



LUND UNIVERSITY

Microwave theory

Karlsson, Anders; Kristensson, Gerhard

2026

Document Version:

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Karlsson, A., & Kristensson, G. (2026). *Microwave theory*. Eget förlag.

Total number of authors:

2

Creative Commons License:

Other

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

ANDERS KARLSSON
and
GERHARD KRISTENSSON

MICROWAVE THEORY

Rules for the ∇ -operator

$$\begin{aligned}
 (1) \quad & \nabla(\varphi + \psi) = \nabla\varphi + \nabla\psi \\
 (2) \quad & \nabla(\varphi\psi) = \psi\nabla\varphi + \varphi\nabla\psi \\
 (3) \quad & \nabla(\mathbf{a} \cdot \mathbf{b}) = (\mathbf{a} \cdot \nabla)\mathbf{b} + (\mathbf{b} \cdot \nabla)\mathbf{a} + \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) \\
 (4) \quad & \nabla(\mathbf{a} \cdot \mathbf{b}) = -\nabla \times (\mathbf{a} \times \mathbf{b}) + 2(\mathbf{b} \cdot \nabla)\mathbf{a} + \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) + \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a})
 \end{aligned}$$

$$\begin{aligned}
 (5) \quad & \nabla \cdot (\mathbf{a} + \mathbf{b}) = \nabla \cdot \mathbf{a} + \nabla \cdot \mathbf{b} \\
 (6) \quad & \nabla \cdot (\varphi\mathbf{a}) = \varphi(\nabla \cdot \mathbf{a}) + (\nabla\varphi) \cdot \mathbf{a} \\
 (7) \quad & \nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b})
 \end{aligned}$$

$$\begin{aligned}
 (8) \quad & \nabla \times (\mathbf{a} + \mathbf{b}) = \nabla \times \mathbf{a} + \nabla \times \mathbf{b} \\
 (9) \quad & \nabla \times (\varphi\mathbf{a}) = \varphi(\nabla \times \mathbf{a}) + (\nabla\varphi) \times \mathbf{a} \\
 (10) \quad & \nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a}) + (\mathbf{b} \cdot \nabla)\mathbf{a} - (\mathbf{a} \cdot \nabla)\mathbf{b} \\
 (11) \quad & \nabla \times (\mathbf{a} \times \mathbf{b}) = -\nabla(\mathbf{a} \cdot \mathbf{b}) + 2(\mathbf{b} \cdot \nabla)\mathbf{a} + \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) + \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a})
 \end{aligned}$$

$$\begin{aligned}
 (12) \quad & \nabla \cdot \nabla\varphi = \nabla^2\varphi = \Delta\varphi \\
 (13) \quad & \nabla \times (\nabla \times \mathbf{a}) = \nabla(\nabla \cdot \mathbf{a}) - \nabla^2\mathbf{a} \\
 (14) \quad & \nabla \times (\nabla\varphi) = \mathbf{0} \\
 (15) \quad & \nabla \cdot (\nabla \times \mathbf{a}) = 0 \\
 (16) \quad & \nabla^2(\varphi\psi) = \varphi\nabla^2\psi + \psi\nabla^2\varphi + 2\nabla\varphi \cdot \nabla\psi
 \end{aligned}$$

$$\begin{aligned}
 (17) \quad & \nabla r = \hat{\mathbf{r}} \\
 (18) \quad & \nabla \times \mathbf{r} = \mathbf{0} \\
 (19) \quad & \nabla \times \hat{\mathbf{r}} = \mathbf{0} \\
 (20) \quad & \nabla \cdot \mathbf{r} = 3 \\
 (21) \quad & \nabla \cdot \hat{\mathbf{r}} = \frac{2}{r} \\
 (22) \quad & \nabla(\mathbf{a} \cdot \mathbf{r}) = \mathbf{a}, \quad \mathbf{a} \text{ constant vector} \\
 (23) \quad & (\mathbf{a} \cdot \nabla)\mathbf{r} = \mathbf{a} \\
 (24) \quad & (\mathbf{a} \cdot \nabla)\hat{\mathbf{r}} = \frac{1}{r}(\mathbf{a} - \hat{\mathbf{r}}(\mathbf{a} \cdot \hat{\mathbf{r}})) = \frac{\mathbf{a}_\perp}{r} \\
 (25) \quad & \nabla^2(\mathbf{r} \cdot \mathbf{a}) = 2\nabla \cdot \mathbf{a} + \mathbf{r} \cdot (\nabla^2\mathbf{a})
 \end{aligned}$$

$$\begin{aligned}
 (26) \quad & \nabla u(f) = (\nabla f) \frac{du}{df} \\
 (27) \quad & \nabla \cdot \mathbf{F}(f) = (\nabla f) \cdot \frac{d\mathbf{F}}{df} \\
 (28) \quad & \nabla \times \mathbf{F}(f) = (\nabla f) \times \frac{d\mathbf{F}}{df} \\
 (29) \quad & \nabla = \hat{\mathbf{r}}(\hat{\mathbf{r}} \cdot \nabla) - \hat{\mathbf{r}} \times (\hat{\mathbf{r}} \times \nabla)
 \end{aligned}$$

Important vector identities

- (1) $(\mathbf{a} \times \mathbf{c}) \times (\mathbf{b} \times \mathbf{c}) = \mathbf{c}((\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c})$
- (2) $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$
- (3) $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$
- (4) $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$

Integration formulas

Stoke's theorem and related theorems

- (1) $\iint_S (\nabla \times \mathbf{A}) \cdot \hat{\mathbf{n}} \, dS = \int_C \mathbf{A} \cdot d\mathbf{r}$
- (2) $\iint_S \hat{\mathbf{n}} \times \nabla \varphi \, dS = \int_C \varphi \, d\mathbf{r}$
- (3) $\iint_S (\hat{\mathbf{n}} \times \nabla) \times \mathbf{A} \, dS = \int_C d\mathbf{r} \times \mathbf{A}$

Gauss' theorem (divergence theorem) and related theorems

- (1) $\iiint_V \nabla \cdot \mathbf{A} \, dv = \iint_S \mathbf{A} \cdot \hat{\mathbf{n}} \, dS$
- (2) $\iiint_V \nabla \varphi \, dv = \iint_S \varphi \hat{\mathbf{n}} \, dS$
- (3) $\iiint_V \nabla \times \mathbf{A} \, dv = \iint_S \hat{\mathbf{n}} \times \mathbf{A} \, dS$

Green's formulas

- (1) $\iiint_V (\psi \nabla^2 \varphi - \varphi \nabla^2 \psi) \, dv = \iint_S (\psi \nabla \varphi - \varphi \nabla \psi) \cdot \hat{\mathbf{n}} \, dS$
- (2)
$$\begin{aligned} \iiint_V (\psi \nabla^2 \mathbf{A} - \mathbf{A} \nabla^2 \psi) \, dv \\ = \iint_S (\nabla \psi \times (\hat{\mathbf{n}} \times \mathbf{A}) - \nabla \psi (\hat{\mathbf{n}} \cdot \mathbf{A}) - \psi (\hat{\mathbf{n}} \times (\nabla \times \mathbf{A})) + \hat{\mathbf{n}} \psi (\nabla \cdot \mathbf{A})) \, dS \end{aligned}$$

Karlsson & Kristensson: Microwave theory

Microwave theory

Anders Karlsson and Gerhard Kristensson

Contents

Preface	v
1 Historic notes	vii
1.1 Luigi Galvani and the first electric circuit	vii
1.2 Hans-Christian Ørsted	ix
1.3 Samuel Morse	ix
1.4 Alexander Graham Bell	x
1.5 Wireless telegraphy	xi
2 The Maxwell equations	1
2.1 Boundary conditions at interfaces	3
2.1.1 Impedance boundary conditions	7
2.2 Energy conservation and Poynting's theorem	8
Problems in Chapter 2	10
3 Time harmonic fields and Fourier transform	13
3.1 The Maxwell equations	14
3.2 Constitutive relations	15
3.3 Poynting's theorem	16
Problems in Chapter 3	17
4 Transmission lines	19
Transmission lines	19
4.1 Time and frequency domain	20
4.1.1 Phasors ($j\omega$ method)	20
4.1.2 Fourier transformation	21
4.1.3 Fourier series	21
4.1.4 Laplace transformation	23
4.2 Two-ports	23
4.2.1 The impedance matrix	23
4.2.2 The chain matrix ($ABCD$ -matrix)	24
4.2.3 The hybrid matrix	24
4.2.4 Reciprocity	24
4.2.5 Transformation between matrices	25
4.2.6 Circuit models for two-ports	27

4.2.7	Combined two-ports	28
4.2.8	Cascad coupled two-ports	30
4.3	Transmission lines in time domain	31
4.3.1	Wave equation	31
4.3.2	Wave propagation in the time domain	33
4.3.3	Reflection from a load on a lossless line	34
4.4	Transmission lines in frequency domain	36
4.4.1	Input impedance	37
4.4.2	Standing wave ratio	39
4.4.3	Waves on lossy transmission lines in the frequency domain	39
4.4.4	Distortion free lines	40
4.5	Wave propagation in terms of $\mathbf{E}$ and $\mathbf{H}$	41
4.6	Transmission line parameters	44
4.6.1	Explicit expressions	47
4.6.2	Determination of R, L, G, C with the finite element method	48
4.6.3	Transverse inhomogeneous region	49
4.7	The scattering matrix $[S]$	50
4.7.1	S-matrix with different characteristic impedances at the ports	51
4.7.2	Relation between $[S]$ and $[Z]$	52
4.7.3	Matching of load impedances	53
4.7.4	Matching with stub	54
4.8	Smith chart	55
4.8.1	Matching a load by using the Smith chart	57
4.8.2	Frequency Response in the Smith Chart	58
	Problems in Chapter 4	59
	Summary of chapter 4	61
5	Electromagnetic fields with a preferred direction	63
	Electromagnetic fields with a preferred direction	63
5.1	Decomposition of vector fields	63
5.2	Decomposition of the Maxwell equations	64
5.3	Specific z -dependence of the fields	65
	Problems in Chapter 5	66
	Summary of chapter 5	67
6	Waveguides at fix frequency	69
	Waveguides at fix frequency	69
6.1	Boundary conditions	69
6.2	TM and TE modes	70
6.2.1	The longitudinal components of the fields	72
6.2.2	Transverse components of the fields	78
6.3	TEM modes	79
6.3.1	Waveguides with several conductors	81
6.4	Vector basis functions in hollow waveguides	81

6.4.1	The fundamental mode	83
6.5	Examples	84
6.5.1	Planar waveguide	84
6.5.2	Waveguide with rectangular cross-section	85
6.5.3	Waveguide with circular cross-section	89
6.6	Analyzing waveguides with FEM	90
6.7	Normalization integrals	92
6.8	Power flow density	95
6.9	Losses in walls	99
6.9.1	Losses in waveguides with FEM: method 1	103
6.9.2	Losses in waveguides with FEM: method 2	104
6.10	Sources in waveguides	105
6.11	Mode-Matching Technique	109
6.11.1	Cascading	113
6.11.2	Waveguides with bends	114
6.12	Transmission lines in inhomogeneous media by FEM	115
6.13	Substrate integrated waveguides	119
	Problems in Chapter 6	120
	Summary of chapter 6	125
7	Resonance cavities	129
	Resonance cavities	129
7.1	General cavities	129
7.1.1	The resonances in a lossless cavity	129
7.1.2	The resonances in a cavity with losses	131
7.1.3	The Quality factor (Q-factor) for a mode in a cavity	132
7.1.4	Slater's theorem	135
7.1.5	Measuring electric and magnetic fields in cavities	137
7.2	Example: Cylindrical cavities	138
7.3	Example: Spherical cavities	142
7.3.1	Vector spherical harmonics	142
7.3.2	Regular spherical vector waves	143
7.3.3	Resonance frequencies in a spherical cavity	143
7.3.4	Q-factors for spherical cavity	145
7.3.5	Two concentric spheres	146
7.4	Analyzing resonance cavities with FEM	148
7.5	Excitation of modes in a cavity	151
	Problems in Chapter 7	153
	Summary of chapter 7	154
8	Transients in waveguides	155
	Transients in waveguides	155
	Problems in Chapter 8	158
	Summary of chapter 8	159

9	Dielectric waveguides	161
	Dielectric waveguides	161
9.1	Planar dielectric waveguides	162
9.2	Cylindrical dielectric waveguides	164
9.2.1	The electromagnetic fields	164
9.2.2	Boundary conditions	165
9.3	Circular dielectric waveguide	165
9.3.1	Waveguide modes	166
9.3.2	HE modes	169
9.3.3	EH modes	171
9.3.4	TE and TM modes	172
9.4	Optical fibers	173
9.4.1	Effective index of refraction and phase velocity	176
9.4.2	Dispersion	177
9.4.3	Attenuation in optical fibers	179
9.4.4	Dielectric waveguides analyzed with FEM	179
9.4.5	Dielectric resonators analyzed with FEM	181
	Problems in Chapter 9	183
	Summary of chapter 9	185
A	Bessel functions	189
A.1	Bessel and Hankel functions	189
A.1.1	Useful integrals	193
A.2	Modified Bessel functions	194
A.3	Spherical Bessel and Hankel functions	195
B	∇ in curvilinear coordinate systems	199
B.1	Cartesian coordinate system	199
B.2	Circular cylindrical (polar) coordinate system	199
B.3	Spherical coordinates system	200
C	Units and constants	203
D	Notation	205
	Literature	209
	Answers	211
	Index	217

Preface

This book explores wave propagation along guiding structures, including transmission lines, hollow waveguides, and optical fibers—all of which underpin a vast range of modern applications. Optical fiber systems serve as the backbone of the internet and global telecommunications. While transmission lines have largely been superseded by optical and wireless systems over long distances, they remain essential for short-range communication, measurement equipment, and high-frequency circuitry. Hollow waveguides, on the other hand, are critical for radar systems and high-frequency instrumentation, as well as for delivering high-power microwaves in particle accelerators. Finally, an entire chapter is dedicated to electromagnetic fields inside metallic-walled cavities. These cavities are fundamental to particle accelerators, as they supply the electric fields that accelerate particle bunches as they travel down the beam pipe.

Electromagnetic fields in waveguides and cavities are governed by the Maxwell equations—a system of partial differential equations (PDEs). For several key geometries, these equations can be solved analytically. Throughout this book, we derive analytical solutions for the most fundamental geometries using the method of separation of variables. For more complex structures, numerical methods are required. While various commercial software packages exist for these evaluations, our examples primarily reference COMSOL Multiphysics, which employs the finite element method (FEM). While modern commercial solvers can analyze highly intricate structures, a firm understanding of fundamental analytical solutions remains essential for validating and interpreting numerical results.

Readers are assumed to have a foundational background in vector calculus, electromagnetic theory, and circuit theory. The nabla operator (∇) is used extensively to express results in coordinate-independent form.

Each chapter concludes with a selection of problems to reinforce the material, with more challenging exercises marked by an asterisk (*). Answers to most problems are provided at the end of the book.

Chapter 1

Historic notes

Between 1781 and 1902, a series of key discoveries laid the foundation for modern telegraphy, telephony, and long-distance wireless communication. This period of rapid innovation opened with a remarkable finding by Luigi Galvani in 1786.

1.1 Luigi Galvani and the first electric circuit

In 1781, Luigi Galvani began investigating the nervous system of frogs by dissecting dead specimens to expose the main crural nerve. During one dissection, he observed a striking phenomenon: when static electricity discharged near the exposed nerve via his scalpel, the frog's leg twitched violently. This observation sparked ten years of rigorous research into the link between electrical currents and muscle stimulation, driven by Galvani's ambition to uncover the biological mechanisms underlying animal and human life.

Initially, Galvani focused his research on what he termed "artificial electricity," applying electrical discharges from Leyden jars, that were primitive capacitors charged by hand-cranked electrostatic generators, to stimulate muscle movement.

In September 1786, Galvani turned his attention to atmospheric electricity. Inspired by Benjamin Franklin's demonstration that lightning consisted of electrical discharges, Galvani hypothesized that atmospheric electricity might share the same fundamental nature as the biological impulses governing animal muscles and neural function.

Galvani's experimental setup was very simple: an iron wire extended across the roof of a building, designed to lead atmospheric electricity down to the exposed nerve of a dissected frog. Connected to the frog was also a secondary, shorter wire made of copper. Unintentionally, the copper wire touched both the iron wire and the frog simultaneously, causing the leg muscle to contract violently.

Galvani's explanation was that all living organisms contain an intrinsic electrical fluid, which he termed animal electricity, that serves as the vital force of life. He posited that this fluid flows from the brain through the nervous system to trigger muscle movement. According to Galvani, connecting two dissimilar metals to a specimen completed a circuit that allowed this innate animal electricity to flow freely.

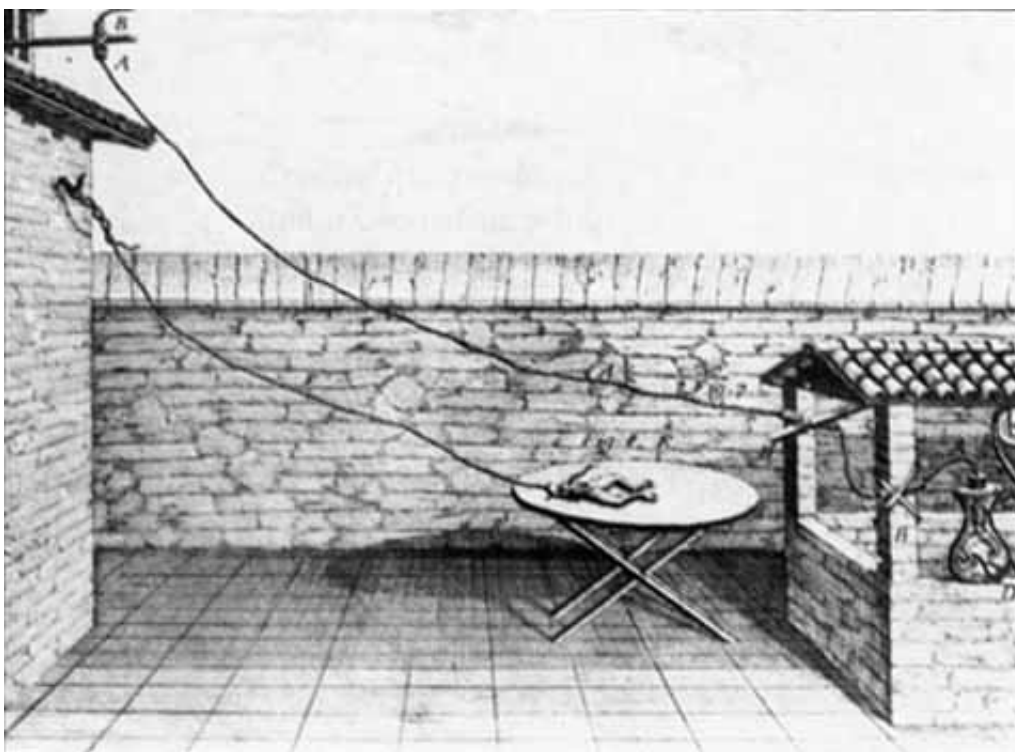


Figure 1.1: Galvani's experiment with atmospheric electricity 1786.

Today, we know that Galvani's explanation was only partially correct. He was right in asserting that electricity, specifically in the form of ionic currents, governs bodily movement and neural activity. However, the electrical current generated in his experiment had a different origin. The circuit formed by the frog, iron wire, and copper wire operated as an electrochemical cell; the frog leg served not only as an extraordinarily sensitive current detector, but also as the electrolyte in a primitive galvanic element.

Despite his flawed physiological interpretation, Galvani successfully constructed the first functional electric circuit, a monumental achievement for which he deserves full credit. In 1791, he published his findings in a landmark treatise. Once this work entered the scientific community, it was only a matter of time before subsequent researchers unraveled the true electrochemical mechanisms behind his experiments.

It was Galvani's friend Alessandro Volta who found the explanation. Volta demonstrated that animal tissue was not required to generate current; he replaced the frog leg with a simple saline solution and immersed copper and iron wires directly into the liquid. With an electroscope he proved that there was a voltage between the wires. He published his results in 1800. This ignited a revolution in chemistry, enabling researchers to use electrolysis to isolate new elements and fundamentally expand the understanding of chemical reactions.

It is surprising that it wasn't until 1820, decades after Galvani's discovery, that his work led to a major breakthrough in physics.

1.2 Hans-Christian Ørsted

On April 21, 1820, during a lecture at the University of Copenhagen, the Danish physicist Hans Christian Ørsted (1777–1851) discovered that a compass needle deflected when brought near a current-carrying wire. The phenomenon could only be explained by the fact that the electric current generated a magnetic field.

To investigate the discovery further, Ørsted acquired a powerful voltaic battery capable of delivering high currents. Through subsequent experiments, he demonstrated that the magnetic field forms concentric circles around a long, straight conductor, that the field strength decreases in inverse proportion to the distance from the conductor, and that materials such as glass, wood, and certain metals do not shield the field. On July 21, 1820, he published his findings in a four-page article written in Latin, which was quickly translated into several languages. The English is "Experiments on the Effect of a Current of Electricity on the Magnetic Needle".

Ørsted's discovery was, in a way, fortunate. Had Earth's magnetic field been much weaker, compasses might never have been invented in the first place. Conversely, had it been much stronger, the magnetic field generated by his electric current would have been too weak to noticeably deflect the needle. Luckily, Earth's magnetic field is roughly $50\text{ }\mu\text{T}$, which proved to be just right for uncovering the link between electricity and magnetism.

Ørsted's discovery was a major step in physics. It led in 1821 to Ampère's theory for magnetostatics and in 1831 to Michael Faraday's discovery of electric induction.

1.3 Samuel Morse

In electrical technology the electromagnet, the galvanometer, and the relay were important inventions that directly emerged from Ørsted's discovery. They made it possible to develop systems for telegraphy over long distances.

During the period 1830–1840, some sixty different systems were developed, see [18, p. 68]. Only three of these achieved commercial success. The first commercial telegraph system was installed in St. Petersburg in 1832. The system, developed by Baron Pavel Schilling, used six cables connected to 36 galvanometers, enabling the transmission of 26 letters and 10 digits.

In 1837 William Fothergill Cooke and Charles Wheatstone patented a telegraph with a simpler design than Schilling's telegraph. The receiver was equipped with five needles that, using electric coils, could be rotated to point to letters of the alphabet. The system required five wires to operate. Cooke and Wheatstone's telegraph was installed along railways in England, and some of these systems remained in use until the 1930s. With the help of relay stations, Cooke and Wheatstone's telegraph could effectively transmit messages over long distances. However, the system had the disadvantage of requiring five wires, making installation expensive and complex.

Samuel Morse was primarily a portrait painter. In 1825, while working in Washington, D.C., his wife fell ill and died in Connecticut. Due to slow horse-drawn mail deliver, he received the news days after she was already buried. Devastated,

he dedicated himself to solving the problem of long-distance communication.

In 1832 he conceives the idea of an electric telegraph and made a sketch of a design for a system that uses short and long electrical pulses to mark paper. The design of the system developed further with help from two partners, Leonard Gale and Alfred Vail. They filed for a patent in 1837.

For years, Morse struggled to secure financial support for his telegraph, a process described in detail in [22, Chapter 3]. He presented his telegraph to the US Congress as early as 1838, but this did not result in any government funding. Four years later, he made another attempt; following a vote of 89 in favour and 83 against, Morse was granted \$30,000 to install telegraph lines between Washington, D.C., and Baltimore. The voting figures reflect the politicians' lack of understanding: in addition to the 83 members who evidently did not believe in the technology, 70 members were at least clever enough to abstain, reasoning that they wished to avoid responsibility for spending public funds on a machine they could not comprehend.

Morse's original telegraph featured a receiver that used an electromagnet to move a pen up and down. The pen recorded the Morse-coded messages as a long sequence of dashes on a strip of paper that was wound past the pen, see [18, p. 79]. Later versions of the system employed audible receivers, enabling telegraph operators to hear the messages and simultaneously transcribe them into plain text.

Samuel Morse solution to long distance communication laid the groundwork for a global communications network, culminating in the interconnection of much of the industrialized world by 1858. The final link in this infrastructure was the transatlantic telegraph cable, completed in August 1858. That cable was unfortunately damaged soon after it was installed but was replaced by a better one in 1864.

1.4 Alexander Graham Bell

The telephone was invented by Alexander Graham Bell in 1876. He grew up in Scotland but moved to Canada at the age of 24. Bell's recipe for success was a combination of determination, persistence, and enthusiasm. He conducted all the research that eventually led to the telephone in his spare time. By profession, Bell was a teacher of the deaf and mute; in 1871, he was appointed professor of teaching techniques for the deaf and mute at Boston University.

Bell's early research focused on the harmonic telegraph, an early form of frequency division multiplexing that allowed multiple channels to share a single telegraph line, each tuned to a distinct frequency. The device's receiver relied on a flexible diaphragm driven by an electromagnet. Upon noticing the diaphragm's visible mechanical vibrations, Bell realized the technology could serve as an assistive communication tool for the hearing-impaired. However, he soon expanded his scope toward a more ambitious objective: the transmission of human speech over telegraph lines. To achieve this, he needed both a transmitter and a receiver. The transmitter had to link sound to an electric current so that the current varied in step with the sound, while the receiver had to convert the current back into sound. A long telegraph line would carry the current between the transmitter and the receiver. Bell

and his assistant, Watson, used a so-called water microphone as the transmitter; this device created resistance that varied in rhythm with the sound. For the receiver, they used an electromagnet combined with a thin metal disc.

It was through a mistake that Bell and his assistant Watson got their first indication that they were on the right track. Watson accidentally bumped into a metal blade belonging to the transmitter, which caused Bell to hear a metallic ping in the receiver. After forty weeks of experimentation, they finally succeeded in transmitting speech on March 10, 1876. At the moment of success Bell was on the upper floor of his house and said to the transmitter "MR. WATSON, COME HERE, I WANT YOU." Watson, who was in the basement, at the lower end of the wire, heard Bell's voice from the receiver and rushed up the three stairs and shouted "I can hear you! I can hear the WORDS."

With this, the telephone was invented. Bell received a patent for his telephone and founded the Bell Telephone Company, which developed and sold telephones and telephone exchanges. This company formed the basis of AT&T, which is today the world's largest telecommunications company. Bell registered patents in several countries, but for some reason not in the Nordic countries. This was of great importance to Sweden where Lars Magnus Ericsson started a company that developed and sold telephones. Today this company is known as Ericsson.

1.5 Wireless telegraphy

In 1902, Guglielmo Marconi achieved a major breakthrough by transmitting a wireless signal across the Atlantic Ocean, connecting station sites in Ireland and Newfoundland. This historic milestone was made possible by several key technological advances:

- 1831: Michael Faraday discovered electromagnetic induction and introduced the concept of electric and magnetic field lines to describe electromagnetic phenomena.
- 1855: Wilhelm Weber and Rudolf Kohlrausch conducted a landmark experiment to measure the vacuum permittivity ϵ_0 . Since the vacuum permeability μ_0 had been established around 1825 by measuring forces between parallel current-carrying conductors, they noted that

$$c_0 = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \approx 3 \times 10^8 \text{ m/s}$$

which remarkably matched the known speed of light in vacuum.

- 1857: Gustav Kirchhoff formulated a theory for signal propagation along two-conductor telegraph lines with circular cross-sections [10]. By integrating circuit theory with electromagnetism, Kirchhoff demonstrated that signal voltages propagate along lossless conductors at the speed of light. Although his analysis relied on Weber's electrodynamics—which omitted displacement current (the induction of a magnetic field by a time-varying electric field) and

could not yield the general wave equation—Kirchhoff successfully obtained the correct wave behavior by applying the continuity equation to couple the time-dependent voltage and current along the line.

- 1865: James Clerk Maxwell publishes his complete theory of electromagnetism [15]. Building on the work of Weber, Kohlrausch, and Kirchhoff, Maxwell embraced Faraday’s field-centric perspective, formulating his theory in terms of electric and magnetic fields rather than action-at-a-distance forces used in Weber’s theory. This paradigm shift enabled Maxwell to realize that a time-varying electric field generates a magnetic field, which prompted him to introduce the displacement current to Ampère’s law and ultimately unify electromagnetic wave theory.
- 1885: Oliver Heaviside reformulates Maxwell’s original set of 20 coupled scalar equations into the compact vector form used today. By employing vector calculus, Heaviside cast the fundamental equations of electrodynamics in vacuum as:

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}\end{aligned}$$

- 1887: Heinrich Hertz experimentally demonstrated the existence of electromagnetic waves, providing the first definitive validation of Maxwell’s theory, see Figure 1.2. Eight years prior, while completing his doctorate, Hertz had been encouraged by his advisor, Hermann von Helmholtz, to test Maxwell’s predictions. Hertz finally pursued the problem after becoming a professor of physics at the University of Karlsruhe.

Hertz’s findings gained rapid acclaim in Great Britain, where physicists immediately recognized them as a confirmation of Maxwell’s field theory. In contrast, many German physicists initially resisted this interpretation, incorrectly claiming the results could be accounted for by Weber’s force-based electrodynamics.

Despite his monumental discovery, Hertz remained remarkably modest regarding its practical implications. When discussing the utility of the waves he had generated, he noted:

”It’s of no use whatsoever. This is just an experiment that proves Master Maxwell was right—we just have these mysterious electromagnetic waves that we cannot see with the naked eye. But they are there.”

When asked whether his discovery held any potential applications, he famously replied:

”Nothing, I think.”

- 1894: As a 20-year-old student at the University of Bologna, Guglielmo Marconi attended a lecture on Heinrich Hertz’s electromagnetic experiments. Inspired by the demonstration, he envisioned using these waves for wireless telegraphy and began experimenting at his family’s home. Through clever refinements to both the transmitter and receiver, he steadily increased the range of his transmissions.

By 1895, after successfully transmitting signals over 3.2 kilometers, Marconi sought to scale up his efforts. He wrote to the Italian Ministry of Post and Telegraphs requesting financial support, but his proposal was ignored. The Minister ultimately dismissed the letter, writing ”to the Longara” across it—a derogatory reference to Rome’s famous insane asylum on Via della Lungara. Undeterred, Marconi relocated to England, where he secured backing and established his wireless telegraphy company.

Prior to 1900, the technology behind wired and wireless communications was sufficiently straightforward for an engineer or physicist to grasp an entire system. Since then, system complexity has escalated continuously each year. Consequently, a comprehensive description of modern communication systems lies beyond the scope of this book. It is remarkable to reflect that all modern electrical technology owes its inception to the twitching of a frog’s leg in September 1786.

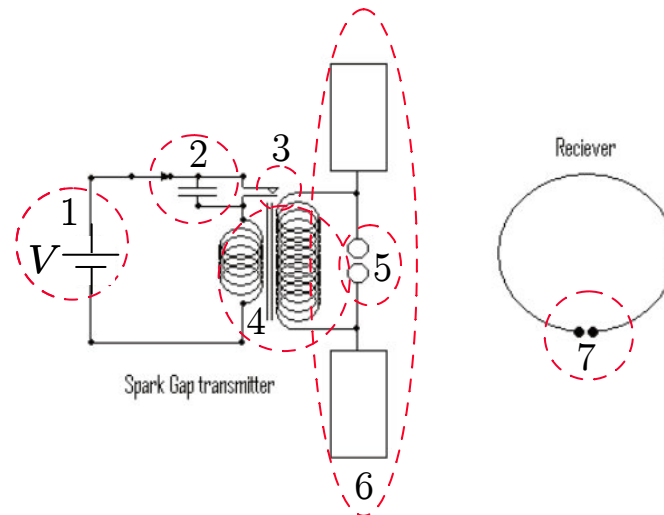


Figure 1.2: The equipment Hertz used to detect electromagnetic waves.

- The voltage source 1 provides a constant voltage V .
- The capacitor 2 is charged by the current from the voltage source. That creates an electric force on the switch 3.
- When the force is large enough the switch 3 closes, the capacitor is then short-circuited, causing the voltage V to be applied across the left coil of the transformer 4. The right coil of the transformer has significantly more turns than the left one. Through induction, a voltage much higher than V is thereby generated across the right coil of the transformer.
- The voltage across the right coil is high enough to create a spark between the two spheres 5.
- The spark constitutes a current pulse that excites the antenna 6. The antenna consists of two rectangular plates and the wires leading down to the spheres 5. This excitation causes the antenna to emit a powerful electromagnetic wave with a wavelength equal to half the length of the antenna. The antenna is approximately 1.5 meters long, resulting in a wave frequency of 100 MHz.
- The wave strikes the circular receiving antenna. The wave's magnetic field is perpendicular to the circle, inducing a voltage between the two spheres 7. This voltage is high enough to produce clearly visible sparks between the spheres.
- Everything repeats in a cyclical process, the period of which is determined by how quickly capacitor 2 charges.

Chapter 2

The Maxwell equations

The Maxwell equations constitute the fundamental mathematical model for all theoretical analysis of macroscopic electromagnetic phenomena. James Clerk Maxwell¹ published these famous equations in 1865 [15]. 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

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (2.1)$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \mathbf{J}(\mathbf{r}, t) + \frac{\partial \mathbf{D}(\mathbf{r}, t)}{\partial t} \quad (2.2)$$

The equation (2.1) (or the corresponding integral formulation) is the *Faraday's law of induction*³, and the equation (2.2) is the *Ampère's (generalized) law*.⁴ The vector fields in the Maxwell equations are⁵:

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

All of these fields are functions of the space coordinates $\mathbf{r}$ and time t , but these arguments are often suppressed. We give the arguments only when equations or expressions can be misinterpreted

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

²A detailed derivation of these macroscopic equations from a microscopic formulation is found in [8, 19].

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

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

⁵For simplicity we sometimes use the names **E**-field, **D**-field, **B**-field and **H**-field.

2 The Maxwell equations

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

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (2.3)$$

where q is the electric charge of the particle and $\mathbf{v}$ 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 $\mathbf{J}$. 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 $\frac{\partial \mathbf{D}}{\partial t}$. In Chapter 3 we address the differences between the electric flux density $\mathbf{D}$ and the electric field $\mathbf{E}$, as well as the differences between the magnetic field $\mathbf{H}$ and the magnetic flux density $\mathbf{B}$.

One of the fundamental principles of physics is the conservation of electric charge:

$$\nabla \cdot \mathbf{J}(\mathbf{r}, t) + \frac{\partial \rho(\mathbf{r}, t)}{\partial t} = 0 \quad (2.4)$$

Here $\rho(\mathbf{r}, t)$ is the charge density (charge/unit volume) that is associated with the current density $\mathbf{J}(\mathbf{r}, t)$.

Two additional equations are usually associated with the Maxwell equations.

$$\nabla \cdot \mathbf{B} = 0 \quad (2.5)$$

$$\nabla \cdot \mathbf{D} = \rho \quad (2.6)$$

The equation (2.5) tells us that there are no magnetic charges and that the magnetic flux is conserved while (2.6) is *Gauss law*. Under suitable assumptions, both of these equations can be derived from (2.1), (2.2) and (2.4). To see this, we take the divergence of (2.1) and (2.2). This implies

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

since $\nabla \cdot (\nabla \times \mathbf{A}) \equiv 0$. We interchange the order of differentiation and use (2.4) and get

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

These equations imply

$$\begin{cases} \nabla \cdot \mathbf{B}(\mathbf{r}, t) = f_1(\mathbf{r}) \\ \nabla \cdot \mathbf{D}(\mathbf{r}, t) - \rho(\mathbf{r}, t) = f_2(\mathbf{r}) \end{cases}$$

where $f_1(\mathbf{r})$ and $f_2(\mathbf{r})$ do not explicitly depend on time t . If the fields $\mathbf{B}$, $\mathbf{D}$ and ρ are identically zero before a finite time, τ i.e.,

$$\begin{cases} \mathbf{B}(\mathbf{r}, t) = \mathbf{0} \\ \mathbf{D}(\mathbf{r}, t) = \mathbf{0} \\ \rho(\mathbf{r}, t) = 0 \end{cases} \quad t < \tau \quad (2.7)$$

then (2.5) and (2.6) follow. Static or time-harmonic fields do not satisfy (2.7), since there is no finite time τ before the fields are all zero.⁶ We assume that (2.7) is valid for time-dependent fields and then it is sufficient to use the equations (2.1), (2.2) and (2.4).

Equations (2.1) and (2.2) yield six scalar component equations. For a given current density $\mathbf{J}$, Maxwell's equations leave 12 unknowns distributed among $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$, and $\mathbf{H}$. Closing this six-equation gap requires the *constitutive relations*, introduced in the next chapter.

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

$$\begin{cases} \mathbf{D} = \epsilon_0 \mathbf{E} \\ \mathbf{B} = \mu_0 \mathbf{H} \end{cases} \quad (2.8)$$

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 $\mu_0 \approx 4\pi \cdot 10^{-7}$ Vs/Am.

