$, $\sqrt{\epsilon_2} = n_2 = 1$ and $k_0 a = 10$. The points of intersection give the solutions and we see that for this frequency there are three propagating EH-modes for $m = 1$. The vertical dashed lines indicate the values of ha for which the left hand side goes to infinity, while the vertical dotted line indicates the point where $q \to 0$, i.e., where the right hand sides approach $-\infty$. When the frequency decreases such that $k_0 a\sqrt{n_1^2 - n_2^2} < 3.832$ then there are no points of intersection for the EH-equation and hence no propagating EH-modes. The dispersion curves for the lowest EH-modes are depicted in figure 8.3. The graphs give the normalized longitudinal wavenumber, k_z/k_0, plotted as a function of the normalized frequency $V = k_0 a\sqrt{n_1^2 - n_2^2}$. We have chosen $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.5$ and $\sqrt{\epsilon_2} = n_2 = 1$ as was the case also in figures 8.5 and 8.4 .

The cut-off frequencies for the EH-modes are the frequencies for which the curves for the left and right hand sides intersect when $q = 0$, i.e., $ha = k_0 a\sqrt{n_1^2 - n_2^2}$. By using the

asymptotics in eq. 8.23 we see that the right hand side goes to infinity when $q \to 0$ and hence $J_m(ha)$ has to be zero when $q = 0$. The cut-off frequencies for the EH-modes are then given by

$$f_{mn}^{EH} = \frac{c_0}{2\pi a} \frac{\xi_{mn}}{\sqrt{n_1^2 - n_2^2}}$$

where $J_m(\xi_{mn}) = 0$. In figure 8.3 the dispersion curves start at the cut-off frequency V, where $J_m(V) = 0$ and $k_z = k_2$.

8.3.4 TE- and TM-modes

The axially symmetric solutions have $m = 0$, where the odd modes have $E_z = 0$ and $H_z \neq 0$, i.e., they are TE-modes. The even modes have $E_z \neq 0$ and $H_z = 0$ and are TM-modes. For these modes the system of equations (8.11) is split up into two independent systems

$$\begin{pmatrix} J_0(ha) & -K_0(qa) \\ \frac{\epsilon_1}{ha} J_0'(ha) & \frac{\epsilon_2}{qa} K_0'(qa) \end{pmatrix} \begin{pmatrix} A_0 \\ C_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{8.19}$$

for TM-modes and

$$\begin{pmatrix} J_0(ha) & -K_0(qa) \\ \frac{\mu_1}{ha} J_0'(ha) & \frac{\mu_2}{qa} K_0'(qa) \end{pmatrix} \begin{pmatrix} B_0 \\ D_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{8.20}$$

for TE-modes. If we utilize that $J_0' = -J_1$ and $K_0' = -K_1$ the determinant condition gives the characteristic equation

$$\frac{J_1(ha)}{haJ_0(ha)} = -\frac{\epsilon_2}{\epsilon_1} \frac{K_1(qa)}{qaK_0(qa)} \tag{8.21}$$

that determines k_z for the TM-modes and

$$\frac{J_1(ha)}{haJ_0(ha)} = -\frac{\mu_2}{\mu_1} \frac{K_1(qa)}{qaK_0(qa)} \tag{8.22}$$

that determines k_z for the TE-modes. The left and right hand sides are plotted as functions of ha in figure 8.5. The core has in this figure $\mu_1 = 1$ and $\sqrt{\epsilon_1} = n_1 = 1.5$ (the values of silica) and a radius given by $k_0 a = 10$. The cladding is air, i.e., $\epsilon_2 = \mu_2 = 1$. The points of intersection are the solutions and in this case there are three propagating TE-modes and three propagating TM-modes.

The cut-off frequencies for the TE- and TM-modes are obtained by letting $q = 0$ in the characteristic equations. For small qa the asymptotic formulas (see appendix A and equation (A.6) on page 200)

$$\begin{aligned} K_0(qa) &\sim -\ln(qa/2) - \gamma + O((qa)^2) \\ K_m(qa) &\sim 2^{m-1}(m-1)!(qa)^{-m} \end{aligned} \tag{8.23}$$

The right hand side goes to $-\infty$ when qa goes to zero. This point corresponds to $ha = k_0 a\sqrt{n_1^2 - n_2^2}$ and is indicated by a dotted vertical line in figure 8.5. In order for the left hand side to go to infinity $J_0(ha)$ has to go to zero. The cut-off frequency for the TE- and TM-modes are then determined by

$$f_{0n} = \frac{c_0}{2\pi a} \frac{\xi_{0n}}{\sqrt{n_1^2 - n_2^2}}$$
$$\text{where } J_0(\xi_{0n}) = 0$$

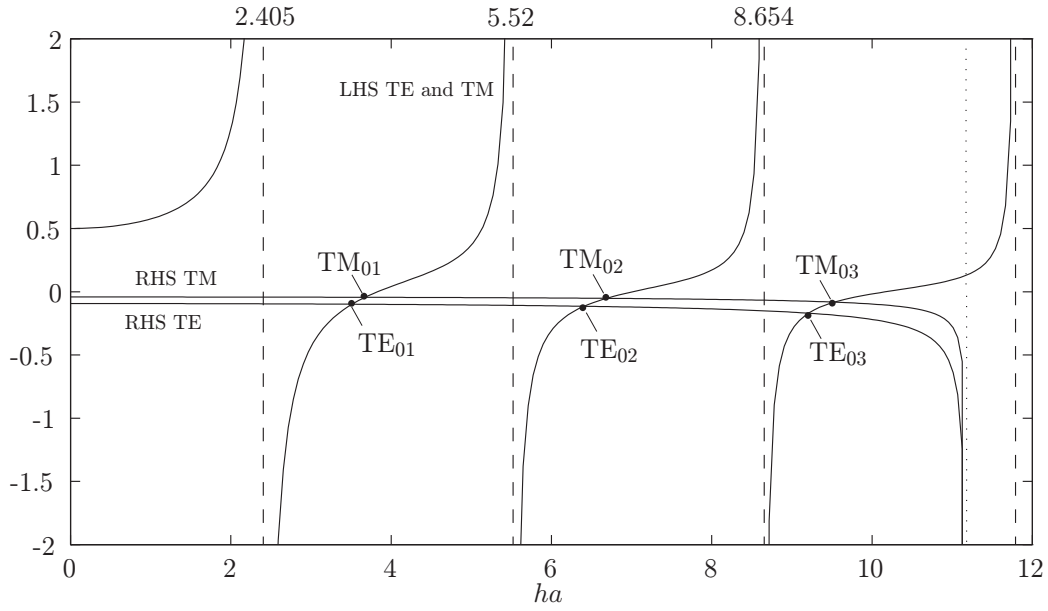


Figure 8.5: Left and right hand sides of the characteristic equations for the TE- and TM-modes, i.e., equations (8.22) and (8.21), when $n_1 = 1.5$, $n_2 = 1$ and $k_0 a = 10$. The points of intersection give the transverse wavenumber $h = k_{1t}$ for the propagating modes. The dotted vertical line indicates the highest frequency for which the right hand side is real. The dashed vertical lines are placed at the zeros of the Bessel functions J_0. At these zeros the left hand side is singular.

The first zeros to $J_0(x)$ are given in table A.1 on page 196. The smallest zero is $\xi_{01} = 2.405$ which means that for frequencies $f < 2.405 c_0/2\pi a\sqrt{n_1^2 - n_2^2}$ there are no propagating TE- and TM-modes.

8.4 Optical fibers

An optical fiber is a dielectric waveguide with very small losses at the frequency of operation. The most common type of optical fiber is a circular silica fiber where the core is doped. Due to the doping the core has a slightly larger index of refraction than the cladding, see figure 8.6. The wavelength is often the one for which the losses are minimized. In figure 8.6 we see the losses in a lightly doped core as a function of the vacuum wavelength. Some of the phenomena that give rise to losses are indicated. We see that the sum of the losses is at a minimum at the vacuum wavelength $\lambda = 1.55\,\mu$m, which is in the infrared region. This is the most common wavelength used in optical fiber communication. In order to minimize dispersion the single mode fiber is used, which means that the HE_{11} mode is the only propagating mode. The frequency then has to be below the cut-off frequency for the next modes which are the TE_{01} and TM_{01} modes. The cut-off frequencies for the TE_{01} and TM_{01} modes are given by $k_0 a\sqrt{n_1^2 - n_2^2} = 2.405$, where a is the radius of the core. In order to operate a fiber as a single mode fiber we have to have $k_0 a\sqrt{n_1^2 - n_2^2} < 2.405$.

Since 1980 communication systems based on optical fibers are invaluable for both long

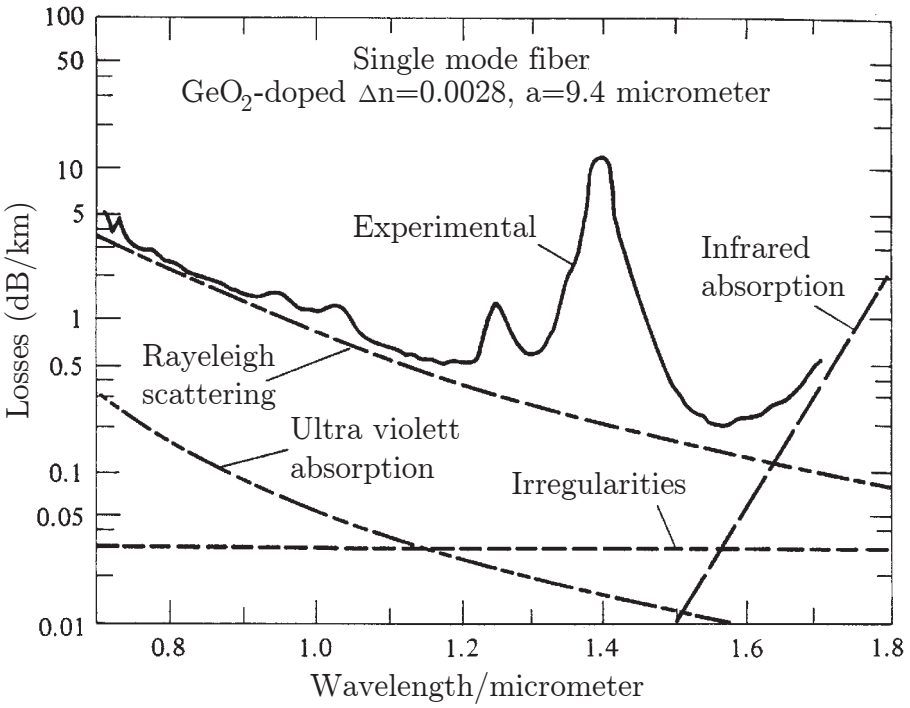


Figure 8.6: Spectrum for the losses for a single mode fiber made out of silica and with the core doped with GeO_2.

distance as well as short distance communication. They are gradually replacing systems based on communication with electric wires and they complement the wireless systems based on microwaves. The short but intense history of optical fiber systems is interesting and some of the important achievements are given below [4]:

1960 The He-Ne-laser is presented.

1963 The laser diode (GaAs) is presented.

1966 Silica is suggested by Charles Kao as a candidate for an optical waveguide. Since the best glass materials at this time have an attenuation of 1 dB/m, the suggestion does not seem to be feasible. One consider an attenuation of 20 dB/km to be the largest attenuation that is acceptable for communication.

1970 Glass fibers with an attenuation of 17 dB/km are presented by a research group at Corning Inc. The research in the area of optical fibers grows rapidly.

1975 Glass fibers with an attenuation of 4 dB/km are presented by the research group at Corning Inc.

1977 The first optical fiber systems are tested and one manages to transfer 140 Mbit/s on a 9 km long fiber. The fiber is a multi-mode fiber. i.e., the light pulses consist of a large number of propagating modes.

[4] A more detailed historic overview is given in eg., J. E. Midwinter and Y. L. Guo, *Optoelectronics and Lightwave Technology*, John Wiley & Sons, New York, 1992.

1980 Extremely pure silica glass with an attenuation of 0.2 dB/km at the wavelength 1550 nm can be manufactured. Diode lasers and detectors for the 1550 nm wavelength are developed. In order to reduce dispersion one develops the single mode fiber. A bit rate of 140 Mbit/s over a distance of 200 km is reported.

1990 Erbium doped fibers that can be used as optical amplifiers are developed. Before the development of the optical amplifiers the signal had to be transformed into an electric signal in order to be amplified. The attenuation of the signal is now not a problem. The phenomena that sets the limit for the bit rate is the dispersion.

today Fibers with 10 Gbit/s per channel are used in commercial systems. By using wavelength dispersion multiplexing (WDM) each fiber can carry several channels and bit rates of 40 Gbit/s are common in these systems. In laboratories bit rates of 1 Tbit/s have been achieved.

Example 8.1
If we assume a fiber with $n_1 = 1.5028$ and $n_2 = 1.5$ where infrared light with vacuum wavelength $\lambda = 1.55\,\mu$m is used the condition for the single mode fiber is that

$$a < 2.405/(k_0\sqrt{n_1^2 - n_2^2}) = 2.405\lambda/(2\pi\sqrt{n_1^2 - n_2^2})$$

That corresponds to $a < 6.5\,\mu$m. ∎

Figure 8.7 shows the left and right hand sides in the characteristic equations for the TE- and TM-modes, see equations (8.22) and (8.21), plotted as functions of the transverse wavenumber in the core, h, multiplied with the radius s a. The fiber has $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.52$, $\sqrt{\epsilon_2} = n_2 = 1.50$ and $k_0 a = 40$. Since $n_1/n_2 \approx 1$ the characteristic equations for the TE- and TM-modes are almost identical, as seen from the figure. The points of intersection give the solutions to the characteristic equations, and in this case there are three propagating modes each of the TE and TM-modes. When $q \to 0$ the right hand side goes to $-\infty$. This point corresponds to $ha = k_0 a\sqrt{n_1^2 - n_2^2}$ and is indicated by a vertical line in the figures. The corresponding curves for the EH- and HE-modes with $m = 1$ are given in figure 8.8. At $k_0 a = 40$ we see that the modes HE_{11}, HE_{12}, HE_{13}, EH_{11} and EH_{12} propagate. In figure 8.9 there are dispersion curves for the lowest EH- and HE-modes. The curves show the normalized longitudinal wavenumber k_z/k_0 as a function of the normalized frequency $V = k_0 a\sqrt{n_1^2 - n_2^2}$.