Inside a material there is a difference between the field $\epsilon_0 \mathbf{E}$ and the electric flux density $\mathbf{D}$, and between the magnetic flux density $\mathbf{B}$ and the field $\mu_0 \mathbf{H}$. 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 $\mathbf{P}$ and the magnetization $\mathbf{M}$. The definitions of these fields are

$$\mathbf{P} = \mathbf{D} - \epsilon_0 \mathbf{E} \quad (2.9)$$

$$\mathbf{M} = \frac{1}{\mu_0} \mathbf{B} - \mathbf{H} \quad (2.10)$$

The polarization $\mathbf{P}$ 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 $\mathbf{M}$ is the volume density of magnetic dipole moment and hence a measure of the net (bounded) currents in the material. The origin of $\mathbf{M}$ can also be *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.

2.1 Boundary conditions at interfaces

This section derives the boundary conditions for electromagnetic fields at an interface between two distinct materials, focusing exclusively on stationary surfaces.

The Maxwell equations in (2.1)–(2.2), assume that all field quantities are differentiable functions of space and time. At an interface between two materials, the fields 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.

⁶We will return to the derivation of equations (2.5) and (2.6) for time-harmonic fields in Chapter 3 on Page 15.

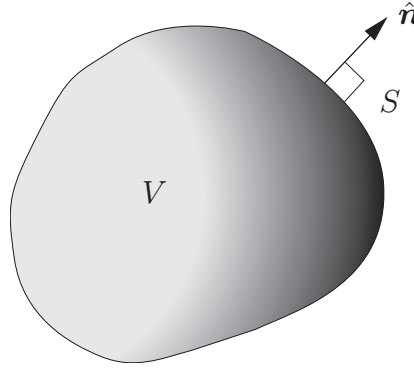


Figure 2.1: Geometry of integration.

We let V be an arbitrary (simply connected) volume, bounded by the surface S with unit outward normal vector $\hat{\mathbf{n}}$, see Figure 2.1. We integrate the Maxwell equations, (2.1)–(2.2) and (2.5)–(2.6), over the volume V and obtain

$$\begin{aligned}
\iiint_V \nabla \times \mathbf{E} \, dv &= - \iiint_V \frac{\partial \mathbf{B}}{\partial t} \, dv \\
\iiint_V \nabla \times \mathbf{H} \, dv &= \iiint_V \mathbf{J} \, dv + \iiint_V \frac{\partial \mathbf{D}}{\partial t} \, dv \\
\iiint_V \nabla \cdot \mathbf{B} \, dv &= 0 \\
\iiint_V \nabla \cdot \mathbf{D} \, dv &= \iiint_V \rho \, dv
\end{aligned} \tag{2.11}$$

where dv is the volume measure ($dv = dx \, dy \, dz$).

The following two integration theorems for vector fields are useful:

$$\begin{aligned}
\iiint_V \nabla \cdot \mathbf{A} \, dv &= \iint_S \mathbf{A} \cdot \hat{\mathbf{n}} \, dS \\
\iiint_V \nabla \times \mathbf{A} \, dv &= \iint_S \hat{\mathbf{n}} \times \mathbf{A} \, dS
\end{aligned}$$

Here $\mathbf{A}$ is a continuously differentiable vector field in V , and dS 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 2.1).

After interchanging the differentiation w.r.t. time t and integration (volume V

⁷Distinguish between the Gauss law, (2.6), and the Gauss theorem.

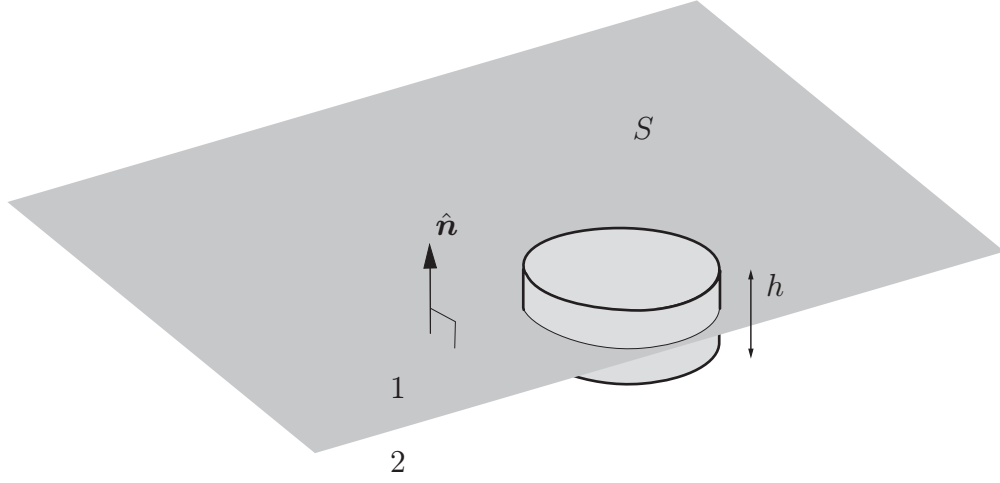


Figure 2.2: Interface between two different media 1 and 2.

is fixed in time and we assume all fields to be sufficiently regular) (2.11) reads

$$\iint_S \hat{\mathbf{n}} \times \mathbf{E} \, dS = -\frac{d}{dt} \iiint_V \mathbf{B} \, dv \quad (2.12)$$

$$\iint_S \hat{\mathbf{n}} \times \mathbf{H} \, dS = \iiint_V \mathbf{J} \, dv + \frac{d}{dt} \iiint_V \mathbf{D} \, dv \quad (2.13)$$

$$\iint_S \mathbf{B} \cdot \hat{\mathbf{n}} \, dS = 0 \quad (2.14)$$

$$\iint_S \mathbf{D} \cdot \hat{\mathbf{n}} \, dS = \iiint_V \rho \, dv \quad (2.15)$$

In a domain V where the fields $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$ and $\mathbf{H}$ are continuously differentiable, these integral expressions are equivalent to the differential equations (2.1) and (2.6). 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, (2.12)–(2.15), 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 (2.1)–(2.2). The fields $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$ and $\mathbf{H}$, that satisfy the equations (2.12)–(2.15) are called weak solutions to the Maxwell equations, in the case the fields are not continuously differentiable and (2.1)–(2.2) lack meaning.

The integral expressions (2.12)–(2.15) are applied to a volume V_h that cuts the interface between two different media, see Figure 2.2. The unit normal $\hat{\mathbf{n}}$ of the interface S is directed from medium 2 into medium 1. We assume that all electromagnetic fields $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$ and $\mathbf{H}$, 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 $\mathbf{E}_1$ and $\mathbf{E}_2$, respectively, and a similar notation, with

6 The Maxwell equations

index 1 or 2, is adopted for the other fields. The current density $\mathbf{J}$ and the charge density ρ can adopt infinite values at the interface for perfectly conducting (metal) surfaces.⁸ It is convenient to introduce a *surface current density* $\mathbf{J}_S$ and *surface charge density* ρ_S as a limit process

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

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 $\mathbf{J}$ and ρ go to infinity in such a way that $\mathbf{J}_S$ and ρ_S have well defined values in this process. The surface current density $\mathbf{J}_S$ 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 $\frac{d}{dt} \iiint_{V_h} \mathbf{B} dv$ and $\frac{d}{dt} \iiint_{V_h} \mathbf{D} dv$ approach zero as $h \rightarrow 0$, since the fields $\mathbf{B}$ and $\mathbf{D}$ 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 (2.12)–(2.15) approach zero as $h \rightarrow 0$. The contribution from the upper part (unit normal $\hat{\mathbf{n}}$) and lower part (unit normal $-\hat{\mathbf{n}}$) 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 $h \rightarrow 0$ are

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

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

$$\begin{cases} \hat{\mathbf{n}} \times (\mathbf{E}_1 - \mathbf{E}_2) = \mathbf{0} \\ \hat{\mathbf{n}} \times (\mathbf{H}_1 - \mathbf{H}_2) = \mathbf{J}_S \\ \hat{\mathbf{n}} \cdot (\mathbf{B}_1 - \mathbf{B}_2) = 0 \\ \hat{\mathbf{n}} \cdot (\mathbf{D}_1 - \mathbf{D}_2) = \rho_S \end{cases} \quad (2.16)$$

These boundary conditions prescribe how the electromagnetic fields on each side of the interface are related to each other. 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 $\mathbf{J}_S$. If the surface current density

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

is zero, eg., when the material has finite conductivity⁹, 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 2.3 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 (2.16) 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 (2.16) is reduced to

$$\begin{cases} \hat{\mathbf{n}} \times \mathbf{E}_1 = \mathbf{0} \\ \hat{\mathbf{n}} \times \mathbf{H}_1 = \mathbf{J}_S \\ \hat{\mathbf{n}} \cdot \mathbf{B}_1 = 0 \\ \hat{\mathbf{n}} \cdot \mathbf{D}_1 = \rho_S \end{cases} \quad (2.17)$$

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

2.1.1 Impedance boundary conditions

While the approximate boundary condition in (2.17) often suffices at a dielectric-metal interface, higher precision requires a more rigorous approach. One option is to treat the dielectric and metal as separate domains and apply the exact boundary conditions in (2.16). However, this requires solving for the fields in both regions. In finite element methods (FEM), this approach quickly becomes computationally prohibitive: because the wavelength inside a conductor is drastically shorter than in free space (see Section 6.9), the metallic region demands a significantly finer mesh, driving up memory consumption and execution time. To bypass this issue, a third option is to employ an impedance boundary condition (IBC) (derived in Section 6.9). The IBC relates the electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$ evaluated on the dielectric side of the surface via:

$$\begin{aligned} \mathbf{E} - \hat{\mathbf{n}}(\mathbf{E} \cdot \hat{\mathbf{n}}) &= \eta_s \hat{\mathbf{n}} \times \mathbf{H} \\ \eta_s &= (1 - i) \sqrt{\frac{\omega \mu_0}{2\sigma}} = \frac{1 - i}{\sigma \delta} \end{aligned} \quad (2.18)$$

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

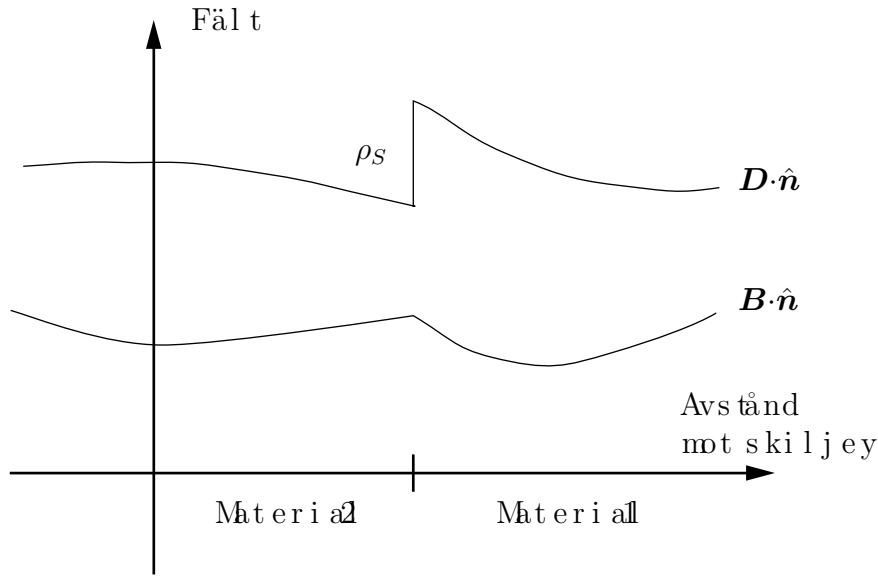


Figure 2.3: The variation of the normal components B_n and D_n at the interface between two different media.

Here, η_s is the wave impedance of the metal, and $\delta = \sqrt{2/(\omega\mu_0\sigma)}$ denotes its skin depth (cf. Section 6.9). Note that $\mathbf{E} - \hat{\mathbf{n}}(\mathbf{E} \cdot \hat{\mathbf{n}})$ is the tangential component of the electric field at the boundary.

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

2.2 Energy conservation and Poynting's theorem

Energy conservation is shown from the Maxwell equations (2.1) and (2.2).

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

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

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

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

$$\nabla \cdot \mathbf{S} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{E} \cdot \mathbf{J} = 0 \quad (2.19)$$

We have here introduced the *Poynting's vector*,¹⁰ $\mathbf{S} = \mathbf{E} \times \mathbf{H}$, which gives the power flow per unit area in the direction of the vector $\mathbf{S}$. The energy conservation is made

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

visible if we integrate equation (2.19) over a volume V , bounded by the surface S with unit outward normal vector $\hat{\mathbf{n}}$, see Figure 2.1, and use the divergence theorem. We get

$$\begin{aligned} \iint_S \mathbf{S} \cdot \hat{\mathbf{n}} \, dS &= \iiint_V \nabla \cdot \mathbf{S} \, dv \\ &= - \iiint_V \left[\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \right] dv - \iiint_V \mathbf{E} \cdot \mathbf{J} \, dv \end{aligned} \quad (2.20)$$

The terms are interpreted in the following way:

- The left hand side:

$$\iint_S \mathbf{S} \cdot \hat{\mathbf{n}} \, dS$$

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

$$\iiint_V \left[\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \right] dv$$

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 (2.20)

$$\iiint_V \mathbf{E} \cdot \mathbf{J} \, dv$$

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

To this end, (2.20) expresses energy balance or more correctly power balance in the volume V , i.e.,

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

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 2.2, Poynting's vector in medium 1 close to the interface is

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

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

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

The boundary condition at the interface is given by (2.16).

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

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

$$\iint_S \mathbf{S}_1 \cdot \hat{\mathbf{n}} \, dS = \iint_S \mathbf{S}_2 \cdot \hat{\mathbf{n}} \, dS - \iint_S \mathbf{E}_2 \cdot \mathbf{J}_S \, dS \quad (2.21)$$

where the surface S is an arbitrary part of the interface. Note that the unit normal $\hat{\mathbf{n}}$ 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 (2.21) since the surface current density $\mathbf{J}_S$ is parallel to the interface S and the tangential components of the electric field are continuous across the interface, i.e.,

$$\iint_S \mathbf{E}_1 \cdot \mathbf{J}_S \, dS = \iint_S \mathbf{E}_2 \cdot \mathbf{J}_S \, dS$$

Equation (2.21) is easily proved by a cyclic permutation of the vectors and the use of the boundary conditions.

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

By integration of this expression over the interface S we obtain power conservation over the surface S as expressed in equation (2.21).

Problems in Chapter 2

2.1 Show the following analogous theorem of Gauss theorem:

$$\iiint_V \nabla \times \mathbf{A} \, dv = \iint_S \hat{\mathbf{n}} \times \mathbf{A} \, dS$$

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

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

$$\iiint \mathbf{B} \cdot \mathbf{H} \, dv = 0$$

where the integration is over all space.

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

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

Use the divergence theorem to prove the problem.

- 2.3** 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 $\mathbf{J}$ 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 $\mathbf{S} = -\hat{\rho}_{\frac{1}{2}}a\sigma E^2$ 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

$$\begin{aligned} \iint_S \mathbf{S} \cdot \hat{\mathbf{n}} \, dS &= -\pi a^2 l \sigma E^2 \\ \iiint_V \mathbf{E} \cdot \mathbf{J} \, dv &= \pi a^2 l \sigma E^2 \end{aligned}$$

Chapter 3

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 $\mathbf{E}(\mathbf{r}, t)$, is defined as

$$\mathbf{E}(\mathbf{r}, \omega) = \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, t) e^{i\omega t} dt$$

with its inverse transform

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

The Fourier transform for all other time dependent fields are defined in the same way. We use the same symbol for the physical field $\mathbf{E}(\mathbf{r}, t)$, as for the Fourier transformed field $\mathbf{E}(\mathbf{r}, \omega)$ —only the argument differs.

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 $\mathbf{E}$ to be real.

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

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

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

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

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

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,

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
	$4.2\text{--}7.9 \cdot 10^{14}$ Hz	0.38–0.72 μm	Visible light

$\omega \rightarrow -\omega$, and the use of the condition above, we have

$$\begin{aligned}
\mathbf{E}(\mathbf{r}, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^0 \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega + \frac{1}{2\pi} \int_0^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega \\
&= \frac{1}{2\pi} \int_0^{\infty} [\mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} + \mathbf{E}(\mathbf{r}, -\omega) e^{i\omega t}] d\omega \\
&= \frac{1}{2\pi} \int_0^{\infty} [\mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} + \mathbf{E}^*(\mathbf{r}, \omega) e^{i\omega t}] d\omega = \frac{1}{\pi} \operatorname{Re} \int_0^{\infty} \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} d\omega
\end{aligned} \tag{3.1}$$

where $\operatorname{Re} z$ 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 3. 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 $\mathbf{E}(\mathbf{r}, \omega)$ is related to the time harmonic field $\mathbf{E}(\mathbf{r}, t)$ via the rule

$$\mathbf{E}(\mathbf{r}, t) = \operatorname{Re} \{ \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t} \} \tag{3.2}$$

3.1 The Maxwell equations

To begin analyzing time-harmonic fields, we transform Maxwell's equations (2.1) and (2.2) into the frequency domain using either phasors or a Fourier transform. Replacing the time derivative $\frac{\partial}{\partial t}$ with $-i\omega$ yields

$$\nabla \times \mathbf{E}(\mathbf{r}, \omega) = i\omega \mathbf{B}(\mathbf{r}, \omega) \tag{3.3}$$

$$\nabla \times \mathbf{H}(\mathbf{r}, \omega) = \mathbf{J}(\mathbf{r}, \omega) - i\omega \mathbf{D}(\mathbf{r}, \omega) \tag{3.4}$$

Note that the electromagnetic fields $\mathbf{E}(\mathbf{r}, \omega)$, $\mathbf{B}(\mathbf{r}, \omega)$, $\mathbf{D}(\mathbf{r}, \omega)$ and $\mathbf{H}(\mathbf{r}, \omega)$, and the current density $\mathbf{J}(\mathbf{r}, \omega)$ are complex vector fields.

The continuity equation (2.4) is transformed in a similar way

$$\nabla \cdot \mathbf{J}(\mathbf{r}, \omega) - i\omega\rho(\mathbf{r}, \omega) = 0 \quad (3.5)$$

The remaining two equations from Chapter 2, (2.5) and (2.6), are transformed into

$$\nabla \cdot \mathbf{B}(\mathbf{r}, \omega) = 0 \quad (3.6)$$

$$\nabla \cdot \mathbf{D}(\mathbf{r}, \omega) = \rho(\mathbf{r}, \omega) \quad (3.7)$$

These equations are a consequence of (3.3) and (3.4) and the continuity equation (3.5) (c.f., Chapter 2 on Page 2). To see this we take the divergence of the Maxwell equations (3.3) and (3.4), and get $(\nabla \cdot (\nabla \times \mathbf{A}) = 0)$

$$\begin{aligned} i\omega \nabla \cdot \mathbf{B}(\mathbf{r}, \omega) &= 0 \\ i\omega \nabla \cdot \mathbf{D}(\mathbf{r}, \omega) &= \nabla \cdot \mathbf{J}(\mathbf{r}, \omega) = i\omega\rho(\mathbf{r}, \omega) \end{aligned}$$

Division by $i\omega$ (provided $\omega \neq 0$) gives (3.6) and (3.7).

In a homogenous non-magnetic source free medium we obtain the vector Helmholtz equation for the electric field by eliminating the magnetic field from (3.3) and (3.4). This is done by taking the rotation of (3.3) and utilizing (3.4). The result is

$$\nabla^2 \mathbf{E}(\mathbf{r}, \omega) + k(\omega)^2 \mathbf{E}(\mathbf{r}, \omega) = \mathbf{0} \quad (3.8)$$

where

$$k(\omega) = \omega \sqrt{\epsilon_0 \mu_0 (\epsilon + i\sigma/(\omega\epsilon_0))} \quad (3.9)$$

is the wavenumber. Also the magnetic field satisfies the vector Helmholtz equation

$$\nabla^2 \mathbf{H}(\mathbf{r}, \omega) + k(\omega)^2 \mathbf{H}(\mathbf{r}, \omega) = \mathbf{0} \quad (3.10)$$

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

$$\begin{cases} \nabla \times \mathbf{E}(\mathbf{r}, \omega) = ik_0 (c_0 \mathbf{B}(\mathbf{r}, \omega)) \\ \nabla \times (\eta_0 \mathbf{H}(\mathbf{r}, \omega)) = -ik_0 (c_0 \eta_0 \mathbf{D}(\mathbf{r}, \omega)) \end{cases} \quad (3.11)$$

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 (3.11) the field quantities $\mathbf{E}$, $c_0 \mathbf{B}$, $\eta_0 \mathbf{H}$, $c_0 \eta_0 \mathbf{D}$ have the same unit V/m. The standard formulation (3.11) is used throughout this book.

3.2 Constitutive relations

The constitutive relations are the relations between the fields $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$ and $\mathbf{H}$. 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

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

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 $\mathbf{J}(\mathbf{r}, \omega) = \sigma(\omega)\mathbf{E}(\mathbf{r}, \omega)$, always can be included in the constitutive relations by redefining the permittivity .

$$\epsilon_c = \epsilon + i \frac{\sigma}{\omega \epsilon_0}$$

The right hand side in Ampère's law (3.4) is

$$\mathbf{J} - i\omega \mathbf{D} = \sigma \mathbf{E} - i\omega \epsilon_0 \epsilon \mathbf{E} = -i\omega \epsilon_0 \epsilon_c \mathbf{E}$$

3.3 Poynting's theorem

In Chapter 2 we derived Poynting's theorem, see (2.19) on Page 8,

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

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 period¹. We denote the time average as $\langle \cdot \rangle$ and for Poynting's theorem we get

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

The different terms in this quantity are

$$\langle \mathbf{S}(t) \rangle = \frac{1}{2} \text{Re} \{ \mathbf{E}(\omega) \times \mathbf{H}^*(\omega) \} \quad (3.12)$$

and

$$\begin{aligned} \langle \mathbf{H}(t) \cdot \frac{\partial \mathbf{B}(t)}{\partial t} \rangle &= \frac{1}{2} \text{Re} \{ i\omega \mathbf{H}(\omega) \cdot \mathbf{B}^*(\omega) \} \\ \langle \mathbf{E}(t) \cdot \frac{\partial \mathbf{D}(t)}{\partial t} \rangle &= \frac{1}{2} \text{Re} \{ i\omega \mathbf{E}(\omega) \cdot \mathbf{D}^*(\omega) \} \\ \langle \mathbf{E}(t) \cdot \mathbf{J}(t) \rangle &= \frac{1}{2} \text{Re} \{ \mathbf{E}(\omega) \cdot \mathbf{J}^*(\omega) \} \end{aligned}$$

¹The 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 $T = 2\pi/\omega$.

$$\begin{aligned} \langle f_1(t)f_2(t) \rangle &= \frac{1}{T} \int_0^T f_1(t)f_2(t) dt = \frac{1}{T} \int_0^T \text{Re} \{ f_1(\omega)e^{-i\omega t} \} \text{Re} \{ f_2(\omega)e^{-i\omega t} \} dt \\ &= \frac{1}{4T} \int_0^T \{ f_1(\omega)f_2(\omega)e^{-2i\omega t} + f_1^*(\omega)f_2^*(\omega)e^{2i\omega t} + f_1(\omega)f_2^*(\omega) + f_1^*(\omega)f_2(\omega) \} dt \\ &= \frac{1}{4} \{ f_1(\omega)f_2^*(\omega) + f_1^*(\omega)f_2(\omega) \} = \frac{1}{2} \text{Re} \{ f_1(\omega)f_2^*(\omega) \} \end{aligned}$$

Poynting's theorem (balance of power) for time harmonic fields becomes ($\nabla \cdot \mathbf{S}(t) = \nabla \cdot \langle \mathbf{S}(t) \rangle$):

$$\begin{aligned} \nabla \cdot \langle \mathbf{S}(t) \rangle + \frac{1}{2} \operatorname{Re} \{ i\omega [\mathbf{H}(\omega) \cdot \mathbf{B}^*(\omega) + \mathbf{E}(\omega) \cdot \mathbf{D}^*(\omega)] \} \\ + \frac{1}{2} \operatorname{Re} \{ \mathbf{E}(\omega) \cdot \mathbf{J}^*(\omega) \} = 0 \end{aligned} \quad (3.13)$$

Of special interest is the case without currents² $\mathbf{J} = \mathbf{0}$. Poynting's theorem is then simplified to

$$\begin{aligned} \nabla \cdot \langle \mathbf{S}(t) \rangle &= -\frac{1}{2} \operatorname{Re} \{ i\omega [\mathbf{H}(\omega) \cdot \mathbf{B}^*(\omega) + \mathbf{E}(\omega) \cdot \mathbf{D}^*(\omega)] \} \\ &= -\frac{i\omega}{4} \left\{ \mathbf{H}(\omega) \cdot \mathbf{B}^*(\omega) - \mathbf{H}^*(\omega) \cdot \mathbf{B}(\omega) \right. \\ &\quad \left. + \mathbf{E}(\omega) \cdot \mathbf{D}^*(\omega) - \mathbf{E}(\omega)^* \cdot \mathbf{D}(\omega) \right\} \end{aligned}$$

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

Problems in Chapter 3

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

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

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

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

where ξ is an arbitrary real number.

3.2 For real vectors $\mathbf{A}$ and $\mathbf{B}$ we have

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

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

²Conducting currents can, as we have seen, be included in the permittivity dyadic ϵ .

Chapter 4

Transmission lines

To accurately analyze circuit signals and apply appropriate modeling approximations, we must consider both their frequency band and the physical size of the circuit. We demonstrate this by examining signals ranging from DC to very high frequencies within a circuit composed of linear elements, specifically resistors, capacitors, inductors, and sources.

Definition: A circuit is considered discrete (or lumped) if wave propagation effects can be safely neglected during analysis. In practice, a circuit meets this criterion when its physical dimensions are significantly smaller than the free-space wavelength $\lambda = c_0/f$.

- We begin with zero frequency ($f = 0$), where the wavelength is infinite and the circuits are strictly lumped. Here, capacitors act as open circuits and inductors as short circuits. Along a wire with negligible resistance, current is uniform in both time and space, and the voltage drop vanishes. The voltages and currents in the circuits are governed by Ohm's and Kirchhoff's laws, which follow directly from electrostatic and stationary current relations:

$$\nabla \times \mathbf{E}(\mathbf{r}) = \mathbf{0}$$

$$\mathbf{J}(\mathbf{r}) = \sigma \mathbf{E}(\mathbf{r})$$

$$\nabla \cdot \mathbf{J}(\mathbf{r}) = 0$$

- At non-zero frequencies where the wavelength remains much larger than the circuit's physical dimensions, the circuits still behave as lumped-element networks. However, time variations now induce energy storage. Applying Faraday's law (2.1) and the continuity equation (2.4) yields the time-dependent terminal relations between current i and voltage v for capacitors and inductors:

$$i = C \frac{dv}{dt}$$

$$v = L \frac{di}{dt}$$

where C denotes the capacitance and L denotes the inductance. Together with Ohm's law and Kirchhoff's laws, these relations form a complete set of equations sufficient for determining all voltages and currents across the circuits.

The wires that connect circuit elements are assumed to have negligible resistance, inductance, and capacitance. This ensures that the current and voltage in each wire are constant in space, but not in time.

- Next, we consider the frequency band where the wavelength is no longer much larger than the physical dimensions of the circuit. In this regime, wave propagation effects become significant, causing the phase and amplitude of currents and voltages along the conductors to vary with both time and space. Consequently, lumped-element circuit theory is no longer valid, and we must transition to transmission line theory—the central focus of this chapter. While rooted in the full set of the Maxwell equations, this theory is conveniently formulated in terms of distributed voltages and currents.
- At even higher frequencies, we reach a regime where transmission line theory alone is insufficient to fully characterize the system. This occurs when circuit components and interconnects begin to act as antennas, resulting in the unintended radiation of electromagnetic waves. Accurately describing the circuit under these conditions requires a combined framework of both electromagnetic field theory and transmission line theory.

4.1 Time and frequency domain

Analyzing signals in linear circuits is often significantly more advantageous when conducted in the frequency domain. To establish the necessary foundation, we briefly review the core time-to-frequency domain transformation rules introduced in Chapter 3, supplemented by a concise overview of transformations utilizing Fourier series and the Laplace transform. Crucially, the resulting algebraic relationships between voltages and currents remain identical in the frequency domain, regardless of which specific transformation method is applied. Throughout this book, we will primarily employ phasors and the Fourier transform to transition between the time and frequency domains.

4.1.1 Phasors ($j\omega$ method)

For time-harmonic signals, we employ phasors. The transformation between the time domain and the frequency domain is defined as:

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

where V is the complex voltage. This convention is equivalent to the mapping $v(t) = \text{Re}\{V e^{j\omega t}\}$, as introduced in Chapter 3. In general, the linear relationship between complex voltage and current is expressed as Ohm's law in the frequency

domain, $V = ZI$, where Z denotes the impedance. Consequently, the impedance is R for a resistor, $j\omega L$ for an inductor, $1/j\omega C$ for a capacitor. The admittance, defined as the reciprocal of impedance ($Y = 1/Z$), is also utilized extensively throughout this chapter.

4.1.2 Fourier transformation

If the time-domain signal $v(t)$ is absolutely integrable, that is, it satisfies the condition $\int_{-\infty}^{\infty} |v(t)| dt < \infty$. Its forward and inverse Fourier transforms are given by:

$$\begin{aligned} V(\omega) &= \int_{-\infty}^{\infty} v(t) e^{-j\omega t} dt \\ v(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} V(\omega) e^{j\omega t} d\omega \end{aligned} \quad (4.1)$$

Note that the definition of the Fourier transform used here differs from the convention adopted in Chapter 3, where $e^{-i\omega t}$ is replaced by $e^{j\omega t}$ (see the discussion below). As established in Chapter 3, the presence of negative angular frequencies poses no physical issue for real-valued signals, as they can be accounted for via the conjugate symmetry relation:

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

In the frequency domain the relations between current and voltage are identical with the corresponding relations obtained by phasors

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

Remark: (Comment on j and i) *The electrical engineering literature typically adopts the time convention $e^{j\omega t}$ for both the phasor method and the Fourier transform, whereas the physics literature generally uses $e^{-i\omega t}$. Expressions can be translated from one convention to the other by taking the complex conjugate of all equations and mapping i to j . In this chapter, we adhere to the electrical engineering convention ($e^{j\omega t}$), whereas the remainder of the book employs the physics convention ($e^{-i\omega t}$). This shift is motivated by literature conventions: transmission lines are predominantly analyzed within electrical engineering frameworks, while hollow metallic and dielectric waveguides are more traditionally treated using physics-based conventions.*

4.1.3 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

$\{e^{jn\omega_0 t}\}_{n=-\infty}^{n=\infty}$ is a complete orthogonal system of functions on an interval of length T and we may expand $f(t)$ in a Fourier series as

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

The Fourier coefficients c_m are determined by exploiting the orthogonality of the basis functions. Multiplying both sides of the expansion by $e^{-jm\omega_0 t}$ and integrating over one full period yields:

$$c_m = \frac{1}{T} \int_0^T f(t) e^{-jm\omega_0 t} dt$$

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

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

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

By expanding the time-domain current and voltage into their respective Fourier series,

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

the governing relations between the spectral coefficients V_n and I_n for the fundamental circuit elements become:

$$\begin{cases} V_n = RI_n & \text{resistor} \\ V_n = jn\omega_0 LI_n & \text{inductor} \\ V_n = \frac{I_n}{jn\omega_0 C} & \text{capacitor} \end{cases}$$

Consequently, determining the Fourier coefficients for the currents and voltages within a linear circuit is a straightforward extension of phasor analysis. The explicit Fourier series expansions will not be heavily utilized in this chapter.

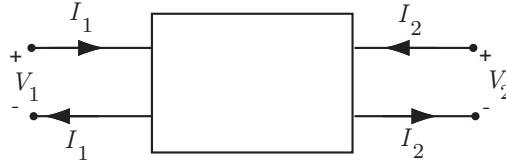


Figure 4.1: A two-port. Notice that the total current entering each port is always zero.

4.1.4 Laplace transformation

For a time-domain signal $v(t)$ defined for $t \geq 0$, we may apply the unilateral Laplace transform:

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

In practice, lookup tables are typically utilized to perform the inverse mapping from $V(s)$ back to $v(t)$. Substituting the complex frequency s with $j\omega$ directly yields the corresponding expressions for the steady-state $j\omega$ -method and the Fourier transform. The Laplace transform is exceptionally well suited for analyzing transient behaviors, system stability, and overall frequency responses. Under this transform, the algebraic relationships between the complex voltage and current variables are given by:

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

4.2 Two-ports

A two-port network is an electrical circuit with two pairs of terminals, shown in Figure 4.1. By definition, the net current entering each port is zero. This book focuses exclusively on passive, linear two-ports—meaning the network contains no independent sources. In the frequency domain, we represent a two-port using a 2×2 matrix of complex values, where the choice of matrix elements depends on how we pair the port voltages (V_1, V_2) and currents (I_1, I_2) . We now present some different type of pairings

4.2.1 The impedance matrix

The voltages and currents in Figure 4.1 can be related via the impedance matrix $[Z]$

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = [Z] \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} \quad (4.2)$$

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

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

4.2.2 The chain matrix (*ABCD*-matrix)

Another key representation relates the input variables directly to the output variables

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = [K] \begin{pmatrix} V_2 \\ -I_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_2 \\ -I_2 \end{pmatrix} \quad (4.3)$$

The matrix $[K]$ is commonly referred to as the *ABCD* matrix (or chain matrix). Notice the negative sign assigned to I_2 ; this convention simplifies the analysis when cascading multiple two-port networks.

Inverting this relationship yields the reverse representation

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

Note that the inverse matrix $[K']$ can be derived directly from $[K]^{-1}$ simply by flipping the signs of its off-diagonal elements.

4.2.3 The hybrid matrix

Alternatively, the two-port parameters can be expressed using the hybrid matrix $[H]$, which mixes voltage and current variables

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

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

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

4.2.4 Reciprocity

Consider a two-port network where a signal source (voltage or current) is applied to one port, and the resulting response is measured at the other port. If we swap the positions of the source and the detector, and the measured response remains identical, 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

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

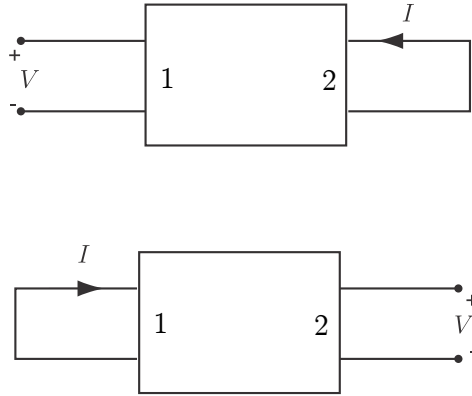


Figure 4.2: Reciprocal two-port. If voltage V at port 1 gives the shortening current I in port 2 then the voltage V at port 2 gives the shortening current I at port 1.

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 4.2. 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.

4.2.5 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.

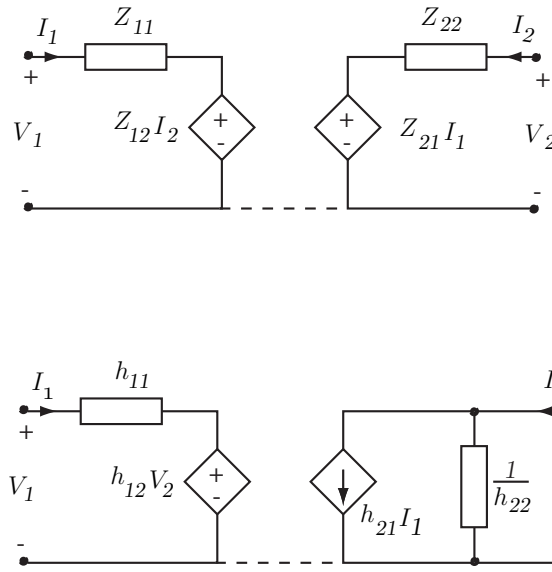


Figure 4.3: Equivalent circuits for a passive two-port. The upper corresponds to the impedance representation and the lower to the hybrid representation.

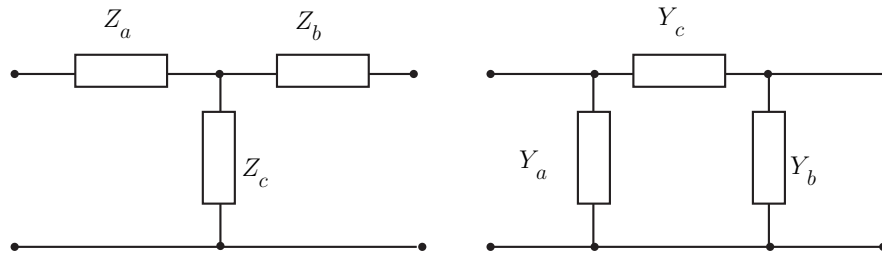


Figure 4.4: Equivalent T - and Π -circuits for a reciprocal passive two-port.

4.2.6 Circuit models for two-ports

In general, a linear two-port network is fully specified by four complex parameters. It can be modeled using an equivalent circuit containing two impedances and two dependent sources. For instance, Figure 4.3 shows the equivalent networks derived directly from the Z - and H -parameter representations.

If the two-port network is reciprocal, it requires only three independent complex parameters, since $Z_{12} = Z_{21}$. While the dependent-source models remain valid, a reciprocal network can also be represented purely with passive components using equivalent T - or Π -circuits. The T -circuit configuration (Figure 4.4) consists of three passive impedances (Z_a , Z_b , and Z_c), giving the 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 -circuit, c.f., figure 4.4, is obtained by shortening port 1 and port 2, respectively.

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

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

$$\begin{aligned} Y_a &= Y_{11} + Y_{21} \\ Y_b &= Y_{22} + Y_{21} \\ Y_c &= -Y_{21} \end{aligned}$$

4.2.7 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, $V \rightarrow V$, current amplifier, $I \rightarrow I$, transimpedance amplifier, $I \rightarrow V$, and transadmittance amplifier $V \rightarrow I$. 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 output.

A two-port network can be connected to another two-port network using four distinct feedback topologies. These configurations form the foundation for designing different types of feedback amplifiers. Each topology corresponds to a specific amplifier type, defined by how the input is sensed and how the output is sampled

- Voltage Amplifier ($V \rightarrow V$): Voltage-controlled voltage source.
- Current Amplifier ($I \rightarrow I$): Current-controlled current source.
- Transimpedance Amplifier ($I \rightarrow V$): Current-controlled voltage source.
- Transadmittance Amplifier ($V \rightarrow I$): Voltage-controlled current source.

To achieve maximum efficiency and prevent signal loading, the ideal input and output impedances depend entirely on whether the signal is a voltage or a current.

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

According to Figure 4.5 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 $[H] = \sum_{n=1}^N [H_n]$.

The ideal $V \rightarrow V$ amplifier has infinite input impedance (to minimize input current draw) and zero output impedance (to ensure maximum voltage transfer to the load).

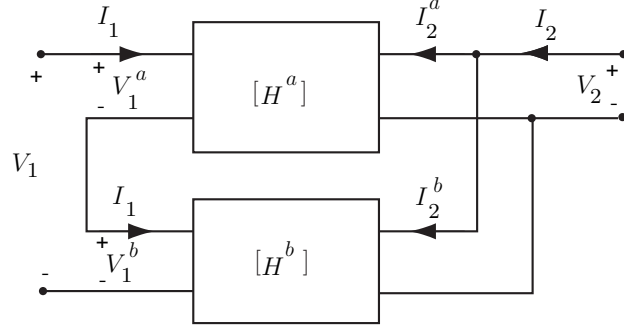


Figure 4.5: Series-parallel coupling: $[H] = [H^a] + [H^b]$

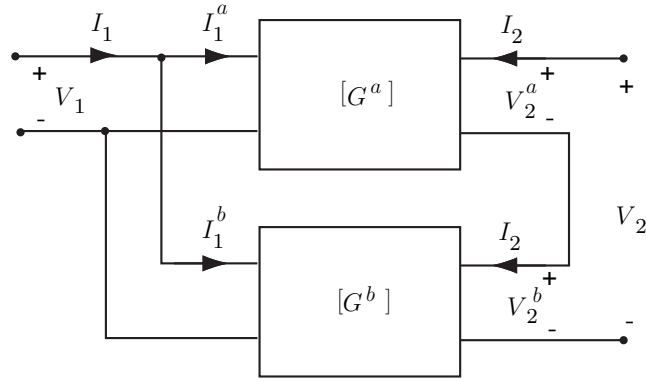


Figure 4.6: Parallel-series coupling: $[G] = [G^a] + [G^b]$

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

According to Figure 4.6 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 $[G] = \sum_{n=1}^N [G_n]$.

The ideal $I \rightarrow I$ amplifier has zero input impedance (to ensure all source current enters the amplifier) and infinite output impedance (to behave as an ideal current source).

Series coupling ($V \rightarrow I$ coupling)

According to Figure 4.7 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 $[Z] = \sum_{n=1}^N [Z_n]$.

The ideal $V \rightarrow I$ amplifier has infinite input impedance and infinite output impedance.

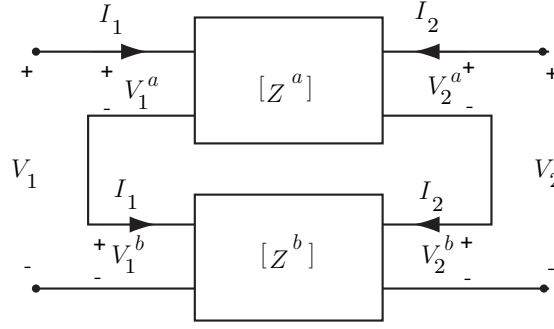


Figure 4.7: Series coupling: $[Z] = [Z^a] + [Z^b]$.

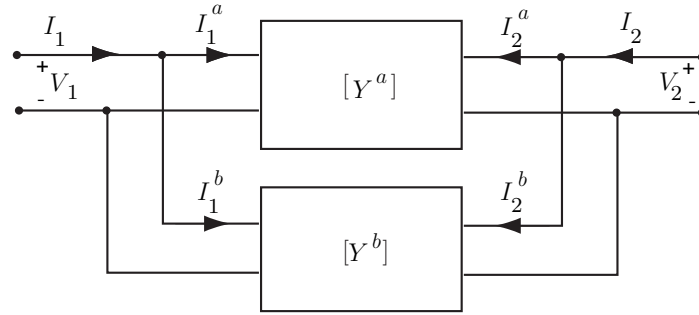


Figure 4.8: Parallel coupling: $[Y] = [Y^a] + [Y^b]$.

Parallel coupling ($I \rightarrow V$ coupling)

According to Figure 4.8 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 $[Y] = \sum_{n=1}^N [Y_n]$.

The ideal $I \rightarrow V$ amplifier has zero input impedance and zero output impedance.

4.2.8 Cascad coupled two-ports

We cascade two two-ports according to Figure 4.9 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

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

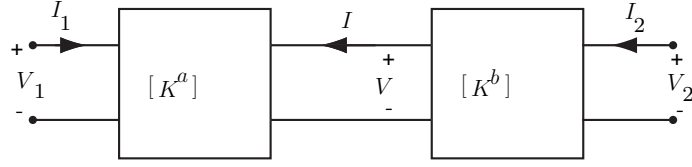


Figure 4.9: Cascade coupling: $[K] = [K^a][K^b]$.

4.3 Transmission lines in time domain

Transmission lines are wires that are not short compared to the signal's wavelength ($\lambda = c/f$, where c is the wave speed in the medium and f is frequency).

Signals on a transmission line can be characterized through two equivalent frameworks:

- Voltage and Current: Governed by scalar wave equations derived from circuit theory.
- Electric and Magnetic Fields: Governed by Maxwell's equations subject to boundary conditions.

While both approaches yield identical results, using voltage and current equations is generally far more convenient for practical analysis.

4.3.1 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 4.10.

The transmission line is specified 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 ($1/(\Omega\text{m})$)

R : resistance per unit length (Ω/m)

The line parameters are shown in Figure 4.11. For a homogeneous transmission line, the line parameters are independent of both the position z and time t .

- The voltage $v(z, t)$ between the conductors gives rise to line charges $\rho_\ell(z, t)$ and $-\rho_\ell(z, t)$ on the two conductors. The capacitance per unit length is defined as

$$C = \frac{\rho_\ell(z, t)}{v(z, t)} \quad (4.4)$$

- The current $i(z, t)$ flowing through the conductors generates a magnetic flux $\Phi(z, t)$ per unit length between the conductors. The inductance per unit length is defined as

$$L = \frac{\Phi(z, t)}{i(z, t)} \quad (4.5)$$

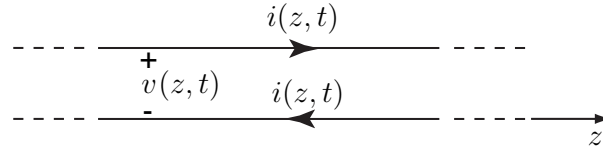


Figure 4.10: Voltage and current for the transmission line.

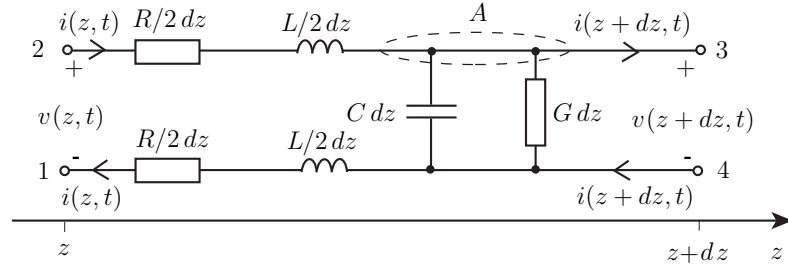


Figure 4.11: Equivalent circuit model for a transmission line segment of infinitesimal length dz .

- The voltage $v(z, t)$ between the conductors can cause a leakage current between them. If $i_{\text{leak}}(z, t)$ represents the leakage current per unit length, the conductance per unit length is defined as

$$G = \frac{i_{\text{leak}}(z, t)}{v(z, t)} \quad (4.6)$$

- The inherent resistance of the conductors causes a voltage drop along the line. If the voltage drop per unit length is $v_{\text{drop}}(z, t)$, the total resistance per unit length for the two-conductor system is defined as

$$R = \frac{2v_{\text{drop}}(z, t)}{i(z, t)} \quad (4.7)$$

We let the piece of the line shown in Figure 4.11 be short enough to be a discrete circuit. The Kirchhoff's voltage law for the loop $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$ and the Kirchhoff's current law for the node A give the transmission line equations

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

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

$$-\frac{\partial v(z, t)}{\partial z} = Ri(z, t) + L \frac{\partial i(z, t)}{\partial t} \quad (4.8)$$

$$-\frac{\partial i(z, t)}{\partial z} = Gv(z, t) + C \frac{\partial v(z, t)}{\partial t} \quad (4.9)$$

We eliminate the current by operating with $\frac{\partial}{\partial z}$ on (4.8) and with $R + L \frac{\partial}{\partial t}$ on (4.9). We combine the two equations to get the scalar wave equation for the voltage

$$\frac{\partial^2 v(z, t)}{\partial z^2} - LC \frac{\partial^2 v(z, t)}{\partial t^2} - (LG + RC) \frac{\partial v(z, t)}{\partial t} - RGv(z, t) = 0 \quad (4.10)$$

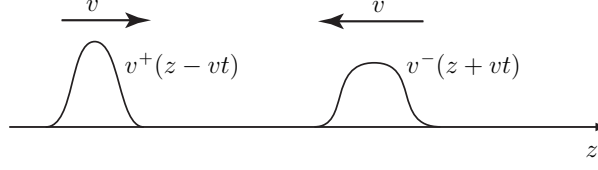


Figure 4.12: Wave propagation of pulses

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

4.3.2 Wave propagation in the time domain

For simplicity we start with a lossless transmission line. Thus R and G are zero and (4.8)–(4.10) read

$$-\frac{\partial v(z, t)}{\partial z} = L \frac{\partial i(z, t)}{\partial t} \quad (4.11)$$

$$-\frac{\partial i(z, t)}{\partial z} = C \frac{\partial v(z, t)}{\partial t} \quad (4.12)$$

$$\frac{\partial^2 v(z, t)}{\partial z^2} - LC \frac{\partial^2 v(z, t)}{\partial t^2} = 0 \quad (4.13)$$

The general solution to (4.13) is

$$v(z, t) = f(z - v_p t) + g(z + v_p t)$$

where

$$v_p = \frac{1}{\sqrt{LC}}$$

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

$$\begin{aligned} \xi &= z - v_p t \\ \eta &= z + v_p t \end{aligned}$$

that transforms the wave equation (4.13) to the equation

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

The function $f(z - v_p t)$ has constant argument when $z = v_p t + \text{konstant}$. That means that $f(z - v_p t)$ is a wave that propagates with the speed $v_p = 1/\sqrt{LC}$ in the positive z -direction, see Figure 4.12. In the same way we can argue that $g(z + v_p t)$ 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

$$v(z, t) = v^+(z - v_p t) + v^-(z + v_p t) \quad (4.14)$$

The shapes 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 (4.14) into (4.11) or (4.12)

$$i(z, t) = i^+(z - v_p t) + i^-(z + v_p t)$$

where

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

Here, 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. One can prove that the phase speed $v_p = 1/\sqrt{LC}$ is equal to the wave speed of electromagnetic waves in the material surrounding the conductors, i.e., $v_p = 1/\sqrt{\epsilon_0 \epsilon \mu_0} = c/\sqrt{\epsilon}$ 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 $c \approx 3 \cdot 10^8$ m/s is the speed of light in vacuum. We summarize the results for wave propagation on a lossless transmission line:

$$\begin{aligned} v(z, t) &= v^+(z - v_p t) + v^-(z + v_p t) \\ i(z, t) &= i^+(z - v_p t) + i^-(z + v_p t) = \frac{1}{Z_0} (v^+(z - v_p t) - v^-(z + v_p t)) \\ v_p &= 1/\sqrt{LC}, \quad Z_0 = \sqrt{L/C} \end{aligned} \quad (4.15)$$

4.3.3 Reflection from a load 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 $v_i(z, t) = v^+(z - v_p t)$ has been generated at $z = 0$ and it generates a reflected wave $v_r(z, t) = v^-(z + v_p t)$ once it reaches $z = \ell$. The total voltage and current on the line are

$$\begin{aligned} v(z, t) &= v_i(z, t) + v_r(z, t) \\ i(z, t) &= i_i(z, t) + i_r(z, t) = \frac{1}{Z_0} (v_i(z, t) - v_r(z, t)) \end{aligned} \quad (4.16)$$

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

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

We insert (4.16) 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

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

Thus, at $z = \ell$ the reflected wave is a scaled version of the incident wave. The scaling constant is the dimensionless reflection coefficient

$$\Gamma = (R_L - Z_0)/(R_L + Z_0) \quad (4.17)$$

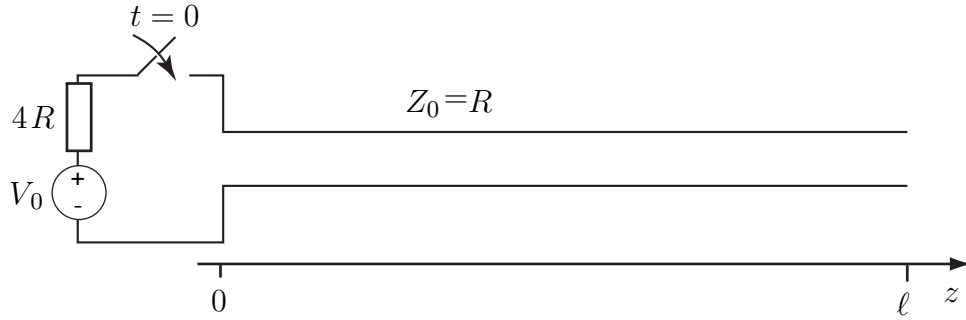
Example 4.1

- 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.
- If the line is open at $z = \ell$ then $R_L = \infty$. We get $\Gamma = 1$ and $v_r(\ell, t) = v_i(\ell, t)$. All power sent to the load is reflected.
- If the line is shortened at $z = \ell$ then $R_L = 0$ and we get $\Gamma = -1$ and $v_r(\ell, t) = -v_i(\ell, t)$. All power sent to the load is reflected.

■

Example 4.2

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.

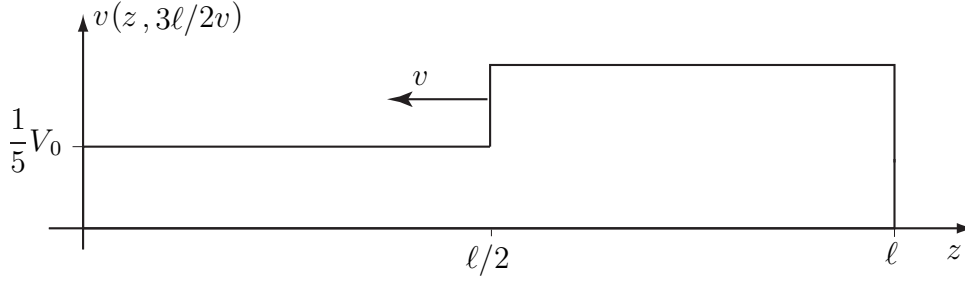
$t = 0^+$: A step pulse $v_i(z, t)$ starts to propagate along the line. The voltage at $z = 0$ is given by $v(0, t) = v_i(0, t) = Z_0 i(0, t) = R i(0, t)$. The line is equivalent to a resistance R and the voltage division formula gives

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

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

$t = \frac{\ell}{v_p}$: 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 $z = 0$.

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



$t = \frac{2\ell}{v_p}$: 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

$$v(0, 2\ell/v_p^+) = v_1 + v_2 + v_3 = V_0/5 + V_0/5 + 3V_0/25 = 13V_0/25$$

$t > \frac{2\ell}{v_p}$: 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.

■

4.4 Transmission lines in frequency domain

Shifting our focus to time-harmonic signals, we assume the time-harmonic voltage source was energized long ago, allowing all transient effects to fully decay. The resulting incident and reflected sinusoidal waves interfere along the line to establish a standing wave pattern.

By applying phasor representation, we transform the time-domain voltages and currents into the frequency domain. Here, the complex voltage $V(z)$ and current $I(z)$ satisfy the frequency-domain transmission line equations, c.f., (4.8) and (4.9)

$$-\frac{dV(z)}{dz} = (R + j\omega L)I(z) \quad (4.18)$$

$$-\frac{dI(z)}{dz} = (G + j\omega C)V(z) \quad (4.19)$$

We begin by consider lossless lines ($R = 0, G = 0$). By differentiating (4.18) w.r.t. z and using (4.19) the voltage $V(z)$ is seen to satisfy

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

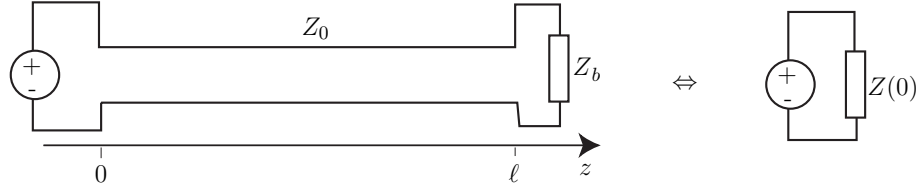


Figure 4.13: A transmission line with length ℓ and load Z_L can be replaced by its input impedance $Z(0)$ give by (4.23).

where $\beta = \omega\sqrt{LC} = \omega/v_p$ is the phase coefficient. The general solution consists of two linearly independent waves traveling in opposite directions:

$$V(z) = V^+(z) + V^-(z) = V_p e^{-j\beta z} + V_n e^{j\beta z} \quad (4.20)$$

where $V_p = |V_p|e^{j\phi_p}$ and $V_n = |V_n|e^{j\phi_n}$ are complex amplitudes. The time dependent voltage is

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

Since $\omega t - \beta z = \omega(t - z/v_p)$ we see that $v^+(z, t)$ is a wave propagating in the positive z -direction with speed v_p . 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

$$\lambda = 2\pi/\beta = 2\pi v_p/\omega = v_p/f$$

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

$$I(z) = I^+(z) + I^-(z) = I_p e^{-j\beta z} + I_n e^{j\beta z} \quad (4.21)$$

We insert (4.20) and (4.21) in (4.18) 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., (4.15).

4.4.1 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 with load is equivalent to an input impedance $Z(0)$ at $z = 0$ where

$$Z(0) = \frac{V(0)}{I(0)} = Z_0 \frac{V_p + V_n}{V_p - V_n} = Z_0 \frac{1 + V_n/V_p}{1 - V_n/V_p} \quad (4.22)$$

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

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

Thus

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

which together with (4.22) gives

$$Z(0) = Z_0 \frac{Z_L \cos(\beta\ell) + jZ_0 \sin(\beta\ell)}{Z_0 \cos(\beta\ell) + jZ_L \sin(\beta\ell)} \quad (4.23)$$

This is the input impedance of the line. Lumped circuit theory would incorrectly predict $Z(0) = Z_L$ by neglecting wave propagation effects, leading to significant errors whenever the line length ℓ is comparable to or larger than the wavelength.

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., (4.17)

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

It is convenient to introduce the reflection coefficient at the load as

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} \quad (4.24)$$

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

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

Example 4.3

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 load impedance is matched to the line. ■

Example 4.4

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

$$Z_L = \infty \quad \Rightarrow \quad Z(0) = -jZ_0 \cot \beta\ell \quad (4.25)$$

$$Z_L = 0 \quad \Rightarrow \quad Z(0) = jZ_0 \tan \beta\ell \quad (4.26)$$

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 4.5

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

$$Z(0) = \frac{Z_0^2}{Z_L} \quad (4.27)$$

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 4.6

Matching a load by $\lambda/4$ transformer: Assume that we like to match load resistance R_L to a lossless line with characteristic impedance Z_1 . This can be done with a $\lambda/4$ transformer with characteristic impedance $Z_0 = \sqrt{Z_1 R_L}$. ■

4.4.2 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 root-mean-square (rms) voltage.

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

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

The maximum voltage is where the waves $V_p e^{-j\beta z}$ and $V_n e^{j\beta z}$ are in phase and the minimum voltage where they are out of phase. 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

$$e^{2j\beta\Delta z} = e^{j2\pi}, \quad \beta\Delta z = \pi$$

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

4.4.3 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 transported 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 $G \ll 1/(\omega C)$.

We derive the equation for the voltage from the general transmission line equations (4.18) and (4.19)

$$\frac{d^2 V(z)}{dz^2} - \gamma^2 V(z) = 0 \quad (4.28)$$

where

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} = \text{propagation constant}. \quad (4.29)$$

The general solution to (4.28) is

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

where $V_p = |V_p| e^{j\phi_p}$ and $V_n = |V_n| e^{j\phi_n}$ are complex constants. 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

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \quad (4.30)$$

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

$$\gamma = \alpha + j\beta$$

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

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

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 $z = \omega t / \beta + \text{constant}$ and this leads us to the definition of the phase speed

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

A line with length ℓ , characteristic impedance Z_0 and load impedance Z_L , has the input impedance

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

The derivation is almost identical to the one for the input impedance in (4.23) for the lossless line.

4.4.4 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

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

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 distortion. 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.

These early cables suffered from severely limited transmission rates, primarily due to high conductor resistance. The cable design consisted of a single copper core surrounded by an insulating layer, with seawater serving as the return conductor. For the 1865 and 1866 cables, the copper wire radius was approximately 1.6 mm, and the overall cable radius was 15 mm. Based on these known values we can estimate the line parameters to be $R = 2.2 \cdot 10^{-3} \Omega/\text{m}$, $L = 0.4 \mu\text{H}/\text{m}$, $C = 80 \text{ pF}/\text{m}$ and $G = 10^{-13} (\Omega\text{m})^{-1}$. 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

$$\frac{\partial^2 v(z, t)}{\partial z^2} = RC \frac{\partial v(z, t)}{\partial t} \quad (4.31)$$

This is the diffusion equation.

Assuming a step-voltage excitation $v(0, t) = V_0 H(t)$ applied at $z = 0$, where $H(t)$ is the Heaviside step function, and initial conditions $v(z, 0) = 0$ for $z > 0$, the exact solution to (4.31) is

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

The step response at the receiving station in America (for a unit step input in England) is shown in Figure 4.14. Detection posed a significant challenge: for a short-circuited termination and an input voltage of $V_0 = 1 \text{ V}$, the received current remains below 0.1 mA, necessitating relatively high driving voltages to obtain a readable signal. Despite these physical constraints, the 1865 and 1866 cables achieved a transmission rate of eight words per minute using Morse code (roughly one character per second)—a major improvement over the 1858 cable, which managed only a few words per hour.

4.5 Wave propagation in terms of E and H

While we have previously described wave propagation using voltage and current, we can also analyze it through the lens of electromagnetic fields. This field-based approach is generally more complex because it requires working with the vector fields $\mathbf{E}$ and $\mathbf{H}$ rather than the scalar quantities v and i . Additionally, linking this field analysis back to standard circuit theory concepts can be challenging.

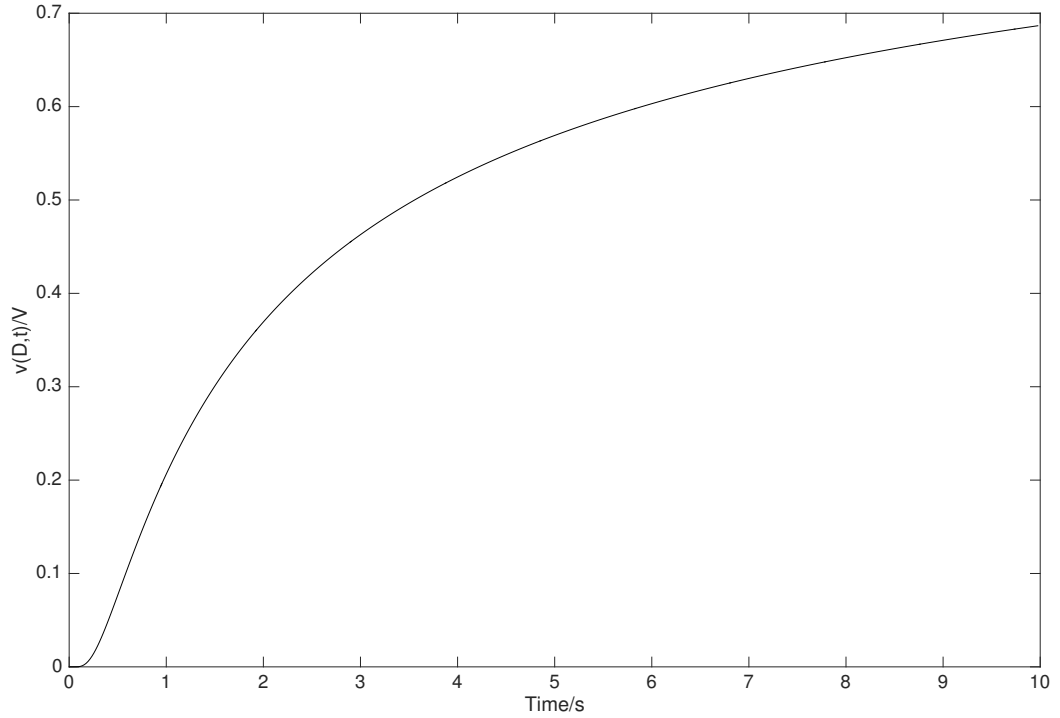


Figure 4.14: The received voltage $v(D, t)$ for a unit step voltage $v(0, t) = H(t)$ V, where $D = 4300$ km is the length of the Atlantic cable. The line parameters are $R = 2.2 \cdot 10^{-3}$ Ω/m , $L = 0.4$ $\mu\text{H}/\text{m}$, $C = 80$ pF/m and $G = 10^{-13}$ $(\Omega\text{m})^{-1}$. The line is open ($Z_L = \infty$) at $z = D$.

Despite these drawbacks, analyzing wave propagation via electromagnetic fields is highly valuable—and, as we will see, there are scenarios where electromagnetic theory is absolutely necessary.

We begin by considering transmission lines with a homogeneous medium between the conductors, assuming negligible resistance per unit length. Under these conditions, the propagating wave is a transverse electromagnetic (TEM) wave, meaning that both the electric and magnetic fields lack a longitudinal (z -) component.

In the frequency domain, the current, voltage, and electromagnetic fields must share the same propagation factor, $e^{\pm\gamma z}$. Within the region between the two conductors, these TEM fields are expressed as:

$$\begin{aligned} \mathbf{E}(\mathbf{r}, \omega) &= \mathbf{E}_T(x, y)e^{\pm\gamma z} \\ \mathbf{H}(\mathbf{r}, \omega) &= \mathbf{H}_T(x, y)e^{\pm\gamma z} \end{aligned} \quad (4.32)$$

From the z -component of Faraday's law of induction, it follows that the transverse electric field is irrotational:

$$\nabla \times \mathbf{E}_T(x, y) = \mathbf{0}$$

Consequently, there exists a scalar potential $\Psi(x, y)$ such that:

$$\mathbf{E}_T(x, y) = -\nabla\Psi(x, y)$$

Because the region between the conductors is free of charge, Gauss's law simplifies to Laplace's equation for the potential:

$$\nabla \cdot \mathbf{E}_T(x, y) = -\nabla^2\Psi(x, y) = 0$$

Furthermore, Ampère's law demonstrates that the relationship between the transverse electric and magnetic fields is identical to that of a plane wave.

The potential $\Psi(x, y)$ is constant on each conductor. We let V_1 and V_2 denote the potentials on conductor 1 and conductor 2, respectively. Once we determine the potential distribution $\Psi(x, y)$ in the region surrounding the conductors, we can readily calculate the primary line parameters R , L , G , and C .

The conductors are assumed to have high conductivity, thus the currents and charges are confined to an extremely thin skin layer at the surface of the conductors. This allows us to define the surface current density and surface charge density as:

$$\hat{\mathbf{z}}J_s(x, y) = \hat{\mathbf{n}} \times \mathbf{H}_T(x, y) \quad (4.33)$$

$$\rho_s(x, y) = \epsilon_0\epsilon\hat{\mathbf{n}} \cdot \mathbf{E}_T(x, y) \quad (4.34)$$

where ϵ is the relative permittivity of the material between the conductors and $\hat{\mathbf{n}}$ is the normal unit vector as in Figure 4.15.

We consider the two types of transmission lines illustrated in Figure 4.15: the open transmission line on the left and the closed (shielded) transmission line on the right.

To find the electrostatic potential $\Psi(x, y)$ for the closed line, we solve Laplace's equation subject to Dirichlet boundary conditions on the conductors:

$$\begin{aligned} \nabla^2\Psi(x, y) &= 0, \text{ outside the conductors} \\ \Psi(x, y) &= \begin{cases} V_1 & \text{on conductor 1} \\ V_2 & \text{on conductor 2} \end{cases} \end{aligned} \quad (4.35)$$

Here, we can assign arbitrary values to the potentials V_1 and V_2 , provided that $V_1 \neq V_2$. This boundary value problem yields a unique potential distribution $\Psi(x, y)$, from which the line parameters R , L , G , and C can be uniquely determined.

The boundary value problem for the open transmission line is slightly more complex. Because the transmission line must maintain charge neutrality, the total charge per unit length on the two conductors must sum to zero:

$$\oint_{\Gamma_1} \rho_s(x, y) d\ell = - \oint_{\Gamma_2} \rho_s(x, y) d\ell \quad (4.36)$$

where Γ_1 and Γ_2 are the boundaries of conductor 1 and 2, respectively, as shown in Figure 4.15. This charge-neutrality condition implies that the potential must vanish at infinity: $\Psi(x, y) \rightarrow 0$ as $|(x, y)| \rightarrow \infty$.

Consequently, we can formulate the boundary value problem for the open transmission line as:

$$\begin{aligned} \nabla^2\Psi(x, y) &= 0, \text{ outside the conductors} \\ \Psi(x, y) &= \begin{cases} V_1 & \text{on conductor 1} \\ V_2 & \text{on conductor 2} \end{cases} \\ \Psi(x, y) &\rightarrow 0 \text{ as } |(x, y)| \rightarrow \infty \end{aligned} \quad (4.37)$$

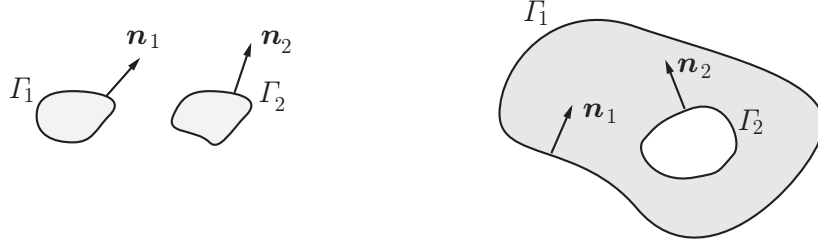


Figure 4.15: The cross-section of two types of transmission lines. The structure to the left is open and the structure to the right is closed. The curves Γ_1 and Γ_2 define the boundaries of the conductors. The normal vectors $\hat{n}_1$ and $\hat{n}_2$ are directed out from the conductors.

Here, we can set V_1 to an arbitrary reference value (e.g., 1 V). The potential V_2 on conductor 2 is an unknown constant that is uniquely determined as part of the solution when enforcing the charge-neutrality condition. Under these conditions, the boundary value problem yields a unique solution for the potential distribution $\Psi(x, y)$.

Once (4.35) and (4.37) are solved the electric and magnetic fields are obtained from

$$\mathbf{E}_T(x, y) = -\nabla\Psi(x, y) \quad (4.38)$$

$$\mathbf{H}_T(x, y) = \pm j \frac{\gamma}{\omega\mu_0} \hat{\mathbf{z}} \times \mathbf{E}_T(x, y) \quad (4.39)$$

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

Under the standard assumption that the line losses (R and G) are sufficiently small, the propagation constant can be approximated as $\gamma \approx j\beta$. Substituting this approximation yields the simplified relations:

$$\begin{aligned} \mathbf{E}_T(x, y) &= -\nabla\Psi(x, y) \\ \mathbf{H}_T(x, y) &\approx \mp \frac{\beta}{\omega\mu_0} \hat{\mathbf{z}} \times \mathbf{E}_T(x, y) \end{aligned} \quad (4.40)$$

4.6 Transmission line parameters

In this subsection it is shown how R, L, G, C are determined from the solutions of (4.35) and (4.37).

We assume that the two transmission line conductors have a very high conductivity, σ_c , and are embedded in a homogeneous dielectric medium characterized by a relative permittivity, ϵ , and a very low conductivity, σ_d . Specifically, we assume that σ_c is sufficiently high and σ_d is sufficiently low such that they satisfy the conditions:

$$\sigma_c \gg \omega\epsilon_0 \quad \text{and} \quad \sigma_d \ll \omega\epsilon_0\epsilon$$

Furthermore, the skin depth of the conductors, defined as:

$$\delta = \sqrt{\frac{2}{\omega\sigma_c\mu_0}}$$

is assumed to be much smaller than the cross-sectional dimensions (such as the diameters) of the conductors.

We now proceed under the assumption that the static boundary value problem—either (4.35) or (4.37)—has been solved, yielding the potential $\Psi(x, y)$ everywhere outside the conductors. For the open transmission line, this solution also determines the constant potential V_2 on conductor 2.

For simplicity we assume that $V_1 - V_2 > 0$. We first determine the charge per unit length as

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

and the capacitance per unit length as

$$C = \frac{\rho_\ell}{V_1 - V_2}$$

The inductance is obtained from the phase speed $v_p = (\epsilon_0 \epsilon \mu_0)^{-1/2} = (LC)^{-1/2}$

$$L = \frac{1}{C v_p^2} = \frac{\epsilon_0 \epsilon \mu_0}{C} \quad (4.41)$$

The conductance G follows directly from the fundamental relationship between capacitance and dielectric loss:

$$G = C \frac{\sigma_d}{\epsilon_0 \epsilon} \quad (4.42)$$

To determine the resistance R , we first relate the surface current density $J_s(x, y)$ to the surface charge density $\rho_s(x, y)$ using the continuity equation. In the frequency domain, the charge conservation law along the propagation direction (z) is expressed as:

$$\frac{dJ_s(x, y)e^{\pm\gamma z}}{dz} = -j\omega\rho_s(x, y)e^{\pm\gamma z}$$

We use the fact that $\gamma \approx j\beta$ to obtain

$$J_s(x, y) = \mp v_p \rho_s(x, y) \quad (4.43)$$

where the phase speed remains $v_p = 1/\sqrt{LC}$.

The time average of 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

$$P_d = \frac{1}{2} R_s \oint_{\Gamma_1 + \Gamma_2} |J_s(x, y)|^2 d\ell = \frac{1}{2} v_p^2 R_s \oint_{\Gamma_1 + \Gamma_2} |\rho_s(x, y)|^2 d\ell$$

where R_s is the surface resistance

$$R_s = \frac{1}{\sigma_c \delta} = \sqrt{\frac{\omega \mu_0}{2 \sigma_c}} \quad (4.44)$$

The expression for the surface resistance is derived in chapter 6. The resistance per unit length is thus given by

$$R = \frac{1}{|I|^2} R_s \oint_{\Gamma_1 + \Gamma_2} |J_s(x, y)|^2 d\ell = \frac{1}{|I|^2} v_p^2 R_s \oint_{\Gamma_1 + \Gamma_2} |\rho_s(x, y)|^2 d\ell \quad (4.45)$$

The current is either obtained from $|I| = Z_0^{-1}(V_1 - V_2) = v_p|\rho_\ell|$, or from integration over one of the conductors

$$|I| = \oint_{\Gamma_1} |J_s(x, y)| d\ell = \oint_{\Gamma_2} |J_s(x, y)| d\ell$$

Example 4.7

We exemplify by considering wave propagation in the positive z -direction in 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 and a small conductivity σ_d . 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

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

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 of the current along the inner conductor. The Ampère's law gives

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

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

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

the capacitance per unit length is

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

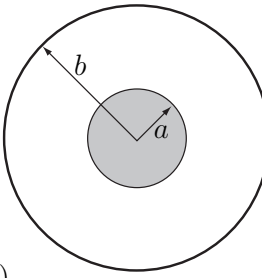
The inductance L , conductance G , and resistance R follow from (4.41), (4.42) and (4.45). The expressions are listed in subsection 4.6.1.