8.4.1 Effective index of refraction and phase velocity

For a propagating mode we can write the longitudinal wavenumber k_z as:

$$k_z = n_e k_0 = n_e(n_1, n_2, \omega)\frac{\omega}{c_0} \tag{8.24}$$

were n_e is the *effective index of refraction*. The *Phase velocity* for a propagating mode is defined by (c.f., the definition on page 79)

$$v_f = \frac{\omega}{k_z} = \frac{c_0}{n_e} \tag{8.25}$$

For a propagating mode we have that $k_z = \sqrt{k_1^2 - h^2} = \sqrt{k_2^2 + q^2}$ where $h^2 > 0$ and $q^2 > 0$ and hence $k_2 < k_z < k_1$, $n_2 < n_e < n_1$ and $c_1 < v_f < c_2$.

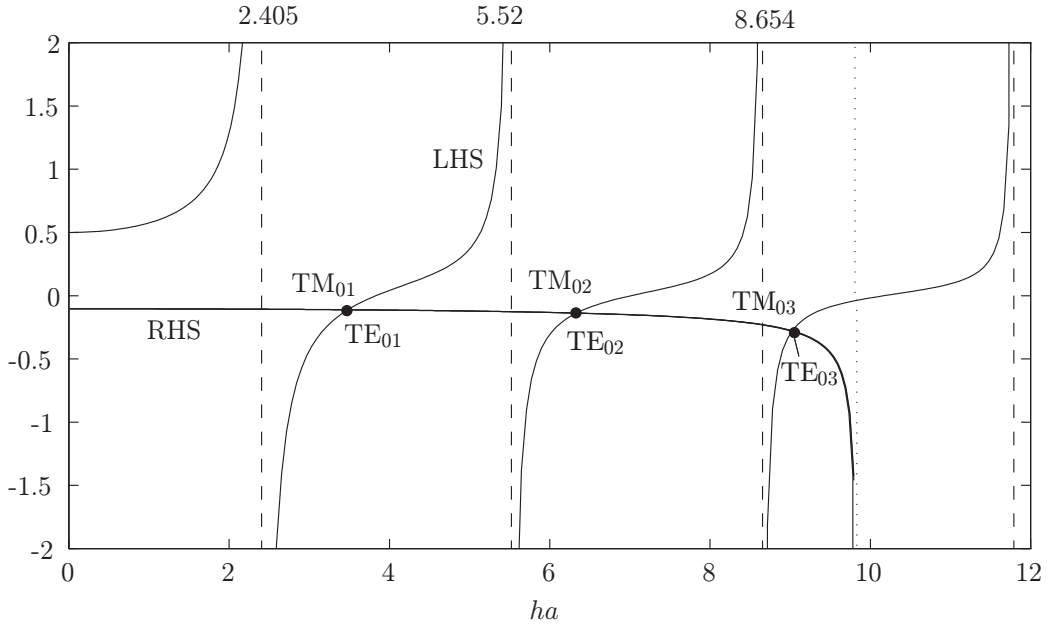


Figure 8.7: The left and right hand sides of the characteristic equations for the TE- and TM-modes when $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.52$, $\sqrt{\epsilon_2} = n_2 = 1.50$ and $k_0 a = 40$. The points of intersection give the transverse wavenumber $h = k_{1t}$ for the propagating modes. The dotted vertical line indicates the highest frequency for which the right hand side is real. The dashed vertical lines are located at the zeros of the Bessel function J_0, where the left hand side is singular.

8.4.2 Dispersion

The dispersion is a measure of the frequency dependence of the wave propagation. In a waveguide there are three different phenomena that lead to dispersion:

- Multi-mode dispersion: Since different modes have different phase speeds a pulse that consists of several modes is dispersive.
- Waveguide dispersion: For a single mode the phase velocity is a function of frequency. That means that a pulse consisting of a single mode changes shape when it propagates.
- Material dispersion: The material in a waveguide might have an index of refraction that is frequency dependent. This is the case for an optical fiber where the index of refraction for the pure silica and the doped silica is frequency dependent in the infrared region.

Since we are mainly interested in single mode fibers we do not consider multi-mode dispersion in this section. The sum of the waveguidedispersion and the material dispersion is called the chromatic dispersion. The parameter that indicates the waveguide dispersion is the dispersion parameter

$$D = \frac{d\tau_g}{d\lambda} \tag{8.26}$$

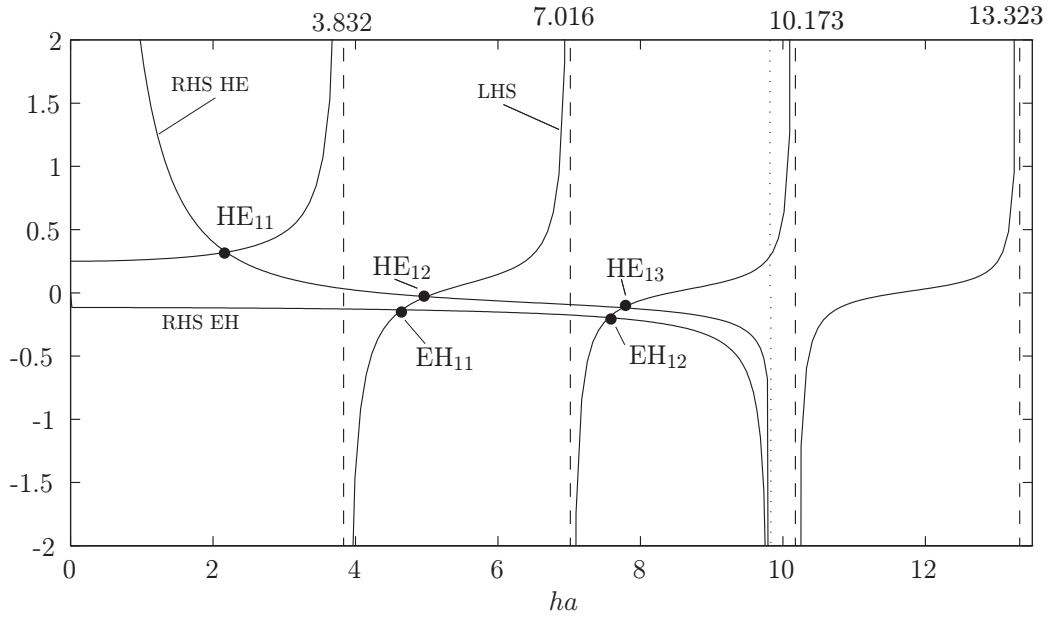


Figure 8.8: The left and right hand sides of the characteristic equations for the EH- and HE-modes with $m = 1$. The optical fiber has $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.52$, $\sqrt{\epsilon_2} = n_2 = 1.50$ ands $k_0 a = 40$. The points of intersection give the transverse wavenumber h for the propagating modes.

where $\tau_g = \frac{dk_z}{d\omega}$ is the inverse of the group velocity and λ is the wavelength in vacuum. We see that

$$\tau_g = \frac{n_e}{c_0} - \frac{\lambda}{c_0}\frac{dn_e}{d\lambda}$$
$$D = \frac{d\tau_g}{d\lambda} = -\frac{\lambda}{c_0}\frac{d^2 n_e}{d\lambda^2}$$

If D is less than zero, the waveguide has positive, or also called normal, dispersion. If D is greater than zero, the waveguide has negative, or anomalous dispersion. If a light pulse is propagated through a normally dispersive medium, the higher frequency components travel slower than the lower frequency components. The pulse therefore becomes positively chirped, or up-chirped, increasing in frequency with time. Conversely, if a pulse travels through an anomalously dispersive medium, high frequency components travel faster than the lower ones, and the pulse becomes negatively chirped, or down-chirped, decreasing in frequency with time. In optics materials have normal dispersion in most frequency bands, but close to resonances there are small frequency bands where the dispersion can be anomalous. The hollow waveguides described in chapter 5, with vacuum or air, always have anomalous since the group velocity increases with frequency and the phase velocity decreases with velocity. That means that high frequencies in a pulse travels faster than low frequencies. This is seen from the analysis in chapter 7.

For a waveguide that has no chromatic dispersion, the group velocity, and then also τ, is frequency independent. That means that D is zero. A pulse propagating in a waveguide

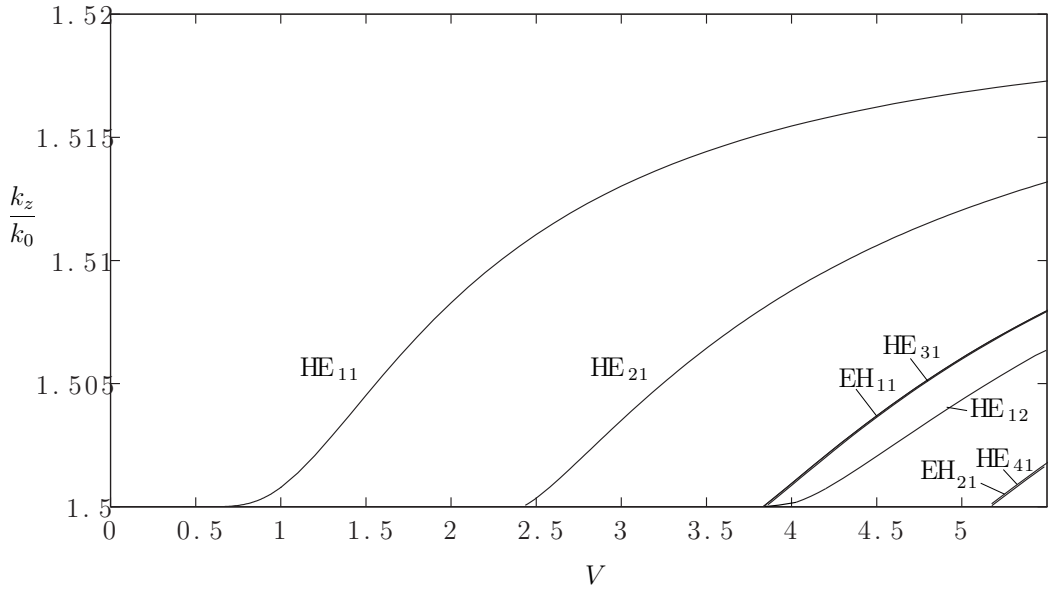


Figure 8.9: The normalized longitudinal wavenumber k_z/k_0 as a function of the normalized frequency $V = k_0 a\sqrt{(n_1^2 - n_2^2)}$ for the lowest HE and EH modes when $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.52$ and $\sqrt{\epsilon_2} = n_2 = 1.50$. We see that the dispersion curves for HE_{31} and EH_{11} are almost identical. These form together a linearly polarized mode.

without chromatic dispersion does not change its shape, but might be attenuated. If D is large it means that τ_g and also the group velocity varies rapidly with λ. The large difference in group velocity between different frequencies gives rise to a rapid pulse broadening. The chromatic dispersion can be decomposed in the contribution from the material dispersion D_m and the waveguide dispersion D_v as

$$D = D_m + D_v \tag{8.27}$$

When $n_1 - n_2 \ll 1$ we can determine the material dispersion by utilizing the approximation $n_e \approx n_1$ which gives

$$D_m = -\frac{\lambda}{c_0}\frac{d^2 n_1}{d\lambda^2} \tag{8.28}$$

When we consider the waveguide dispersion we keep n_1 and n_2 constant, i.e.,

$$D_v = -\frac{\lambda}{c_0}\frac{d^2 n_e}{d\lambda^2}|_{n_1, n_2 \text{ konst.}}$$

Example 8.2
Figure 8.10 shows the material dispersion as a function of the wavelength. We have already seen from figure 8.6 that the most suitable wavelength for a single mode fiber made out of silica is $\lambda = 1.55\,\mu$m. At this wavelength the material dispersion is positive. In order to obtain a dispersion free fiber at the wavelength $\lambda = 1.55\,\mu$m we have to compensate the positive material dispersion with a negative waveguide dispersion. Since the wavelength for the fiber is given we can obtain the desired waveguide dispersion by choosing a suitable

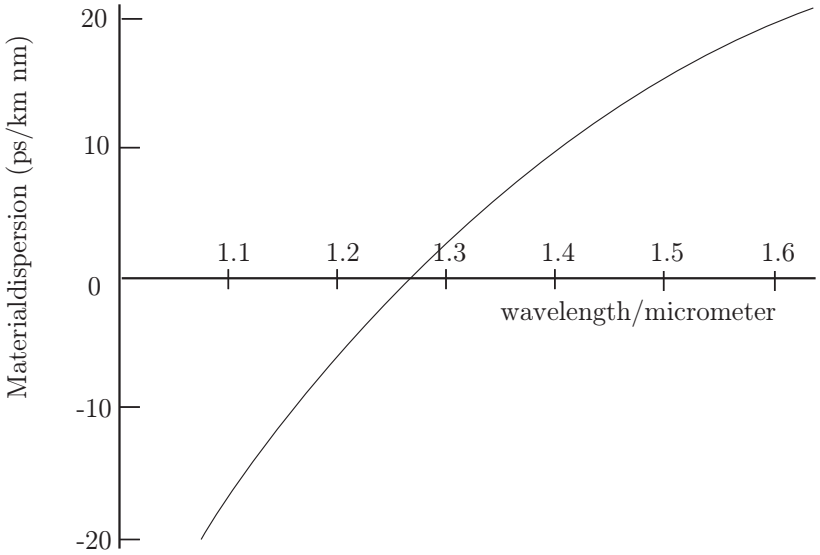


Figure 8.10: Material dispersion for silica.

value for the radius of the core a. The method to choose a radius of the fiber such that the dispersion is zero is called to dispersion-shift the fiber.

■

8.4.3 Attenuation in optical fibers

Sofar we have assumed that the material of the fiber has been lossless. This is an approximation since there are small losses in all fibers. These losses come from scattering from impurities in the material and from losses in the silica glass itself, see figure 8.6. One can purify the material such that the scattering from the impurities are negligible, but the losses in the silica glass cannot go below a certain value. The absorption is, as we have seen, very dependent on the wavelength. The smallest attenuation in silica is at the wavelength $1.55\,\mu$m. The attenuation is so small that it does not affect the expressions and equations that have been derived in this chapter. For a fiber with $n_1 - n_2 \ll 1$ the only change is that we replace the factor $e^{\mathrm{i}k_z z}$ with $e^{\mathrm{i}\,\mathrm{Re}\{k_z\}z}e^{-\alpha z}$. The attenuation α is approximately the same as for plane wave propagation in a dielectric material with small losses

$$\alpha = \mathrm{Im}\{k_z\} = k_0\,\mathrm{Im}\{n_e\} \simeq k_0\,\mathrm{Im}\{\sqrt{\epsilon_1}\} \simeq \frac{k_0\,\mathrm{Im}\{\epsilon_1\}}{2\sqrt{\mathrm{Re}\{\epsilon_1\}}} \tag{8.29}$$

where $\epsilon_1 = \mathrm{Re}\{\epsilon_1\} + i\,\mathrm{Im}\{\epsilon_1\}$ is the complex permittivity.