■

4.6.1 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



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

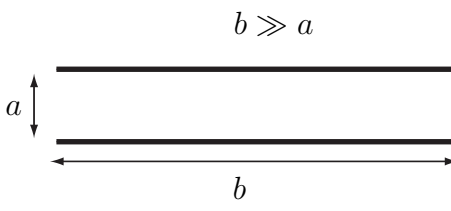
$$L = \frac{\mu_0}{2\pi} \ln \frac{b}{a}$$

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

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

Parallel plate

Notice that $b \gg a$.



$$C \approx \frac{\epsilon_0\epsilon b}{a}$$

$$L \approx \frac{\mu_0 a}{b}$$

$$Z_0 = \sqrt{\frac{L}{C}} \approx \sqrt{\frac{\mu_0}{\epsilon_0\epsilon}} \frac{a}{b} \quad (\text{lossless line})$$

$$R \approx \frac{2R_s}{b}$$

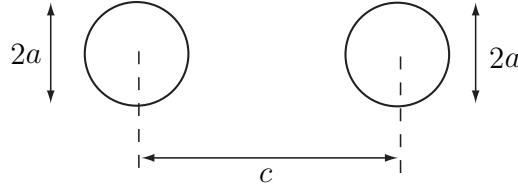
Two-wire line

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

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

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

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



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

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

Determining line parameters numerically is straightforward when the surrounding medium is homogeneous. Throughout this book, we use the commercial software COMSOL Multiphysics, a finite element method (FEM) platform with specialized application modules. Specifically, we rely on the AC/DC and RF modules, utilizing the former in this section. To demonstrate the process, we provide step-by-step instructions in COMSOL for a two-wire line in air, where each wire has a radius $a = 1$ cm and an axis-to-axis distance $c = 3$ cm.

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 cm².
4. Next we draw the circular wires and put them in the center of the square, c.f., Figure 4.16.
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. 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., $\text{es.normE}^2 * 8.854e - 12$ in **expression**).
11. From C we get L and G from (4.41) and (4.42).
12. Finally we get the resistance by letting the FEM program numerically evaluate the integral in (4.45).

Figure 4.16 illustrates the geometry and resulting field distribution. Using the default mesh, COMSOL yields a capacitance of $C = 2.885 \times 10^{-11}$ F/m, which agrees closely with the exact analytical value (see subsection 4.6.1) of $C = 2.8901 \times 10^{-11}$ F/m; mesh refinement provides only a minor improvement. This demonstrates that the finite element method is well suited for these calculations. However, reducing the surrounding square domain to 20×20 cm² shifts the capacitance to $C = 2.95 \times 10^{-11}$ F/m, even under mesh refinement. This example highlights the necessity of validating numerical results in COMSOL by systematically varying both mesh density and the extent of the computational domain.

Example 4.8

In this example we let COMSOL calculate the capacitance of a plate transmission line with 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 8. These are very accurate and we see that the error in the approximate analytic result is not negligible unless b/a is very large.

■

4.6.3 Transverse inhomogeneous region

Thus far, the line parameters R , L , G , and C have been determined for transmission lines embedded in media where the conductivity and permittivity are independent of the x and y coordinates. However, radio-frequency (RF) circuits frequently employ microstrip transmission lines consisting of a strip conductor situated above a ground plane and separated by a dielectric substrate. The waves propagating along these structures are classified as quasi-TEM waves because their longitudinal (z -directed) electric and magnetic field components are small but non-zero.

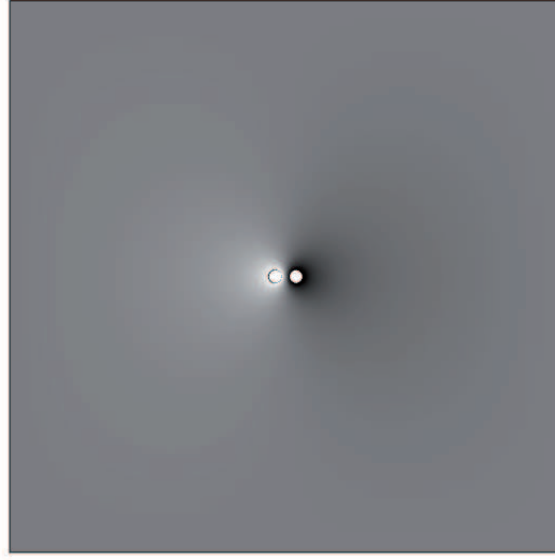


Figure 4.16: The geometry and voltage for the two-wire calculation in COMSOL. The right wire has the potential $V = -0.5$ V and the left the potential $V = 0.5$ V. The outer square has the size 80×80 cm² and potential 0 V.

Section 6.12 presents an accurate method for analyzing wave propagation on such structures. This approach solves the Maxwell equations using the finite element method (FEM).

4.7 The scattering matrix $[S]$

The scattering matrix (S -matrix) is fundamental to high-frequency measurements. Instead of dealing directly with total voltages, it relates incoming and outgoing voltage waves at the ports.

Assuming identical characteristic impedances for the transmission lines at both ports, the S -matrix is defined as:

$$\begin{pmatrix} V_1^- \\ V_2^- \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} V_1^+ \\ V_2^+ \end{pmatrix} = [S] \begin{pmatrix} V_1^+ \\ V_2^+ \end{pmatrix} \quad (4.46)$$

Here, the superscript $+$ denotes a wave traveling toward a port, while $-$ denotes a wave traveling away from it. In this framework:

- S_{11} and S_{22} represent the reflection coefficients at Ports 1 and 2, respectively.
- S_{21} is the transmission coefficient from Port 1 to Port 2.
- S_{12} is the transmission coefficient from Port 2 to Port 1.

When an incident wave with power P_i enters Port 1, the resulting transmitted power

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 4.1: The capacitance of a parallel plate line as a function of b/a . The approximate analytic values are from subsection (4.6.1). The FEM values are from a 2D-solver and are very accurate.

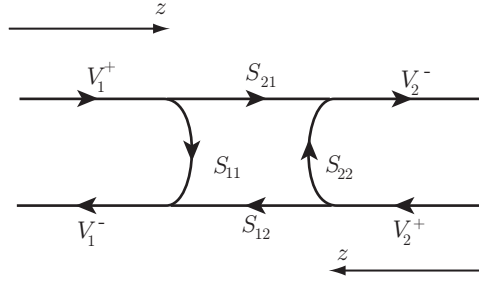


Figure 4.17: A flow diagram of the scattering matrix. The incident waves V_1^+ and V_2^+ are assumed to travel in the positive z -direction and hence two coordinate systems are used.

P_t at Port 2 and reflected power P_r at Port 1 are given by:

$$\begin{aligned} P_t &= |S_{21}|^2 P_i \\ P_r &= |S_{11}|^2 P_i \end{aligned} \quad (4.47)$$

4.7.1 S-matrix with different characteristic impedances at the ports

Now consider a two-port system where the transmission lines connected to Port 1 and Port 2 have different characteristic impedances, Z_1 and Z_2 , respectively.

To preserve the power expressions defined in (4.47), the transmission parameters S_{21} and S_{12} are scaled to account for the impedance mismatch:

$$\begin{pmatrix} V_1^- \\ V_2^- \end{pmatrix} = \begin{pmatrix} S_{11} & \sqrt{\frac{Z_1}{Z_2}} S_{12} \\ \sqrt{\frac{Z_2}{Z_1}} S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} V_1^+ \\ V_2^+ \end{pmatrix} \quad (4.48)$$

With this adjustment, the definitions for the reflection coefficients S_{11} and S_{22} in (4.46) remain unchanged, and the power relationships in (4.47) still hold.

4.7.2 Relation between $[S]$ and $[Z]$

The scattering matrix $[S]$ and impedance matrix $[Z]$, introduced in subsection 4.2.1, can be directly related. Assuming both ports are connected to transmission lines with identical characteristic impedance Z_0 , we define the total voltage vector $\mathbf{V} = (V_1 \ V_2)^T$ and current vector $\mathbf{I} = (I_1 \ I_2)^T$ in terms of their incident (+) and reflected (−) wave components:

$$\begin{aligned}\mathbf{V} &= \mathbf{V}^+ + \mathbf{V}^- = ([U] + [S]) \mathbf{V}^+ \\ \mathbf{I} &= \mathbf{I}^+ + \mathbf{I}^- = \frac{1}{Z_0}(\mathbf{V}^+ - \mathbf{V}^-) = \frac{1}{Z_0}([U] - [S]) \mathbf{V}^+\end{aligned}$$

where $[U]$ denotes the 2×2 identity matrix.

Using the impedance matrix relationship $\mathbf{V} = [Z]\mathbf{I}$, substituting the expressions above yields:

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

Solving for $[Z]$ and $[S]$, respectively, gives the conversion formulas:

$$[Z] = Z_0([U] + [S])([U] - [S])^{-1} \quad (4.49)$$

$$[S] = ([Z] + Z_0[U])^{-1}([Z] - Z_0[U]) \quad (4.50)$$

For a reciprocal two-port network, the impedance matrix $[Z]$ is symmetric ($[Z] = [Z]^T$), which directly implies that the corresponding scattering matrix $[S]$ is also symmetric ($[S] = [S]^T$).

For a lossless two-port network, the net real power entering the network must be zero. Using the total voltage vector $\mathbf{V}$ and current vector $\mathbf{I}$, this power balance condition is expressed as

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

Because the cross-terms $(\mathbf{V}^{-T} \mathbf{V}^{+*} - \mathbf{V}^{+T} \mathbf{V}^{-*})$ evaluate to a purely imaginary value, taking the real part eliminates them. The remaining terms are purely real, leaving

$$\mathbf{V}^{+T} \mathbf{V}^{+*} - \mathbf{V}^{-T} \mathbf{V}^{-*} = 0$$

Substituting $\mathbf{V}^- = [S]\mathbf{V}^+$ yields:

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

Since this relation must hold for any arbitrary incident wave vector $\mathbf{V}^+$, the matrix inside the parentheses must vanish

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

Taking the conjugate transpose (Hermitian conjugate, superscript H) of both sides shows that the scattering matrix of a lossless network is unitary

$$[S]^H [S] = [U] \implies [S]^{-1} = [S]^H$$

If we use the tables in subsection 4.2.5 we can relate the S -matrix to the hybrid and cascade matrices. The S -matrix can be obtained from measurements.

4.7.3 Matching of load impedances

Impedance matching a load to a transmission line optimizes signal transmission, most commonly by maximizing power transfer. For a single-frequency signal, exact matching is easily achieved through several standard techniques. However, when the signal spans a frequency band and the load impedance is frequency-dependent, exact matching across the entire bandwidth is generally impossible. In such cases, the goal shifts to finding an optimal matching network that minimizes reflections over the target frequency band.

Conjugate matching

In basic circuit theory, the maximum power transfer theorem dictates that to maximize the power delivered from a source with internal impedance Z_i , the load impedance Z_L must be the complex conjugate of the source impedance

$$Z_L = Z_i^*$$

In practice, however, the load impedance Z_L is usually fixed by the application. To achieve conjugate matching and maximize power transfer to such a fixed load, a lossless matching network must be inserted between the source and the load.

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 seen in subsection 4.4.1 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., (4.27)

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

If the load is complex ($Z_L = R_L + jX_L$), its reactive component must be canceled before applying the quarter-wave transformer. This is accomplished by adding either a series reactance of $-X_L$, or a parallel susceptance of $B = -\frac{X_L}{R_L^2 + X_L^2}$.

Below, we present two additional methods for single-frequency impedance matching. While these techniques achieve a perfect match at a single operating frequency, broadband impedance matching across a range of frequencies presents a greater challenge and is beyond the scope of this section.

Two-port matching

A load impedance Z_L can be matched to the characteristic impedance Z_0 of a lossless transmission line using an L-section matching network consisting of a capacitor and an inductor. The two possible lumped-element two-ports are illustrated in Figure 4.18.

For the two-port on the left, the matching condition is

$$Z_0 = jX + \frac{1}{G_L + j(B + B_L)} \quad (4.51)$$

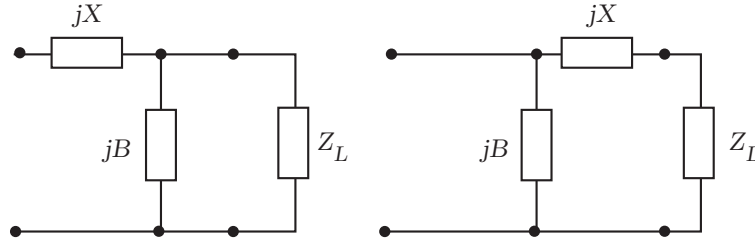


Figure 4.18: Two-ports for matching of Z_L .

and for the two-port to the right, it is

$$\frac{1}{Z_0} = jB + \frac{1}{R_L + j(X + X_L)} \quad (4.52)$$

where $Z_L = R_L + jX_L$ and $Y_L = G_L + jB_L$. Because the transmission line is lossless, Z_0 is purely real. Equating the real and imaginary parts in (4.51) allows us to solve for the reactive components X and B :

$$X = \pm Z_0 \sqrt{\frac{1 - Z_0 G_L}{Z_0 G_L}}$$

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

To yield real-valued solutions, we must satisfy the condition $Z_0 G_L \leq 1$, which is equivalent to $R_L + X_L^2/R_L \geq Z_0$. This requirement is strictly fulfilled whenever $R_L \geq Z_0$.

Similarly, separating the real and imaginary parts of Equation (4.52) yields:

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

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

This requires that $R_L \leq Z_0$. Consequently, the left two-port in Figure 4.18 must be employed when $R_L \geq Z_0$, while the right two-port is selected when $R_L \leq Z_0$.

Example 4.9

When $R_L = Z_0$ we have two choices. We can choose the left circuit in Figure 4.18 with $X = 0$ and $B = -B_L$, or we can choose the right circuit with $B = 0$ and $X = -X_L$. ■

4.7.4 Matching with stub

A short transmission line is often referred to as 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 an impedance match.

Since the load impedance of a shortened stub is zero its input impedance is, according to (4.23), purely reactive and given by

$$Z_{st} = jZ_0 \tan \beta \ell$$

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

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

Connecting the stub in parallel with the line at position d yields an impedance match when

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

Normalizing by Z_0 leads to the complex matching condition

$$1 = -\frac{j}{\tan \beta \ell} + \frac{1 + j z_L \tan \beta d}{z_L + j \tan \beta d} \quad (4.53)$$

Equating the real parts of (4.53) determines the stub placement distance 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 (4.53)

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

Out of the infinitely many periodic solutions, we select the smallest positive length ℓ . Note that these expressions simplify considerably when the load impedance Z_L is purely resistive.

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.

4.8 Smith chart

Developed in 1939 by Phillip H. Smith at Bell Telephone Laboratories, the Smith chart remains a fundamental tool in microwave engineering. Modern RF measurement systems, such as vector network analyzers, routinely display input impedance spectra directly on a Smith chart, providing a comprehensive representation of device behavior. Beyond visualization, the chart provides an intuitive graphical framework for designing impedance-matching networks on transmission lines with known characteristic impedance. This subsection introduces the basic principles of the Smith chart and illustrates its practical application through representative examples.

We assume a lossless transmission line with length ℓ , characteristic impedance Z_0 , and a load impedance Z_L at $z = \ell$. The reflection coefficient at the load is given by (4.24)

$$\Gamma = \frac{V^-(\ell)}{V^+(\ell)} = \frac{V_n}{V_p} e^{2j\beta \ell} = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{z_L - 1}{z_L + 1} \quad (4.54)$$

where we introduced the normalized impedance

$$z_L = Z_L / Z_0$$

We now express z_L as a function of Γ

$$z_L = \frac{1 + \Gamma}{1 - \Gamma} \quad (4.55)$$

The relation in (4.54) is a so called Möbius transformation from the complex plane $z_L = r_L + jx_L$ to the complex plane $\Gamma = \Gamma_r + j\Gamma_i$. The inverse transformation is given by (4.55) and is a Möbius transformation from the Γ -plane to the z_L -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_L -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_L -plane to the Γ -plane implies:

- A passive load, i.e., a load impedance $z_L = r_L + jx_L$ with $r_L \geq 0$, is mapped on a point inside the circle $|\Gamma| \leq 1$. Load impedances with $|z_L| = \infty$ are mapped to the point $\Gamma = 1$.
- The imaginary axis, $z_L = jx_L$, is mapped to a circle $|\Gamma| = |jx_L - 1|/|jx_L + 1| = 1$. Loads $z_L = \pm j\infty$ are mapped on $\Gamma = 1$ and $z_L = 0$ is mapped on $\Gamma = -1$.
- The line $z_L = r_L + jx_L$, $-\infty < x_L < \infty$ is mapped on a circle that goes through the points $\Gamma = 1$ and $\Gamma = (r_L - 1)/(r_L + 1)$ and has its centre on the real axis.
- The positive real axis $z_L = r_L$, $r_L \geq 0$ is mapped on the line $\Gamma = \Gamma_r$, $-1 \leq \Gamma_r \leq 1$.
- The line $z_L = r_L + jx_L$, $0 \leq r_L < \infty$ is mapped on the circle segment from the point $\Gamma = (jx_L - 1)/(jx_L + 1) = \exp(j(\pi - 2\arctan x_L))$ to the point $\Gamma = 1$. The circle segment is perpendicular to the circle $\Gamma = 1$.
- An inductive load, i.e., a load impedance z_L with $x_L > 0$, is mapped on a point with $\Gamma_i > 0$.
- A capacitive load, i.e., a load impedance z_L with $x <_L 0$, is mapped on a point with $\Gamma_i < 0$.
- If an impedance z_L is mapped on the point Γ then z_L^* is mapped on Γ^* .

When we draw the mapping of the straight lines $z_L = r_L + jx_L$, $-\infty < x_L < \infty$ and $z_L = r_L + jx_L$, $0 \leq r_L < \infty$ in the Γ -plane we get the Smith-chart. Some important properties of the Smith chart are:

1. A matched impedance load has $Z_L = Z_0$, i.e., $z_L = 1$, 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 $z_p < \ell$ it is

$$\Gamma(z_p) = \Gamma_0 e^{-2j\beta(\ell - z_p)}$$

Hence we move clock wise an angle $2\beta(\ell - z_p) = 4\pi(\ell - z_p)/\lambda$ along the circle with radius $|\Gamma_0|$ in the Smith chart when the distance to the load increases by a distance $(\ell - z_p)$.

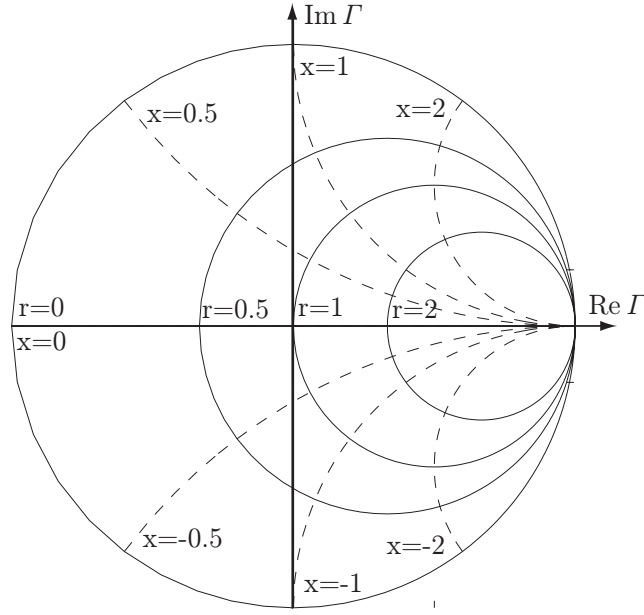


Figure 4.19: The Smith chart for the impedance z_L . When the admittance y_L is displayed x_L is exchanged for b_L and r_L for g_L . In the chart we use r, x instead of r_L, x_L , which is in accordance with the standard Smith chart.

3. When we move quarter of a wavelength, $(\ell - z_p) = \lambda/4$, 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 $y_L = 1/z_L$ is on the same circle with radius $|\Gamma|$ as z_L but at the point rotated 180° around $\Gamma = 0$. Then the values x_L and r_L in the Smith chart should be read as the normalized susceptance b_L and the normalized conductance g_L .

4.8.1 Matching a load by using the Smith chart

The Smith chart provides a highly effective graphical method for matching a complex load $Z_L = R_L + jX_L$ to a transmission line with characteristic impedance Z_0 at a single operating frequency. Graphically, matching can be conceptualized as navigating from the normalized load impedance $z_L = Z_L/Z_0$ to the center of the chart, $z_{in} = 1$. The permissible operations in this graphical construction follow two fundamental rules:

1. Line propagation: Clockwise rotation along a circle of constant reflection coefficient magnitude $|\Gamma|$ (centered at $\Gamma = 0$) corresponds to a movement towards the generator by a distance Δz , with phase rotation $\phi = 2\beta\Delta z = 4\pi\Delta z/\lambda$.
2. Reactive tuning: Motion along a constant-resistance circle r_L in the impedance (z) chart corresponds to adding a series reactance x . Motion along a constant-conductance circle g_L in the admittance (y) chart corresponds to adding a parallel susceptance b .

Adding resistive or conductive elements is prohibited, as lumped resistors dissipate power and degrade efficiency.

4.8.2 Frequency Response in the Smith Chart

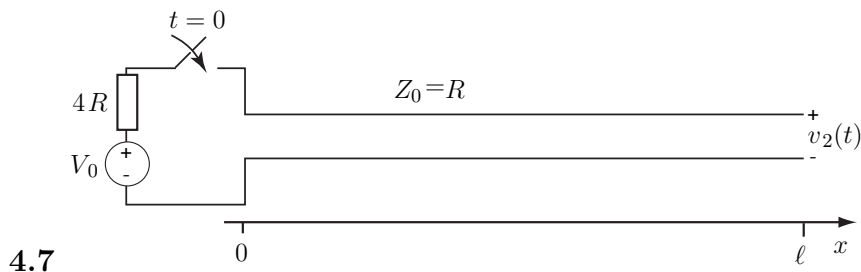
Consider a frequency-dependent input impedance $Z_{\text{in}}(f)$. Tracing $Z_{\text{in}}(f)$ across a continuous frequency band yields a locus on the Smith chart that completely characterizes the wideband behavior and matching performance of the device.

Problems in Chapter 4

- 4.1** 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 $V = V_0(1 - j)/\sqrt{2}$ where V_0 is real. Determine $v(t)$.
- 4.2** 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
- 4.3** Derive the relation between $[Z]$ and $[H]$ for a general two-port.
- 4.4** Show that Z and Y are symmetric matrices for a reciprocal N -port.
- 4.5** Show that the impedance matrix for a lossless N -port is purely imaginary.
- 4.6** A lossy transmission line has the following data at 100 MHz:

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

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



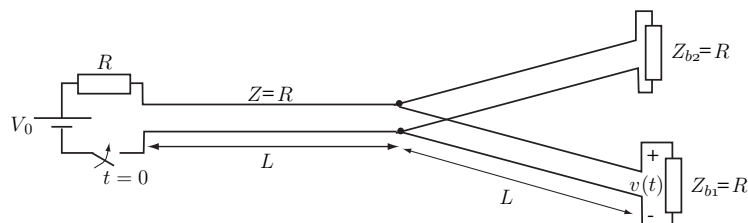
A lossless transmission line has the length ℓ . The line has the characteristic impedance R and the phase speed v_p . In the left end the line is connected to an ideal voltage source in series with a resistance $4R$, see figure. The other end of the line is open. That means that $i(\ell, t) = 0$. The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

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

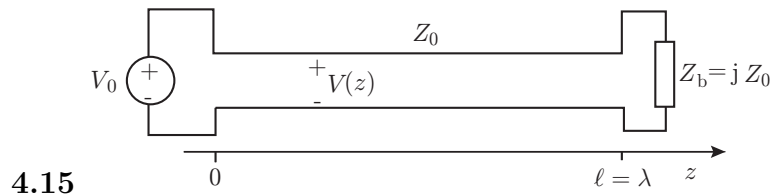
where V_0 is constant voltage. We are interested in the voltage $v_2(t)$ in the right end of the line. The source and the resistor are connected to the line via short wires.

- a) Determine $v_2(t)$ when $t \rightarrow +\infty$.
 b) Determine $v_2(\ell/2v_p)$.
 c) Determine $v_2(3\ell/2v_p)$.

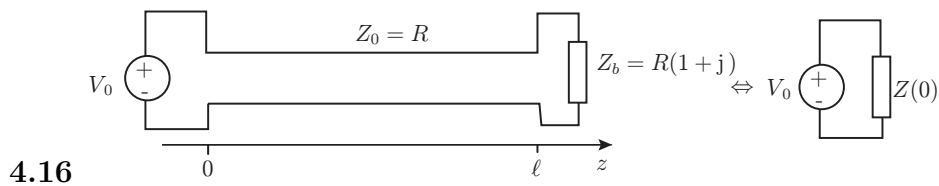
- d) Determine $v_2(5\ell/2v_p)$.
- e) Determine $v_2(7\ell/2v_p)$.
- 4.8** A lossy transmission line is terminated by a matched load. The voltage along the line is measured at two positions 20 m apart. The result was 2,8 V and 2,1 V.
- a) Determine the attenuation constant α .
- b) Determine the attenuation in dB/km.
- 4.9** 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 $-j360 \text{ V/A}$ and when it was shortened the impedance was $j10 \text{ V/A}$. The wavelength along the line was larger than 1 km. Determine the characteristic impedance Z_0 and the phase coefficient β .
- 4.10** An antenna with the purely resistive impedance 300Ω is to be matched to a coaxial cable with the characteristic impedance 60Ω . 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.
- 4.11** 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.
- 4.12** A lossless line with the characteristic impedance $Z_0 = 50 \Omega$ has a resistive load R_b . The standing wave ratio is $\text{SWR} = 3$. Determine R_b .
- 4.13** A lossless transmission line with characteristic impedance $Z_0 = 60 \Omega$ is terminated by a load impedance $Z_b = (60 + j60) \Omega$. The length of the line is $\lambda/8$, where λ is the wavelength along the line. Determine the input impedance of the line.

**4.14**

- 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.



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.



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 choose the length ℓ such that the input impedance $Z(0)$ is real.

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

Summary of chapter 4

Transmission lines

Time domain line equations

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

Frequency domain line equations

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

Characteristic impedance

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

Reflection coefficient at the load

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

Input impedance

$$Z(0) = \begin{cases} Z_0 \frac{Z_L \cos(\beta\ell) + jZ_0 \sin(\beta\ell)}{Z_0 \cos(\beta\ell) + jZ_L \sin(\beta\ell)}, & \text{lossless line} \\ Z_0 \frac{Z_L \cosh \gamma\ell + Z_0 \sinh \gamma\ell}{Z_L \sinh \gamma\ell + Z_0 \cosh \gamma\ell} & \text{lossy line} \end{cases}$$

Chapter 5

Electromagnetic fields with a preferred direction

In this chapter, we decompose an arbitrary vector field into its longitudinal component along the z -direction and its transverse vector component in the xy -plane. Applying this field decomposition to the Maxwell equations, we analyze the resulting field solutions for structures with a uniform cross section in the z -direction. The relations formulated here establish the foundation for all subsequent chapters on guided waves.

Note that from this chapter onward, we adopt the time-harmonic convention $e^{-i\omega t}$, shifting from the $e^{j\omega t}$ convention used in Chapter 4. This aligns with the standard convention found in the majority of waveguide literature. For brevity, explicit dependence on ω and t is omitted from the field arguments throughout this chapter.

5.1 Decomposition of vector fields

An arbitrary vector field $\mathbf{F}(\mathbf{r})$ can always be decomposed in two perpendicular vectors. 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:

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

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

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

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

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 $\mathbf{r}$

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

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

$$\begin{aligned}\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}(\mathbf{r})) &= \hat{\mathbf{z}} \cdot \left[\left(\nabla_T + \hat{\mathbf{z}} \frac{\partial}{\partial z} \right) \times (\mathbf{F}_T(\mathbf{r}) + \hat{\mathbf{z}} F_z(\mathbf{r})) \right] \\ &= \hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{F}_T(\mathbf{r})) = -\nabla_T \cdot (\hat{\mathbf{z}} \times \mathbf{F}_T(\mathbf{r}))\end{aligned}\quad (5.1)$$

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

$$\begin{aligned}\nabla \times \mathbf{F}(\mathbf{r}) - \hat{\mathbf{z}} (\hat{\mathbf{z}} \cdot (\nabla \times \mathbf{F}(\mathbf{r}))) \\ &= \left[\left(\nabla_T + \hat{\mathbf{z}} \frac{\partial}{\partial z} \right) \times (\mathbf{F}_T(\mathbf{r}) + \hat{\mathbf{z}} F_z(\mathbf{r})) \right] - \hat{\mathbf{z}} (\hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{F}_T(\mathbf{r}))) \\ &= \nabla_T \times \hat{\mathbf{z}} F_z(\mathbf{r}) + \hat{\mathbf{z}} \frac{\partial}{\partial z} \times \mathbf{F}_T(\mathbf{r}) = \hat{\mathbf{z}} \times \frac{\partial}{\partial z} \mathbf{F}_T(\mathbf{r}) - \hat{\mathbf{z}} \times \nabla_T F_z(\mathbf{r})\end{aligned}\quad (5.2)$$

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

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

5.2 Decomposition of the Maxwell equations

The Maxwell equations for a dispersive material are

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

We use the decomposition of a vector field that was introduced in section 5.1. The longitudinal components of the equations become, see (5.1),

$$\begin{cases} \hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{E}_T(\mathbf{r}, \omega)) = i\omega\mu_0\mu(\mathbf{r}, \omega)H_z(\mathbf{r}, \omega) \\ \hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{H}_T(\mathbf{r}, \omega)) = -i\omega\epsilon_0\epsilon(\mathbf{r}, \omega)E_z(\mathbf{r}, \omega) \end{cases}\quad (5.3)$$

The transverse parts are, see (5.2),

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

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

$$\begin{cases} \frac{\partial}{\partial z} \mathbf{E}_T(\mathbf{r}, \omega) = \nabla_T E_z(\mathbf{r}, \omega) - i\omega\mu_0\mu(\mathbf{r}, \omega)\hat{\mathbf{z}} \times \mathbf{H}_T(\mathbf{r}, \omega) \\ \frac{\partial}{\partial z} \mathbf{H}_T(\mathbf{r}, \omega) = \nabla_T H_z(\mathbf{r}, \omega) + i\omega\epsilon_0\epsilon(\mathbf{r}, \omega)\hat{\mathbf{z}} \times \mathbf{E}_T(\mathbf{r}, \omega) \end{cases}\quad (5.4)$$

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

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

where $\eta_0 = \sqrt{\mu_0/\epsilon_0}$ is the wave impedance of vacuum. The operator $\hat{\mathbf{z}} \times$ on the right hand side gives a rotation by an angle $\pi/2$ in the xy -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 6 and 9.

5.3 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 (5.3) that the transverse components of the fields have the same z -dependence and thus all fields can be written as

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

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

$$\begin{cases} \hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{E}_T(\boldsymbol{\rho}, k_z, \omega)) = i\omega\mu_0\mu(\omega)H_z(\boldsymbol{\rho}, k_z, \omega) \\ \hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{H}_T(\boldsymbol{\rho}, k_z, \omega)) = -i\omega\epsilon_0\epsilon(\omega)E_z(\boldsymbol{\rho}, k_z, \omega) \end{cases} \quad (5.5)$$

and

$$\begin{cases} ik_z \hat{\mathbf{z}} \times \mathbf{E}_T(\boldsymbol{\rho}, k_z, \omega) - i\omega\mu_0\mu(\omega)\mathbf{H}_T(\boldsymbol{\rho}, k_z, \omega) = \hat{\mathbf{z}} \times \nabla_T E_z(\boldsymbol{\rho}, k_z, \omega) \\ ik_z \hat{\mathbf{z}} \times \mathbf{H}_T(\boldsymbol{\rho}, k_z, \omega) + i\omega\epsilon_0\epsilon(\omega)\mathbf{E}_T(\boldsymbol{\rho}, k_z, \omega) = \hat{\mathbf{z}} \times \nabla_T H_z(\boldsymbol{\rho}, k_z, \omega) \end{cases} \quad (5.6)$$

We observe that the transverse components of the vectors $\mathbf{E}$ and $\mathbf{H}$, i.e. $\mathbf{E}_T$ and $\mathbf{H}_T$ can be expressed in terms of the longitudinal components E_z and H_z . To demonstrate this, we operate with $\hat{\mathbf{z}} \times$ on the first of the transverse equations in (5.6), utilizing $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ (BAC-CAB-rule), and by eliminating $\hat{\mathbf{z}} \times \mathbf{H}_T(\boldsymbol{\rho}, k_z, \omega)$ using the second equation in (5.6). If we treat the second equation in (5.6) in the same manner we get

$$\begin{cases} -ik_z \mathbf{E}_T(\boldsymbol{\rho}, k_z, \omega) - \frac{\omega\mu_0\mu(\omega)}{k_z} \left[\hat{\mathbf{z}} \times \nabla_T H_z(\boldsymbol{\rho}, k_z, \omega) \right. \\ \quad \left. - i\omega\epsilon_0\epsilon(\omega)\mathbf{E}_T(\boldsymbol{\rho}, k_z, \omega) \right] = -\nabla_T E_z(\boldsymbol{\rho}, k_z, \omega) \\ -ik_z \mathbf{H}_T(\boldsymbol{\rho}, k_z, \omega) + \frac{\omega\epsilon_0\epsilon(\omega)}{k_z} \left[\hat{\mathbf{z}} \times \nabla_T E_z(\boldsymbol{\rho}, k_z, \omega) \right. \\ \quad \left. + i\omega\mu_0\mu(\omega)\mathbf{H}_T(\boldsymbol{\rho}, k_z, \omega) \right] = -\nabla_T H_z(\boldsymbol{\rho}, k_z, \omega) \end{cases}$$

or

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

We rewrite these equations

$$\begin{cases} \mathbf{E}_T(\boldsymbol{\rho}, k_z, \omega) = \frac{i}{k_t^2} \{k_z \nabla_T E_z(\boldsymbol{\rho}, k_z, \omega) - \omega \mu_0 \mu(\omega) \hat{\mathbf{z}} \times \nabla_T H_z(\boldsymbol{\rho}, k_z, \omega)\} \\ \mathbf{H}_T(\boldsymbol{\rho}, k_z, \omega) = \frac{i}{k_t^2} \{k_z \nabla_T H_z(\boldsymbol{\rho}, k_z, \omega) + \omega \epsilon_0 \epsilon(\omega) \hat{\mathbf{z}} \times \nabla_T E_z(\boldsymbol{\rho}, k_z, \omega)\} \end{cases} \quad (5.7)$$

where we introduced the *transverse wavenumber* k_t

$$k_t^2 = \frac{\omega^2}{c_0^2} \epsilon(\omega) \mu(\omega) - k_z^2 \quad (5.8)$$

The relation between the wavenumbers $k(\omega)$, 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 (3.8) and (3.10) on page 15. The result is

$$\begin{cases} \nabla_T^2 E_z(\boldsymbol{\rho}, k_z, \omega) + k_t^2 E_z(\boldsymbol{\rho}, k_z, \omega) = 0 \\ \nabla_T^2 H_z(\boldsymbol{\rho}, k_z, \omega) + k_t^2 H_z(\boldsymbol{\rho}, k_z, \omega) = 0 \end{cases} \quad (5.9)$$

where the transverse wavenumber k_t is defined in (5.8). The transverse components also satisfy a system of partial differential equations. They are less useful since they are vectorial.

Problems in Chapter 5

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

5.2 Show that one can relate $\mathbf{E}_T$ and $\mathbf{H}_T$ in the following way:

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

and

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

We have here assumed that all fields have the specific z -dependence $\exp(ik_z z)$.

Summary of chapter 5

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

Maxwell equations

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

Transverse components

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

Equations for the longitudinal components

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

Chapter 6

Waveguides at fix frequency

Waveguides are structures that guide waves along a given direction. Figure 6.1 gives an example of geometry for a waveguide. The surface of the waveguide is denoted S and the normal to the surface $\hat{\mathbf{n}}$. Note that the normal $\hat{\mathbf{n}}$ 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 has the generating curve Γ , c.f., figure 6.4. Figure 6.4a shows a waveguide with a simply connected cross section Ω , while the figure 6.4b 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 are studied in this book. The first type is the hollow waveguide, which has metallic walls that enclose the guiding region. The other type is the open waveguide, for which parts of the enclosing surface is not metallic.

6.1 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

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

The sufficient boundary conditions on a perfectly conducting surface are, c.f., (2.17) on page 7,

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

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

We express the boundary conditions in terms of the decomposed fields in Section 5.1

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

In the first equation, $\hat{\mathbf{n}} \times \mathbf{E}_T$ is directed strictly along the z -axis, whereas $\hat{\mathbf{n}} \times \hat{\mathbf{z}} E_z$ lies purely in the transverse plane (tangential to the boundary curve Γ). Because these two vector

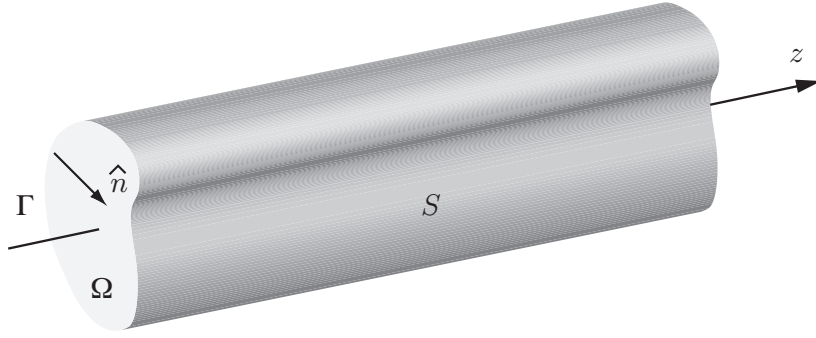


Figure 6.1: Geometry for waveguide.

contributions are orthogonal, each must vanish independently. In the second equation, the term $\hat{\mathbf{n}} \cdot \hat{\mathbf{z}} H_z$ vanishes identically due to the orthogonality of $\hat{\mathbf{n}}$ and $\hat{\mathbf{z}}$.

Consequently, the boundary conditions reduce to the equivalent set:

$$\begin{cases} E_z(\mathbf{r}, \omega) = 0 \\ \hat{\mathbf{z}} \cdot (\hat{\mathbf{n}} \times \mathbf{E}_T(\mathbf{r}, \omega)) = \hat{\mathbf{n}} \cdot (\mathbf{E}_T \times \hat{\mathbf{z}}) = 0 \\ \hat{\mathbf{n}} \cdot \mathbf{H}_T(\mathbf{r}, \omega) = 0 \end{cases} \quad \mathbf{r} \text{ on } S \quad (6.1)$$

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

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

We utilize (5.4)

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

in order to eliminate the z -derivative, and get

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

By utilizing the original boundary condition (6.1), the boundary conditions on the surface S are reduced to

$$\begin{cases} E_z(\mathbf{r}, \omega) = 0 \\ \frac{\partial}{\partial n} H_z(\mathbf{r}, \omega) = 0 \end{cases} \quad \mathbf{r} \text{ on } S \quad (6.2)$$

where $\frac{\partial}{\partial n} H_z(\mathbf{r}, \omega) = \hat{\mathbf{n}} \cdot \nabla_T H_z(\mathbf{r}, \omega)$. 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.

6.2 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 (6.2) separate the z -component of

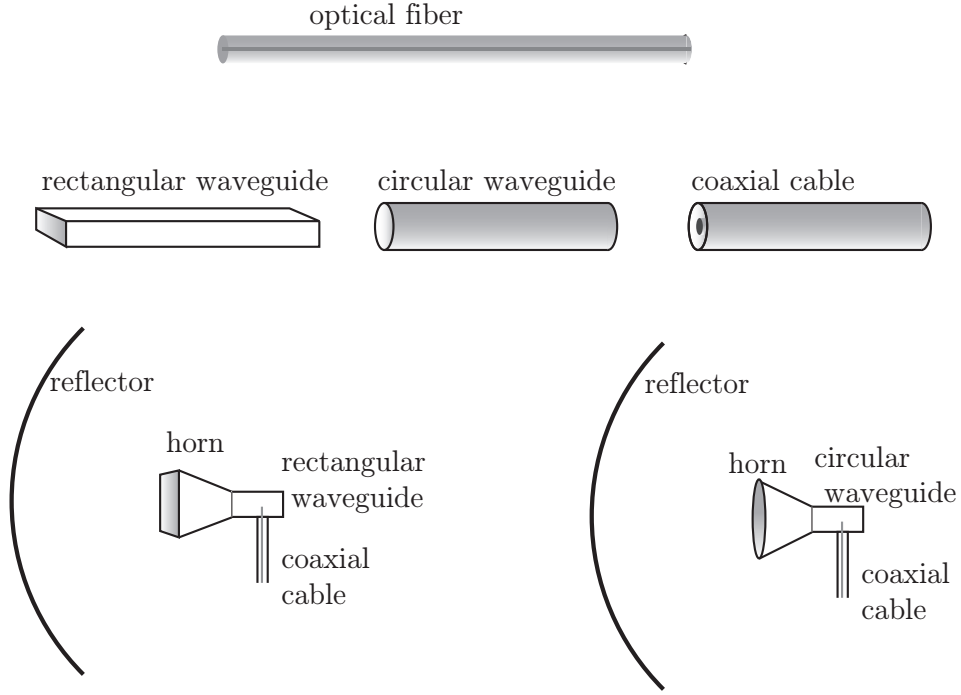


Figure 6.2: Examples of waveguides. Reflector antenna with a rectangular and circular feed horn.

the electric field, E_z from the z -component of the magnetic field, H_z . Consequently, we seek field solutions where either E_z or H_z vanishes identically:

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

The first scenario defines the transverse magnetic (TM) mode, in which the magnetic field has no axial component. The second represents the transverse electric (TE) mode, where the electric field lacks an axial component. Because there is no cross-coupling within the governing differential equations or along the boundary S , these two modes decouple completely and can be studied independently.

Later, we will also address the specific conditions required to support transverse electromagnetic (TEM) solutions, where both E_z and H_z are zero.

We let the region $z_1 < z < z_2$ be source free, i.e., $\mathbf{J} = \mathbf{0}$, c.f., Figure 6.5 and determine the waves that can exist in this region. Regions with sources are treated in Section 6.10.

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 (3.8) and (3.10) on page 15, and the boundary conditions for $E_z(\mathbf{r}, \omega)$ and $H_z(\mathbf{r}, \omega)$ are summarized as

$$\begin{cases} \nabla^2 E_z(\mathbf{r}, \omega) + k^2(\omega) E_z(\mathbf{r}, \omega) = 0 \\ E_z(\mathbf{r}, \omega) = 0 \quad \mathbf{r} \text{ on } S \end{cases} \quad z \in [z_1, z_2], \boldsymbol{\rho} \in \Omega \text{ (TM)} \\ \begin{cases} \nabla^2 H_z(\mathbf{r}, \omega) + k^2(\omega) H_z(\mathbf{r}, \omega) = 0 \\ \frac{\partial H_z}{\partial n}(\mathbf{r}, \omega) = 0 \quad \mathbf{r} \text{ on } S \end{cases} \quad z \in [z_1, z_2], \boldsymbol{\rho} \in \Omega \text{ (TE)} \end{cases} \quad (6.3)$$

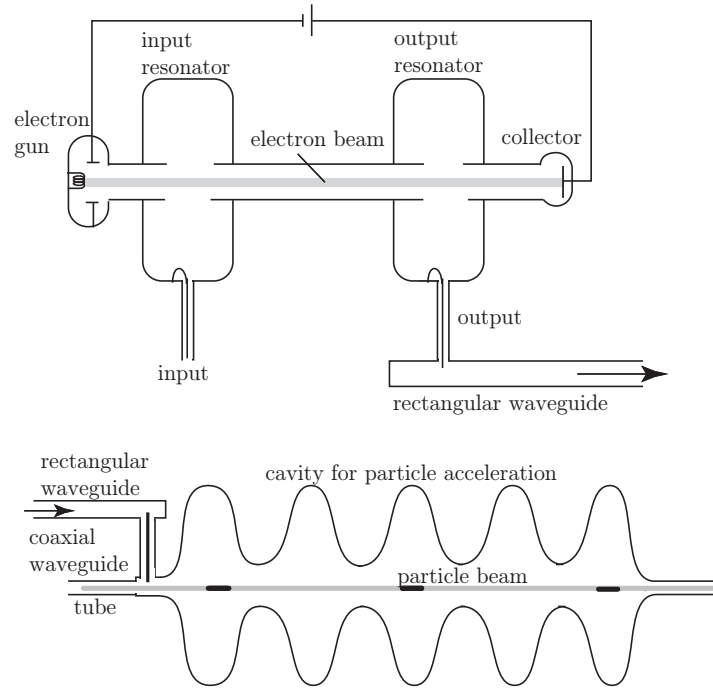


Figure 6.3: Upper: Klystron that generates the field for a cavity in an accelerator. Lower: Feed of a cavity in an accelerator.

where the wavenumber 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 $\mathbf{E}$ and $\mathbf{H}$ from (5.7) on page 66. The first system in (6.3) corresponds to the transverse magnetic (TM) mode, where $H_z = 0$. Conversely, the second system represents the transverse electric (TE) mode, characterized by $E_z = 0$.

We solve the two boundary-value problems in (6.3) using the method of separation of variables. As a fundamental technique in mathematical physics, this method provides a complete set of transverse basis functions in the x and y coordinates. The longitudinal field component, E_z for TM modes or H_z for TE modes, is then expanded in terms of these eigenfunctions, from which the remaining transverse field components are readily evaluated via their relations to the longitudinal fields.

6.2.1 The longitudinal components of the fields

We make the following ansatz

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

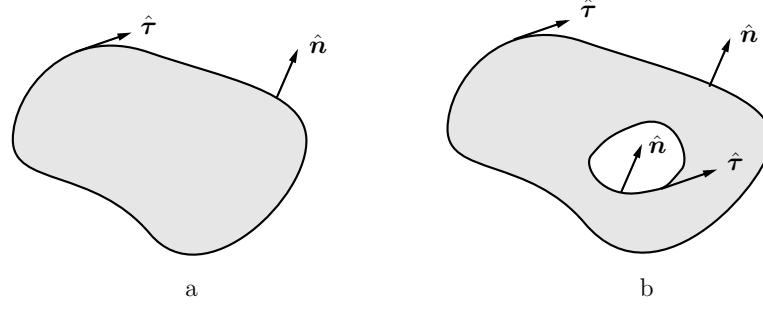


Figure 6.4: Cross sections of waveguides. The z -axis is directed perpendicular to the plane of the paper and $\hat{\tau} = \hat{z} \times \hat{n}$.

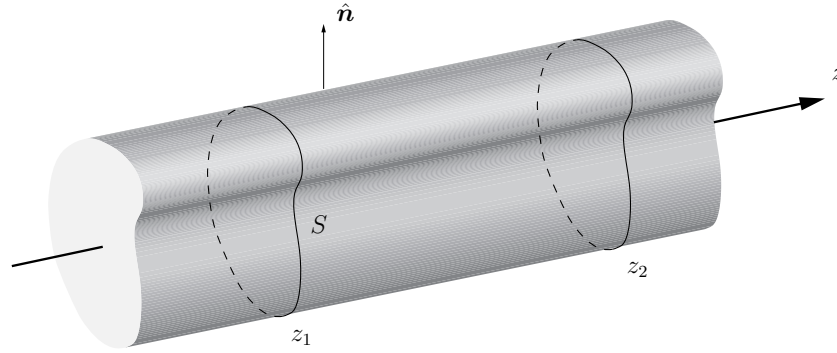


Figure 6.5: The source free region in the waveguide.

where $\boldsymbol{\rho} = \hat{x}x + \hat{y}y$ and η_0 is the wave impedance for vacuum. We insert this into (6.3)

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

and

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

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

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

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

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

and

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

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 TM and TE 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 wavenumber) 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 TE and TM modes 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 6.1, 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 TM and $w_n(\boldsymbol{\rho})$ for TE. They are solutions to ¹

$$\begin{cases} \nabla_T^2 v_n(\boldsymbol{\rho}) + k_{tn}^2 v_n(\boldsymbol{\rho}) = 0 & \boldsymbol{\rho} \in \Omega \\ v_n(\boldsymbol{\rho}) = 0 & \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \quad (\text{TM}) \quad (6.4)$$

and

$$\begin{cases} \nabla_T^2 w_n(\boldsymbol{\rho}) + k_{tn}^2 w_n(\boldsymbol{\rho}) = 0 & \boldsymbol{\rho} \in \Omega \\ \frac{\partial w_n(\boldsymbol{\rho})}{\partial n} = 0 & \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \quad (\text{TE}) \quad (6.5)$$

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 $\epsilon(\omega)$ 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 $E_z(\mathbf{r}, \omega)$ and $H_z(\mathbf{r}, \omega)$ are expanded as

$$\begin{cases} E_z(\mathbf{r}, \omega) = \sum_{n=1}^{\infty} a_n(z, \omega) v_n(\boldsymbol{\rho}) \\ H_z(\mathbf{r}, \omega) = \sum_{n=1}^{\infty} b_n(z, \omega) w_n(\boldsymbol{\rho}) \end{cases} \quad (6.6)$$

¹In 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.

Example 6.1

All eigenvalues for TM and TE modes 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

$$\oint_{\Gamma} f(\boldsymbol{\rho}) \frac{\partial}{\partial n} f(\boldsymbol{\rho}) dl = \iint_{\Omega} (\nabla_T f(\boldsymbol{\rho}))^2 dxdy + \iint_{\Omega} f(\boldsymbol{\rho}) \nabla_T^2 f(\boldsymbol{\rho}) dxdy$$

Notice that if the function is independent of the z -coordinate, then $\nabla f = \nabla_T f$.

We first consider TM modes 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 (6.4) to get the equality

$$\iint_{\Omega} (\nabla_T v_n(\boldsymbol{\rho}))^2 dxdy = k_{tn}^2 \iint_{\Omega} (v_n(\boldsymbol{\rho}))^2 dxdy \quad (6.7)$$

and the inequality

$$k_{tn}^2 \iint_{\Omega} (v_n(\boldsymbol{\rho}))^2 dxdy \geq 0$$

If v_n is not identically zero, the inequality implies that the eigenvalue for TM is non-negative, $k_{tn}^2 \geq 0$.

By letting $f(\boldsymbol{\rho}) = w_n(\boldsymbol{\rho})$, and by using (6.5) and the boundary condition $\frac{\partial}{\partial n} w_n(\boldsymbol{\rho}) = 0$, the corresponding relation for the TE modes

$$\iint_{\Omega} (\nabla_T w_n(\boldsymbol{\rho}))^2 dxdy = k_{tn}^2 \iint_{\Omega} (w_n(\boldsymbol{\rho}))^2 dxdy \quad (6.8)$$

and the inequality

$$k_{tn}^2 \iint_{\Omega} (w_n(\boldsymbol{\rho}))^2 dxdy \geq 0$$

are obtained. Unless w_n is not identically zero, the eigenvalue for the TE modes 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 TM and TE modes². From (6.7) and (6.8) we find that $k_{tn}^2 = 0$ implies that $\nabla_T v_n = \nabla_T w_n = 0$ in Ω , i.e., $v_n = \text{constant}$ and $w_n = \text{constant}$ in Ω . For TM 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_{tn}^2 for TM modes are all positive. To prove the same result for TE modes we use a result from page 64 and (5.3).

$$H_z(\mathbf{r}, \omega) = \frac{1}{i\omega\mu_0\mu(\omega)} \hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{E}_T(\mathbf{r}, \omega)) = \frac{1}{i\omega\mu_0\mu(\omega)} \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{E}_T(\mathbf{r}, \omega))$$

Since $w_n = \text{constant}$ in Ω , H_z in the left hand side cannot depend on $\boldsymbol{\rho}$. Stoke's theorem on the cross section Ω gives

$$\iint_{\Omega} H_z dxdy = \frac{1}{i\omega\mu_0\mu(\omega)} \iint_{\Omega} \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{E}_T(\mathbf{r}, \omega)) dxdy = \frac{1}{i\omega\mu_0\mu(\omega)} \oint_{\Gamma} \mathbf{E}_T(\mathbf{r}, \omega) \cdot d\mathbf{r} = 0$$

²The eigenvalue $k_{tn}^2 = 0$ is possible for the TEM-case.

due to the boundary condition $\hat{\mathbf{n}} \times \mathbf{E} = \mathbf{0}$. We get

$$H_z \iint_{\Omega} dx dy = 0$$

which implies that $H_z = 0$ or $w_n = 0$, which is a contradiction, and as above it follows that all eigenvalues $k_{t_n}^2$ for TE modes are positive. ■

Example 6.2

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 $k_{t_m}^2$ for TM and TE modes are orthogonal. We start with Gauss' theorem in the plane ($f_n = v_n$ for TM and $f_n = w_n$ for TE modes)

$$\begin{aligned} 0 &= \oint_{\Gamma} (f_n \nabla_T f_m - f_m \nabla_T f_n) \cdot \hat{\mathbf{n}} dl = \iint_{\Omega} \nabla_T \cdot (f_n \nabla_T f_m - f_m \nabla_T f_n) dx dy \\ &= \iint_{\Omega} (\nabla_T f_n \cdot \nabla_T f_m - \nabla_T f_m \cdot \nabla_T f_n + f_n \nabla_T^2 f_m - f_m \nabla_T^2 f_n) dx dy \\ &= (k_{t_n}^2 - k_{t_m}^2) \iint_{\Omega} f_n f_m dx dy \end{aligned}$$

where we have used the eigenvalue equation (6.4) or (6.5). If the eigenvalues are different

$$\iint_{\Omega} f_n f_m dx dy = 0$$

Thus the eigenfunctions v_n and v_m or w_n and w_m that belong to the eigenvalues $k_{t_n}^2$ and $k_{t_m}^2$ are orthogonal. ■

We insert the expansions in (6.6) into the original equation (6.3). 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 :

$$\begin{cases} \frac{\partial^2 a_n}{\partial z^2}(z, \omega) + \left(\frac{\omega^2}{c_0^2} \epsilon(\omega) \mu(\omega) - k_{t_n}^2 \right) a_n(z, \omega) = 0 \\ \frac{\partial^2 b_n}{\partial z^2}(z, \omega) + \left(\frac{\omega^2}{c_0^2} \epsilon(\omega) \mu(\omega) - k_{t_n}^2 \right) b_n(z, \omega) = 0 \end{cases}$$

The general solutions to these equations are

$$\begin{cases} a_n(z, \omega) = a_n^{\pm} e^{\pm i k_{z_n}(\omega) z} \\ b_n(z, \omega) = b_n^{\pm} e^{\pm i k_{z_n}(\omega) z} \end{cases}$$

where the *longitudinal wavenumber* k_{z_n} is

$$k_{z_n}(\omega) = \left(\frac{\omega^2}{c_0^2} \epsilon(\omega) \mu(\omega) - k_{t_n}^2 \right)^{\frac{1}{2}} \quad (6.9)$$

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 5.3. 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

$$\begin{cases} E_z(\mathbf{r}, \omega) = \sum_{n=1}^{\infty} v_n(\boldsymbol{\rho}) \left(a_n^+(\omega) e^{ik_{zn}z} + a_n^-(\omega) e^{-ik_{zn}z} \right) \\ H_z(\mathbf{r}, \omega) = \sum_{n=1}^{\infty} w_n(\boldsymbol{\rho}) \left(b_n^+(\omega) e^{ik_{zn}z} + b_n^-(\omega) e^{-ik_{zn}z} \right) \end{cases} \quad (6.10)$$

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. 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

$$\begin{cases} E_z(\mathbf{r}, \omega) = \sum_{n=1}^{\infty} v_n(\boldsymbol{\rho}) a_n^+(\omega) e^{ik_{zn}z} \\ H_z(\mathbf{r}, \omega) = \sum_{n=1}^{\infty} w_n(\boldsymbol{\rho}) b_n^+(\omega) e^{ik_{zn}z} \end{cases} \quad (6.11)$$

For a lossless waveguide, a mode with index n is propagating if $k_{tn} < k = \omega/c$, as k_{zn} is then purely real and the wave propagates without attenuation in the z -direction. Conversely, if $k_{tn} > k$, the mode is evanescent (non-propagating) and then $\text{Re } k_{zn} = 0$ and $\text{Im } k_{zn} < 0$, causing the wave to decay exponentially along z .

When $k_{tn} = k$, the mode is at its cutoff frequency. Here, $k_{zn} = 0$, resulting in a standing wave pattern in the xy -plane without power transfer along the guide. Since the sequence of transverse wavenumbers k_{tn} increases monotonically, only a finite number of modes propagate at any given frequency in (6.11). The cutoff frequency $f_{cn} = \omega_{cn}/(2\pi)$ for mode n is defined by setting $k_{zn} = 0$, yielding:

$$f_{cn} = \frac{k_{tn} c_0}{2\pi \sqrt{\epsilon \mu}} \quad (6.12)$$

The relation between the frequency f , the longitudinal wavenumber k_{zn} and the cut-off frequency f_{cn} is

$$k_{zn} = \frac{2\pi}{c_0} \sqrt{\epsilon \mu} (f^2 - f_{cn}^2)^{\frac{1}{2}}$$

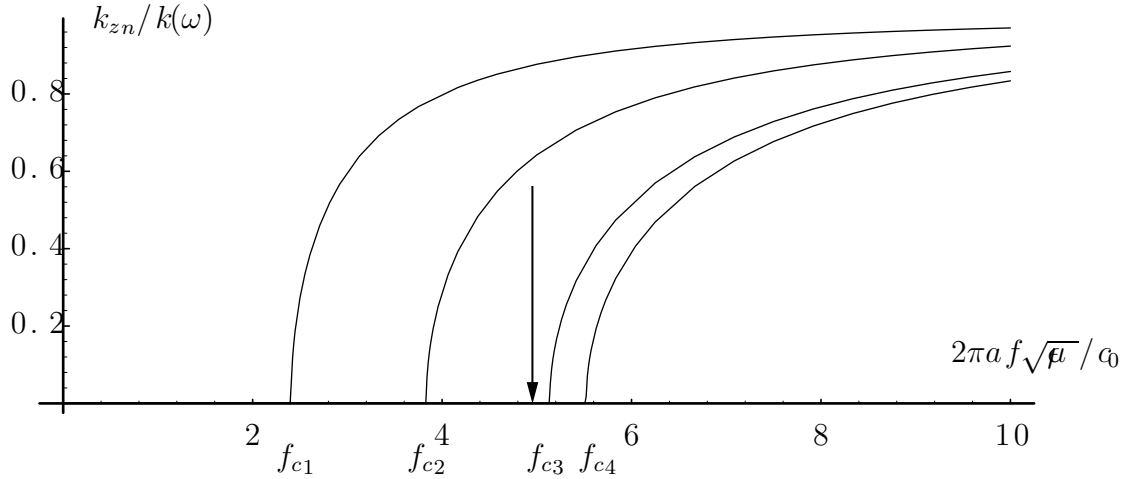


Figure 6.6: Waveguide dispersion for a circular waveguide (radius a) as a function of frequency f . The dispersion relations for the first four modes are shown. The arrow shows a frequency where only two of the modes are propagating. The longitudinal wavenumber k_{zn} is normalized with the wavenumber $k(\omega)$ and the frequency f is normalized with $c_0/(2\pi a\sqrt{\epsilon\mu})$, c.f., the explicit expressions of eigenvalues in table 6.4.

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 6.6. 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 material parameters $\epsilon(\omega)$ or $\mu(\omega)$ is complex, we define the cut-off frequency to be the frequency where $k_{tn} = \text{Re } k(\omega)$. 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.

6.2.2 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 $\exp\{\pm i k_z z\}$. To generate the transverse components of the fields we use the results from section 5.3. Equation (5.7) on page 66 provides us with the relations we need

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

where the relation (6.9) has been used. The result is

$$\left\{ \begin{array}{l} \mathbf{E}_T(\mathbf{r}, \omega) = \sum_{n=1}^{\infty} \frac{i}{k_{tn}^2} (a_n^+ e^{ik_{zn}z} - a_n^- e^{-ik_{zn}z}) k_{zn} \nabla_T v_n(\boldsymbol{\rho}) \\ \quad - \sum_{n=1}^{\infty} \frac{i}{k_{tn}^2} (b_n^+ e^{ik_{zn}z} + b_n^- e^{-ik_{zn}z}) \omega \mu_0 \mu(\omega) \hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho}) \\ \mathbf{H}_T(\mathbf{r}, \omega) = \sum_{n=1}^{\infty} \frac{i}{k_{tn}^2} (b_n^+ e^{ik_{zn}z} - b_n^- e^{-ik_{zn}z}) k_{zn} \nabla_T w_n(\boldsymbol{\rho}) \\ \quad + \sum_{n=1}^{\infty} \frac{i}{k_{tn}^2} (a_n^+ e^{ik_{zn}z} + a_n^- e^{-ik_{zn}z}) \omega \epsilon_0 \epsilon(\omega) \hat{\mathbf{z}} \times \nabla_T v_n(\boldsymbol{\rho}) \end{array} \right. \quad (6.13)$$

The transverse components are well defined quantities since $k_t^2 > 0$. 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.

6.3 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 6.2. 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 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 4 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 5.

If the z -components of the fields are zero we obtain from (5.3) and (5.4)

$$\left\{ \begin{array}{l} \hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{E}_T(\mathbf{r}, \omega)) = 0 \\ \hat{\mathbf{z}} \cdot (\nabla_T \times \mathbf{H}_T(\mathbf{r}, \omega)) = 0 \end{array} \right.$$

and

$$\left\{ \begin{array}{l} \frac{\partial}{\partial z} \mathbf{E}_T(\mathbf{r}, \omega) = -i\omega \mu_0 \mu(\omega) \hat{\mathbf{z}} \times \mathbf{H}_T(\mathbf{r}, \omega) \\ \frac{\partial}{\partial z} \mathbf{H}_T(\mathbf{r}, \omega) = i\omega \epsilon_0 \epsilon(\omega) \hat{\mathbf{z}} \times \mathbf{E}_T(\mathbf{r}, \omega) \end{array} \right. \quad (6.14)$$

Since $\nabla_T \times \mathbf{E}_T(\mathbf{r}, \omega)$ only has a z -component we see that

$$\nabla_T \times \mathbf{E}_T(\mathbf{r}, \omega) = \mathbf{0}$$

everywhere. Then there exists a scalar potential $\psi(\mathbf{r}, \omega)$ such that

$$\mathbf{E}_T(\mathbf{r}, \omega) = -\nabla_T \Psi(\mathbf{r}, \omega) \quad (6.15)$$

The potential Ψ is not uniquely defined, since every potential that differs by a function of z from Ψ gives the same field $\mathbf{E}_T$. Equations (3.7) and (3.8) on page 15 give

$$\begin{cases} \nabla^2 \mathbf{E}_T(\mathbf{r}, \omega) + k^2(\omega) \mathbf{E}_T(\mathbf{r}, \omega) = \mathbf{0} \\ \nabla_T \cdot \mathbf{E}_T(\mathbf{r}, \omega) = 0 \end{cases}$$

where the wavenumber, as usual, is $k^2(\omega) = \omega^2 \epsilon(\omega) \mu(\omega) / c_0^2$. When we insert $\mathbf{E}_T(\mathbf{r}, \omega) = -\nabla_T \Psi(\mathbf{r}, \omega)$ into this equation we get

$$\begin{cases} \nabla_T (\nabla^2 \Psi(\mathbf{r}, \omega) + k^2(\omega) \Psi(\mathbf{r}, \omega)) = \mathbf{0} \\ \nabla_T^2 \Psi(\mathbf{r}, \omega) = 0 \end{cases}$$

We eliminate the derivatives w.r.t. the transverse coordinates by using the lower equation. The result is

$$\nabla_T \left\{ \frac{d^2}{dz^2} \Psi(\mathbf{r}, \omega) + k^2(\omega) \Psi(\mathbf{r}, \omega) \right\} = \mathbf{0}$$

which gives

$$\frac{d^2}{dz^2} \Psi(\mathbf{r}, \omega) + k^2(\omega) \Psi(\mathbf{r}, \omega) = C(z)$$

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

$$\frac{d^2}{dz^2} \Psi(\mathbf{r}, \omega) + k^2(\omega) \Psi(\mathbf{r}, \omega) = 0$$

The solutions to this equation are

$$\Psi(\mathbf{r}, \omega) = \psi^\pm(\boldsymbol{\rho}) e^{\pm i k(\omega) z}$$

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 $\mathbf{E}_T(\mathbf{r}, \omega)$ and $\mathbf{H}_T(\mathbf{r}, \omega)$ using (6.14) and (6.15). The result is

$$\begin{cases} \mathbf{E}_T(\mathbf{r}, \omega) = -\nabla_T \psi^+(\boldsymbol{\rho}) e^{i k(\omega) z} - \nabla_T \psi^-(\boldsymbol{\rho}) e^{-i k(\omega) z} \\ \mathbf{H}_T(\mathbf{r}, \omega) = -(\eta_0 \eta(\omega))^{-1} \hat{\mathbf{z}} \times \left(\nabla_T \psi^+(\boldsymbol{\rho}) e^{i k(\omega) z} - \nabla_T \psi^-(\boldsymbol{\rho}) e^{-i k(\omega) z} \right) \end{cases} \quad (6.16)$$

where we have introduced the relative wave impedance for the material in the waveguide $\eta(\omega)$ as

$$\eta(\omega) = \sqrt{\frac{\mu(\omega)}{\epsilon(\omega)}} \quad (6.17)$$

Thus the relation between the $\mathbf{E}$ and $\mathbf{H}$ 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 $\nabla \cdot \mathbf{E} = 0$ and (6.15)

$$\nabla_T^2 \psi(\boldsymbol{\rho}) = 0 \quad (6.18)$$

The boundary condition is $\psi = \text{constant}$ on every simply connected part of the curve Γ (this is obtained from $\hat{\mathbf{n}} \times \mathbf{E} = \mathbf{0}$ 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 6.4b. We also see that the solution ψ is independent of ω . The potential is normalized such that

$$\iint_{\Omega} \nabla_T \psi(\boldsymbol{\rho}) \cdot \nabla_T \psi(\boldsymbol{\rho}) \, dx dy = 1 \quad (6.19)$$

If we compare the longitudinal wavenumber k_z for the TEM mode with the corresponding wavenumbers for TM and TE modes, we see that the TEM mode corresponds to the forbidden case $k_t = 0$.

In chapter 4 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 $\psi^{\pm}(\boldsymbol{\rho})e^{\pm ik(\omega)z}$ by

$$V^{\pm}(z) = \psi^{\pm}(\boldsymbol{\rho})e^{\pm ik(\omega)z} \big|_{\boldsymbol{\rho} \text{ on conductor one}} - \psi^{\pm}(\boldsymbol{\rho})e^{\pm ik(\omega)z} \big|_{\boldsymbol{\rho} \text{ on conductor two}}$$

6.3.1 Waveguides with several conductors

Consider a waveguide with N separated simply connected conductors. We denote the border of conductor number p by Γ_p and consider the problems

$$\begin{aligned} \nabla_T^2 \psi_p(\boldsymbol{\rho}) &= 0 \\ \psi_p(\boldsymbol{\rho}) &= \begin{cases} V_p & \text{on } \Gamma_p \\ V_0 & \text{on } \Gamma_q, q = 1 \dots N, q \neq p \end{cases} \end{aligned} \quad (6.20)$$

where $p = 1 \dots N - 1$ and where $V_p \neq V_0$. 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 $p = N$ 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 6.7 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 (6.20) with $N = 5$.

6.4 Vector basis functions in hollow waveguides

Based on (6.10) and (6.13) we define a set of vector valued basis functions $\mathbf{E}_{n\nu}^{\pm}(\mathbf{r}, \omega)$, that are suitable for the expansion of the electric field in a hollow waveguide.

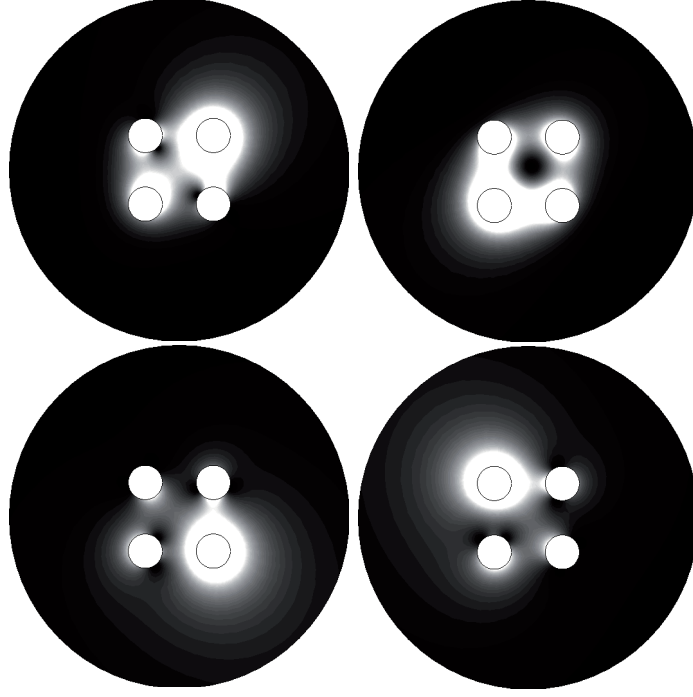


Figure 6.7: The power flow density of the four TEM modes in a waveguide consisting of four metallic circular conductors in a hollow circular metallic cylinder. The light areas have large power flow density.

The vector basis functions for TM modes that propagate in $\pm z$ -direction are defined as ³

$$\begin{cases} \mathbf{E}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \{\mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) \pm v_n(\boldsymbol{\rho})\hat{\mathbf{z}}\} e^{\pm ik_{zn}z} \\ \mathbf{H}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \pm \mathbf{H}_{Tn\nu}(\boldsymbol{\rho}, \omega) e^{\pm ik_{zn}z} \end{cases} \quad \nu = \text{TM} \quad (6.21)$$

The TE vector basis functions are defined by

$$\begin{cases} \mathbf{E}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) e^{\pm ik_{zn}z} \\ \mathbf{H}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \{\pm \mathbf{H}_{Tn\nu}(\boldsymbol{\rho}, \omega) + w_n(\boldsymbol{\rho})\hat{\mathbf{z}}\} e^{\pm ik_{zn}z} \end{cases} \quad \nu = \text{TE} \quad (6.22)$$

We define the vector basis for TEM modes as

$$\begin{cases} \mathbf{E}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) e^{\pm ikz} \\ \mathbf{H}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \pm \mathbf{H}_{Tn\nu}(\boldsymbol{\rho}, \omega) e^{\pm ikz} \end{cases} \quad \nu = \text{TEM} \quad (6.23)$$

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, $\nu = \text{TM}, \text{TE}, \text{TEM}$.

The connections between w_n , v_n , ψ_n and the transverse components $\mathbf{E}_{Tn\nu}$ and $\mathbf{H}_{Tn\nu}$

³Note that the functions are normalized such that $E_z = v_n$ and $H_z = w_n$ are dimensionless. The $\mathbf{E}$ and $\mathbf{H}$ are to be multiplied with an amplitude with dimension A/m for TE modes and V/m for TM modes.

are obtained from (5.7) and (6.16):

$$\begin{cases} \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) = \begin{cases} \frac{i}{k_{zn}^2} k_{zn} \nabla_T v_n(\boldsymbol{\rho}), & \nu = \text{TM} \\ -\frac{i\omega}{k_{zn}^2} \mu_0 \mu(\omega) \hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho}), & \nu = \text{TE} \\ -\nabla_T \psi_n(\boldsymbol{\rho}, \omega), & \nu = \text{TEM} \end{cases} \\ \mathbf{H}_{Tn\nu}(\boldsymbol{\rho}, \omega) = Z_{n\nu}^{-1} \hat{\mathbf{z}} \times \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) \end{cases} \quad (6.24)$$

where $Z_{n\nu}$ is the mode impedance defined by

$$Z_{n\nu} = \begin{cases} \frac{k_{zn}}{\omega \epsilon_0 \epsilon}, & \nu = \text{TM} \\ \frac{\omega \mu_0 \mu}{k_{zn}}, & \nu = \text{TE} \\ \eta_0 \eta, & \nu = \text{TEM} \end{cases} \quad (6.25)$$

As before, $v_n(\boldsymbol{\rho})$ and $w_n(\boldsymbol{\rho})$ are the real eigenfunctions to the eigenvalue problems in (6.4) and (6.5). We normalize them such that

$$\begin{cases} \iint_{\Omega} v_n(\boldsymbol{\rho}) v_{n'}(\boldsymbol{\rho}) \, dx dy = \delta_{n,n'} \\ \iint_{\Omega} w_n(\boldsymbol{\rho}) w_{n'}(\boldsymbol{\rho}) \, dx dy = \delta_{n,n'} \end{cases} \quad (6.26)$$

where δ_{ij} is the Kronecker delta (c.f., Appendix D).