8.4.4 Dielectric waveguides analyzed with FEM

We have seen that the dielectric waveguides are much harder to analyze analytically than the hollow waveguides. Fortunately the numerical methods can handle a large number of problems. We give three examples here, the circular dielectric waveguide, a dielectric waveguide with elliptic cross section and finally a dielectric resonator.

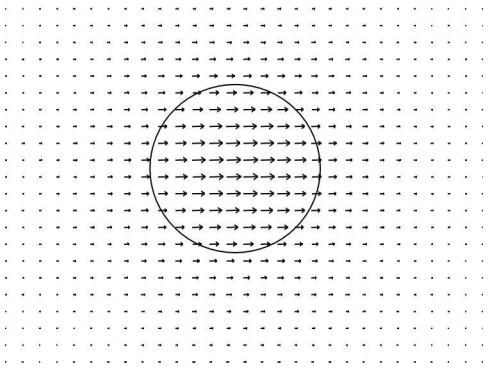


Figure 8.11: Electric field for a circular optical fiber.

Example 8.3
Consider a circular dielectric waveguide with radius $a = 4\,\mu$m and with $n = 1.503$ in the core and $n = 1.5$ in the cladding. The wavelength in vacuum is $\lambda = 1.55\,\mu$m. Find k_z and the electric field distribution. We do the following steps:

- Choose 2D and **2D> Electromagnetic waves> Mode analysis**.
- Draw a circle with radius around 40 μm (cladding) and a circle with radius $a = 4\,\mu$m (core).
- We define materials with the permittivity $\epsilon = n^2 = 1.503^2$ in the core and $\epsilon = 1.5^2$ in the cladding. The conductivity is 0 and the relative permeability 1 in both materials.
- We let the frequency be $f = 3e8/1.55e - 6$ in **Mode analysis**.
- We add 1.5 in the **Search for effective mode index at**.
- Solve. We look at the power flow density and a surface array plot of the different solutions that COMSOL has calculated.
- Under **Plot 2D** the different mode indices are given. There should be just two modes with mode index between 1.5 and 1.503.

■

In figure 8.11 the electric field for the circular waveguide is depicted. The other mode has the same mode index but with the electric field rotated 90° compared to the one in the figure. In figure 8.12 we have exchanged the circle for an ellipse with half axes $a = 5\,\mu$m and $b = 3\,\mu$m. In that case the two propagating modes do not have the same effective mode indices and the electric fields are different. The mode indices COMSOL gives are 1.500543 for the mode with the electric field along the minor axis and 1.500544 for the mode with the electric field along the major axis.

8.4.5 Dielectric resonators analyzed with FEM

A dielectric resonator is a piece of dielectric material. Placed in air, or vacuum, such a piece has resonance frequencies. At these frequencies the electromagnetic field is confined inside the dielectric piece. Outside the piece the field attenuates exponentially. Dielectric

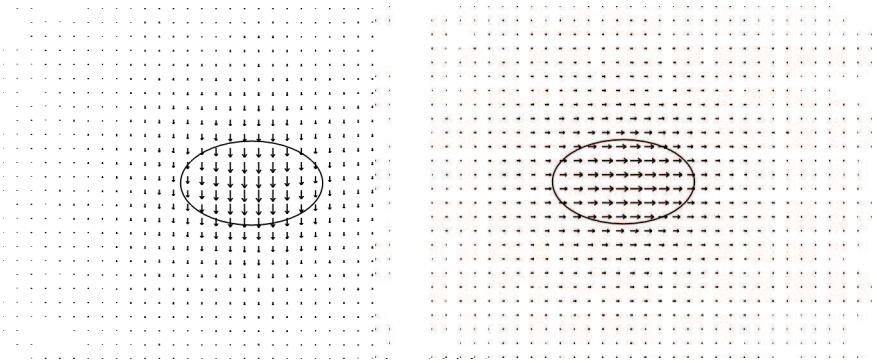


Figure 8.12: The electric fields the two modes for an elliptic optical fiber. COMSOL gives the mode index 1.500543 for the mode to the left, i.e., with the electric field along the minor axis and 1.500544 for the mode to the right, i.e., with the electric field along the major axis.

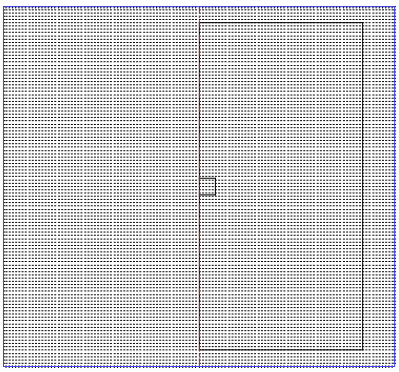


Figure 8.13: The geometry in COMSOL for the cylindric dielectric resonator. Notice that we only draw that half of the cylinders that are to the right of the symmetry line.

resonators are used as bandpassfilters with very high $Q-$values, i.e., they are very narrow banded. The resonators are also used as resonating antennas and lately these dielectric antennas have become quite popular. The dielectric resonators can be analyzed with FEM. The resonant frequencies of a resonators are determined by the shape of the resonator and the the permittivity of the material. As a rule of thumb the first resonant frequency occurs when the diameter of the resonator is on the order of half of the wavelength in the material. Below we describe how the resonance frequencies can be calculated using COMSOL for a cylindric resonator with radius 1 mm and height 1 mm. It is quite time-consuming to find and specify the resonances. Since we have an outer region there are resonances that do not belong to the resonator and it is important that we disregard these. One way to filter out all of the resonances that do not belong to the resonator is to do the calculations with two different sizes of the outer region. Only the frequencies that do not change can be resonances of the resonator. In the example below two different sizes were used.

First we do it by considering the axially symmetric resonances.

- We choose **2D axisymmetric>Electromagnetic waves>Eigenfrequency**.

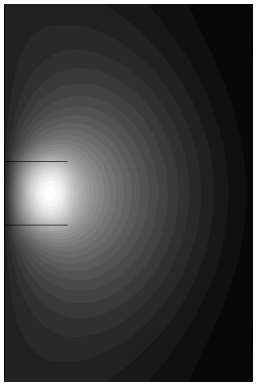


Figure 8.14: The electric field of the lowest resonant frequency $f_r = 46.2$ GHz for the cylindric resonantor with $\epsilon = 10$, radius $a = 1$ mm and heght $h = 1$ mm. The mode is a TE-mode, i.e., the electric field is directed ion the azimuthal direction.

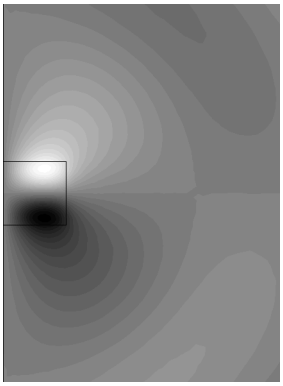


Figure 8.15: The electric field of the second lowest resonant TE-mode. The frequency is $f_r = 75.7$ GHz for the cylindric resonantor with $\epsilon = 10$, radius $a = 1$ mm and heght $h = 1$ mm.

- The axis $r = 0$ is the symmetry line. We draw an outer rectangle that defines the computational domain and a smaller rectangle that is the resonator. This is done as in figure 8.13. The outer rectangle is ten times as large as the resonator.
- Next the boundary conditions are chosen. The boundary lines on the axis $r = 0$ has boundary condition **axial symmetry** and at the outer boundary lines we may use scattering boundary condition, or perfect electric conductor. Notice that we do not need to specify the boundary conditions between the outer and inner rectangle.
- In **Material>+Material** we specify the materials for the two regions. In this example we let the outer region be air $\epsilon = 1$ and the inner be a material with $\epsilon = 10$.
- In **Study>Eigenfrequency** we specify that COMSOL should find 6 frequencies and that it should search for frequencies around 40 GHz. This a value that we have to guess and we base our guess on that the radius should be approximately a half

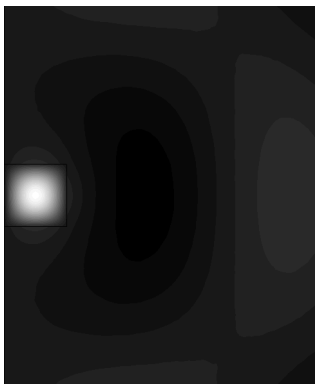


Figure 8.16: The magnetic field of the lowest resonant TM-mode. The frequency is $f_r = 73.4$ GHz for the cylindric resonator with $\epsilon = 10$, radius $a = 1$ mm and heght $h = 1$ mm.

of a wavelength, i.e., $\lambda \sim 2$ mm which gives $f \sim c_0/(\sqrt{\epsilon}\lambda) \approx 50$ GHz. To be on the safe side we start the search 10 GHz below this value.

- We solve and look at the different resonances. There are a number of resonances that are resonances in the outer cylinder. The ones we are looking for should have their electric field confined to the resonator. In figures 8.14 the electric field of the lowest resonance is depicted. The frequency is 46.3 GHz. This mode has its electric field directed in the azimuthal direction, i.e., $\boldsymbol{E} = E(r_c, z)\hat{\boldsymbol{\phi}}$. In figure 8.15 the next resonance is depicted. The frequency is 75.7 GHz. The lowest TM-mode is depicted in 8.16.

Problems in Chapter 8

8.1 Prove that the boundary conditions

$$\begin{cases} \mu_1 \hat{\boldsymbol{n}} \cdot \boldsymbol{H}_1 = \mu_2 \hat{\boldsymbol{n}} \cdot \boldsymbol{H}_2 \\ \epsilon_1 \hat{\boldsymbol{n}} \cdot \boldsymbol{E}_1 = \epsilon_2 \hat{\boldsymbol{n}} \cdot \boldsymbol{E}_2 \end{cases} \qquad \boldsymbol{r} \text{ on } S$$

follow from the boundary conditions in section 8.2.2.

*__8.2__ Determine the TE-modes in a planar waveguide $0 < y < b$ that for $0 < y < a$ is filled with a dielctric material with permittivity ϵ and for $a < y < b$ is filled with air. Determine also the transcendental equation that determines the longitudinal wavenumber.

8.3 Determine $v_m(\boldsymbol{\rho})$ and $w_m(\boldsymbol{\rho})$ in the cladding for a circular dielectric waveguide at the cut-off frequency and show that these functions satisfy the Laplace equation $\nabla_T^2 \psi(\boldsymbol{\rho}) = 0$.

*__8.4__ Show that there are no propagating modes in a circular waveguide when $n_2 > n_1$.

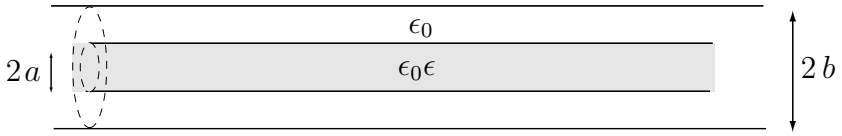


Figure 8.17: Geometry for problem 8.5.

*__8.5__ A circular hollow waveguide with radius b, is for $0 < \rho < a$ filled with a dielectric material with relative permittivity ϵ, see figure 8.17. Determine all axially symmetric modes.

Summary of chapter 8

Expressions for the fields in a circular dielectric waveguide

$$\begin{cases} v_m(\rho,\phi) = AJ_m(h\rho)\sin m\phi \\ w_m(\rho,\phi) = BJ_m(h\rho)\cos m\phi \end{cases} \qquad \rho < a$$

$$\begin{cases} v_m(\rho,\phi) = CK_m(q\rho)\sin m\phi \\ w_m(\rho,\phi) = DK_m(q\rho)\cos m\phi \end{cases} \qquad \rho > a$$

$$\frac{C}{A} = \frac{D}{B} = \frac{J_m(ha)}{K_m(qa)}$$

$$\frac{B}{A} = \frac{D}{C} = \frac{mk_z}{k_0}\left(\frac{1}{h^2a^2} + \frac{1}{q^2a^2}\right)\left(\frac{\mu_1 J'_m(ha)}{haJ_m(ha)} + \frac{\mu_2 K'_m(qa)}{qaK_m(qa)}\right)^{-1}$$

$$h = k_{1t}$$

$$q = -ik_{2t} = |k_{2t}|$$

TE- and TM-modes (m=0)

Cut-off frequencies (both cases)

$$f_{0n} = \frac{c_0}{2\pi a}\frac{x_{0n}}{\sqrt{n_1^2 - n_2^2}}$$

where $J_0(x_{0n}) = 0$

Characteristic equation for the TE-modes

$$\boxed{\begin{array}{l} A = 0,\ B \neq 0 \\ \dfrac{J_1(ha)}{haJ_0(ha)} = -\dfrac{\mu_2}{\mu_1}\dfrac{K_1(qa)}{qaK_0(qa)} \end{array}}$$

Characteristic equation for the TM-modes

$$\boxed{\begin{array}{l} B = 0,\ A \neq 0 \\ \dfrac{J_1(ha)}{haJ_0(ha)} = -\dfrac{\epsilon_2}{\epsilon_1}\dfrac{K_1(qa)}{qaK_0(qa)} \end{array}}$$

EH-modes ($m > 0$)

Cut-off frequencies:

$$\boxed{f_{mn}^{EH} = \frac{c_0}{2\pi a}\frac{x_{mn}}{\sqrt{n_1^2 - n_2^2}}, \quad J_m(x_{mn}) = 0}$$

Characteristic equation:

$$\boxed{\begin{array}{l} \dfrac{J_{m+1}(ha)}{haJ_m(ha)} = \dfrac{K_m'(qa)}{2qaK_m(qa)}\dfrac{\mu_1\epsilon_2 + \mu_2\epsilon_1}{\mu_1\epsilon_1} + \dfrac{m}{(ha)^2} \\ -\left[\left(\dfrac{K_m'(qa)}{2qaK_m(qa)}\dfrac{\mu_1\epsilon_2 - \mu_2\epsilon_1}{\mu_1\epsilon_1}\right)^2 + \left(\dfrac{mk_z}{k_1}\right)^2\left(\dfrac{1}{q^2a^2} + \dfrac{1}{h^2a^2}\right)^2\right]^{1/2} \end{array}}$$