The vector basis functions in (6.24) constitute a complete set of vector basis functions for the Maxwell equations in a source free region of the waveguide. With this notation a general solution can be written in a very condensed form

$$\begin{cases} \mathbf{E}(\mathbf{r}, \omega) = \sum_{\nu=\text{TM,TE,TEM}}^n (a_{n\nu}^+ \mathbf{E}_{n\nu}^+(\mathbf{r}, \omega) + a_{n\nu}^- \mathbf{E}_{n\nu}^-(\mathbf{r}, \omega)) \\ \mathbf{H}(\mathbf{r}, \omega) = \sum_{\nu=\text{TM,TE,TEM}}^n (a_{n\nu}^+ \mathbf{H}_{n\nu}^+(\mathbf{r}, \omega) + a_{n\nu}^- \mathbf{H}_{n\nu}^-(\mathbf{r}, \omega)) \end{cases} \quad \begin{cases} z \in [z_1, z_2] \\ \boldsymbol{\rho} \in \Omega \end{cases} \quad (6.27)$$

where summation is over the mode indices n and $\nu = \text{TM, TE, TEM}$. An advantage with this definition of the vector basis functions is that the expansion coefficients $a_{n\nu}^{\pm}$ are the same for the expansions of the electric and magnetic fields. The expansion coefficients $a_{n\nu}^{\pm}$ are determined by the sources of the fields c.f., section 6.10. 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.

6.4.1 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

$$b_f = \frac{\text{BW}}{f_{\text{center}}} = 2 \frac{f_1 - f_0}{f_1 + f_0} \quad (6.28)$$

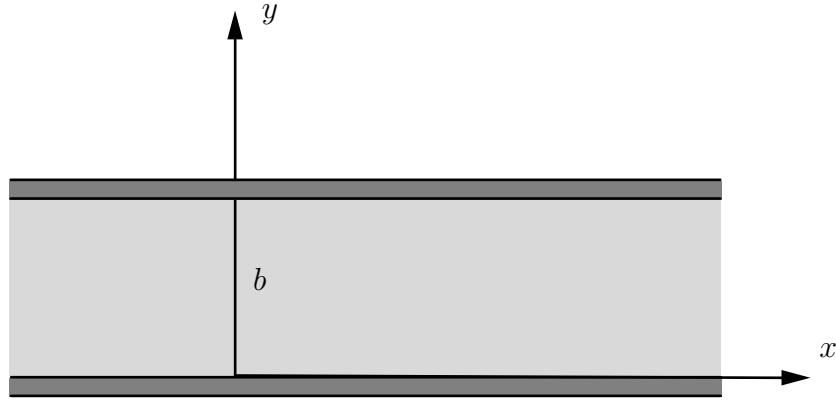


Figure 6.8: Geometry for a planar waveguide.

where $BW = f_1 - f_0$ is the bandwidth and $f_{\text{center}} = 0.5(f_0 + f_1)$ 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.

6.5 Examples

We now give examples of very important, cross-sections for which we can derive explicit expressions of the vector basis functions.

6.5.1 Planar waveguide

We start with the simplest case, which is the planar waveguide. The geometry of this waveguide is depicted in figure 6.8.

The solution is obtained from the following one-dimensional eigenvalue problems:

$$\begin{cases} \frac{d^2 Y(y)}{dy^2} + \gamma Y(y) = 0, & 0 \leq y \leq b \\ Y(y) = 0, & y = 0, b \end{cases}$$

and

$$\begin{cases} \frac{d^2 Y(y)}{dy^2} + \gamma Y(y) = 0, & 0 \leq y \leq b \\ \frac{dY}{dy}(y) = 0, & y = 0, b \end{cases}$$

The solutions to these problems are given by

$$Y_n(y) = \sin\left(\frac{n\pi y}{b}\right), \quad n = 1, 2, 3, \dots$$

and

$$Y_n(y) = \cos\left(\frac{n\pi y}{b}\right), \quad n = 0, 1, 2, 3, \dots$$

	Eigenfunctions $v_n, w_n, \nabla_T \psi$	Eigenvalues $k_{t_n}^2$
TM _{<i>n</i>}	$v_n = \sqrt{\frac{2}{b}} \sin\left(\frac{n\pi y}{b}\right)$	$\pi^2 \frac{n^2}{b^2}$
TE _{<i>n</i>}	$w_n = \sqrt{\frac{2}{b}} \cos\left(\frac{n\pi y}{b}\right)$	$\pi^2 \frac{n^2}{b^2}$
TEM	$\nabla_T \psi = \sqrt{\frac{1}{b}} \hat{\mathbf{y}}$	0

Table 6.1: A table of the normalized eigenfunctions to (6.4), (6.5) and (6.18) for the planar waveguide, c.f., figure 6.8 for the definition of the geometry. The integer n has the values $n = 1, 2, 3, \dots$

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 basis functions for TM are

$$v_n(y) = \sqrt{\frac{2}{b}} \sin\left(\frac{n\pi y}{b}\right), \quad n = 1, 2, 3, \dots$$

and the basis functions for TE are

$$w_n(y) = \sqrt{\frac{2}{b}} \cos\left(\frac{n\pi y}{b}\right), \quad n = 1, 2, 3, \dots$$

The index $n = 0$ does not correspond to a TM mode, as seen from page 76.

The planar waveguide has two separated surfaces which means that a TEM mode exists. We use (6.18) to determine the TEM mode basis functions

$$\begin{cases} \frac{d^2 \psi(y)}{dy^2} = 0, & 0 \leq y \leq b \\ \psi(y) = \begin{cases} C_1, & y = 0 \\ C_2, & y = b \end{cases} \end{cases}$$

The solution is

$$\psi(y) = C_1 + \frac{C_2 - C_1}{b} y$$

and the TEM mode has the normalized basis function (see (6.16))

$$\nabla_T \psi^\pm(y) = \sqrt{\frac{1}{b}} \hat{\mathbf{y}}$$

The results are collected in table 6.1.

6.5.2 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 6.9. The

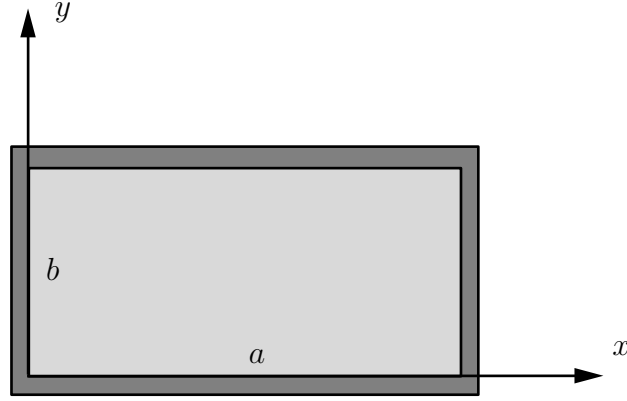


Figure 6.9: The geometry for a waveguide with rectangular cross-section.

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

$$\begin{cases} \frac{\partial^2 v(\boldsymbol{\rho})}{\partial x^2} + \frac{\partial^2 v(\boldsymbol{\rho})}{\partial y^2} + k_t^2 v(\boldsymbol{\rho}) = 0 \\ v(\boldsymbol{\rho}) = 0 \quad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \quad (\text{TM})$$

and

$$\begin{cases} \frac{\partial^2 w(\boldsymbol{\rho})}{\partial x^2} + \frac{\partial^2 w(\boldsymbol{\rho})}{\partial y^2} + k_t^2 w(\boldsymbol{\rho}) = 0 \\ \frac{\partial w}{\partial n}(\boldsymbol{\rho}) = 0 \quad \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \quad (\text{TE})$$

The solution is based on the following one-dimensional eigenvalue problems:

$$\begin{cases} \frac{d^2 X(x)}{dx^2} + \gamma X(x) = 0, & 0 \leq x \leq a \\ X(x) = 0, & x = 0, a \end{cases}$$

and

$$\begin{cases} \frac{d^2 \tilde{X}(x)}{dx^2} + \gamma \tilde{X}(x) = 0, & 0 \leq x \leq a \\ \frac{d\tilde{X}}{dx}(x) = 0, & x = 0, a \end{cases}$$

The solutions to these two problems are

$$X_m(x) = \sin\left(\frac{m\pi x}{a}\right), \quad m = 1, 2, 3, \dots$$

and

$$\tilde{X}_m(x) = \cos\left(\frac{m\pi x}{a}\right), \quad m = 0, 1, 2, 3, \dots,$$

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

	Eigenfunctions v_{mn}, w_{mn}	Eigenvalues $k_{t mn}^2$
TM _{mn}	$v_{mn} = \frac{2}{\sqrt{ab}} \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right)$	$\pi^2 \left(\frac{m^2}{a^2} + \frac{n^2}{b^2}\right)$
TE _{mn}	$w_{mn} = \sqrt{\frac{\varepsilon_m \varepsilon_n}{ab}} \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right)$	$\pi^2 \left(\frac{m^2}{a^2} + \frac{n^2}{b^2}\right)$

Table 6.2: Table of normalized eigenfunctions to (6.4) and (6.5) for rectangular waveguides, see figure 6.9. The integers m and n can have values $m, n = 0, 1, 2, 3, \dots$, 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 76. 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 TE₁₀ mode. This mode is called the fundamental mode and is very important.

are obtained as a product of these sets of one-dimensional eigenfunctions ⁴, i.e.,

$$\begin{cases} \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right), & \text{TM} \\ \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right), & \text{TE} \end{cases}$$

The eigenvalues in the two cases are the same $k_t^2 = \pi^2 (m^2/a^2 + n^2/b^2)$. The normalized functions are

$$\begin{cases} v_{mn} = \frac{2}{\sqrt{ab}} \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right), & \text{TM} \\ w_{mn} = \sqrt{\frac{\varepsilon_m \varepsilon_n}{ab}} \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right), & \text{TE} \end{cases}$$

where the Neumann-factor is $\varepsilon_m = 2 - \delta_{m,0}$. The results are collected in table 6.2.

Example 6.3

The fundamental mode of a rectangular waveguide with $a > b$ is the TE₁₀ mode. It has the cut-off frequency $f_{c10} = \frac{c_0}{2a}$ and $w_{10} = \sqrt{\frac{2}{ab}} \cos\left(\frac{\pi x}{a}\right)$. The normalized electric field is

$$\mathbf{E}_{10\text{TE}}(x, \omega) = \hat{\mathbf{y}} \frac{i\omega\mu_0}{\pi} \sqrt{\frac{2a}{b}} \sin\left(\frac{\pi x}{a}\right) \quad (6.29)$$

If $a > 2b$ then the second mode is TE₂₀ that has cut-off frequency $f_{c20} = \frac{c_0}{a}$. If $b < a < 2b$ then TE₀₁ is the second mode with cut-off frequency $f_{c01} = \frac{c_0}{2b}$. In order to maximize the

⁴A 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} (m ⁻¹) ^a	43.3	107i	119i	136i	179i	184i	233i	255i	274i	282i
k_{zmn} (m ⁻¹) ^b	144.6	86.6	70.6	22.6	114i	122i	188i	215i	237i	246i

^aThe frequency is $f = 3.8$ GHz.

^bThe frequency is $f = 7.6$ GHz.

Table 6.3: Table of the lowest cut-off frequencies f_{cmn} and the longitudinal wavenumber k_{zmn} for a rectangular waveguide with dimensions $4.7 \text{ cm} \times 2.2 \text{ cm}$. 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 $\exp(-\text{Im}\{k_{zmn}\}z)$.

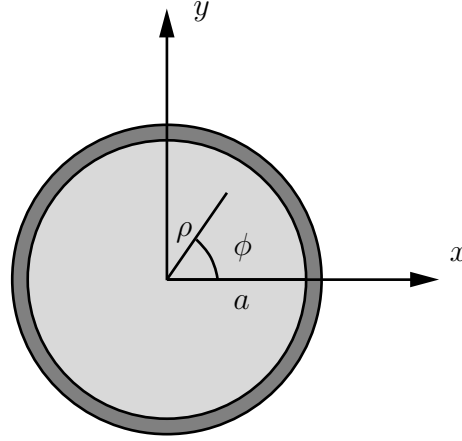


Figure 6.10: Geometry for waveguide with circular cross-section.

bandwidth it is common to have rectangular waveguides with $a > 2b$. Then the bandwidth is $\text{BW} = \frac{c_0}{2a}$ and the fractional bandwidth is $b_f = 2(c_0/2a)/(3c_0/2a) = 2/3 = 0.67$. ■

Example 6.4

A rectangular waveguide has dimensions $4.7 \text{ cm} \times 2.2 \text{ cm}$. The cut-off frequencies f_{cmn} for the different modes are easy to calculate from (6.12) and table 6.2. The longitudinal wavenumbers k_{zmn} , given by (6.9), are related to the frequency f and the cut-off frequency f_{cmn} in the following way

$$k_{zmn} = \frac{2\pi}{c_0} \sqrt{\epsilon\mu} \sqrt{f^2 - f_{cmn}^2}$$

The results are given in table 6.3. The bandwidth is $\text{BW} = 3.19 \text{ GHz}$ and the fractional bandwidth is $b_f = 2/3$. ■

6.5.3 Waveguide with circular cross-section

The geometry of the circular waveguide with radius a is depicted in figure 6.10. 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

$$\begin{cases} \nabla_T^2 v(\rho) + k_t^2 v(\rho) = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial v(\rho)}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 v(\rho)}{\partial \phi^2} + k_t^2 v(\rho) = 0 \\ v(a, \phi) = 0 \end{cases} \quad (\text{TM})$$

and

$$\begin{cases} \nabla_T^2 w(\rho) + k_t^2 w(\rho) = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial w(\rho)}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 w(\rho)}{\partial \phi^2} + k_t^2 w(\rho) = 0 \\ \frac{\partial w}{\partial n}(a, \phi) = 0 \end{cases} \quad (\text{TE})$$

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

$$\begin{cases} \rho \frac{\partial}{\partial \rho} \left(\rho \frac{\partial f(\rho)}{\partial \rho} \right) + (k_t^2 \rho^2 - \gamma) f(\rho) = 0 \\ \frac{\partial^2 g(\phi)}{\partial \phi^2} + \gamma g(\phi) = 0 \end{cases}$$

The solution to the eigenvalue problem in the variable ϕ is

$$g(\phi) = \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}, \quad m = 0, 1, 2, 3, \dots$$

Only integer values of m are allowed since the function must be periodic in ϕ with period 2π , i.e., only $\gamma = m^2$, $m = 0, 1, 2, 3, \dots$ 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 TM and TE modes, respectively, add extra conditions. For these boundary conditions to be satisfied, the transverse wavenumber has to satisfy

$$k_t a = \begin{cases} \xi_{mn}, & (\text{TM}) \\ \eta_{mn}, & (\text{TE}) \end{cases}$$

	Eigenfunctions v_{mn}, w_{mn}	Eigenvalues $k_{t\,mn}^2$
TM _{mn}	$v_{mn} = \frac{\sqrt{\varepsilon_m} J_m(\xi_{mn}\rho/a)}{\sqrt{\pi} a J'_m(\xi_{mn})} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}$	$\frac{\xi_{mn}^2}{a^2}$
TE _{mn}	$w_{mn} = \frac{\sqrt{\varepsilon_m} \eta_{mn} J_m(\eta_{mn}\rho/a)}{\sqrt{\pi} (\eta_{mn}^2 - m^2) a J_m(\eta_{mn})} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}$	$\frac{\eta_{mn}^2}{a^2}$

Table 6.4: Table of the normalized eigenfunctions to (6.4) and (6.5) for waveguides with circular cross-section, see figure 6.10 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 $J'_m(\eta_{mn}) = 0$, $m = 0, 1, 2, 3, \dots$, $n = 1, 2, 3, \dots$ are listed in tables A.1 and A.2 in appendix A. The mode with the lowest cut-off frequency is the TE₁₁ mode.

where ξ_{mn} and η_{mn} , $n = 1, 2, 3, \dots$, 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

$$\begin{cases} v_{mn} = \frac{\sqrt{\varepsilon_m} J_m(\xi_{mn}\rho/a)}{\sqrt{\pi} a J'_m(\xi_{mn})} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}, & \text{TM} \\ w_{mn} = \frac{\sqrt{\varepsilon_m} \eta_{mn} J_m(\eta_{mn}\rho/a)}{\sqrt{\pi} (\eta_{mn}^2 - m^2) a J_m(\eta_{mn})} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix}, & \text{TE} \end{cases}$$

where $\varepsilon_m = 2 - \delta_{m,0}$. The results are collected in table 6.4.

Example 6.5

The fundamental mode is the TE₁₁ mode. The cut-off frequency is given by $f_{c11} = \frac{c_0 \eta_{11}}{2\pi a}$ 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 TM₀₁ mode with cut-off frequency $f_{c01} = \frac{c_0 \xi_{01}}{2\pi a}$ where $\xi_{01} = 2.405$ is the first zero of $J_0(x)$. The bandwidth is $\text{BW} = f_{c01} - f_{c11} = \frac{c_0}{2\pi a} (2.405 - 1.841) = 0.564 \frac{c_0}{2\pi a}$. The fractional bandwidth is $b_f = 2 \frac{0.564}{1.841 + 2.405} = 0.265$. ■

6.6 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 $\iint_{\Omega} |E_z(\boldsymbol{\rho})|^2 dx dy = 1$ for the TM modes and $\iint_{\Omega} |H_z(\boldsymbol{\rho})|^2 dx dy = 1$ for the TE modes. To obtain this normalization we divide all field values with $\iint_{\Omega} |E_z(\boldsymbol{\rho})|^2 dx dy$ for the TM modes and $\iint_{\Omega} |H_z(\boldsymbol{\rho})|^2 dx dy$ for the TE modes. To integrate we right click on **Derived values** and choose integration and surface integral.
- Notice that there are many options of surface graphs to choose from.

Example 6.6

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 6.11. 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 6.12 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 6.11. 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 $a > 2b$.

■

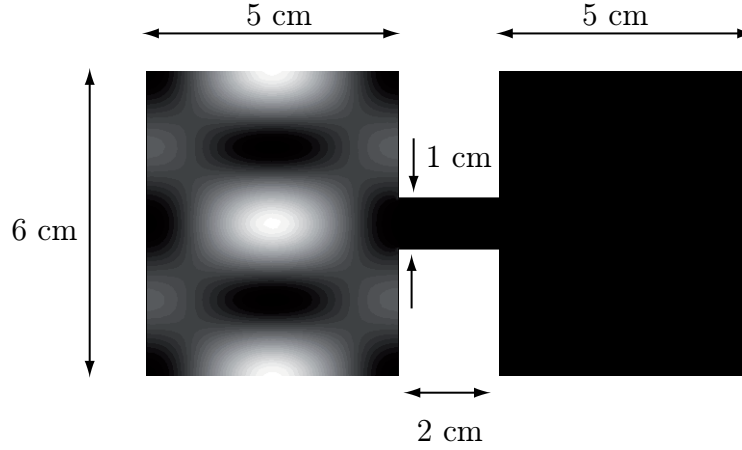


Figure 6.11: Geometry for the ridge waveguide and the power flow density for the TM mode with cut-off frequency 5.83 GHz.

6.7 Normalization integrals

In this section we solve the following integrals

$$\begin{aligned} & \iint_{\Omega} \hat{z} \cdot \{ \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) \times \mathbf{H}_{Tn'\nu'}^*(\boldsymbol{\rho}, \omega) \} \, dx dy \\ & \iint_{\Omega} \hat{z} \cdot \{ \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) \times \mathbf{H}_{Tn'\nu'}(\boldsymbol{\rho}, \omega) \} \, dx dy \end{aligned} \quad (6.30)$$

The real part of the first integral is twice the power that the mode $n\nu$ transports in the waveguide, see section 6.8. The second integral is needed when we treat sources in waveguides in Section 6.10. The values of the integrals are given in (6.37) and (6.39).

The eigenfunctions $v_n(\boldsymbol{\rho})$ and $w_n(\boldsymbol{\rho})$ are real functions that are normalized according to (6.26), i.e.,

$$\begin{cases} \iint_{\Omega} v_n(\boldsymbol{\rho}) v_{n'}(\boldsymbol{\rho}) \, dx dy = \delta_{n,n'} \\ \iint_{\Omega} w_n(\boldsymbol{\rho}) w_{n'}(\boldsymbol{\rho}) \, dx dy = \delta_{n,n'} \end{cases}$$

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

$$\begin{aligned} \iint_{\Omega} \nabla_T g(\boldsymbol{\rho}) \cdot \nabla_T f(\boldsymbol{\rho}) \, dx dy &= - \iint_{\Omega} g(\boldsymbol{\rho}) \nabla_T^2 f(\boldsymbol{\rho}) \, dx dy \\ &+ \int_{\Gamma} g(\boldsymbol{\rho}) \hat{n} \cdot \nabla_T f(\boldsymbol{\rho}) \, dl \end{aligned} \quad (6.31)$$

where dl is the line element along the border curve Γ .

For the TM mode, using the eigenfunctions $f(\boldsymbol{\rho}) = v_n(\boldsymbol{\rho})$ and $g(\boldsymbol{\rho}) = v_{n'}(\boldsymbol{\rho})$, the

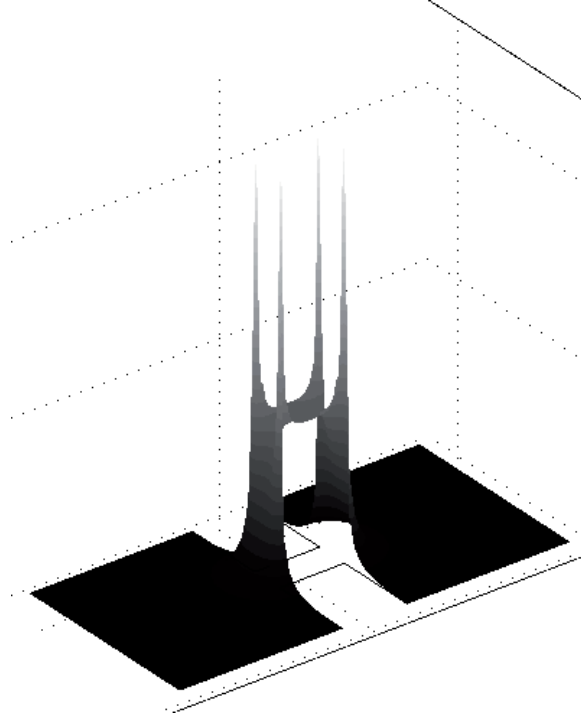


Figure 6.12: The power flow density for the fundamental mode of the ridge waveguide, i.e., the TE mode with cut-off frequency 0.663 GHz. The power flow is concentrated to the narrow section. Notice that the electric field is very strong at edges.

boundary integral vanishes because $v_n(\boldsymbol{\rho}) = 0$ along Γ . This yields:

$$\begin{aligned} \iint_{\Omega} \nabla_T v_n(\boldsymbol{\rho}) \cdot \nabla_T v_{n'}(\boldsymbol{\rho}) \, dx dy &= - \iint_{\Omega} v_n(\boldsymbol{\rho}) \nabla_T^2 v_{n'}(\boldsymbol{\rho}) \, dx dy \\ &= k_{t_{n'}}^2 \iint_{\Omega} v_n(\boldsymbol{\rho}) v_{n'}(\boldsymbol{\rho}) \, dx dy = k_{t_n}^2 \delta_{n,n'} \end{aligned} \quad (6.32)$$

where we have used (6.4) and (6.26).

Similarly, for the TE mode, substituting the eigenfunctions $f(\boldsymbol{\rho}) = w_n(\boldsymbol{\rho})$ and $g(\boldsymbol{\rho}) = w_{n'}(\boldsymbol{\rho})$ causes the boundary integral to vanish, since $\frac{\partial}{\partial n} w_n(\boldsymbol{\rho}) = 0$ on Γ . Consequently

$$\begin{aligned} \iint_{\Omega} \nabla_T w_n(\boldsymbol{\rho}) \cdot \nabla_T w_{n'}(\boldsymbol{\rho}) \, dx dy &= - \iint_{\Omega} w_n(\boldsymbol{\rho}) \nabla_T^2 w_{n'}(\boldsymbol{\rho}) \, dx dy \\ &= k_{t_{n'}}^2 \iint_{\Omega} w_n(\boldsymbol{\rho}) w_{n'}(\boldsymbol{\rho}) \, dx dy = k_{t_n}^2 \delta_{n,n'} \end{aligned} \quad (6.33)$$

where we used (6.5) and (6.26). We see that the gradient of the eigenfunctions are orthogonal, but not normalized.

Four more combinations of integrals are of interest in later sections:

$$\left\{ \begin{array}{l} \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ \nabla_T v_n(\boldsymbol{\rho}) \times \nabla_T w_{n'}(\boldsymbol{\rho}) \} \, dx dy \\ \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ (\hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times (\hat{\mathbf{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho})) \} \, dx dy \end{array} \right. \quad (6.34)$$

and

$$\left\{ \begin{array}{l} \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ \nabla_T v_n(\boldsymbol{\rho}) \times (\hat{\mathbf{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho})) \} \, dx dy \\ \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ (\hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times \nabla_T w_{n'}(\boldsymbol{\rho}) \} \, dx dy \end{array} \right. \quad (6.35)$$

We rewrite the first integral in (6.34) by using formulas for the nabla-operator and Stoke's theorem (the unit normal vector to Ω is $\hat{\mathbf{z}}$).

$$\begin{aligned} \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ \nabla_T v_n(\boldsymbol{\rho}) \times \nabla_T w_{n'}(\boldsymbol{\rho}) \} \, dx dy &= \hat{\mathbf{z}} \cdot \iint_{\Omega} \nabla_T \times (v_n(\boldsymbol{\rho}) \nabla_T w_{n'}(\boldsymbol{\rho})) \, dx dy \\ &= \iint_{\Omega} \nabla \times (v_n(\boldsymbol{\rho}) \nabla_T w_{n'}(\boldsymbol{\rho})) \cdot \hat{\mathbf{n}} \, dx dy = \int_{\Gamma} (v_n(\boldsymbol{\rho}) \nabla_T w_{n'}(\boldsymbol{\rho})) \cdot d\mathbf{r} = 0 \end{aligned}$$

since $v_n(\boldsymbol{\rho}) = 0$ on the boundary curve Γ . The linear element is $d\mathbf{r} = \hat{\mathbf{z}} \times \hat{\mathbf{n}} dl = \hat{\boldsymbol{\tau}} dl$. In the same manner we can show that the second integral in (6.34) is zero by using $(\mathbf{c} \times \mathbf{a}) \times (\mathbf{c} \times \mathbf{b}) = \mathbf{c}((\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c})$ and utilizing

$$\begin{aligned} \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ (\hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times (\hat{\mathbf{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho})) \} \, dx dy \\ = \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ \nabla_T w_n(\boldsymbol{\rho}) \times \nabla_T v_{n'}(\boldsymbol{\rho}) \} \, dx dy = 0 \end{aligned}$$

The integrals in (6.35) are simplified by using the BAC-CAB-rule.

$$\begin{aligned} \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ \nabla_T v_n(\boldsymbol{\rho}) \times (\hat{\mathbf{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho})) \} \, dx dy \\ = \iint_{\Omega} \nabla_T v_n(\boldsymbol{\rho}) \cdot \nabla_T v_{n'}(\boldsymbol{\rho}) \, dx dy \\ \iint_{\Omega} \hat{\mathbf{z}} \cdot \{ (\hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times \nabla_T w_{n'}(\boldsymbol{\rho}) \} \, dx dy \\ = - \iint_{\Omega} \nabla_T w_n(\boldsymbol{\rho}) \cdot \nabla_T w_{n'}(\boldsymbol{\rho}) \, dx dy \end{aligned}$$

By using the orthogonality integrals (6.32) and (6.33) we can summarize the integrals in

(6.34) and (6.35) as

$$\left\{ \begin{array}{l} \iint_{\Omega} \hat{\mathbf{z}} \cdot \{\nabla_T v_n(\boldsymbol{\rho}) \times \nabla_T w_{n'}(\boldsymbol{\rho})\} \, dx dy = 0 \\ \iint_{\Omega} \hat{\mathbf{z}} \cdot \{(\hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times (\hat{\mathbf{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho}))\} \, dx dy = 0 \\ \iint_{\Omega} \hat{\mathbf{z}} \cdot \{\nabla_T v_n(\boldsymbol{\rho}) \times (\hat{\mathbf{z}} \times \nabla_T v_{n'}(\boldsymbol{\rho}))\} \, dx dy = k_{t_n}^2 \delta_{n,n'} \\ \iint_{\Omega} \hat{\mathbf{z}} \cdot \{(\hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho})) \times \nabla_T w_{n'}(\boldsymbol{\rho})\} \, dx dy = -k_{t_n}^2 \delta_{n,n'} \end{array} \right. \quad (6.36)$$

From the definitions of $\mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega)$ and $\mathbf{H}_{Tn\nu}(\boldsymbol{\rho}, \omega)$ in (6.24) we find by using (6.36) that

$$\iint_{\Omega} \hat{\mathbf{z}} \cdot \{\mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) \times \mathbf{H}_{Tn'\nu'}^*(\boldsymbol{\rho}, \omega)\} \, dx dy = 2P_{n\nu}^E \delta_{n,n'} \delta_{\nu,\nu'} \quad (6.37)$$

We have here introduced the mode power $P_{n\nu}^E$ as

$$P_{n\nu}^E = \begin{cases} \frac{\omega}{2k_{t_n}^2} \begin{cases} k_{zn} \epsilon_0 \epsilon^*(\omega), & \nu = \text{TM} \\ k_{zn}^* (\omega) \mu_0 \mu(\omega), & \nu = \text{TE} \end{cases} \\ \frac{1}{2\eta_0 \eta^*}, & \nu = \text{TEM} \end{cases} \quad (6.38)$$

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 6.10 we need the following integral

$$\iint_{\Omega} \hat{\mathbf{z}} \cdot \{\mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) \times \mathbf{H}_{Tn'\nu'}(\boldsymbol{\rho}, \omega)\} \, dx dy = -2U_{n\nu} \delta_{n,n'} \delta_{\nu,\nu'} \quad (6.39)$$

where

$$U_{n\nu} = \begin{cases} \frac{\omega k_{zn}}{2k_{t_n}^2} \begin{cases} \epsilon_0 \epsilon(\omega), & \nu = \text{TM} \\ \mu_0 \mu(\omega), & \nu = \text{TE} \end{cases} \\ \frac{1}{2\eta_0 \eta}, & \nu = \text{TEM} \end{cases} \quad (6.40)$$

6.8 Power flow density

The active power flow density of the electromagnetic field is given by the Poynting vector (see (2.20) on page 9)

$$\langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) = \frac{1}{2} \text{Re} \{ \mathbf{E}(\mathbf{r}, \omega) \times \mathbf{H}^*(\mathbf{r}, \omega) \}$$

The z -component of this vector gives the time average of the power per unit area transported along the waveguide

$$\hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) = \frac{1}{2} \hat{\mathbf{z}} \cdot \text{Re} \{ \mathbf{E}(\mathbf{r}, \omega) \times \mathbf{H}^*(\mathbf{r}, \omega) \}$$

We introduce the decomposition of the vector fields in their longitudinal and transverse components according to section 5.1.

$$\begin{aligned}\hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) &= \frac{1}{2} \hat{\mathbf{z}} \cdot \text{Re} \{ (\mathbf{E}_T(\mathbf{r}, \omega) + \hat{\mathbf{z}} E_z(\mathbf{r}, \omega)) \times (\mathbf{H}_T^*(\mathbf{r}, \omega) + \hat{\mathbf{z}} H_z^*(\mathbf{r}, \omega)) \} \\ &= \frac{1}{2} \hat{\mathbf{z}} \cdot \text{Re} \{ \mathbf{E}_T(\mathbf{r}, \omega) \times \mathbf{H}_T^*(\mathbf{r}, \omega) \}\end{aligned}$$

Using the general waveguide solutions from (6.27), (6.21), and (6.22), we rewrite the expression for power flow in the z -direction

$$\begin{aligned}\hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) &= \sum_{\substack{n, n' \\ \nu, \nu' = \text{TM, TE, TEM}}} \frac{1}{2} \hat{\mathbf{z}} \cdot \text{Re} \left\{ \left(a_{n\nu}^+ \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) e^{ik_{zn}z} + a_{n\nu}^- \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) e^{-ik_{zn}z} \right) \right. \\ &\quad \left. \times \left(a_{n'\nu'}^{+*} \mathbf{H}_{Tn'\nu'}^*(\boldsymbol{\rho}, \omega) e^{-ik_{zn'}^*z} - a_{n'\nu'}^{-*} \mathbf{H}_{Tn'\nu'}^*(\boldsymbol{\rho}, \omega) e^{ik_{zn'}^*z} \right) \right\}\end{aligned}$$

The time average of the total power flow at z is

$$\iint_{\Omega} \hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) \, dx dy$$

In this surface integral over the cross section the integral in (6.37) 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

$$\begin{aligned}\iint_{\Omega} \hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) \, dx dy &= \text{Re} \sum_{\nu = \text{TM, TE, TEM}}^n P_{n\nu}^E \left\{ |a_{n\nu}^+|^2 e^{i(k_{zn} - k_{zn}^*)z} - |a_{n\nu}^-|^2 e^{-i(k_{zn} - k_{zn}^*)z} \right. \\ &\quad \left. + a_{n\nu}^- a_{n\nu}^{+*} e^{-i(k_{zn} + k_{zn}^*)z} - a_{n\nu}^+ a_{n\nu}^{-*} e^{i(k_{zn} + k_{zn}^*)z} \right\} \\ &= \text{Re} \sum_{\nu = \text{TM, TE, TEM}}^n P_{n\nu}^E \left\{ |a_{n\nu}^+|^2 e^{-2 \text{Im} k_{zn} z} - |a_{n\nu}^-|^2 e^{2 \text{Im} k_{zn} z} + 2i \text{Im} \left(a_{n\nu}^- a_{n\nu}^{+*} e^{-2i \text{Re} k_{zn} z} \right) \right\}\end{aligned}$$

where the mode power, $P_{n\nu}^E$, is given by (6.38), 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

$$\iint_{\Omega} \hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) \, dx dy = \text{Re} \sum_{\nu = \text{TM, TE}}^n P_{n\nu}^E |a_{n\nu}^+|^2 e^{-2 \text{Im} k_{zn} z} \quad (6.41)$$

for the waves propagating in the $+z$ -direction and

$$\iint_{\Omega} \hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) \, dx dy = -\text{Re} \sum_{\nu = \text{TM, TE}}^n P_{n\nu}^E |a_{n\nu}^-|^2 e^{2 \text{Im} k_{zn} z}$$

for the waves propagating in the $-z$ -direction. These sums are finite since k_{zn} , and hence $P_{n\nu}^E$, is purely imaginary when $f < f_{cn}$, or equivalently $k_{tn} > k(\omega)$.

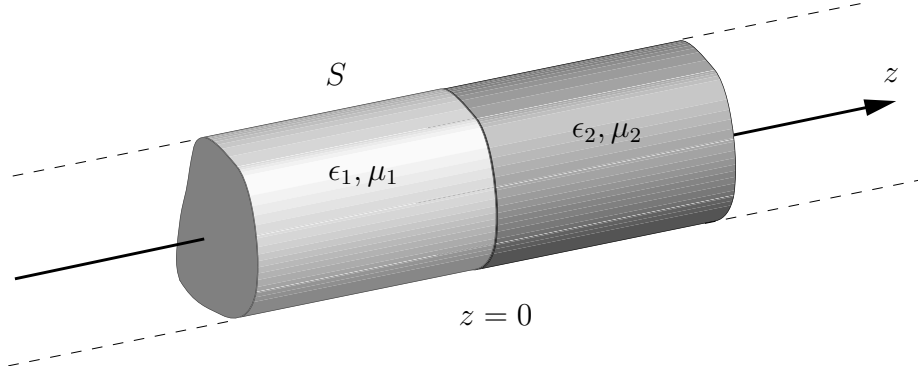


Figure 6.13: Reflection and transmission in a waveguide filled with two isotropic materials with material parameters ϵ_1, μ_1 and ϵ_2, μ_2 .

Example 6.7

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 $f_{c\text{TE}_{11}} = 1.841c_0/2\pi a = 1.76 \cdot 10^{11}$ Hz, see table 6.4. 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 $\alpha_{\text{TE}_{11}} = \frac{4\pi}{c_0} \sqrt{f_{c\text{TE}_{11}}^2 - f^2} = 7363 \text{ m}^{-1}$. With a thickness of 0.5 mm this corresponds to an attenuation of 16 dB of the microwaves. ■

Example 6.8

Consider a source in the region $z < 0$ that 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 6.13. Determine the reflection and transmission coefficients for the waves.

We denote the electric and magnetic fields generated by the sources $\mathbf{E}_{n\nu}^+(\mathbf{r}, \omega)$ and $\mathbf{H}_{n\nu}^+(\mathbf{r}, \omega)$. 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 (6.27) (normalization is $a_{n\nu}^+ = 1$)

$$\begin{cases} \mathbf{E}(\mathbf{r}, \omega) = \mathbf{E}_{n\nu}^+(\mathbf{r}, \omega) + \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} r_{n'\nu'} \mathbf{E}_{n'\nu'}^-(\mathbf{r}, \omega) \\ \mathbf{H}(\mathbf{r}, \omega) = \mathbf{H}_{n\nu}^+(\mathbf{r}, \omega) + \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} r_{n'\nu'} \mathbf{H}_{n'\nu'}^-(\mathbf{r}, \omega) \end{cases} \quad z \leq 0$$

where ϵ and μ have values ϵ_1 and μ_1 .