HE-modes ($m > 0$)

Cut-off frequencies $m = 1$

$$f_{1n} = \frac{c_0}{2\pi a}\frac{x_{1n}}{\sqrt{n_1^2 - n_2^2}}, \quad \text{where} J_1(x_{1n}) = 0$$

Cut-off frequencies $m > 1$

$$f_{mn} = \frac{c_0}{2\pi a}\frac{x_{mn}}{\sqrt{n_1^2 - n_2^2}} \quad \text{where}$$
$$x_{mn}\frac{J_m(x_{mn})}{J_{m-1}(x_{mn})} = (m-1)\frac{n_1^2 + n_2^2}{n_2^2}$$

Characteristic equation

$$\frac{J_{m+1}(ha)}{haJ_m(ha)} = \frac{K'_m(qa)}{2qaK_m(qa)}\frac{\mu_1\epsilon_2 + \mu_2\epsilon_1}{\mu_1\epsilon_1} + \frac{m}{(ha)^2}$$
$$+ \left[\left(\frac{K'_m(qa)}{2qaK_m(qa)}\frac{\mu_1\epsilon_2 - \mu_2\epsilon_1}{\mu_1\epsilon_1}\right)^2 + \left(\frac{mk_z}{k_1}\right)^2\left(\frac{1}{q^2a^2} + \frac{1}{h^2a^2}\right)^2\right]^{1/2}$$

Effective index of refraction

$$n_e = k_z/k_0, \quad n_1 < n_e < n_2$$

Dispersion

$$D = D_m + D_v$$
$$D_m = -\frac{\lambda}{c_0}\frac{d^2n_1}{d\lambda^2}$$
$$D_v = -\frac{\lambda}{c_0}\frac{d^2n_e}{d\lambda^2}|_{n_1,n_2konst.}$$

Appendix A

Bessel functions

In wave propagation problems the Bessel differential equation often appears, especially in problems showing axial or spherical symmetries. This appendix collects some useful and important results for the solution of the Bessel differential equation. Moreover, the modified Bessel functions and the spherical Bessel and Hankel functions are presented. Additional results and some of the derivations are found in eg., Ref. 3.

A.1 Bessel and Hankel functions

The Bessel differential equation is

$$z^2 \frac{d^2}{dz^2} Z_n(z) + z \frac{d}{dz} Z_n(z) + (z^2 - n^2) Z_n(z) = 0 \tag{A.1}$$

where n is assumed integer[1].

There exist two linearly independent solutions to this differential equation. One is regular at the origin, $z = 0$, and this solution is the Bessel function $J_n(z)$ of order n. The argument z is a complex number. These solutions are often called cylindrical Bessel function of order n, which stresses the affinity to problems with the axial symmetry. The Bessel functions $J_n(z)$ are defined real-valued for a real argument z. An everywhere in the complex z-plane convergent power series is

$$J_n(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{z}{2}\right)^{n+2k} \tag{A.2}$$

We notice immediately that $J_n(z)$ is an even function for even n and odd for odd n, i.e.,

$$J_n(-z) = (-1)^n J_n(z)$$

A commonly used integral representation of the Bessel functions is

$$J_n(z) = \frac{1}{\pi} \int_0^{\pi} \cos\left(z \sin t - nt\right) \, \mathrm{d}t = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{\mathrm{i}z \cos t} \mathrm{e}^{\mathrm{i}n\left(t - \frac{1}{2}\pi\right)} \, \mathrm{d}t \tag{A.3}$$

[1] A more general definition with eg., complex-valued n is also possible, but the expressions and the results often differ.

	Root j			
Order n	1	2	3	4
0	2.405	5.520	8.654	11.79
1	3.832	7.016	10.17	13.32
2	5.136	8.417	11.62	14.80
3	6.380	9.761	13.01	16.22
4	7.588	11.06	14.37	17.62

Table A.1: Table of the roots ξ_{nj} to $J_n(z)$.

From this integral representation, we see that the Bessel functions for positive and negative *integer* orders, n, are related to each other.

$$J_{-n}(z) = (-1)^n J_n(z)$$

The power series representation in (A.2) implies that for small arguments we have

$$J_n(z) = \frac{1}{n!}\left(\frac{z}{2}\right)^n + O(z^{n+2})$$

For large arguments hold ($-\pi < \arg z < \pi$)

$$J_n(z) = \left(\frac{2}{\pi z}\right)^{1/2}\left\{P_n(z)\cos\left(z - \frac{n\pi}{2} - \frac{\pi}{4}\right) - Q_n(z)\sin\left(z - \frac{n\pi}{2} - \frac{\pi}{4}\right)\right\}$$

where the functions $P_n(z)$ and $Q_n(z)$ have the following asymptotic expansions ($\nu = 4n^2$)

$$\begin{cases} P_n(z) \sim 1 - \dfrac{(\nu-1)(\nu-9)}{2!(8z)^2} + \dfrac{(\nu-1)(\nu-9)(\nu-25)(\nu-49)}{4!(8z)^4} - \dots \\ Q_n(z) \sim \dfrac{\nu-1}{8z} - \dfrac{(\nu-1)(\nu-9)(\nu-25)}{3!(8z)^3} + \dots \end{cases} \tag{A.4}$$

The roots of the Bessel function $J_n(z)$ are all real, and the first roots, ξ_{nj}, are listed in Table A.1. The derivative of the Bessel function $J_n(z)$ has also only real roots, η_{nj}, and the first ones are listed in Table A.2. Larger roots (larger j values) are asymptotically given by

$$\xi_{nj} = j\pi + \left(n - \frac{1}{2}\right)\frac{\pi}{2}, \qquad \eta_{nj} = j\pi + \left(n - \frac{3}{2}\right)\frac{\pi}{2}$$

Another, linearly independent solution to the Bessel differential equation, which is real-valued for real arguments, is the Neumann function[2] $N_n(z)$. The power series expansion is

$$\begin{aligned} N_n(z) = {} & \frac{2}{\pi}\left(\ln\left(\frac{z}{2}\right) + \gamma - \frac{1}{2}\sum_{k=1}^{n}\frac{1}{k}\right)J_n(z) \\ & - \frac{1}{\pi}\sum_{k=0}^{\infty}(-1)^k\frac{\left(\frac{z}{2}\right)^{n+2k}}{k!(n+k)!}\sum_{l=1}^{k}\left(\frac{1}{l} + \frac{1}{l+n}\right) \\ & - \frac{1}{\pi}\sum_{k=0}^{n-1}\frac{(n-k-1)!}{k!}\left(\frac{z}{2}\right)^{-n+2k} \end{aligned}$$

[2]These solutions are also called Bessel functions of the second kind.

	Root j			
Order n	1	2	3	4
0	3.832	7.016	10.17	13.32
1	1.841	5.331	8.536	11.71
2	3.054	6.706	9.970	13.17
3	4.201	8.015	11.34	14.59
4	5.317	9.282	12.68	15.96

Table A.2: Table of the roots η_{nj} to $J_n'(z)$.

where the Euler constant $\gamma = 0.577\,215\,66\ldots$, and where all sums are dined as zero if the summation index exceeds the upper summation limit. This solution is singular at the origin $z = 0$. For small arguments the dominant contribution is

$$N_0(z) = \frac{2}{\pi}\left(\ln\left(\frac{z}{2}\right) + \gamma\right) + O(z^2)$$
$$N_n(z) = -\frac{(n-1)!}{\pi}\left(\frac{z}{2}\right)^{-n} + \ldots$$

For large arguments the Neumann function has an asymptotic expansion ($-\pi < \arg z < \pi$)

$$N_n(z) = \left(\frac{2}{\pi z}\right)^{1/2}\left(P_n(z)\sin\left(z - \frac{n\pi}{2} - \frac{\pi}{4}\right) + Q_n(z)\cos\left(z - \frac{n\pi}{2} - \frac{\pi}{4}\right)\right)$$

where the functions $P_n(z)$ and $Q_n(z)$ are given by (A.4).

In the solution of scattering problems, linear combinations of Bessel and Neumann functions, i.e., the Hankel functions, $H_n^{(1)}(z)$ and $H_n^{(2)}(z)$ of the first and the second kind, respectively, are natural[3]. These are defined as

$$H_n^{(1)}(z) = J_n(z) + \mathrm{i}N_n(z)$$
$$H_n^{(2)}(z) = J_n(z) - \mathrm{i}N_n(z)$$

The Hankel functions of the first and second kind have integral representations

$$H_n^{(1)}(z) = \frac{2}{\mathrm{i}\pi}\mathrm{e}^{-\mathrm{i}n\frac{\pi}{2}}\int_0^\infty \mathrm{e}^{\mathrm{i}z\cosh s}\cosh ns\,ds, \qquad 0 < \arg z < \pi$$
$$H_n^{(2)}(z) = \frac{2\mathrm{i}}{\pi}\mathrm{e}^{\mathrm{i}n\frac{\pi}{2}}\int_0^\infty \mathrm{e}^{-\mathrm{i}z\cosh s}\cosh ns\,ds, \qquad -\pi < \arg z < 0$$

For large argumens, the Hankel functions have asymptotic expansions

$$\begin{aligned} H_n^{(1)}(z) &= \left(\frac{2}{\pi z}\right)^{1/2}\mathrm{e}^{\mathrm{i}\left(z-\frac{n\pi}{2}-\frac{\pi}{4}\right)}\left(P_n(z) + \mathrm{i}Q_n(z)\right), \qquad -\pi < \arg z < 2\pi \\ H_n^{(2)}(z) &= \left(\frac{2}{\pi z}\right)^{1/2}\mathrm{e}^{-\mathrm{i}\left(z-\frac{n\pi}{2}-\frac{\pi}{4}\right)}\left(P_n(z) - \mathrm{i}Q_n(z)\right), \qquad -2\pi < \arg z < \pi \end{aligned} \tag{A.5}$$

[3]These also called Bessel functions of the third kind.

where the functions $P_n(z)$ and $Q_n(z)$ are given by (A.4).

Solutions to the Bessel differential equation of different order are related to each other by recursion relations. Some of the more important ones are ($n = 0, 1, 2, \ldots, m = 0, 1, 2, \ldots$)[4]

$$
\begin{aligned}
&Z_{n-1}(z) - Z_{n+1}(z) = 2Z_n'(z) \\
&Z_{n-1}(z) + Z_{n+1}(z) = \frac{2n}{z} Z_n(z) \\
&Z_{n+1}(z) = \frac{n}{z} Z_n(z) - Z_n'(z) \\
&Z_n'(z) = Z_{n-1}(z) - \frac{n}{z} Z_n(z) \\
&\left(\frac{d}{z\,dz}\right)^m \left[z^n Z_n(z)\right] = z^{n-m} Z_{n-m}(z) \\
&\left(\frac{d}{z\,dz}\right)^m \left[z^{-n} Z_n(z)\right] = (-1)^m z^{-n-m} Z_{n+m}(z)
\end{aligned}
$$

Here $Z_n(z)$ is a fixed arbitrary linear combination of $J_n(x)$, $N_n(x)$, $H_n^{(1)}(x)$ or $H_n^{(2)}(x)$. Specifically, we have

$$J_1(z) = -J_0'(z)$$

which is frequently used in the analysis in this textbook.

Some useful indefinite integrals with solutions to the Bessel differential equation, which are often used in the text, are ($n = 0, 1, 2, \ldots$)

$$
\begin{aligned}
&\int x^{n+1} Z_n(x)\,dx = x^{n+1} Z_{n+1}(x) = -x^{n+1}\left(Z_n'(x) - \frac{n}{x} Z_n(x)\right) \\
&\int x^{-n+1} Z_n(x)\,dx = -x^{-n+1} Z_{n-1}(x) = -x^{-n+1}\left(Z_n'(x) + \frac{n}{x} Z_n(x)\right) \\
&\int x \left(Z_n(x)\right)^2 dx = \frac{x^2}{2}\left[\left(Z_n(x)\right)^2 - Z_{n-1}(x) Z_{n+1}(x)\right] \\
&\qquad\qquad = \frac{x^2}{2}\left(Z_n'(x)\right)^2 + \frac{1}{2}(x^2 - n^2)\left(Z_n(x)\right)^2
\end{aligned}
$$

As above, $Z_n(x)$ is an arbitrary linear combination of $J_n(x)$, $N_n(x)$, $H_n^{(1)}(x)$ or $H_n^{(2)}(x)$. Some additional — more complex —but useful determined integrals are ($n = 0, 1, 2, \ldots, m = 0, 1, 2, \ldots$)

$$
\begin{aligned}
&\int \left[(\alpha^2 - \beta^2)x - \frac{m^2 - n^2}{x}\right] Z_m(\alpha x) Y_n(\beta x)\,dx = \beta x Z_m(\alpha x) Y_{n-1}(\beta x) \\
&\qquad\qquad - \alpha x Z_{m-1}(\alpha x) Y_n(\beta x) + (m - n) Z_m(\alpha x) Y_m(\beta x) \\
&\int x Z_m(\alpha x) Y_m(\beta x)\,dx = \frac{\beta x Z_m(\alpha x) Y_{m-1}(\beta x) - \alpha x Z_{m-1}(\alpha x) Y_m(\beta x)}{\alpha^2 - \beta^2} \\
&\int \frac{Z_m(\alpha x) Y_n(\alpha x)}{x}\,dx = \alpha x \frac{Z_{m-1}(\alpha x) Y_n(\alpha x) - Z_m(\alpha x) Y_{n-1}(\alpha x)}{m^2 - n^2} - \frac{Z_m(\alpha x) Y_n(\alpha x)}{m + n}
\end{aligned}
$$

[4] These recursion relations hold for non-integer values of n, t.ex. $n = 1/2$. The index m, however, must be an integer.

Here, $Z_n(\alpha x)$ and $Y_n(\beta x)$ is an arbitrary linear combination of $J_n(x)$, $N_n(x)$, $H_n^{(1)}(x)$ or $H_n^{(2)}(x)$.

For Bessel functions, $J_n(z)$, Neumann functions, $N_n(z)$, and Hankel functions, $H_n^{(1)}(z)$ or $H_n^{(2)}(z)$, we have for a complex argument z

$$\begin{cases} J_n(z^*) = (J_n(z))^* \\ N_n(z^*) = (N_n(z))^* \end{cases} \qquad \begin{cases} H_n^{(1)}(z^*) = \left(H_n^{(2)}(z)\right)^* \\ H_n^{(2)}(z^*) = \left(H_n^{(1)}(z)\right)^* \end{cases}$$

The Graf addition theorem for Bessel functions is useful. Let $Z_n(x)$ be any linear combination of $J_n(x)$, $N_n(x)$, $H_n^{(1)}(x)$ and $H_n^{(2)}(x)$. The Graf addition theorem is

$$Z_n(w)\begin{pmatrix}\cos n\phi \\ \sin n\phi\end{pmatrix} = \sum_{k=-\infty}^{\infty} Z_{n+k}(u)J_k(v)\begin{pmatrix}\cos k\alpha \\ \sin k\alpha\end{pmatrix}, \qquad \left|v\mathrm{e}^{\pm \mathrm{i}\alpha}\right| < |u|$$

where w is

$$w = \sqrt{u^2 + v^2 - 2uv\cos\alpha}$$

A.1.1 Useful integrals

Some integrals related to Bessel functions used in this book are derived in this subsection. We start with the integral representation for integer order n, (A.3)

$$\int_0^{2\pi} \mathrm{e}^{\mathrm{i}z\cos\phi}\mathrm{e}^{\mathrm{i}n\phi}\,\mathrm{d}\phi = 2\pi\mathrm{i}^n J_n(z)$$

From this we easily conclude by a simple change of variables that

$$\int_0^{2\pi} \mathrm{e}^{\mathrm{i}z\cos(\phi-\alpha)}\mathrm{e}^{\mathrm{i}n\phi}\,\mathrm{d}\phi = \mathrm{e}^{\mathrm{i}n\alpha}\int_{-\alpha}^{2\pi-\alpha} \mathrm{e}^{\mathrm{i}z\cos\psi}\mathrm{e}^{\mathrm{i}n\psi}\,\mathrm{d}\psi = 2\pi\mathrm{i}^n J_n(z)\mathrm{e}^{\mathrm{i}n\alpha}$$

This integral is a function of the variables z and α.

In scattering problems the following integral is often encountered:

$$\boldsymbol{I}(\boldsymbol{\tau},\rho) = \int_0^{2\pi} \hat{\boldsymbol{\phi}}\,\mathrm{e}^{\mathrm{i}\boldsymbol{\rho}\cdot\boldsymbol{\tau}}\mathrm{e}^{\mathrm{i}n\phi}\,\mathrm{d}\phi = \int_0^{2\pi} \hat{\boldsymbol{\phi}}\,\mathrm{e}^{\mathrm{i}\rho\tau\cos(\phi-\alpha)}\mathrm{e}^{\mathrm{i}n\phi}\,\mathrm{d}\phi$$

where the position vector in the x-y-plane is denoted $\boldsymbol{\rho} = \hat{\boldsymbol{x}}\rho\cos\phi + \hat{\boldsymbol{y}}\rho\sin\phi$ and $\boldsymbol{\tau} = \hat{\boldsymbol{x}}\tau\cos\alpha + \hat{\boldsymbol{y}}\tau\sin\alpha$. This integral can be rewritten as

$$\boldsymbol{I}(\boldsymbol{\tau},\rho) = \frac{1}{\mathrm{i}\rho}\hat{\boldsymbol{z}}\times\nabla_{\boldsymbol{\tau}}\int_0^{2\pi} \mathrm{e}^{\mathrm{i}\rho\tau\cos(\phi-\alpha)}\mathrm{e}^{\mathrm{i}n\phi}\,\mathrm{d}\phi$$

where the nabla-operator in the τ variable is denoted $\nabla_{\boldsymbol{\tau}}$, and where we used

$$\hat{\boldsymbol{z}}\times\nabla_{\boldsymbol{\tau}}\mathrm{e}^{\mathrm{i}\boldsymbol{\rho}\cdot\boldsymbol{\tau}} = \mathrm{i}\hat{\boldsymbol{z}}\times\boldsymbol{\rho}\mathrm{e}^{\mathrm{i}\boldsymbol{\rho}\cdot\boldsymbol{\tau}} = \mathrm{i}\rho\hat{\boldsymbol{z}}\times\hat{\boldsymbol{\rho}}\mathrm{e}^{\mathrm{i}\boldsymbol{\rho}\cdot\boldsymbol{\tau}} = \mathrm{i}\rho\hat{\boldsymbol{\phi}}\mathrm{e}^{\mathrm{i}\boldsymbol{\rho}\cdot\boldsymbol{\tau}}$$

From this we infer

$$\boldsymbol{I}(\boldsymbol{\tau},\rho) = \frac{2\pi\mathrm{i}^{n-1}}{\rho}\hat{\boldsymbol{z}}\times\nabla_{\boldsymbol{\tau}}J_n(\rho\tau)\mathrm{e}^{\mathrm{i}n\alpha} = 2\pi\mathrm{i}^{n-1}\hat{\boldsymbol{z}}\times\left(\hat{\boldsymbol{\tau}}J_n'(\rho\tau)\mathrm{e}^{\mathrm{i}n\alpha} + \mathrm{i}n\hat{\boldsymbol{\alpha}}\frac{J_n(\rho\tau)}{\rho\tau}\mathrm{e}^{\mathrm{i}n\alpha}\right)$$

A.2 Modified Bessel functions

The modified Bessel's differential equation is

$$z^2\frac{d^2}{dz^2}Z_n(z) + z\frac{d}{dz}Z_n(z) - (z^2+n^2)Z_n(z) = 0$$

where n is assumed to be a positive integer. A comparison with Bessel's differential equation in seaction A.1 shows that the solutions to the modified Bessel's equation are Bessel functions with argument iz. It is suitable to introduce a notation for the two independent solutions to the Bessel's modified differential equations:

$$I_n(z) = \mathrm{i}^{-n}J_n(\mathrm{i}z)$$
$$K_n(z) = \frac{\mathrm{i}^{n+1}\pi}{2}H_n^{(1)}(\mathrm{i}z)$$

The factors multiplying the Bessel- and Hankel functions in the right hand sides make both $I_n(z)$ and $K_n(z)$ real for real arguments z. The function $I_n(z)$ is regular (finite) at $z = 0$, while $K_n(z)$ is singular when $z \to 0$.

From the properties of the Bessel functions it is seen that $I_n(z)$ is an even function for even n and an odd function for odd n, i.e.,

$$I_n(-z) = (-1)^n I_n(z)$$

and

$$I_{-n}(z) = (-1)^n I_n(z)$$

This result can also be obtained from the power series expansion of $I_n(z)$ around $z = 0$.

$$I_n(z) = \sum_{k=0}^{\infty}\frac{1}{k!(n+k)!}\left(\frac{z}{2}\right)^{n+2k}$$

The corresponding power series expansion of $K_n(z)$ around $z = 0$ is

$$\begin{aligned}K_n(z) = (-1)^{n+1}&\left(\ln\left(\frac{z}{2}\right) + \gamma - \frac{1}{2}\sum_{k=1}^{n}\frac{1}{k}\right)I_n(z)\\ &+ (-1)^n\frac{1}{2}\sum_{k=0}^{\infty}\frac{\left(\frac{z}{2}\right)^{n+2k}}{k!(n+k)!}\sum_{l=1}^{k}\left(\frac{1}{l}+\frac{1}{l+n}\right)\\ &+\frac{1}{2}\sum_{k=0}^{n-1}(-1)^k\frac{(n-k-1)!}{k!}\left(\frac{z}{2}\right)^{-n+2k}\end{aligned}$$

where Euler's constant is $\gamma = 0.577\,215\,66\ldots$, and where all sums are defined to be zero if the upper summation limit is smaller than the lower summation index. From these power series expansions it is seen that the dominating contributions for small arguments are

$$I_n(z) = \frac{1}{n!}\left(\frac{z}{2}\right)^n + O(z^{n+2})$$

and

$$\begin{cases} K_0(z) = -\left(\ln\left(\frac{z}{2}\right)+\gamma\right) + O(z^2) \\ K_n(z) = \frac{(n-1)!}{2}\left(\frac{z}{2}\right)^{-n} + \ldots, \quad n \neq 0 \end{cases} \tag{A.6}$$

For large arguments

$$I_n(z) = \frac{e^z}{\sqrt{2\pi z}} P_n(z), \quad -\frac{\pi}{2} < \arg z < \frac{\pi}{2}$$

and

$$K_n(z) = \sqrt{\frac{\pi}{2z}} \mathrm{e}^{-z} Q_n(z), \quad -\frac{3\pi}{2} < \arg z < \frac{3\pi}{2}$$

where the functions $P_n(z)$ and $Q_n(z)$ have the following asymptotic expansions ($\nu = 4n^2$):

$$\begin{cases} P_n(z) \sim 1 - \dfrac{\nu-1}{8z} + \dfrac{(\nu-1)(\nu-9)}{2!(8z)^2} - \dfrac{(\nu-1)(\nu-9)(\nu-25)}{3!(8z)^3} + \dots \\ Q_n(z) \sim 1 + \dfrac{\nu-1}{8z} + \dfrac{(\nu-1)(\nu-9)}{2!(8z)^2} + \dfrac{(\nu-1)(\nu-9)(\nu-25)}{3!(8z)^3} + \dots \end{cases}$$

There are recursion relations that relate modified Bessel functions of differnet order. For $I_n(z)$, $n = 0, 1, 2, \dots$, the most important relations are ($m = 0, 1, 2, \dots$)[5]

$$\begin{aligned} I_{n-1}(z) + I_{n+1}(z) &= 2I_n'(z) \\ I_{n-1}(z) - I_{n+1}(z) &= \frac{2n}{z} I_n(z) \\ I_{n+1}(z) &= I_n'(z) - \frac{n}{z} I_n(z) \\ I_n'(z) &= I_{n-1}(z) - \frac{n}{z} I_n(z) \\ \left(\frac{d}{z\,dz}\right)^m [z^n I_n(z)] &= z^{n-m} I_{n-m}(z) \\ \left(\frac{d}{z\,dz}\right)^m \left[z^{-n} I_n(z)\right] &= z^{-n-m} I_{n+m}(z) \end{aligned}$$

och fr $K_n(z)$ r motsvarande ($n = 0, 1, 2, \dots, m = 0, 1, 2, \dots$)[5]

$$\begin{aligned} K_{n-1}(z) + K_{n+1}(z) &= -2K_n'(z) \\ K_{n+1}(z) - K_{n-1}(z) &= \frac{2n}{z} K_n(z) \\ K_{n+1}(z) &= \frac{n}{z} K_n(z) - K_n'(z) \\ K_n'(z) &= -K_{n-1}(z) - \frac{n}{z} K_n(z) \\ \left(\frac{d}{z\,dz}\right)^m [z^n K_n(z)] &= (-1)^m z^{n-m} K_{n-m}(z) \\ \left(\frac{d}{z\,dz}\right)^m \left[z^{-n} K_n(z)\right] &= (-1)^m z^{-n-m} K_{n+m}(z) \end{aligned}$$

A.3 Spherical Bessel and Hankel functions

The spherical Bessel and Hankel functions show up in scattering problems, especially when we express a solution in the spherical coordinate system (r, θ, ϕ).

[5]These recursion formulas are also valid for non-integer values of n, eg., $n = 1/2$. It is required that m is a non-negative integer.

The spherical Bessel differential equation is

$$z^2\frac{d^2}{dz^2}Z_n(z) + 2z\frac{d}{dz}Z_n(z) + (z^2 - n(n+1))Z_n(z) = 0$$

where n is assumed to be a non-negative integer and the argument z a complex number.

A solution to this differential equation is the spherical Bessel functions $j_n(z)$, which are defined by the power series expansion[6]

$$j_n(z) = 2^n z^n \sum_{k=0}^{\infty} \frac{(-1)^k (k+n)!}{k!(2k+2n+1)!} z^{2k} \tag{A.7}$$

This solution is real-valued for real arguments, it is regular at the origin, $z = 0$, and its power series is absolutely convergent in the entire complex plane. From the power series, we also conclude that $j_n(z)$ is an even function for even integers n and an odd function for odd n, i.e.,

$$j_n(-z) = (-1)^n j_n(z)$$

The other, linearly independent solution, which is real-valued for real arguments, is the spherical Neumann function $n_n(z)$. Its power series is

$$n_n(z) = \frac{(-1)^{n+1}2^n\pi^{1/2}}{z^{n+1}} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(k-n-1/2)!} \left(\frac{z}{2}\right)^{2k} = \frac{(-1)^{n+1}}{2^n z^{n+1}} \sum_{k=0}^{\infty} \frac{(-1)^k (k-n)!}{k!(2k-2n)!} z^{2k}$$

We see that this solution is singular at the origin $z = 0$.

In scattering problems, we often use linear combinations of the spherical Bessel and Neumann functions. These are the spherical Hankel functions, $h_n^{(1)}(z)$ and $h_n^{(2)}(z)$ of the first and second kind, respectively. These are defined as

$$h_n^{(1)}(z) = j_n(z) + \mathrm{i}n_n(z)$$
$$h_n^{(2)}(z) = j_n(z) - \mathrm{i}n_n(z)$$

The spherical Bessel, Neumann, and Hankel functions for $l = 0$ are

$$\begin{cases} j_0(z) = \dfrac{\sin z}{z} \\ n_0(z) = -\dfrac{\cos z}{z} \end{cases} \qquad \begin{cases} h_0^{(1)}(z) = -\dfrac{\mathrm{i}}{z}\mathrm{e}^{\mathrm{i}z} \\ h_0^{(2)}(z) = \dfrac{\mathrm{i}}{z}\mathrm{e}^{-\mathrm{i}z} \end{cases}$$

and for $l = 1$ they are

$$\begin{cases} j_1(z) = \dfrac{\sin z}{z^2} - \dfrac{\cos z}{z} \\ n_1(z) = -\dfrac{\cos z}{z^2} - \dfrac{\sin z}{z} \end{cases} \qquad \begin{cases} h_1^{(1)}(z) = \mathrm{e}^{\mathrm{i}z}\left(-\dfrac{1}{z} - \dfrac{\mathrm{i}}{z^2}\right) \\ h_1^{(2)}(z) = \mathrm{e}^{-\mathrm{i}z}\left(-\dfrac{1}{z} + \dfrac{\mathrm{i}}{z^2}\right) \end{cases}$$

[6]Between the spherical Bessel function, $j_l(z)$, and the cylindrical Bessel function, $J_n(z)$, see Appendix A.1, there is a relation

$$j_l(z) = \left(\frac{\pi}{2z}\right)^{1/2} J_{l+1/2}(z)$$

Also between the other type of solutions to the spherical Bessel equation and the corresponding solutions to the cylindrical Bessel equation, (A.1), there are similar relations.