A general ansatz for the region $z > 0$ is

$$\begin{cases} \mathbf{E}(\mathbf{r}, \omega) = \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} t_{n'\nu'} \mathbf{E}_{n'\nu'}^+(\mathbf{r}, \omega) \\ \mathbf{H}(\mathbf{r}, \omega) = \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} t_{n'\nu'} \mathbf{H}_{n'\nu'}^+(\mathbf{r}, \omega) \end{cases} \quad z \geq 0$$

where ϵ and μ have values ϵ_2 and μ_2 . Due to the boundary conditions (continuity of the transverse components of $\mathbf{E}$ and $\mathbf{H}$) and the orthogonality relations (6.36) 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

$$\begin{cases} \mathbf{E}(\mathbf{r}, \omega) = \mathbf{E}_{n\nu}^+(\mathbf{r}, \omega) + r_{n\nu} \mathbf{E}_{n\nu}^-(\mathbf{r}, \omega) \\ \mathbf{H}(\mathbf{r}, \omega) = \mathbf{H}_{n\nu}^+(\mathbf{r}, \omega) + r_{n\nu} \mathbf{H}_{n\nu}^-(\mathbf{r}, \omega) \end{cases} \quad z \leq 0$$

and the fields in the region $z > 0$ reduce to

$$\begin{cases} \mathbf{E}(\mathbf{r}, \omega) = t_{n\nu} \mathbf{E}_{n\nu}^+(\mathbf{r}, \omega) \\ \mathbf{H}(\mathbf{r}, \omega) = t_{n\nu} \mathbf{H}_{n\nu}^+(\mathbf{r}, \omega) \end{cases} \quad z \geq 0$$

The boundary condition on the interface $z = 0$ give (see (6.24), (6.21) and (6.22))

$$\begin{cases} (k_1^2 - k_{t_n}^2)^{1/2} (1 + r_{n\nu}) = t_{n\nu} (k_2^2 - k_{t_n}^2)^{1/2} \\ (1 - r_{n\nu}) \epsilon_1 = t_{n\nu} \epsilon_2 \end{cases} \quad \nu = \text{TM}$$

and

$$\begin{cases} (1 + r_{n\nu}) \mu_1 = t_{n\nu} \mu_2 \\ (k_1^2 - k_{t_n}^2)^{1/2} (1 - r_{n\nu}) = t_{n\nu} (k_2^2 - k_{t_n}^2)^{1/2} \end{cases} \quad \nu = \text{TE}$$

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 wavenumbers in region 1 and 2, respectively. Note that the transverse wavenumber k_{t_n} , 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

$$\begin{cases} r_{n\nu} = \frac{\epsilon_1 (k_2^2 - k_{t_n}^2)^{1/2} - \epsilon_2 (k_1^2 - k_{t_n}^2)^{1/2}}{\epsilon_1 (k_2^2 - k_{t_n}^2)^{1/2} + \epsilon_2 (k_1^2 - k_{t_n}^2)^{1/2}} \\ t_{n\nu} = \frac{2\epsilon_1 (k_1^2 - k_{t_n}^2)^{1/2}}{\epsilon_1 (k_2^2 - k_{t_n}^2)^{1/2} + \epsilon_2 (k_1^2 - k_{t_n}^2)^{1/2}} \end{cases} \quad \nu = \text{TM}$$

and

$$\begin{cases} r_{n\nu} = \frac{\mu_2 (k_1^2 - k_{t_n}^2)^{1/2} - \mu_1 (k_2^2 - k_{t_n}^2)^{1/2}}{\mu_2 (k_1^2 - k_{t_n}^2)^{1/2} + \mu_1 (k_2^2 - k_{t_n}^2)^{1/2}} \\ t_{n\nu} = \frac{2\mu_1 (k_1^2 - k_{t_n}^2)^{1/2}}{\mu_2 (k_1^2 - k_{t_n}^2)^{1/2} + \mu_1 (k_2^2 - k_{t_n}^2)^{1/2}} \end{cases} \quad \nu = \text{TE}$$

■

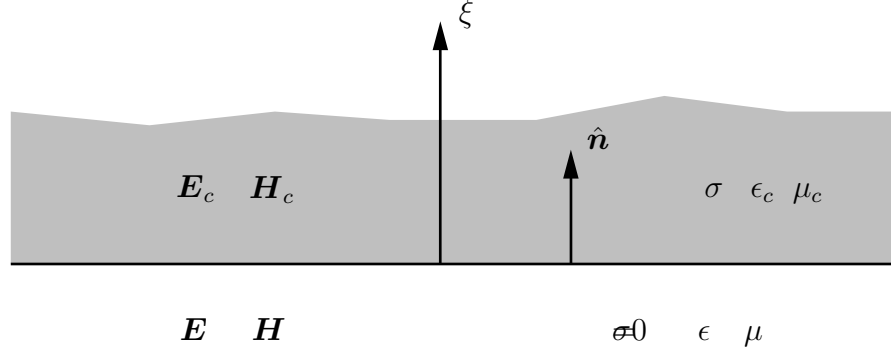


Figure 6.14: The interface between the non-perfectly conducting wall and the inner region of the waveguide.

6.9 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 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 fulfilled. The fields in the metal are denoted $\mathbf{E}_c$, $\mathbf{D}_c$, $\mathbf{B}_c$ and $\mathbf{H}_c$ while the fields outside the metal do not have any index. We let $\hat{\mathbf{n}}_c$ be the normal to the surface, directed into the metal, and introduce the coordinate ξ along the $\hat{\mathbf{n}}_c$ -direction, see figure 6.14. The boundary conditions at the surface are given by (see (2.16) on page 6)

$$\begin{aligned}\hat{\mathbf{n}}_c \times \mathbf{H} &= \hat{\mathbf{n}}_c \times \mathbf{H}_c \\ \hat{\mathbf{n}}_c \times \mathbf{E} &= \hat{\mathbf{n}}_c \times \mathbf{E}_c\end{aligned}$$

i.e., the tangential components of $\mathbf{E}$ and $\mathbf{H}$ 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. The approximation $\sigma \gg \omega \epsilon_0 \epsilon_c$ gives

$$k(\omega) = \frac{\omega \sqrt{\epsilon_c \mu_c}}{c_0} \left(1 + i \frac{\sigma}{\omega \epsilon_c \epsilon_0} \right)^{1/2} \approx \frac{\omega \sqrt{\epsilon_c \mu_c}}{c_0} \left(i \frac{\sigma}{\omega \epsilon_c \epsilon_0} \right)^{1/2} = \frac{1+i}{\sqrt{2}} \sqrt{\sigma \mu_0 \mu_c \omega}$$

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)	$\dagger^a$	$\dagger^a$
Saltwater	4	35.6	0.25 (m)	$\dagger^a$

^aAt this frequency the approximation $\sigma \gg \omega \epsilon_0 \epsilon$ is not valid and $\epsilon \approx 80$.

Table 6.5: 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 $\exp(ik\xi)$ is decomposed as

$$e^{ik\xi} = e^{i\xi/\delta} e^{-\xi/\delta}$$

where

$$\delta = \sqrt{\frac{2}{\omega \mu_0 \mu_c \sigma}} \quad (6.42)$$

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 6.5. 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 $\mathbf{E}_c$ and $\mathbf{H}_c$ as

$$\nabla \simeq \hat{\mathbf{n}}_c \frac{\partial}{\partial \xi} \quad (6.43)$$

Since the normal component of $\mathbf{H}$ at a perfectly conducting surface is zero, we can assume that the normal component of $\mathbf{H}_c$ 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 $\hat{\mathbf{z}}$ -direction, but no component along the ξ -direction. We will soon see that this is in accordance with the approximation in (6.43).

When we utilize the approximation (6.43), the Maxwell equations are simplified to

$$\begin{cases} \mathbf{H}_c = -\frac{i}{\omega \mu_0 \mu_c} \hat{\mathbf{n}}_c \times \frac{\partial \mathbf{E}_c}{\partial \xi} \\ \mathbf{E}_c = \frac{1}{\sigma} \hat{\mathbf{n}}_c \times \frac{\partial \mathbf{H}_c}{\partial \xi} \end{cases} \quad (6.44)$$

Note that the displacement current $-i\omega \epsilon_0 \epsilon_c \mathbf{E}_c$ is negligible compared to the current density $\sigma \mathbf{E}_c$ in Ampère's law, since $\sigma \gg \omega \epsilon_0 \epsilon_c$. We eliminate the electric field from the equations

by operating with $\hat{\mathbf{n}}_c \times \frac{\partial}{\partial \xi}$ on the lower of the two equations in (6.44)

$$\hat{\mathbf{n}}_c \times \frac{\partial \mathbf{E}_c}{\partial \xi} = \frac{1}{\sigma} \hat{\mathbf{n}}_c \times \left(\hat{\mathbf{n}}_c \times \frac{\partial^2 \mathbf{H}_c}{\partial \xi^2} \right) = \frac{1}{\sigma} \left(\hat{\mathbf{n}}_c \left(\hat{\mathbf{n}}_c \cdot \frac{\partial^2 \mathbf{H}_c}{\partial \xi^2} \right) - \frac{\partial^2 \mathbf{H}_c}{\partial \xi^2} \right)$$

When we insert this into Faraday's law we get

$$\mathbf{H}_c = -\frac{i}{\omega \sigma \mu_0 \mu_c} \left(\hat{\mathbf{n}}_c \left(\hat{\mathbf{n}}_c \cdot \frac{\partial^2 \mathbf{H}_c}{\partial \xi^2} \right) - \frac{\partial^2 \mathbf{H}_c}{\partial \xi^2} \right)$$

The normal component of this equation gives $\hat{\mathbf{n}}_c \cdot \mathbf{H}_c = 0$, which we anticipated earlier. Since $\hat{\mathbf{n}}_c \cdot \frac{\partial^2 \mathbf{H}_c}{\partial \xi^2} = \frac{\partial^2}{\partial \xi^2} (\hat{\mathbf{n}}_c \cdot \mathbf{H}_c) = 0$ we obtain the equation

$$\frac{\partial^2 \mathbf{H}_c}{\partial \xi^2} + i\omega \mu_0 \mu_c \sigma \mathbf{H}_c = 0$$

with solution

$$\mathbf{H}_c = \mathbf{H}_{\parallel} e^{-\xi/\delta} e^{i\xi/\delta} \quad (6.45)$$

where δ is the skin depth of the metal, see (6.42), and $\mathbf{H}_{\parallel}$ is the tangential component of the magnetic field at the surface, which can be decomposed in components along the directions $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{z}}$. Under the approximations established in this section, the amplitude of these tangential components matches that of a perfectly conducting surface. The corresponding electric field is obtained by inserting (6.45) into Ampère's law in (6.44).

$$\mathbf{E}_c \simeq \frac{i-1}{\sigma \delta} (\hat{\mathbf{n}}_c \times \mathbf{H}_{\parallel}) e^{-\xi/\delta} e^{i\xi/\delta}$$

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 (2.18) in Section 2.1

$$\mathbf{E} - \hat{\mathbf{n}}(\mathbf{E} \cdot \hat{\mathbf{n}}) = \eta_s \hat{\mathbf{n}} \times \mathbf{H} \quad (6.46)$$

where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic fields at the surface, $\eta_s = \frac{1-i}{\sigma \delta}$ is the impedance of the metal, and $\hat{\mathbf{n}} = -\hat{\mathbf{n}}_c$ is the normal unit vector, directed out from the metal.

We can now determine the ohmic losses in the metal walls. The time-averaged power loss density within the conductor is expressed as:

$$p = \frac{1}{2} \operatorname{Re} \{ \mathbf{J} \cdot \mathbf{E}_c^* \} = \frac{\sigma}{2} \mathbf{E}_c \cdot \mathbf{E}_c^* = \frac{1}{\sigma \delta^2} |\mathbf{H}_{\parallel}|^2 e^{-2\xi/\delta}$$

Integrating this density along the depth variable ξ yields the power loss per unit area:

$$\frac{dP_c}{da} = \int_0^\infty p(\xi) d\xi = \frac{1}{2\sigma \delta} |\mathbf{H}_{\parallel}|^2 = \frac{1}{2} R_s |\mathbf{H}_{\parallel}|^2 \quad (6.47)$$

where R_s is the surface resistance, defined as:

$$R_s = \frac{1}{\sigma \delta} \quad (6.48)$$

c.f., (4.44) in Chapter 4. Although metallic walls have finite thickness, setting the upper integration limit to infinity is mathematically convenient and physically sound. Because the skin depth δ is extremely small relative to the wall thickness, the integrand decays rapidly, making contributions beyond a few skin depths negligible.

Finally, assuming the tangential magnetic field $\mathbf{H}_{\parallel}$ near the boundary closely matches that of a perfect electric conductor (PEC), we can relate it to the surface current density $\mathbf{J}_S$:

$$\mathbf{H}_{\parallel} = -\hat{\mathbf{n}}_c \times \mathbf{J}_S$$

Substituting this relation into (6.47) yields the surface power dissipation in terms of the surface current:

$$\frac{dP_c}{da} = \frac{1}{2} R_s |\mathbf{J}_S|^2 \quad (6.49)$$

We now address the main objective of this section: calculating wave attenuation in a waveguide. While the waveguide walls are highly conductive, they are not perfect. For simplicity, we assume the waveguide interior is filled with a lossless medium characterized by real parameters ϵ and μ . Working above the cut-off frequency, we apply a first-order approximation to determine mode attenuation caused by wall losses. We consider one mode at a time, assuming propagation solely in the positive z -direction.

The E_z and H_z components of the mode are given by

$$\begin{cases} E_z(\mathbf{r}) = a_n^+ v_n(\boldsymbol{\rho}) e^{ik_{zn}z} & \text{TM} \\ H_z(\mathbf{r}) = b_n^+ w_n(\boldsymbol{\rho}) e^{ik_{zn}z} & \text{TE} \end{cases}$$

where the eigenfunctions v_n and w_n are normalized according to (6.26). The corresponding tangential components of the magnetic fields are given by (6.13)

$$|\mathbf{H}_{\parallel T}|^2 = \begin{cases} |a_n^+|^2 \left(\frac{\omega \epsilon_0 \epsilon}{k_{tn}^2} \right)^2 |\hat{\mathbf{z}} \times \nabla_T v_n|^2, & \text{TM} \\ |b_n^+|^2 \frac{k_{zn}^2}{k_{tn}^4} |\nabla_T w_n|^2, & \text{TE} \end{cases}$$

Thus the total magnetic field on the metal surface is given by

$$|\mathbf{H}_{\parallel}|^2 = |\mathbf{H}_{\parallel T}|^2 + |H_z|^2 = \begin{cases} |a_n^+|^2 \frac{(\omega \epsilon_0 \epsilon)^2}{k_{tn}^4} |\hat{\mathbf{z}} \times \nabla_T v_n|^2, & \text{TM} \\ |b_n^+|^2 \left(\frac{k_{zn}^2}{k_{tn}^4} |\nabla_T w_n|^2 + |w_n|^2 \right), & \text{TE} \end{cases} \quad (6.50)$$

According to (6.41), the time average of the power flow, P , in a lossless waveguide with perfectly conducting walls is

$$P = \iint_{\Omega} \mathbf{S} \cdot \hat{\mathbf{z}} \, dx dy = \begin{cases} |a_n^+|^2 \frac{\epsilon_0 \omega k_{zn}}{2 k_{tn}^2} \epsilon, & \text{TM} \\ |b_n^+|^2 \frac{\mu_0 \omega k_{zn}}{2 k_{tn}^2} \mu, & \text{TE} \end{cases} \quad (6.51)$$

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

$$dP = -dz \oint_{\Gamma} \frac{dP_c}{da} dl \quad (6.52)$$

We can express the power loss per unit area, dP_c/da , in the coefficients a_n^+ and b_n^+ by utilizing (6.47) and (6.50). We express a_n^+ and b_n^+ in terms of P by using equation (6.51) and get an equation for the power flow

$$\frac{dP}{dz} = -2\alpha P \quad (6.53)$$

with solution

$$P(z) = P(0)e^{-2\alpha z}$$

where

$$\alpha = \begin{cases} \frac{R_s \omega \epsilon \epsilon_0}{2k_{tn}^2 k_{zn}} \oint_{\Gamma} |\hat{\mathbf{z}} \times \nabla_T v_n|^2 dl, & \text{TM} \\ \frac{R_s k_{tn}^2}{2\omega \mu_0 \mu k_{zn}} \oint_{\Gamma} \left(\frac{k_{zn}^2}{k_{tn}^4} |\nabla_T w_n|^2 + |w_n|^2 \right) dl, & \text{TE} \end{cases}$$

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.

6.9.1 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 (6.49), (6.52) and (6.53)

$$\alpha = -\frac{1}{2P(z)} \frac{dP(z)}{dz} = \frac{1}{2P(z)} \oint_{\Gamma} \frac{dP_c(z)}{da} da = \frac{R_s}{4P(z)} \oint_{\Gamma} |J_s|^2 da \quad (6.54)$$

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.normJs²**. 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 (6.54).

6.9.2 Losses in waveguides with FEM: method 2

Another option to determine the attenuation in a waveguide with COMSOL is to use the impedance boundary condition (6.46). We then do the following steps:

- We choose the, **2D>Electromagnetic waves>Mode analysis**.
- 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 $\alpha = \omega \text{Im}\{n\}/c$.

Example 6.9

Applying impedance boundary conditions to a rectangular copper waveguide ($a = 0.3$ m, $b = 0.15$ m) operating in the TE_{10} mode at 704 MHz yields an effective refractive index of $n = 0.7045 + i1.74 \times 10^{-5}$, corresponding to an attenuation constant of $\alpha = 0.0002565$ Np/m. In the European Spallation Source (ESS) architecture, where the transmission distance between klystrons and cavities is approximately 20 m, power transmission efficiency is given by $e^{-2\alpha L} = 0.9898$. Consequently, roughly 1% of the transmitted RF power is dissipated. At an average delivered power of 5 MW, these waveguide losses amount to approximately 50 kW.

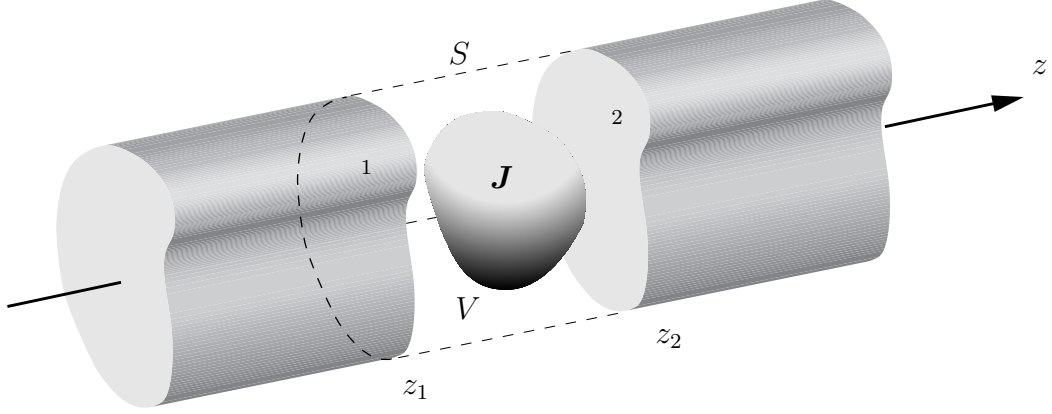


Figure 6.15: Waveguide with a source region V , confined by the cross-sections Ω_1 , Ω_2 and the lateral surface S .

Upsizing the waveguide cross-section to $0.35 \times 0.175 \text{ m}^2$ modifies the effective mode index to $n = 0.7937 + i 1.208 \times 10^{-5}$. This configuration improves the transmission factor to 0.9925, reducing the average RF power loss to 38 kW. ■

6.10 Sources in waveguides

In this section, we analyze the electromagnetic fields generated by known current densities inside a waveguide with a uniform cross-section and perfectly conducting walls. We assume the cross-sectional boundary is simply connected, excluding the presence of TEM modes. This formulation can be readily extended to include waveguides with non-perfectly conducting walls—using the impedance conditions established in the previous section—as well as structures that support TEM modes.

Consider a finite region of the waveguide bounded by cross-sections Ω_1 and Ω_2 (Figure 6.15), containing a known time-harmonic current density

$$\mathbf{J}(\mathbf{r}, t) = \text{Re } \mathbf{J}(\mathbf{r}, \omega) e^{-i\omega t}$$

Because the modal solutions developed in Section 6.4 form a complete functional basis in a source-free region, the total fields excited by the source can be expressed as a superposition of propagating and evanescent (non-propagating) modes. The vector expansion functions are defined in (6.27), with corresponding basis vectors given by (6.21), (6.22), and (6.24).

As in previous sections, we classify the waveguide modes using two indices: $\nu \in \{\text{TM}, \text{TE}\}$ to denote the mode type, and n to index individual modes within each type. The scalar eigenfunctions v_n and w_n satisfy the eigenvalue problems defined in (6.4) and (6.5), subject to the normalization conditions in (6.26). Correspondingly, the normalized vector mode functions fulfill the orthogonality relations in (6.37).

We let V be the volume between the cross-sections Ω_1 and Ω_2 at $z = z_1$ and $z = z_2$, where $z_1 < z_2$, see figure 6.15. 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

$$\begin{cases} \mathbf{E}^+(\mathbf{r}, \omega) = \sum_{\nu=\text{TM,TE}}^n a_{n\nu}^+ \mathbf{E}_{n\nu}^+(\mathbf{r}, \omega) \\ \mathbf{H}^+(\mathbf{r}, \omega) = \sum_{\nu=\text{TM,TE}}^n a_{n\nu}^+ \mathbf{H}_{n\nu}^+(\mathbf{r}, \omega) \end{cases} \quad z \geq z_2 \quad (6.55)$$

Analogously the fields are propagating in the negative z -direction when $z \leq z_1$ and then

$$\begin{cases} \mathbf{E}^-(\mathbf{r}, \omega) = \sum_{\nu=\text{TM,TE}}^n a_{n\nu}^- \mathbf{E}_{n\nu}^-(\mathbf{r}, \omega) \\ \mathbf{H}^-(\mathbf{r}, \omega) = \sum_{\nu=\text{TM,TE}}^n a_{n\nu}^- \mathbf{H}_{n\nu}^-(\mathbf{r}, \omega) \end{cases} \quad z \leq z_1 \quad (6.56)$$

To determine the coefficients $a_{n\nu}^\pm$ we use the Lorentz' reciprocity theorem, which we first derive. Let $\mathbf{E}$ and $\mathbf{H}$ be the fields generated by the current density $\mathbf{J}$, i.e., $\mathbf{E}$ and $\mathbf{H}$ satisfy

$$\begin{cases} \nabla \times \mathbf{E} = i\omega\mu_0\mu\mathbf{H} \\ \nabla \times \mathbf{H} = \mathbf{J} - i\omega\epsilon_0\epsilon\mathbf{E} \end{cases} \quad (6.57)$$

where the electric field $\mathbf{E}$ satisfies the boundary condition $\hat{\mathbf{n}} \times \mathbf{E} = \mathbf{0}$ on the perfectly conducting surface of the waveguide. By definition the vector fields $\mathbf{E}_{n\nu}^\pm$, $\mathbf{H}_{n\nu}^\pm$ for the modes satisfy the homogenous Maxwell equations (no current densities), i.e.,

$$\begin{cases} \nabla \times \mathbf{E}_{n\nu}^\pm = i\omega\mu_0\mu\mathbf{H}_{n\nu}^\pm \\ \nabla \times \mathbf{H}_{n\nu}^\pm = -i\omega\epsilon_0\epsilon\mathbf{E}_{n\nu}^\pm \end{cases} \quad (6.58)$$

and $\hat{\mathbf{n}} \times \mathbf{E}_{n\nu}^\pm = 0$ on the perfectly conducting surfaces.

The following identity is obtained from the Maxwell equations and vector formulas for the nabla-operator

$$\begin{aligned} \nabla \cdot (\mathbf{E}_{n\nu}^\pm \times \mathbf{H} - \mathbf{E} \times \mathbf{H}_{n\nu}^\pm) &= \mathbf{H} \cdot (\nabla \times \mathbf{E}_{n\nu}^\pm) - \mathbf{E}_{n\nu}^\pm \cdot (\nabla \times \mathbf{H}) \\ &\quad - \mathbf{H}_{n\nu}^\pm \cdot (\nabla \times \mathbf{E}) + \mathbf{E} \cdot (\nabla \times \mathbf{H}_{n\nu}^\pm) \end{aligned}$$

With (6.57) and (6.58) we obtain

$$\nabla \cdot (\mathbf{E}_{n\nu}^\pm \times \mathbf{H} - \mathbf{E} \times \mathbf{H}_{n\nu}^\pm) = -\mathbf{J} \cdot \mathbf{E}_{n\nu}^\pm$$

Integrating this relation over the volume V and applying the divergence theorem yields

$$\iint_{S_0} (\mathbf{E}_{n\nu}^\pm \times \mathbf{H} - \mathbf{E} \times \mathbf{H}_{n\nu}^\pm) \cdot \hat{\mathbf{n}} \, dS = - \iiint_V \mathbf{J} \cdot \mathbf{E}_{n\nu}^\pm \, dV \quad (6.59)$$

Here, S_0 denotes the boundary surface enclosing V , comprising the two transverse cross-sections Ω_1 and Ω_2 alongside the segment of the waveguide boundary S spanning $z = z_1$ to $z = z_2$ (Figure 6.15); $\hat{\mathbf{n}}$ is the outward unit normal to S_0 .

On the perfectly conducting side wall S , the tangential electric field vanishes ($\hat{\mathbf{n}} \times \mathbf{E}_{n\nu}^\pm = \hat{\mathbf{n}} \times \mathbf{E} = \mathbf{0}$), causing the side-wall contribution to the surface integral to drop out identically.

The remaining surface integrals over Ω_1 and Ω_2 are then evaluated by substituting the mode expansions (6.55) and (6.56) and invoking the orthogonality relations (6.37).

For the field $\mathbf{E}_{n\nu}^+$ on the surface Ω_1 we use the expansions in (6.56), (6.21) and (6.22). The orthogonality relation (6.39) gives

$$\begin{aligned}
& \iint_{\Omega_1} (\mathbf{E}_{n\nu}^+ \times \mathbf{H} - \mathbf{E} \times \mathbf{H}_{n\nu}^+) \cdot \hat{\mathbf{n}} \, dS \\
&= - \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} a_{n'\nu'}^- \iint_{\Omega_1} (\mathbf{E}_{n\nu}^+ \times \mathbf{H}_{n'\nu'}^- - \mathbf{E}_{n'\nu'}^- \times \mathbf{H}_{n\nu}^+) \cdot \hat{\mathbf{z}} \, dS \\
&= \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} a_{n'\nu'}^- e^{i(k_{zn} - k_{zn'})z_1} \iint_{\Omega_1} (\mathbf{E}_{Tn\nu} \times \mathbf{H}_{Tn'\nu'} + \mathbf{E}_{Tn'\nu'} \times \mathbf{H}_{Tn\nu}) \cdot \hat{\mathbf{z}} \, dS \\
&= -4a_{n\nu}^- U_{n\nu}
\end{aligned}$$

Note that on the surface Ω_1 , the outward directed normal is $-\hat{\mathbf{z}}$. On the surface Ω_2 we use the expansion (6.55) and get

$$\begin{aligned}
& \iint_{\Omega_2} (\mathbf{E}_{n\nu}^+ \times \mathbf{H} - \mathbf{E} \times \mathbf{H}_{n\nu}^+) \cdot \hat{\mathbf{n}} \, dS \\
&= \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} a_{n'\nu'}^+ \iint_{\Omega_2} (\mathbf{E}_{n\nu}^+ \times \mathbf{H}_{n'\nu'}^+ - \mathbf{E}_{n'\nu'}^+ \times \mathbf{H}_{n\nu}^+) \cdot \hat{\mathbf{z}} \, dS \\
&= \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} a_{n'\nu'}^+ e^{i(k_{zn} + k_{zn'})z_2} \iint_{\Omega_2} (\mathbf{E}_{Tn\nu} \times \mathbf{H}_{Tn'\nu'} - \mathbf{E}_{Tn'\nu'} \times \mathbf{H}_{Tn\nu}) \cdot \hat{\mathbf{z}} \, dS = 0
\end{aligned}$$

By using (6.59) we obtain the expression

$$a_{n\nu}^- = \frac{1}{4U_{n\nu}} \iiint_V \mathbf{J} \cdot \mathbf{E}_{n\nu}^+ \, dv$$

where $U_{n\nu}$ is given by (6.40). In the same manner we get the coefficients for $\mathbf{E}_{n\nu}^+$

$$\begin{aligned}
& \iint_{\Omega_1} (\mathbf{E}_{n\nu}^- \times \mathbf{H} - \mathbf{E} \times \mathbf{H}_{n\nu}^-) \cdot \hat{\mathbf{n}} \, dS \\
&= - \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} a_{n'\nu'}^- \iint_{\Omega_1} (\mathbf{E}_{n\nu}^- \times \mathbf{H}_{n'\nu'}^- - \mathbf{E}_{n'\nu'}^- \times \mathbf{H}_{n\nu}^-) \cdot \hat{\mathbf{z}} \, dS \\
&= \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} a_{n'\nu'}^- e^{-i(k_{zn} + k_{zn'})z_1} \iint_{\Omega_1} (\mathbf{E}_{Tn\nu} \times \mathbf{H}_{Tn'\nu'} - \mathbf{E}_{Tn'\nu'} \times \mathbf{H}_{Tn\nu}) \cdot \hat{\mathbf{z}} \, dS = 0
\end{aligned}$$

and

$$\begin{aligned}
& \iint_{\Omega_2} (\mathbf{E}_{n\nu}^- \times \mathbf{H} - \mathbf{E} \times \mathbf{H}_{n\nu}^-) \cdot \hat{\mathbf{n}} \, dS \\
&= \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} a_{n'\nu'}^+ \iint_{\Omega_2} (\mathbf{E}_{n\nu}^- \times \mathbf{H}_{n'\nu'}^+ - \mathbf{E}_{n'\nu'}^+ \times \mathbf{H}_{n\nu}^-) \cdot \hat{\mathbf{z}} \, dS \\
&= \sum_{\substack{n' \\ \nu'=\text{TM,TE}}} a_{n'\nu'}^+ e^{-i(k_{zn} - k_{zn'})z_2} \iint_{\Omega_2} (\mathbf{E}_{Tn\nu} \times \mathbf{H}_{Tn'\nu'} + \mathbf{E}_{Tn'\nu'} \times \mathbf{H}_{Tn\nu}) \cdot \hat{\mathbf{z}} \, dS \\
&= -4a_{n\nu}^+ U_{n\nu}
\end{aligned}$$

This leads to

$$a_{n\nu}^+ = \frac{1}{4U_{n\nu}} \iiint_V \mathbf{J} \cdot \mathbf{E}_{n\nu}^- \, dv$$

The final result is:

$$a_{n\nu}^\pm = \frac{1}{4U_{n\nu}} \iiint_V \mathbf{J} \cdot \mathbf{E}_{n\nu}^\mp \, dv \quad (6.60)$$

Example 6.10

Consider a current density $\mathbf{J}$ that is confined to a thin wire along the curve C with tangential unit vector $\hat{\boldsymbol{\tau}}$. The volume integrals in (6.60) are reduced to line integrals.

$$a_{n\nu}^\pm = \frac{1}{4U_{n\nu}} \int_C I(\boldsymbol{\rho}) \mathbf{E}_{n\nu}^\mp \cdot d\mathbf{r} \quad (6.61)$$

where $d\mathbf{r} = \hat{\boldsymbol{\tau}} dl$. ■

Example 6.11

If the curve C is in a transverse plane $z = z_0$ we see from (6.61), (6.21), (6.22) and (6.23) that $a_{n\nu}^+ = a_{n\nu}^- \exp(-2ik_z z_0)$. ■

Example 6.12

A small, planar, closed wire with $\mathbf{r}$ -independent time-harmonic current I has its center in $\mathbf{r}_0$. The wire is represented by a magnetic elementary dipole $\mathbf{m} = IA\hat{\mathbf{n}}$, where A is the area of the planar surface spanned by the wire and $\hat{\mathbf{n}}$ is the unit normal to A directed according to the right hand rule. Since the loop is small $\mathbf{H}_{n\nu}^\pm(\mathbf{r}) \simeq \mathbf{H}_{n\nu}^\pm(\mathbf{r}_0)$ in A . From (6.61), Stoke's theorem, and the induction law we get

$$\begin{aligned}
a_{n\nu}^\pm &= \frac{I}{4U_{n\nu}} \oint_C \mathbf{E}_{n\nu}^\mp \cdot d\mathbf{r} = \frac{I}{4U_{n\nu}} \iint_\Omega (\nabla \times \mathbf{E}_{n\nu}^\mp) \cdot \hat{\mathbf{n}} \, dS \\
&\simeq \frac{I}{4U_{n\nu}} i\omega\mu_0\mu \mathbf{H}_{n\nu}^\mp(\mathbf{r}_0) \cdot \iint_\Omega \hat{\mathbf{n}} \, dS = \frac{i\omega\mu_0\mu}{4U_{n\nu}} \mathbf{H}_{n\nu}^\mp(\mathbf{r}_0) \cdot \mathbf{m}
\end{aligned}$$

■

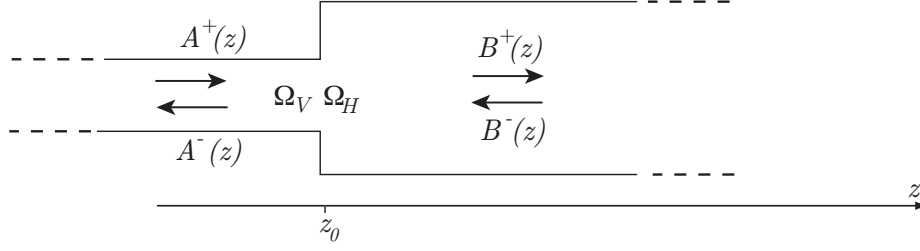


Figure 6.16: The geometry for the Mode-Matching Technique.

Example 6.13

Consider an electric elementary dipole, $\mathbf{p} = iI \, d\mathbf{r}/\omega$ where the vector $d\mathbf{r}$ is much smaller than the wavelength and the dimensions of the waveguide. When this dipole is placed in a waveguide we see from (6.61) that

$$a_{n\nu}^{\pm} = -\frac{i\omega}{4U_{n\nu}} \mathbf{p} \cdot \mathbf{E}_{n\nu}^{\mp}$$

Assume that the dipole is placed in the point $\mathbf{r}_0$. The field from the dipole is given by

$$\mathbf{E}(\mathbf{r}) = \begin{cases} -\frac{i\omega}{4} \sum_n \frac{1}{U_{n\nu}} (\mathbf{p} \cdot \mathbf{E}_{n\nu}^-(\mathbf{r}_0)) \mathbf{E}_{n\nu}^+(\mathbf{r}), & z > z_0 \\ -\frac{i\omega}{4} \sum_n \frac{1}{U_{n\nu}} (\mathbf{p} \cdot \mathbf{E}_{n\nu}^+(\mathbf{r}_0)) \mathbf{E}_{n\nu}^-(\mathbf{r}), & z < z_0 \end{cases}$$

This is rewritten as

$$\mathbf{E}(\mathbf{r}) = \frac{i\omega}{4} \mathbf{p} \cdot \mathcal{G}(\mathbf{r}, \mathbf{r}_0)$$

The function $\mathcal{G}$ is given by

$$\mathcal{G}(\mathbf{r}, \mathbf{r}_0) = \begin{cases} -\sum_n \frac{1}{U_{n\nu}} \mathbf{E}_{n\nu}^-(\mathbf{r}_0) \mathbf{E}_{n\nu}^+(\mathbf{r}), & z > z_0 \\ -\sum_n \frac{1}{U_{n\nu}} \mathbf{E}_{n\nu}^+(\mathbf{r}_0) \mathbf{E}_{n\nu}^-(\mathbf{r}), & z < z_0 \end{cases}$$

and is called the Green dyadic, which is a vectorial analogue of the Green function for scalar fields. ■

6.11 Mode-Matching Technique

The Mode-Matching Technique (MMT) 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 6.16.