The roots of the spherical Hankel functions play an important role in the interpretation of scattering contributions. These are the SEM-poles (Singular Expansion Method). Specifically, the only root in the finite complex plane of $h_1^{(1)}(z)$ is $z = -\mathrm{i}$, and the roots in the finite complex plane of $(zh_1^{(1)}(z))'$ are $z = (\pm\sqrt{3} - \mathrm{i})/2$.

From the power series representation in (A.7) we obtain for small arguments

$$j_n(z) = \frac{2^n n! z^n}{(2n+1)!} + O(z^{n+2})$$

and

$$n_n(z) = -\frac{(2n)!}{2^n n! z^{n+1}} + \dots$$

Specifically, for $l = 1$

$$\begin{cases} j_1(z) \approx \dfrac{z}{3}, \text{ as } z \to 0 \\ h_1^{(1)}(z) \approx \dfrac{1}{\mathrm{i}z^2}, \text{ as } z \to 0 \end{cases} \qquad \begin{cases} \dfrac{(zj_1(z))'}{z} \approx \dfrac{2}{3}, \text{ as } z \to 0 \\ \dfrac{(zh_1^{(1)}(z))'}{z} \approx -\dfrac{1}{\mathrm{i}z^3}, \text{ as } z \to 0 \end{cases}$$

An alternative expansion of the spherical Bessel and Neumann functions in finite trigonometric series are:

$$\begin{cases} j_n(z) = \dfrac{1}{z}\left\{ p_n(z)\sin\left(z - \dfrac{n\pi}{2}\right) + q_n(z)\cos\left(z - \dfrac{n\pi}{2}\right)\right\} \\ n_n(z) = \dfrac{1}{z}\left\{ q_n(z)\sin\left(z - \dfrac{n\pi}{2}\right) - p_n(z)\cos\left(z - \dfrac{n\pi}{2}\right)\right\} \end{cases} \tag{A.8}$$

The explicit for of the finite sums are

$$\begin{cases} p_n(z) = \displaystyle\sum_{k=0}^{[n/2]} (-1)^k \frac{(n+2k)!}{(2k)!(n-2k)!}\frac{1}{(2z)^{2k}} = 1 - \frac{b_n}{z^2} + O\left(z^{-4}\right) \\ q_n(z) = \displaystyle\sum_{k=0}^{[(n-1)/2]} (-1)^k \frac{(n+2k+1)!}{(2k+1)!(n-2k-1)!}\frac{1}{(2z)^{2k+1}} = \frac{a_n}{z} + O\left(z^{-3}\right) \end{cases} \tag{A.9}$$

where $[\cdot]$ denotes the integer part of the argument, and where $a_n = (n+1)n/2$ and $b_n = (n+2)(n+1)n(n-1)/8$. Note that these series are finite and that the series, by definition, is zero if the upper limit of summation is negative.

For the spherical Hankel functions we have

$$\begin{cases} h_n^{(1)}(z) = \dfrac{\mathrm{e}^{\mathrm{i}z - \mathrm{i}(n+1)\pi/2}}{z}\displaystyle\sum_{k=0}^{n} \frac{(n+k)!}{k!(n-k)!}\frac{1}{(-2\mathrm{i}z)^k} = \frac{\mathrm{e}^{\mathrm{i}z - \mathrm{i}(n+1)\pi/2}}{z}\left(p_n(z) + \mathrm{i}q_n(z)\right) \\ h_n^{(2)}(z) = \dfrac{\mathrm{e}^{-\mathrm{i}z + \mathrm{i}(n+1)\pi/2}}{z}\displaystyle\sum_{k=0}^{n} \frac{(n+k)!}{k!(n-k)!}\frac{1}{(2\mathrm{i}z)^k} = \frac{\mathrm{e}^{-\mathrm{i}z + \mathrm{i}(n+1)\pi/2}}{z}\left(p_n(z) - \mathrm{i}q_n(z)\right) \end{cases} \tag{A.10}$$

In scattering problems another combination of spherical Bessel and Hankel functions occurs, i.e.,

$$\begin{cases} \dfrac{(zj_n(z))'}{z} = \dfrac{1}{z}\left\{ P_n(z)\sin\left(z-\dfrac{n\pi}{2}\right) + Q_n(z)\cos\left(z-\dfrac{n\pi}{2}\right)\right\} \\ \dfrac{\left(zh_n^{(1)}(z)\right)'}{z} = \dfrac{\mathrm{e}^{\mathrm{i}z-\mathrm{i}(n+1)\pi/2}}{z}\left(P_n(z)+\mathrm{i}Q_n(z)\right) \\ \dfrac{\left(zh_n^{(2)}(z)\right)'}{z} = \dfrac{\mathrm{e}^{-\mathrm{i}z+\mathrm{i}(n+1)\pi/2}}{z}\left(P_n(z)-\mathrm{i}Q_n(z)\right) \end{cases} \tag{A.11}$$

where

$$\begin{cases} P_n(z) = p_n'(z) - q_n(z) = -\dfrac{a_n}{z} + O\left(z^{-3}\right) \\ Q_n(z) = q_n'(z) + p_n(z) = 1 - \dfrac{a_n+b_n}{z^2} + O\left(z^{-4}\right) \end{cases} \tag{A.12}$$

Between different solutions of spherical Bessel functions there are recursion relations. The most important ones are ($n = 0, 1, 2, \ldots, m = 0, 1, 2, \ldots$)

$$\begin{cases} f_{n-1}(z) + f_{n+1}(z) = \dfrac{2n+1}{z} f_n(z) \\ nf_{n-1}(z) - (n+1)f_{n+1}(z) = (2n+1)f_n'(z) \\ nf_n(z) - zf_{n+1}(z) = zf_n'(z) \\ \left(\dfrac{d}{z\,dz}\right)^m \left[z^{n+1}f_n(z)\right] = z^{n-m+1}f_{n-m}(z) \\ \left(\dfrac{d}{z\,dz}\right)^m \left[z^{-n}f_n(z)\right] = (-1)^m z^{-n-m}f_{n+m}(z) \end{cases} \tag{A.13}$$

Here $f_n(z)$ is a spherical Bessel function, $j_n(z)$, a spherical Neumann function, $n_n(z)$, or the Hankel functions $h_n^{(1)}(z)$ or $h_n^{(2)}(z)$. Useful, especially in the analysis of scattering by vector spherical waves, is also:

$$\frac{(zf_n(z))'}{z} = \frac{f_n(z)}{z} + f_n'(z) = \frac{(n+1)f_{n-1}(z) - nf_{n+1}(z)}{2n+1} = \frac{(n+1)f_n(z)}{z} - f_{n+1}(z) \tag{A.14}$$

The Wronskian relation for $j_n(z)$ and $n_n(z)$ is

$$j_n(z)n_n'(z) - j_n'(z)n_n(z) = \frac{1}{z^2} \tag{A.15}$$

and for $j_n(z)$ and $h_n^{(1)}(z)$ it is

$$j_n(z)h_n^{(1)\prime}(z) - j_n'(z)h_n^{(1)}(z) = \frac{\mathrm{i}}{z^2} \tag{A.16}$$

Appendix B

∇ in curvilinear coordinate systems

In this appendix some important expressions with the ∇-operator in two curvilinear coordinate systems, cylindrical and spherical, are collected. For completeness we start with the Cartesian coordinate system.

B.1 Cartesian coordinate system

The Cartesian coordinates (x, y, z) is the most basic coordinate system. The gradient and the Laplace-operator of a scalar field $\psi(x, y, z)$ in this coordinate system are

$$\nabla\psi = \hat{\boldsymbol{x}}\frac{\partial\psi}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial\psi}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial\psi}{\partial z}$$
$$\nabla^2\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}$$

The divergence, the curl, and the Laplace-operator of a vector field $\boldsymbol{A}(x, y, z) = \hat{\boldsymbol{x}}A_x(x, y, z) + \hat{\boldsymbol{y}}A_y(x, y, z) + \hat{\boldsymbol{z}}A_z(x, y, z)$ are

$$\nabla \cdot \boldsymbol{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$
$$\nabla \times \boldsymbol{A} = \hat{\boldsymbol{x}}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) + \hat{\boldsymbol{y}}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) + \hat{\boldsymbol{z}}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)$$
$$\nabla^2\boldsymbol{A} = \hat{\boldsymbol{x}}\nabla^2 A_x + \hat{\boldsymbol{y}}\nabla^2 A_y + \hat{\boldsymbol{z}}\nabla^2 A_z$$

B.2 Circular cylindrical (polar) coordinate system

We now treat the first curvilinear coordinate system, and start with the circular cylindrical coordinate system (ρ, ϕ, z) defined by

$$\begin{cases} \rho = \sqrt{x^2+y^2} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2+y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2+y^2}} & y < 0 \end{cases} \\ z = z \end{cases}$$

The gradient and the Laplace-operator of a scalar field $\psi(\rho, \phi, z)$ in this coordinate system are

$$\begin{aligned} \nabla\psi &= \hat{\boldsymbol{\rho}}\frac{\partial\psi}{\partial\rho} + \hat{\boldsymbol{\phi}}\frac{1}{\rho}\frac{\partial\psi}{\partial\phi} + \hat{\boldsymbol{z}}\frac{\partial\psi}{\partial z} \\ \nabla^2\psi &= \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\psi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\psi}{\partial\phi^2} + \frac{\partial^2\psi}{\partial z^2} \end{aligned}$$

The divergence, the curl, and the Laplace-operator of a vector field $\boldsymbol{A}(\rho, \phi, z) = \hat{\boldsymbol{\rho}}A_\rho(\rho, \phi, z) + \hat{\boldsymbol{\phi}}A_\phi(\rho, \phi, z) + \hat{\boldsymbol{z}}A_z(\rho, \phi, z)$ are

$$\begin{aligned} \nabla \cdot \boldsymbol{A} &= \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho A_\rho) + \frac{1}{\rho}\frac{\partial A_\phi}{\partial\phi} + \frac{\partial A_z}{\partial z} \\ \nabla \times \boldsymbol{A} &= \hat{\boldsymbol{\rho}}\left(\frac{1}{\rho}\frac{\partial A_z}{\partial\phi} - \frac{\partial A_\phi}{\partial z}\right) + \hat{\boldsymbol{\phi}}\left(\frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial\rho}\right) + \hat{\boldsymbol{z}}\frac{1}{\rho}\left(\frac{\partial}{\partial\rho}(\rho A_\phi) - \frac{\partial A_\rho}{\partial\phi}\right) \\ \nabla^2\boldsymbol{A} &= \hat{\boldsymbol{\rho}}\left(\nabla^2 A_\rho - \frac{A_\rho}{\rho^2} - \frac{2}{\rho^2}\frac{\partial A_\phi}{\partial\phi}\right) + \hat{\boldsymbol{\phi}}\left(\nabla^2 A_\phi - \frac{A_\phi}{\rho^2} + \frac{2}{\rho^2}\frac{\partial A_\rho}{\partial\phi}\right) + \hat{\boldsymbol{z}}\nabla^2 A_z \end{aligned}$$

B.3 Spherical coordinates system

The spherical coordinate system (r, θ, ϕ) (polar angle θ and the azimuth angle ϕ) is defined by

$$\begin{cases} r = \sqrt{x^2+y^2+z^2} \\ \theta = \arccos \frac{z}{\sqrt{x^2+y^2+z^2}} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2+y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2+y^2}} & y < 0 \end{cases} \end{cases}$$

The gradient and the Laplace-operator of a scalar field $\psi(r, \theta, \phi)$ in this coordinate system are

$$\begin{aligned} \nabla\psi &= \hat{\boldsymbol{r}}\frac{\partial\psi}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial\psi}{\partial\theta} + \hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\phi} \\ \nabla^2\psi &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} \\ &= \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} \end{aligned}$$

and the divergence, the curl, and the Laplace-operator of a vector field $\boldsymbol{A}(r,\theta,\phi) = \hat{\boldsymbol{r}}A_r(r,\theta,\phi) + \hat{\boldsymbol{\theta}}A_\theta(r,\theta,\phi) + \hat{\boldsymbol{\phi}}A_\phi(r,\theta,\phi)$ are

$$\begin{aligned}
\nabla \cdot \boldsymbol{A} &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 A_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta A_\theta\right) + \frac{1}{r\sin\theta}\frac{\partial A_\phi}{\partial\phi} \\
\nabla \times \boldsymbol{A} &= \hat{\boldsymbol{r}}\frac{1}{r\sin\theta}\left(\frac{\partial}{\partial\theta}\left(\sin\theta A_\phi\right) - \frac{\partial A_\theta}{\partial\phi}\right) \\
&\quad + \hat{\boldsymbol{\theta}}\frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial A_r}{\partial\phi} - \frac{\partial}{\partial r}\left(rA_\phi\right)\right) + \hat{\boldsymbol{\phi}}\frac{1}{r}\left(\frac{\partial}{\partial r}\left(rA_\theta\right) - \frac{\partial A_r}{\partial\theta}\right) \\
\nabla^2\boldsymbol{A} &= \hat{\boldsymbol{r}}\left(\nabla^2 A_r - \frac{2A_r}{r^2} - \frac{2}{r^2}\frac{\partial A_\theta}{\partial\theta} - \frac{2\cot\theta}{r^2}A_\theta - \frac{2}{r^2\sin\theta}\frac{\partial A_\phi}{\partial\phi}\right) \\
&\quad + \hat{\boldsymbol{\theta}}\left(\nabla^2 A_\theta - \frac{A_\theta}{r^2\sin^2\theta} + \frac{2}{r^2}\frac{\partial A_r}{\partial\theta} - \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial A_\phi}{\partial\phi}\right) \\
&\quad + \hat{\boldsymbol{\phi}}\left(\nabla^2 A_\phi - \frac{A_\phi}{r^2\sin^2\theta} + \frac{2}{r^2\sin\theta}\frac{\partial A_r}{\partial\phi} + \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial A_\theta}{\partial\phi}\right)
\end{aligned}$$

Appendix C

Units and constants

The explicit form of the equations in electromagnetism varies depending on the system of units that we use. The SI-system is the one that is used in most literature nowadays, and this textbook is no exception. The relevant constant in the SI-system that is used in the text is collected in this appendix.

The speed of light in vacuum c_0 has the value (exact value)

$$c_0 = 299\,792\,458 \text{ m/s}$$

μ_0 and ϵ_0 denote the permeability and the permittivity of vacuum, respectively. Their exact values are

$$\mu_0 = 4\pi \cdot 10^{-7} \text{ N/A}^2$$
$$\epsilon_0 = \frac{1}{c_0^2 \mu_0} \text{ F/m}$$

Approximative values of these constants are

$$\mu_0 \approx 12.566\,370\,614 \cdot 10^{-7} \text{ N/A}^2$$
$$\epsilon_0 \approx 8.854\,187\,817 \cdot 10^{-12} \text{ F/m}$$

The wave impedance of vacuum is denoted

$$\eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} = c_0 \mu_0 = 299\,792\,458 \cdot 4\pi \cdot 10^{-7}\ \Omega \approx 376.730\,314\ \Omega$$

The charge of the electron, $-e$, and its mass, m, have the values

$$e \approx 1.602\,177\,33 \cdot 10^{-19} \text{ C}$$
$$m \approx 9.109\,389\,8 \cdot 10^{-31} \text{ kg}$$
$$e/m \approx 1.758\,819\,63 \cdot 10^{11} \text{ C/kg}$$

Appendix D

Notation

Appropriate notion leads to a more easy understood, systematic, and structured text, and, in the same token, implies a tendency of making less errors and slips. Most of the notation is explained at the place in the text were they are introduced, but some more general notion that is often used is collected in this appendix.

- Vector-valued quantities (mostly in $\mathbb{R}^3$) is denoted in slanted bold face, eg., $\boldsymbol{a}$ and $\boldsymbol{b}$, and vectors of unit length have a "hat" or caret ($\hat{}$) over a symbol, eg., $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{\rho}}$.

- The (Euclidean) scalar product between two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$, is denoted in the usual standard way by a dot ($\cdot$), i.e., $\boldsymbol{a} \cdot \boldsymbol{b}$. If the vectors are complex-valued the appropriate scalar product is $\boldsymbol{a}^* \cdot \boldsymbol{b}$, where the star $*$ denotes the complex conjugate of the vector. For complex-valued functions ($L_2(\Omega)$-functions defined on a domain Ω) the scalar product is defined in the standard way

$$(f, g) = \iiint_{\Omega} f(\boldsymbol{r}) g^*(\boldsymbol{r}) \, \mathrm{d}v$$

- We make a distinction between a vector $\boldsymbol{a}$ and its representation in components in a specific coordinate system, and denote the components as a column vector or with brackets around the vector, i.e.,

$$[\boldsymbol{a}] = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}$$

  where

$$\boldsymbol{a} = \hat{\boldsymbol{x}} a_x + \hat{\boldsymbol{y}} a_y + \hat{\boldsymbol{z}} a_z$$

- Linear vector-valued transformations (dyadics) are denoted in bold roman fonts, eg., $\mathbf{A}$. A linear transformation $\mathbf{A}$ acting on a vector field $\boldsymbol{a}$ gives a new vector field $\boldsymbol{b}$ and we use the notation

$$\boldsymbol{b} = \mathbf{A} \cdot \boldsymbol{a}$$

  In a specific coordinate system the linear transformation $\mathbf{A}$ is represented by a 3×3 matrix $[\mathbf{A}]$, where we again use brackets around $\mathbf{A}$ to emphasize that we refer to its components. The components of the vector $\boldsymbol{b}$ is then

$$[\boldsymbol{b}] = [\mathbf{A}] \cdot [\boldsymbol{a}]$$

or

$$\begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix} \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}$$

- The unity and the null dyadics in n dimensions are denoted $\mathbf{I}_n$ and $\mathbf{0}_n$ respectively, and the corresponding matrix representations are denoted $[\mathbf{I}]_n$ and $[\mathbf{0}]_n$, respectively. In three dimensions we have

$$[\mathbf{I}]_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [\mathbf{0}]_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

  or in two dimensions

$$[\mathbf{I}]_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad [\mathbf{0}]_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

- The dyadic $\mathbf{J}$ (matrix $[\mathbf{J}]$) performs a rotation of a projection on the x-y-plane followed by a rotation of $\pi/2$ in the x-y-plane,

$$[\mathbf{J}] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

- The transpose of a matrix is denoted by a superscript $(^t)$ and the Hermitian of a matrix with the superscript dagger $(^\dagger)$, i.e.,

$$A^t_{ij} = A_{ji}$$
$$A^\dagger_{ij} = A^*_{ji}$$

- We use the symbols o and O defined by

$$\begin{cases} f(x) = o\left(g(x)\right), \quad x \to a \quad \Leftrightarrow \quad \lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = 0 \\ f(x) = O\left(g(x)\right), \quad x \to a \quad \Leftrightarrow \quad \dfrac{f(x)}{g(x)} \quad \text{bounded in a neighborhood of } a \end{cases}$$

- The symbol ∎ denotes the end of an example.

- The real and the imaginary part of a complex number $z = x + \mathrm{i}y$ are denoted $\operatorname{Re} z$ and $\operatorname{Im} z$, respectively, dvs.

$$\operatorname{Re} z = x = \frac{1}{2}(z + z^*)$$
$$\operatorname{Im} z = y = \frac{1}{2\mathrm{i}}(z - z^*)$$

  A star $(^*)$ is used to denote the complex conjugate of a complex number, i.e., $z^* = x - \mathrm{i}y$.

- The real and imaginary part of a dyadic can also be defined in terms of its Hermitian transpose. We adopt

$$\operatorname{Re}\mathbf{A} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^\dagger)$$
$$\operatorname{Im}\mathbf{A} = \frac{1}{2\mathrm{i}}(\mathbf{A} - \mathbf{A}^\dagger)$$

  where the dagger (†) denotes the Hermitian transpose. We see that the real and imaginary parts satisfy

$$\operatorname{Re}\mathbf{A} = \operatorname{Re}\mathbf{A}$$
$$\operatorname{Im}\mathbf{A}^\dagger = \operatorname{Im}\mathbf{A}$$

  We see that $\operatorname{Re}\mathbf{A}$ and $\operatorname{Im}\mathbf{A}$ are both Hermitian dyadics and that

$$\mathbf{A} = \operatorname{Re}\mathbf{A} + \mathrm{i}\operatorname{Im}\mathbf{A}$$

- The Heaviside's step function, $H(t)$, is defined in the usual way as

$$H(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}$$

- The Kronecker's delta (function) symbol, δ_{ij}, is defined as

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

- The cylindrical coordinate system (ρ, ϕ, z) is defined by

$$\begin{cases} \rho = \sqrt{x^2 + y^2} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2+y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2+y^2}} & y < 0 \end{cases} \\ z = z \end{cases}$$

  The domain of the coordinates are $\rho \in [0, \infty)$, $\phi \in [0, 2\pi)$, and $z \in (-\infty, \infty)$.

- The spherical coordinate system (r, θ, ϕ) is defined as

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \theta = \arccos \dfrac{z}{\sqrt{x^2 + y^2 + z^2}} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2+y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2+y^2}} & y < 0 \end{cases} \end{cases}$$

  The domain of the coordinates are $r \in [0, \infty)$, $\theta \in [0, \pi]$, and $\phi \in [0, 2\pi)$.

Bibliography

[1] G. P. Agrawal. *Fiber-optic communication systems.* John Wiley & Sons, New York, 1992.

[2] G. P. Agrawal. *Nonlinear fiber optics.* Academic Press, San Diego, 1995.

[3] G. Arfken. *Mathematical Methods for Physicists.* Academic Press, Orlando, third edition, 1985.

[4] D.K. Cheng. *Field and wave electromagnetics.* Addison-Wesley, Reading, 1983.

[5] R.E. Collin. *Field Theory of Guided Waves.* IEEE Press, New York, second edition, 1991.

[6] E. Hallén. *Elektricitetslära.* Almqvist & Wiksells, Uppsala, 1968.

[7] A. Ishimaru. *Electromagnetic Wave Propagation, Radiation, and Scattering.* Prentice Hall, Englewood Cliffs, 1991.

[8] J. D. Jackson. *Classical Electrodynamics.* John Wiley & Sons, New York, third edition, 1999.

[9] D.S. Jones. *Methods in Electromagnetic Wave Propagation.* IEEE Press, Piscataway, second edition, 1994.

[10] J.A. Kong. *Electromagnetic Wave Theory.* John Wiley & Sons, New York, 1986.

[11] J.D. Kraus. *Electromagnetics.* McGraw-Hill, New York, fourth edition, 1992.

[12] G. Kristensson. *Elektromagnetisk vågutbredning.* Studentlitteratur, Lund, 1999.

[13] S.Y. Liao. *Engineering Applications of Electromagnetic Theory.* West Publishing Company, St. Paul, 1988.

[14] B.D. Popović. *Introductory Engineering Electromagnetics.* Addison-Wesley, Reading, 1971.

[15] D.M. Pozar. *Microwave Engineering.* Addison-Wesley, Reading, 1990.

[16] G. Russakoff. A Derivation of the Macroscopic Maxwell Equations American Journal of Physics, vol. 38, 1188-1195, 1970

[17] SIS-Standardiseringskommissionen i Sverige. *SI måttenheter*, September 1988.

[18] J.C. Slater Microwave Electronics Reviews of Modern Physics, Vol 18, 441-514, 1946

[19] J.A. Stratton. *Electromagnetic Theory*. McGraw-Hill, New York, 1941.

[20] J. Van Bladel. *Electromagnetic Fields*. Hemisphere Publishing Corporation, New York, 1985.

[21] A. Yariv. *Optical Electronics*. Saunders College Publishing, Fort Worth, 1991.

Answers

1.1

1.2

1.3

2.1

3.1 a) $V = V_0 e^{j\pi/4}$ b) $v(t) = V_0 \cos(\omega t - \pi/4)$

3.2

3.3

3.6 $L = 0{,}24\,\mu\text{H/m}$ $C = 95\,\text{pF}/m$ $R = 50\,\text{m}\Omega\text{/m}$ $G = 20 \cdot 10^{-6}\,\text{S/m}$

3.7 a) V_0 b) 0 c) $\frac{2}{5}V_0$ d) $\frac{2}{5}V_0$ e) $\frac{16}{25}V_0$

3.8 a) $\alpha = 0.014\text{m}^{-1}$ b) 125 dB/km

3.9 $Z_0 = 60\,\Omega$ $\beta = 5{,}5 \cdot 10^{-3}\,\text{rad/m}$

3.10 $\ell = 0.265\,\text{m}$ $Z_0 = 134\,\Omega$

3.11 $\Gamma = 0{,}5$ $\text{SWR} = 3$

3.12 $R_b = 150\,\Omega$ or $R_b = 17\,\Omega$

3.13 $Z(0) = (120 - \text{j}60)\,\Omega$

3.14 a) $v(t) = \begin{cases} 0 & t < 2L/v \\ \frac{V_0}{3} & t \geq 2L/v \end{cases}$ b) $Z_{\text{in}} = 2R$

3.15 $x = 3\lambda/8$ or $x = 7\lambda/8$

3.16 a) $\beta\ell = 0.5 \arctan 2 \approx 0.55$ and $\beta\ell = 0.5(\arctan 2 + \pi) \approx 2.12$

b) $Z(0) = 2.61R$ and $Z(0) = 0.38R$

4.1

$$\begin{aligned} \boldsymbol{A}_T =& \nabla_T \times (\nabla_T \times \boldsymbol{F}_T) + \nabla_T \frac{\partial}{\partial z} F_z - \frac{\partial^2}{\partial z^2} \boldsymbol{F}_T \\ =& \nabla_T (\nabla_T \cdot \boldsymbol{F}_T) - \nabla_T^2 \boldsymbol{F}_T + \nabla_T \frac{\partial}{\partial z} F_z - \frac{\partial^2}{\partial z^2} \boldsymbol{F}_T \\ A_z =& \frac{\partial}{\partial z} (\nabla_T \cdot \boldsymbol{F}_T) - \nabla_T^2 F_z \end{aligned}$$

5.3 TM-modes:

$$\begin{cases} v_{mn}(\rho,\phi) = C_{mn} \underbrace{\left\{J_m(\zeta_{mn}\rho/a)N_m(\zeta_{mn}) - J_m(\zeta_{mn})N_m(\zeta_{mn}\rho/a)\right\}}_{Z_{mn}(\zeta_{mn}\rho/a)} \begin{pmatrix}\cos m\phi \\ \sin m\phi\end{pmatrix} \\ k_{tmn} = \zeta_{mn}/a \end{cases}$$

where $\{\zeta_{mn}\}_{n=1}^{\infty}$ are the solutions to, $m = 0, 1, 2, 3, \ldots$

$$Z_{mn}(\zeta_{mn}b/a) = J_m(\zeta_{mn}b/a)N_m(\zeta_{mn}) - J_m(\zeta_{mn})N_m(\zeta_{mn}b/a) = 0$$

and the normalization constant C_{mn} is determined by ($\epsilon_m = 2 - \delta_{m,0}$)

$$C_{mn}^{-2} = \frac{2\pi}{\epsilon_m}\int_a^b Z_{mn}^2(\zeta_{mn}\rho/a)\rho\,\mathrm{d}\rho = \frac{\pi}{\epsilon_m}\left\{b^2\left(Z'_{mn}(\zeta_{mn}b/a)\right)^2 - a^2\left(Z'_{mn}(\zeta_{mn})\right)^2\right\}$$

TE-modes:

$$\begin{cases} w_{mn}(\rho,\phi) = D_{mn} \underbrace{\left\{J_m(\gamma_{mn}\rho/a)N'_m(\gamma_{mn}) - J'_m(\gamma_{mn})N_m(\gamma_{mn}\rho/a)\right\}}_{Y_{mn}(\gamma_{mn}\rho/a)} \begin{pmatrix}\cos m\phi \\ \sin m\phi\end{pmatrix} \\ k_{tmn} = \gamma_{mn}/a \end{cases}$$

where $\{\gamma_{mn}\}_{n=1}^{\infty}$ are the solutions to, $m = 0, 1, 2, 3, \ldots$

$$Y'_{mn}(\gamma_{mn}b/a) = J'_m(\gamma_{mn}b/a)N'_m(\gamma_{mn}) - J'_m(\gamma_{mn})N'_m(\gamma_{mn}b/a) = 0$$

and the normalization constant D_{mn} is determined by

$$\begin{aligned} D_{mn}^{-2} =& \frac{2\pi}{\epsilon_m}\int_a^b Y_{mn}^2(\gamma_{mn}\rho/a)\rho\,\mathrm{d}\rho \\ =& \frac{\pi a^2}{\epsilon_m\gamma_{mn}^2}\left\{(\gamma_{mn}^2b^2/a^2 - m^2)\left(Y_{mn}(\gamma_{mn}b/a)\right)^2 - (\gamma_{mn}^2 - m^2)\left(Y_{mn}(\gamma_{mn})\right)^2\right\} \end{aligned}$$

TEM-mode:

$$\nabla_T\psi(\rho,\phi) = \hat{\boldsymbol{\rho}}\frac{1}{\rho}\frac{1}{2\pi\ln\frac{a}{b}}$$

5.4 TM-modes:

$$\begin{cases} v_{mn}(\rho,\phi) = C_{mn}J_{m/2}(\zeta_{mn}\rho/a)\sin\dfrac{m}{2}\phi, \quad m = 1, 2, 3, \ldots \\ k_{tmn} = \zeta_{mn}/a \end{cases}$$

where $\{\zeta_{mn}\}_{n=1}^{\infty}$ are the solutions to, $m = 1, 2, 3, \ldots$

$$J_{m/2}(\zeta_{mn}) = 0$$

and the normalization constant C_{mn} is determined by

$$C_{mn}^{-2} = \frac{\pi a^2}{\zeta_{mn}^2}\int_0^{\zeta_{mn}}\left(J_{m/2}(x)\right)^2 x\,\mathrm{d}x$$

TE-modes:

$$\begin{cases} w_{mn}(\rho,\phi) = D_{mn} J_{m/2}(\gamma_{mn}\rho/a)\cos\dfrac{m}{2}\phi, \quad m = 0,1,2,\ldots \\ k_{tmn} = \gamma_{mn}/a \end{cases}$$

where $\{\gamma_{mn}\}_{n=1}^{\infty}$ are the solution to, $m = 0, 1, 2, \ldots$

$$J'_{m/2}(\gamma_{mn}) = 0$$

and the normalization constant D_{mn} is determined by

$$D_{mn}^{-2} = \frac{\pi a^2}{\gamma_{mn}^2} \int_0^{\gamma_{mn}} \left(J_{m/2}(x)\right)^2 x \, \mathrm{d}x$$

The lowest cut-off frequency is given by the smallest solution to

$$J'_{1/2}(x) = 0 \quad \Leftrightarrow \quad \sin x - 2x\cos x = 0$$

vilken är $x \approx 1.166$. The lowest cut-off frequency is decreased approximately 37%.

5.5 a) $P_r/P_i = 6.1 \cdot 10^{-3}$

b) $P_r = 0$ for the frequency $f = \dfrac{c_0\sqrt{b^2+a^2}}{2ab}\sqrt{\dfrac{\epsilon+1}{\epsilon}} = 7.65$ GHz

Comment: the incident TM-mode can be split up in two plane wave. At the frequency 7.65 GHz the incident angle for these waves is the Brewster angle.

5.6 a) TM_{01}, TE_{11} and TE_{21}

b) $\alpha = \dfrac{\omega\mu_0\sigma}{2(\omega^2\mu_0\epsilon_0\epsilon - k_{tnj}^2)^{1/2}} = \dfrac{\sigma\eta}{2(1-(f_c/f)^2)^{1/2}}$

where $\eta = \eta_0/\sqrt{\epsilon} = 120\pi/\sqrt{\epsilon}\,\Omega$ is the wave impedance in the material, f is the frequency and $f_c$is the cut-off frequency. Since $\epsilon = 3$, $\sigma = 10^{-11}$ and $f_c = ck_{tnj}/(2\pi) = 1.7$ GHz we get

$$\alpha = \frac{1.1\,10^{-9}}{(1-(1.7\,10^9/f)^2)^{1/2}}\ \mathrm{Np/m}$$

The attenuation in dB/km is $20\log(E(0\mathrm{km})/E(1\mathrm{km})) = \alpha\ 8.7\ 10^3$ dB/km with α in Np/m.

5.7 a) $x_0 = 4$ cm

b) $P_r/P_i = 0$

c) $P_r/P_i = 0$

5.8 **TE-modes**:

$$H_z = A_{nj} J_{2n}(\beta_{2nj}\rho)\cos(2n\phi)\exp(ik_{znj}z)$$

where $k_{znj}^2 = (\omega/c_0)^2 - (k_{t2nj})^2$ and $J'_{2n}(k_{t2nj}R) = 0$

TM-modes:

$$E_z = B_{n,j} J_{2n}(k_{t2n,j}\rho)\sin(2n\phi)\exp(ik_{znj}z)$$

where $k_{znj}^2 = (\omega/c_0)^2 - (k_{t2nj})^2$ and $J_{2n}(k_{t2nj}R) = 0$.

5.10 a) Half of the power in 2 half in 3. No power in 4. The fields are directed in the same direction in ports 2 and 3.

b) Half of the power in 2 half in 3. No power in 1. The fields are directed in opposite directions in ports 2 and 3.

6.3 a)

$$Q = \frac{\omega_{101}\mu_0 abd(a^2+d^2)}{2R_S(2b(a^3+d^3)+ad(a^2+d^2))}$$

där $\omega_{101} = c\sqrt{(\pi/a)^2+(\pi/d)^2}$.

b) The resonance frequency is 8,38 GHz and Q=9900.

c) The resonance frequency is 0,838 GHz and Q=31300.

6.4 The resonance wavenumbers (TE or TM) are given by $k_{nml}^2 = \left(\frac{2\pi}{\lambda_{nml}}\right)^2 = (n^2 + m^2 + l^2)\left(\frac{\pi}{a}\right)^2$ where $a = 1$ m. Now $\lambda_{nml} > \lambda = 500$ nm. This means that the n, m, l have to satisfy

$$n^2 + m^2 + l^2 \leq \left(\frac{2a}{\lambda}\right)^2 = 16 \cdot 10^{12}$$

Only positive n, m, l are allowed. The total number of combinations n, m, l that satisfies this relation equals the volume of a quarter of a sphere with radius $4 \cdot 10^6$ units. This means that there are $\frac{1}{4} \cdot \frac{4}{3}\pi(4 \cdot 10^6)^3$ resonances. Since there are both TE and TM modes this should be multiplied by two. The total number is then $1.34 \cdot 10^{20}$.

7.2

$$a(0, t - L/c_0) = H(t) - H(t - t_0) + \int_{t-\min(t,t_0)}^{t} P(-L, t')\, dt'$$

$$\text{where } P(-L,t) = \begin{cases} c_0 k_t L \dfrac{I_1(k_t\sqrt{2c_0Lt - c_0^2t^2})}{\sqrt{2c_0Lt - c_0^2t^2}}, & \text{when } t < 2L/c_0 \\ \\ c_0 k_t L \dfrac{J_1(k_t\sqrt{-2c_0Lt + c_0^2t^2})}{\sqrt{-2c_0Lt + c_0^2t^2}}, & \text{when } t > 2L/c_0 \end{cases}$$

8.2

$$H_z = \begin{cases} H_n \cos(k_{tn}y)e^{\mathrm{i}k_{zn}z}, & \text{då } 0 < y < a \\ H_n \dfrac{\cos(k_{tn}a)\cos(k'_{tn}(b-y))}{\cos(k'_{tn}(b-a))} e^{\mathrm{i}k_{zn}z}, & \text{då } a < y < b \end{cases}$$

where k_{tn}, k'_{tn} and k_{zn} are given by

$$\begin{aligned} k_{tn}^2 &= (\omega/c_0)^2\epsilon - k_{zn}^2 \\ {k'_{tn}}^2 &= (\omega/c_0)^2 - k_{zn}^2 \\ k_{tn}\tan(k'_{tn}(b-a)) &= -k'_{tn}\tan(k_{tn}a) \end{aligned}$$

8.3

$$\begin{aligned} v_m &= AJ_m\left(k_0\sqrt{n_1^2 - n_2^2}\right)\left(\frac{a}{\rho}\right)^m \sin m\phi \\ w_m &= -n_2 AJ_m\left(k_0\sqrt{n_1^2 - n_2^2}\right)\left(\frac{a}{\rho}\right)^m \cos m\phi \end{aligned}$$

8.5

$$\begin{aligned}\boldsymbol{H} =& A_n(\hat{\boldsymbol{z}}J_0(k_{tn}\rho) + \hat{\boldsymbol{\rho}}i\frac{k_{zn}}{k_{tn}}J_0'(k_{tn}\rho))\exp(ik_{zn}z), \quad 0<\rho<a\\ \boldsymbol{H} =& B_n(\hat{\boldsymbol{z}}(J_0(k_{tn}'\rho) - \frac{J_0'(k_{tn}'b)}{Y_0'(k_{tn}'b)}Y_0(k_{tn}'\rho))\\ &+ i\hat{\boldsymbol{\rho}}\frac{k_{zn}}{k_{tn}'}(J_0'(k_{tn}'\rho) - \frac{J_0'(k_{tn}'b)}{Y_0'(k_{tn}'b)}Y_0'(k_{tn}'\rho)))\exp(ik_{zn}z), \quad a<\rho<b\end{aligned}$$

where

$$\begin{aligned}B_n &= A_n\frac{J_0(k_{tn}a)Y_0(k_{tn}'b)}{J_0(k_{tn}'a)Y_0'(k_{tn}'b) - J_0'(k_{tn}'b)Y_0(k_{tn}'a)}\\ k_{tn}^2 &= \left(\frac{\omega}{c_0}\right)^2 - k_{zn}^2\\ {k_{tn}'}^2 &= \left(\frac{\omega}{c_0}\right)^2\epsilon - k_{zn}^2\end{aligned}$$

and k_{zn} are given by the transcendental equation

$$\begin{aligned}&k_{tn}J_0(k_{tn}a)(J_0'(k_{tn}'a)Y_0'(k_{tn}'b) - J_0'(k_{tn}'b)Y_0'(k_{tn}'a))\\ &= k_{tn}'J_0'(\gamma_j a)(J_0(k_{tn}'a)Y_0'(k_{tn}'b) - J_0'(k_{tn}'b)Y_0(k_{tn}'a))\end{aligned}$$

Index

Relations between basis vectors

Cylindrical coordinates (ρ, ϕ, z)

$$\begin{cases} \rho = \sqrt{x^2+y^2} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2+y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2+y^2}} & y < 0 \end{cases} \\ z = z \end{cases}$$

Spherical coordinates (r, θ, ϕ)

$$\begin{cases} r = \sqrt{x^2+y^2+z^2} \\ \theta = \arccos \frac{z}{\sqrt{x^2+y^2+z^2}} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2+y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2+y^2}} & y < 0 \end{cases} \end{cases}$$

$(r, \theta, \phi) \longrightarrow (x, y, z)$

$$\begin{cases} \hat{\boldsymbol{r}} = \hat{\boldsymbol{x}} \sin\theta\cos\phi + \hat{\boldsymbol{y}} \sin\theta\sin\phi + \hat{\boldsymbol{z}} \cos\theta \\ \hat{\boldsymbol{\theta}} = \hat{\boldsymbol{x}} \cos\theta\cos\phi + \hat{\boldsymbol{y}} \cos\theta\sin\phi - \hat{\boldsymbol{z}} \sin\theta \\ \hat{\boldsymbol{\phi}} = -\hat{\boldsymbol{x}} \sin\phi + \hat{\boldsymbol{y}} \cos\phi \end{cases}$$

$(x, y, z) \longrightarrow (r, \theta, \phi)$

$$\begin{cases} \hat{\boldsymbol{x}} = \hat{\boldsymbol{r}} \sin\theta\cos\phi + \hat{\boldsymbol{\theta}} \cos\theta\cos\phi - \hat{\boldsymbol{\phi}} \sin\phi \\ \hat{\boldsymbol{y}} = \hat{\boldsymbol{r}} \sin\theta\sin\phi + \hat{\boldsymbol{\theta}} \cos\theta\sin\phi + \hat{\boldsymbol{\phi}} \cos\phi \\ \hat{\boldsymbol{z}} = \hat{\boldsymbol{r}} \cos\theta - \hat{\boldsymbol{\theta}} \sin\theta \end{cases}$$

$(\rho, \phi, z) \longrightarrow (x, y, z)$

$$\begin{cases} \hat{\boldsymbol{\rho}} = \hat{\boldsymbol{x}} \cos\phi + \hat{\boldsymbol{y}} \sin\phi = (\hat{\boldsymbol{x}} x + \hat{\boldsymbol{y}} y)/\sqrt{x^2+y^2} \\ \hat{\boldsymbol{\phi}} = -\hat{\boldsymbol{x}} \sin\phi + \hat{\boldsymbol{y}} \cos\phi = (-\hat{\boldsymbol{x}} y + \hat{\boldsymbol{y}} x)/\sqrt{x^2+y^2} \\ \hat{\boldsymbol{z}} = \hat{\boldsymbol{z}} \end{cases}$$

$(x, y, z) \longrightarrow (\rho, \phi, z)$

$$\begin{cases} \hat{\boldsymbol{x}} = \hat{\boldsymbol{\rho}} \cos\phi - \hat{\boldsymbol{\phi}} \sin\phi \\ \hat{\boldsymbol{y}} = \hat{\boldsymbol{\rho}} \sin\phi + \hat{\boldsymbol{\phi}} \cos\phi \\ \hat{\boldsymbol{z}} = \hat{\boldsymbol{z}} \end{cases}$$

$(r, \theta, \phi) \longrightarrow (\rho, \phi, z)$

$$\begin{cases} \hat{\boldsymbol{r}} = \hat{\boldsymbol{\rho}} \sin\theta + \hat{\boldsymbol{z}} \cos\theta \\ \hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\rho}} \cos\theta - \hat{\boldsymbol{z}} \sin\theta \\ \hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\phi}} \end{cases}$$

$(\rho, \phi, z) \longrightarrow (r, \theta, \phi)$

$$\begin{cases} \hat{\boldsymbol{\rho}} = \hat{\boldsymbol{r}} \sin\theta + \hat{\boldsymbol{\theta}} \cos\theta \\ \hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\phi}} \\ \hat{\boldsymbol{z}} = \hat{\boldsymbol{r}} \cos\theta - \hat{\boldsymbol{\theta}} \sin\theta \end{cases}$$