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 (6.27) on page 83,

$$\begin{cases} \mathbf{E}(\mathbf{r}) = \sum_{\nu=\text{TM,TE,TEM}}^n (a_{n\nu}^+ \mathbf{E}_{n\nu}^+(\mathbf{r}) + a_{n\nu}^- \mathbf{E}_{n\nu}^-(\mathbf{r})) \\ \mathbf{H}(\mathbf{r}) = \sum_{\nu=\text{TM,TE,TEM}}^n (a_{n\nu}^+ \mathbf{H}_{n\nu}^+(\mathbf{r}) + a_{n\nu}^- \mathbf{H}_{n\nu}^-(\mathbf{r})) \end{cases} \quad z < z_0$$

$$\begin{cases} \mathbf{E}(\mathbf{r}) = \sum_{\nu=\text{TM,TE,TEM}}^n (b_{n\nu}^+ \mathbf{E}_{n\nu}^+(\mathbf{r}) + b_{n\nu}^- \mathbf{E}_{n\nu}^-(\mathbf{r})) \\ \mathbf{H}(\mathbf{r}) = \sum_{\nu=\text{TM,TE,TEM}}^n (b_{n\nu}^+ \mathbf{H}_{n\nu}^+(\mathbf{r}) + b_{n\nu}^- \mathbf{H}_{n\nu}^-(\mathbf{r})) \end{cases} \quad z > z_0$$

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

$$\begin{cases} \mathbf{E}_T^V(\mathbf{r}) = \sum_{\nu=\text{TM,TE,TEM}}^n (a_{n\nu}^+ e^{ik_{zn}^V z} + a_{n\nu}^- e^{-ik_{zn}^V z}) \mathbf{E}_{Tn\nu}^V(\boldsymbol{\rho}) \\ \mathbf{H}_T^V(\mathbf{r}) = \sum_{\nu=\text{TM,TE,TEM}}^n (a_{n\nu}^+ e^{ik_{zn}^V z} - a_{n\nu}^- e^{-ik_{zn}^V z}) \mathbf{H}_{Tn\nu}^V(\boldsymbol{\rho}) \end{cases} \quad z < z_0$$

$$\begin{cases} \mathbf{E}_T^H(\mathbf{r}) = \sum_{\nu=\text{TM,TE,TEM}}^n (b_{n\nu}^+ e^{ik_{zn}^H z} + b_{n\nu}^- e^{-ik_{zn}^H z}) \mathbf{E}_{Tn\nu}^H(\boldsymbol{\rho}) \\ \mathbf{H}_T^H(\mathbf{r}) = \sum_{\nu=\text{TM,TE,TEM}}^n (b_{n\nu}^+ e^{ik_{zn}^H z} - b_{n\nu}^- e^{-ik_{zn}^H z}) \mathbf{H}_{Tn\nu}^H(\boldsymbol{\rho}) \end{cases} \quad z > z_0$$

Note that transverse mode functions are different in the two waveguides since the transverse wavenumber k_{tn} depends on the cross section of the waveguide. We denote the mode functions in the waveguides $\mathbf{E}_{Tn\nu}^V$ and $\mathbf{E}_{Tn\nu}^H$. The same indices are used for the mode functions for the magnetic field and for the longitudinal wavenumbers, i.e., $k_{zn}^{H,V}$. 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

$$\begin{aligned} \mathbf{E}_T^H(\boldsymbol{\rho}, z_0) &= \begin{cases} \mathbf{0}, & \boldsymbol{\rho} \in \Omega_H \text{ and } \boldsymbol{\rho} \notin \Omega_V \\ \mathbf{E}_T^V(\boldsymbol{\rho}, z_0), & \boldsymbol{\rho} \in \Omega_V \end{cases} \\ \mathbf{H}_T^H(\boldsymbol{\rho}, z_0) &= \mathbf{H}_T^V(\boldsymbol{\rho}, z_0), \quad \boldsymbol{\rho} \in \Omega_V \end{aligned} \quad (6.62)$$

We now evaluate the normal surface integral over Ω_H of the vector product between $\mathbf{H}_{Tn'\nu'}^{H*}$ and the upper equation in (6.62). We utilize the normalization integral (6.37) on page 95

$$\iint_{\Omega_H} \hat{\mathbf{z}} \cdot (\mathbf{E}_{Tn\nu}^H(\boldsymbol{\rho}) \times \mathbf{H}_{Tn'\nu'}^{H*}(\boldsymbol{\rho})) \, dS = 2P_{n\nu}^H \delta_{nn'} \delta_{\nu\nu'}$$

where the mode powers in the lossless waveguide are given by, see (6.38),

$$P_{n\nu}^{H,V} = \begin{cases} \frac{\epsilon_0 \epsilon \omega k_{zn}^{H,V}}{2 (k_{tn}^{H,V})^2}, & \nu = \text{TM} \\ \frac{\mu_0 \mu \omega k_{zn}^{H,V*}}{2 (k_{tn}^{H,V})^2}, & \nu = \text{TE} \\ \frac{1}{2\eta_0 \eta}, & \nu = \text{TEM} \end{cases} \quad (6.63)$$

The following relation is obtained

$$P(B^+(z_0) + B^-(z_0)) = Q^t(A^+(z_0) + A^-(z_0)) \quad (6.64)$$

where $A^\pm(z_0)$ and $B^\pm(z_0)$ are column vectors $A_{n\nu}^\pm(z) = a_{n\nu}^\pm e^{\pm i k_{zn}^V z}$ and $B_{n\nu}^\pm(z) = b_{n\nu}^\pm e^{\pm i k_{zn}^H z}$ where t denotes transpose. The matrix Q is given by

$$Q_{n\nu, n'\nu'} = \frac{1}{2} \iint_{\Omega_V} \hat{\mathbf{z}} \cdot (\mathbf{E}_{Tn\nu}^V(\boldsymbol{\rho}) \times \mathbf{H}_{Tn'\nu'}^{H*}(\boldsymbol{\rho})) \, dS$$

while P is the diagonal matrix

$$P_{n\nu, n'\nu'} = P_{n\nu}^H \delta_{nn'} \delta_{\nu\nu'}$$

Next we evaluate the normal surface integral over Ω_V of the vector product of $\mathbf{E}_{Tn'\nu'}^{V*}$ and the lower of the equations in (6.62) and get the relation

$$Q^*(B^+(z_0) - B^-(z_0)) = R^*(A^+(z_0) - A^-(z_0)) \quad (6.65)$$

where R is the diagonal matrix

$$R_{n\nu, n'\nu'} = P_{n\nu}^V \delta_{nn'} \delta_{\nu\nu'}$$

The mode power $P_{n\nu}^V$ is given by (6.63).

The system of equations (6.64) and (6.65) 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

$$\begin{cases} S_{11} = (Q^* P^{-1} Q^t + R^*)^{-1} (R^* - Q^* P^{-1} Q^t) \\ S_{12} = 2(Q^* P^{-1} Q^t + R^*)^{-1} Q^* \\ S_{21} = 2(Q^t R^{*-1} Q^* + P)^{-1} Q^t \\ S_{22} = (Q^t R^{*-1} Q^* + P)^{-1} (Q^t R^{*-1} Q^* - P) \end{cases}$$

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

$$\begin{pmatrix} A^-(z_0) \\ B^+(z_0) \end{pmatrix} = \begin{pmatrix} \tilde{S}_{11} & \tilde{S}_{12} \\ \tilde{S}_{21} & \tilde{S}_{22} \end{pmatrix} \begin{pmatrix} A^+(z_0) \\ B^-(z_0) \end{pmatrix}$$

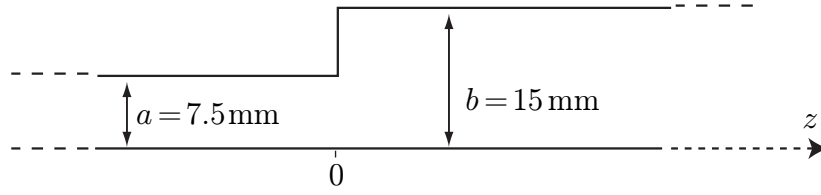


Figure 6.17: Transition between two planar waveguides.

where the scattering matrix is

$$\begin{cases} \tilde{S}_{11} = (R + \tilde{Q}^t P^{*-1} \tilde{Q}^*)^{-1} (\tilde{Q}^t P^{*-1} \tilde{Q}^* - R) \\ \tilde{S}_{12} = 2(R + \tilde{Q}^t P^{*-1} \tilde{Q}^*)^{-1} \tilde{Q}^t \\ \tilde{S}_{21} = 2(\tilde{Q}^* R^{-1} \tilde{Q}^t + P^*)^{-1} \tilde{Q}^* \\ \tilde{S}_{22} = (\tilde{Q}^* R^{-1} \tilde{Q}^t + P^*)^{-1} (P^* - \tilde{Q}^* R^{-1} \tilde{Q}^t) \end{cases}$$

and where

$$\tilde{Q}_{n\nu, n'\nu'} = \frac{1}{2} \iint_{\Omega_H} \hat{\mathbf{z}} \cdot (\mathbf{E}_{T n\nu}^H(\boldsymbol{\rho}) \times \mathbf{H}_{T n'\nu'}^{V*}(\boldsymbol{\rho})) \, dS$$

Example 6.14

Consider a transition between the two planar waveguides depicted in Figure 6.17. An incident TEM wave propagates in the positive z -direction and gives rise to a reflected and transmitted TEM-wave at the transition $z = 0$. The frequency is low enough such 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

$$\begin{aligned} R &= \frac{b - a}{b + a} \\ T &= \frac{2b}{b + a} \end{aligned} \tag{6.66}$$

These coefficients can be compared with the corresponding coefficients obtained from MMT. 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., (4.6.1). In Figure 6.18 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. This despite that the cut-off frequency for the next propagating modes TE_1 and TM_1 is 10 GHz. ■

Example 6.15

To analyze the transition in Figure 6.17 with COMSOL we choose **2D>Electromagnetic**

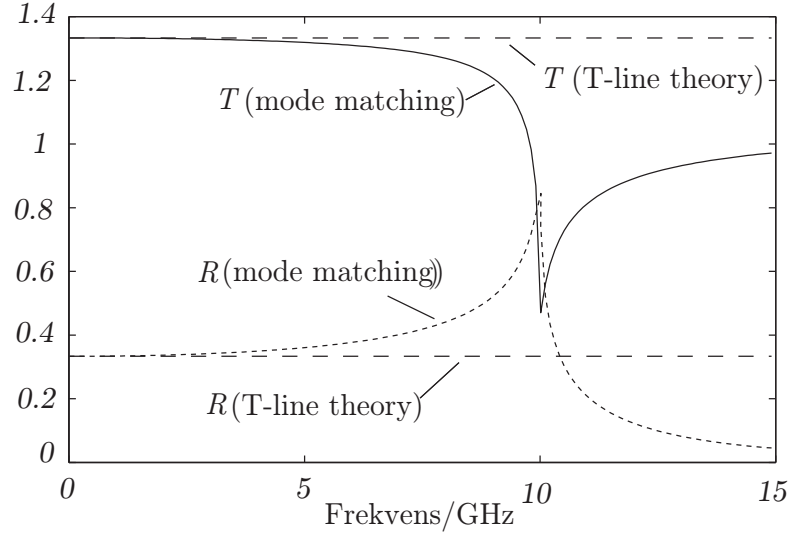


Figure 6.18: Reflection and transmission coefficients for the TEM mode.

waves>Frequency domain. We draw the geometry and define the boundary conditions. All of the surfaces are perfect conductors except the vertical surfaces to the left and right. 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 6.19 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. ■

While FEM can handle waveguide junctions and offers greater flexibility than MMT by avoiding reliance on analytical mode expressions, MMT remains far from redundant. In fact, MMT is significantly superior in several key scenarios. For instance, when junctions are separated by long distances, MMT proves far more efficient. Additionally, MMT inherently decomposes fields into modal sums—a task that is much less straightforward to achieve with FEM.

6.11.1 Cascading

A waveguide with several discontinuities can be treated by a cascading method. Assume the geometry depicted in figure 6.20. The relation between $A^\pm(z_0)$ and $B^\pm(z_0)$ can be written, see (6.64) and (6.65)

$$\begin{pmatrix} P_0 & P_0 \\ Q_0^* & -Q_0^* \end{pmatrix} \begin{pmatrix} B^+(z_0) \\ B^-(z_0) \end{pmatrix} = \begin{pmatrix} Q_0^t & Q_0^t \\ R_0^* & -R_0^* \end{pmatrix} \begin{pmatrix} A^+(z_0) \\ A^-(z_0) \end{pmatrix}$$

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

$$\begin{pmatrix} B^+(z_1) \\ B^-(z_1) \end{pmatrix} = \begin{pmatrix} E^+(z_1 - z_0) & 0 \\ 0 & E^-(z_1 - z_0) \end{pmatrix} \begin{pmatrix} B^+(z_0) \\ B^-(z_0) \end{pmatrix}$$

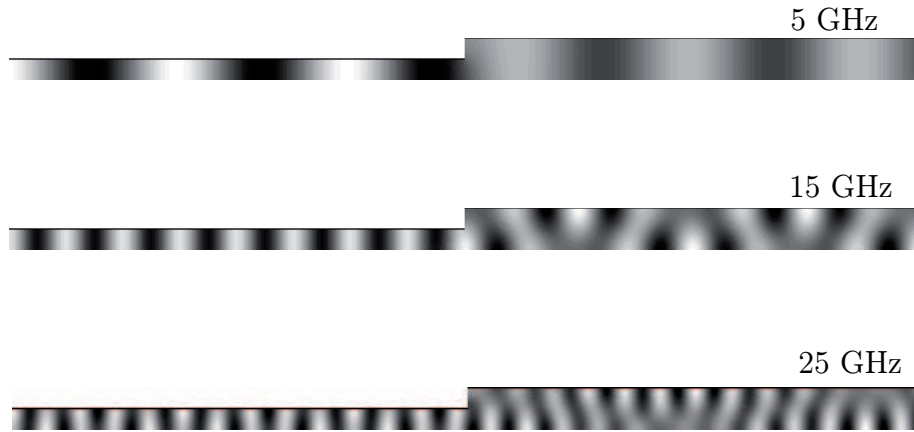


Figure 6.19: The magnetic field distribution for the junction in Figure 6.17 generated by COMSOL at 5, 15, and 25 GHz. The incident TEM-wave enters from the left. The total length of the waveguide is 32 cm. At 5 GHz only the TEM wave propagates. At 15 GHz only the TEM wave propagates in the left waveguide 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.

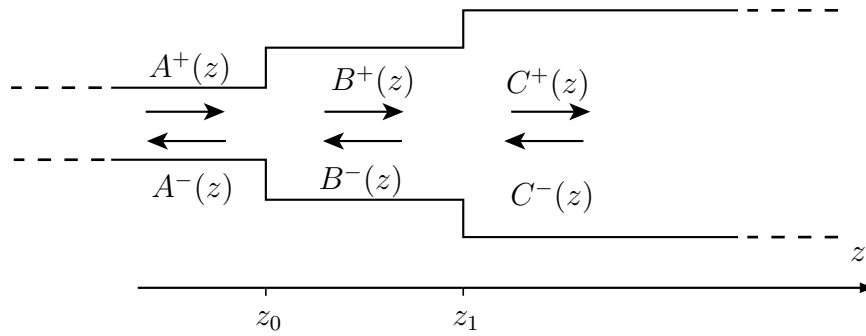


Figure 6.20: Three cascaded waveguides.

where $E_{nv,n'\nu'}^\pm(z) = \delta_{nn'}\delta_{\nu\nu'} \exp(\pm ik_{zn}z)$. 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.

6.11.2 Waveguides with bends

With FEM we can analyze waveguides that are not straight. In the example in Figure 6.21 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.

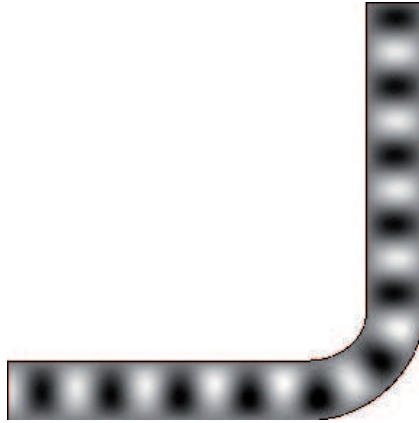


Figure 6.21: The electric field for a TE_{10} mode traveling in a waveguide with a bend. The field is calculated by COMSOL. The frequency is 3 GHz. The cross section of the waveguide in the straight sections is $1\text{cm} \times 0.5\text{ cm}$. The figure shows the xz -plane and the electric field is perpendicular to the plane.

6.12 Transmission lines in inhomogeneous media by FEM

In Chapter 4, we introduced methods for determining the line parameters of transmission lines embedded in media where the permittivity (ϵ) and conductivity (σ) are independent of the x and y coordinates. We now extend this analysis to transmission lines with spatially varying ϵ and σ by treating them as waveguides. Specifically, we first examine a thin microstrip above a substrate, followed by a microstrip with a thickness on the order of the skin depth.

Example 6.16

Consider first the the microstrip in Figure 6.22. 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 necessary, also the ground plane.
- We use impedance boundary condition for the boundaries of the microstrip. The skin depth is $\delta = \sqrt{2/(\omega\sigma\mu_0)} = 6.5\text{ }\mu\text{m}$, which is much smaller than the thickness of the strip, and then the impedance boundary condition is a good approximation.

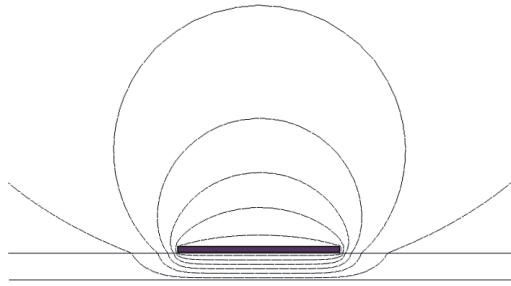


Figure 6.22: The voltage contour lines for the microstrip with substrate $\sigma = 0.01$ S/m and $\epsilon = 4$. Notice the discontinuity in the tangent of the lines at the surface of the substrate.

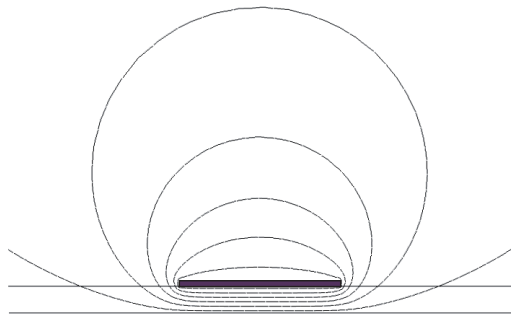


Figure 6.23: The voltage contour lines for the microstrip with vacuum substrate $\sigma = 0$ S/m and $\epsilon = 1$. The voltage contour lines are identical with the magnetic field lines.

- 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 effective 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 $n_{\text{eff}} = 1.8711 + j0.3806$ which corresponds to $\gamma = in_{\text{eff}}k_0 = 0.7977 + j3.922$.

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 integrating the electric field from the ground plane to the microstrip and I is obtained by integrating the z -directed surface current density around the outer

boundary 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 $Z_0 = 21.756 + j4.604$ V/A. We then notice that, c.f., (4.29) and (4.30),

$$\begin{aligned} Z_0 \gamma &= R + j\omega L \\ \frac{\gamma}{Z_0} &= G + j\omega C \end{aligned}$$

Thus

$$\begin{aligned} R &= \text{Real}\{Z_0 \gamma\} \\ L &= \frac{\text{Imag}\{Z_0 \gamma\}}{\omega} \\ G &= \text{Real}\{\gamma Z_0^{-1}\} \\ C &= \frac{\text{Imag}\{\gamma Z_0^{-1}\}}{\omega} \end{aligned} \tag{6.67}$$

We get the values $R = 0.006985$ Ω/m , $L = 141.6$ nH/m, $G = 0.0716$ S/m and $C = 262.8$ pF/m.

Alternatively, the line parameters can be extracted from energy calculations. The inductance per unit length L' is determined by first computing the time-averaged magnetic energy W_m , and then evaluating the total longitudinal current I via a boundary integral of the surface current density over the outer conductor. In COMSOL Multiphysics, W_m is a predefined global quantity obtained directly by integrating the magnetic energy density over the computational domain. Then

$$L = \frac{4W_m}{|I|^2}$$

This gives the value $L = 139.8$ nH/m.

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 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. ■

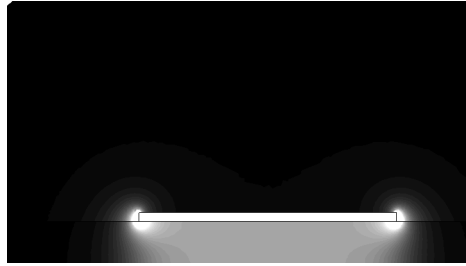


Figure 6.24: The power flow density for the microstrip.

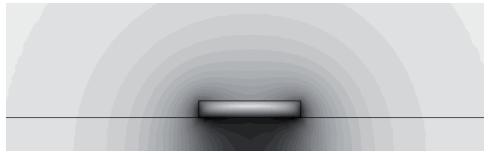


Figure 6.25: The absolute value of the magnetic field for the microstrip in example 17. Notice that the magnetic field penetrates the microstrip.

Example 6.17

Consider the microstrip in Figure 6.25. The dimensions of the microstrip is $3\text{ }\mu\text{m} \times 0.5\text{ }\mu\text{m}$ and it consists of a metal with conductivity $4 \cdot 10^7\text{ S/m}$. The substrate has relative permittivity $\epsilon = 4$, conductivity $\sigma_s = 0.01\text{ S/m}$, and thickness $1\text{ }\mu\text{m}$. 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 $60 \times 50\text{ }\mu\text{m}$ 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 $1 \times 0.5\text{ }\mu\text{m}$ (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 effective 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 microstrip. The effective mode index for that mode is $n_{\text{eff}} = 1.91 - i0.36$, which corresponds to the propagation constant $\gamma = -226 + i1202\text{ m}^{-1}$.

- It is possible to calculate the inductance, capacitance and resistance. We can either use $Z_0 = V/I$ and (6.67) or use

$$\begin{aligned} L &= \frac{4W_m}{|I|^2} \\ C &= \frac{4W_e}{|V|^2} \\ R &= \frac{2P_c}{|I|^2} \\ G &= \frac{2P_s}{|V|^2} \end{aligned}$$

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

$$\begin{aligned} L &= 232.5 \text{ nH/m} \\ C &= 159 \text{ pF/m} \\ R &= 17.1 \text{ k}\Omega/\text{m} \\ G &= 0.0395 \text{ S/m} \end{aligned}$$

The values have been checked by comparing with different meshes and sizes of the computational domain.

■

6.13 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. Figure 6.26 shows a COMSOL simulation of a mode propagating in a SIW having a substrate with $\epsilon = 5$. The frequency is 50 GHz and 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 $500 \mu\text{m}$ and the distance between the centers of two consecutive vias is $1500 \mu\text{m}$. The distance between the centers of two parallel vias is $3000 \mu\text{m}$. The waveguide is similar to a planar waveguide with width $2000 \mu\text{m}$. Such a waveguide has the cut-off frequency 33 GHz. At frequencies below cut-off frequency for the fundamental mode, as in Figure 6.27, no waves propagate. In Figure 6.29 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.

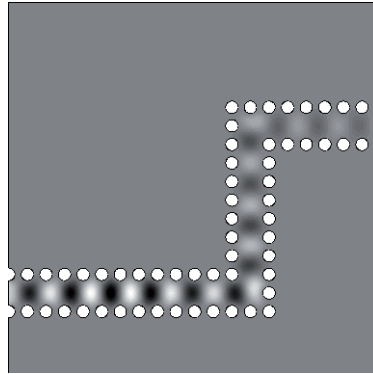


Figure 6.26: A COMSOL simulation of the SIW at 50 GHz. The circles are the cross sections of the circular metallic vias connecting two parallel conducting planes that are separated by a dielectric substrate. Only the fundamental mode is propagating.



Figure 6.27: A COMSOL simulation of the SIW at 30 GHz. No modes are propagating since the frequency is below the cut-off frequency for the fundamental mode.

Problems in Chapter 6

- 6.1** 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?
- 6.2** 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?
- *6.3** 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 6.30.
- 6.4** In a circular waveguide the fundamental mode is the TE_{11} modes. Its cut-off frequency is given by $f_{cTE_{11}} = 1.841c_0/2\pi a$, see table 6.4. By inserting a planar strip

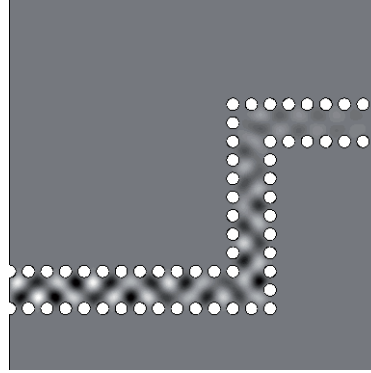


Figure 6.28: A COMSOL simulation of the SIW at 70 GHz. At this frequency both the fundamental and the second mode are propagating.

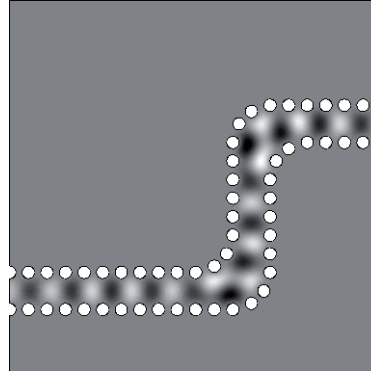


Figure 6.29: A COMSOL simulation of the SIW with smooth bends, at 50 GHz. The reflections at the corners are reduced as compared to the SIW in Figure 6.26.

in the waveguide the frequency of the fundamental mode can be reduced. Determine the modes for the geometry in figure 6.31, 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

$$J_{1/2}(z) = \sqrt{\frac{2}{\pi z}} \sin z, \quad J_{3/2}(z) = \sqrt{\frac{2z}{\pi}} (\sin z - z \cos z)$$

- 6.5** 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 of the total power flow in the waveguide for $z < 0$, for the incident and reflected modes, respectively

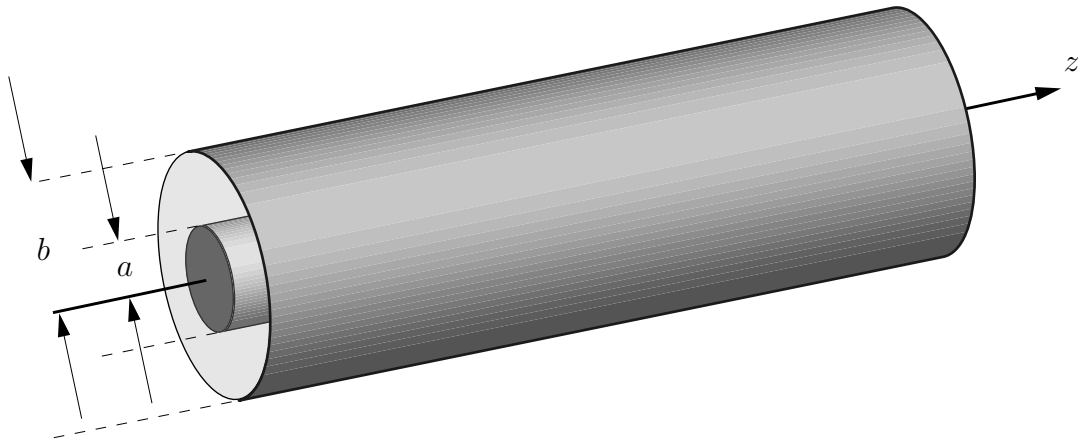


Figure 6.30: The geometry for the coaxial cable in problem 6.3.

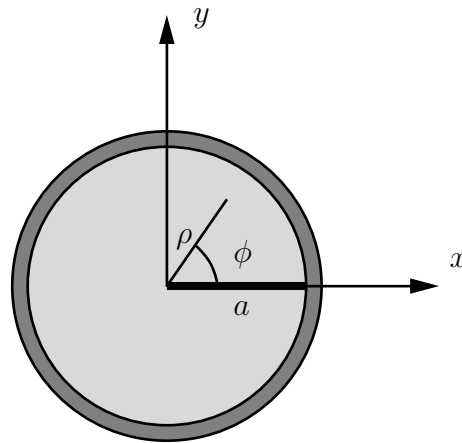


Figure 6.31: Geometry for problem 6.4. The circular waveguide with the metal strip.

- a) Determine the quotient P_r/P_i .
- b) Is there any frequency for which $P_r = 0$?

6.6 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.

6.7 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 6.32. The plate is 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}

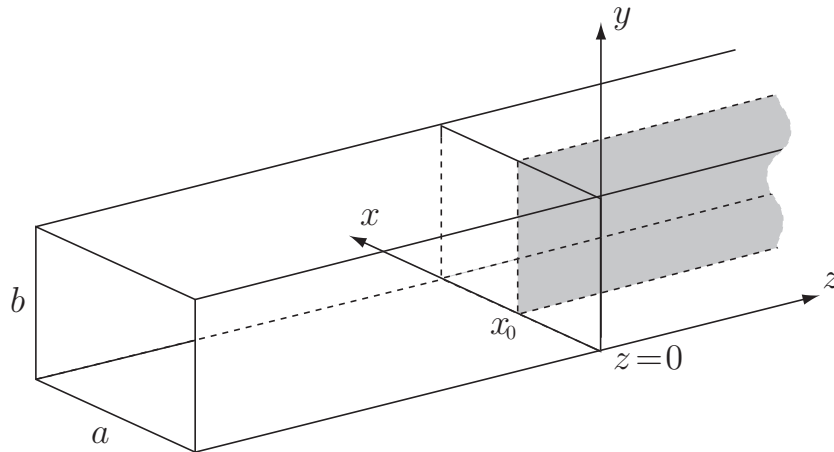


Figure 6.32: Geometry for problem 6.7.

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.

- Determine x_0 .
- 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?
- What is P_r/P_i when a TE_{30} mode propagates in positive z -direction for $z < 0$ at the frequency 10 GHz?

6.8 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.

6.9 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.

6.10 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.

- 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?
- 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?

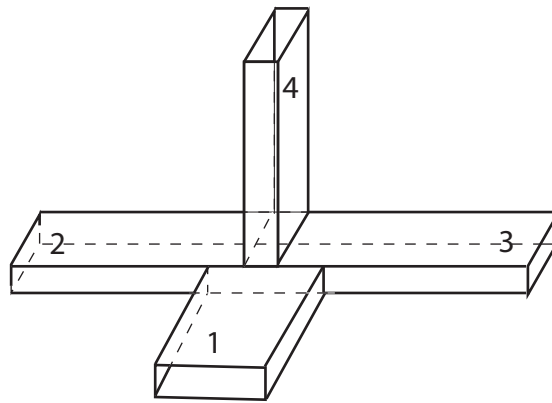


Figure 6.33: Geometry for problem 6.10.

Summary of chapter 6

Mode solutions—TM and TE modes

Eigenvalue problem

$$\boxed{\begin{cases} \nabla_T^2 v_n(\boldsymbol{\rho}) + k_{tn}^2 v_n(\boldsymbol{\rho}) = 0 & \boldsymbol{\rho} \in \Omega \\ v_n(\boldsymbol{\rho}) = 0 & \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \quad (\text{TM})}$$

$$\boxed{\begin{cases} \nabla_T^2 w_n(\boldsymbol{\rho}) + k_{tn}^2 w_n(\boldsymbol{\rho}) = 0 & \boldsymbol{\rho} \in \Omega \\ \frac{\partial w_n(\boldsymbol{\rho})}{\partial n} = 0 & \boldsymbol{\rho} \text{ on } \Gamma \end{cases} \quad (\text{TE})}$$

Normalization

$$\boxed{\begin{aligned} \iint_{\Omega} v_n(\boldsymbol{\rho}) v_{n'}(\boldsymbol{\rho}) \, dx dy &= \delta_{n,n'} \\ \iint_{\Omega} w_n(\boldsymbol{\rho}) w_{n'}(\boldsymbol{\rho}) \, dx dy &= \delta_{n,n'} \end{aligned}}$$

Eigenmodes

$$\boxed{\begin{cases} \mathbf{E}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \{\mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) \pm v_n(\boldsymbol{\rho}) \hat{\mathbf{z}}\} e^{\pm i k_{zn} z} \\ \mathbf{H}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \pm \mathbf{H}_{Tn\nu}(\boldsymbol{\rho}, \omega) e^{\pm i k_{zn} z} \end{cases} \quad \nu = \text{TM}}$$

$$\boxed{\begin{cases} \mathbf{E}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) e^{\pm i k_{zn} z} \\ \mathbf{H}_{n\nu}^{\pm}(\mathbf{r}, \omega) = \{\pm \mathbf{H}_{Tn\nu}(\boldsymbol{\rho}, \omega) + (\omega)_n(\boldsymbol{\rho}) \hat{\mathbf{z}}\} e^{\pm i k_{zn} z} \end{cases} \quad \nu = \text{TE}}$$

Transverse mode components

$$\begin{aligned}
\mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) &= \frac{i}{k_{tn}^2(\omega)} \begin{cases} k_{zn} \nabla_T v_n(\boldsymbol{\rho}), & \nu = \text{TM} \\ -\omega \mu_0 \mu(\omega) \hat{\mathbf{z}} \times \nabla_T w_n(\boldsymbol{\rho}), & \nu = \text{TE} \end{cases} \\
\mathbf{H}_{Tn\nu}(\boldsymbol{\rho}, \omega) &= \frac{i}{k_{tn}^2(\omega)} \begin{cases} \omega \epsilon_0 \epsilon(\omega) \hat{\mathbf{z}} \times \nabla_T v_n(\boldsymbol{\rho}), & \nu = \text{TM} \\ k_{zn} \nabla_T w_n(\boldsymbol{\rho}), & \nu = \text{TE} \end{cases}
\end{aligned}$$

Mode solutions—TEM mode**Potential problem**

$$\begin{aligned}
\nabla_T^2 \psi(\boldsymbol{\rho}) &= 0 \\
\psi(\boldsymbol{\rho}) &= \text{konstant} \quad \boldsymbol{\rho} \text{ on } \Gamma
\end{aligned}$$

Normalization

$$\iint_{\Omega} \nabla_T \psi(\boldsymbol{\rho}) \cdot \nabla_T \psi(\boldsymbol{\rho}) \, dx dy = 1$$

Eigenmodes

$$\begin{cases} \mathbf{E}_{\nu}^{\pm}(\mathbf{r}, \omega) = \mathbf{E}_{T\nu}(\boldsymbol{\rho}, \omega) e^{\pm i k z} \\ \mathbf{H}_{\nu}^{\pm}(\mathbf{r}, \omega) = \pm \mathbf{H}_{T\nu}(\boldsymbol{\rho}, \omega) e^{\pm i k z} \end{cases} \quad \nu = \text{TEM}$$

Transverse components

$$\begin{aligned}\mathbf{E}_{T\text{TEM}}(\boldsymbol{\rho}, \omega) &= -\nabla_T \psi(\boldsymbol{\rho}, \omega) \\ \eta_0 \mathbf{H}_{T\text{TEM}}(\boldsymbol{\rho}, \omega) &= \frac{k(\omega)c_0}{\omega\mu(\omega)} \hat{\mathbf{z}} \times \mathbf{E}_{T\text{TEM}}(\boldsymbol{\rho}, \omega)\end{aligned}$$

General expansions in modes

$$\left\{ \begin{aligned} \mathbf{E}(\mathbf{r}, \omega) &= \sum_{\nu=\text{TM}, \text{TE}, \text{TEM}}^n (a_{n\nu}^+ \mathbf{E}_{n\nu}^+(\mathbf{r}, \omega) + a_{n\nu}^- \mathbf{E}_{n\nu}^-(\mathbf{r}, \omega)) \\ \mathbf{H}(\mathbf{r}, \omega) &= \sum_{\nu=\text{TM}, \text{TE}, \text{TEM}}^n (a_{n\nu}^+ \mathbf{H}_{n\nu}^+(\mathbf{r}, \omega) + a_{n\nu}^- \mathbf{H}_{n\nu}^-(\mathbf{r}, \omega)) \end{aligned} \right\} \begin{cases} z \in [z_1, z_2] \\ \boldsymbol{\rho} \in \Omega \end{cases}$$

Orthogonality relation

$$\iint_{\Omega} \hat{\mathbf{z}} \cdot \{ \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}, \omega) \times \mathbf{H}_{Tn'\nu'}^*(\boldsymbol{\rho}, \omega) \} \, dS = 2P_{n\nu}^E \delta_{n,n'} \delta_{\nu,\nu'}$$

Mode power

$$P_{n\nu}^E = \begin{cases} \frac{\omega}{2k_{tn}^2} \begin{cases} k_{zn} \epsilon_0 \epsilon^*(\omega), & \nu = \text{TM} \\ k_{zn}^*(\omega) \mu_0 \mu(\omega), & \nu = \text{TE} \end{cases} \\ \frac{1}{2\eta_0 \eta^*}, & \nu = \text{TEM} \end{cases}$$

Power flow

General expressions

$$\iint_{\Omega} \hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) \, dx dy$$

$$= \sum_{\nu=\text{TM,TE}}^n \text{Re } P_{n\nu}^E \left\{ |a_{n\nu}^+|^2 e^{-2 \text{Im } k_{zn} z} - |a_{n\nu}^-|^2 e^{2 \text{Im } k_{zn} z} + 2i \text{Im} \left(a_{n\nu}^- a_{n\nu}^{+*} e^{-2i \text{Re } k_{zn} z} \right) \right\}$$

Propagating modes

$$\iint_{\Omega} \hat{\mathbf{z}} \cdot \langle \mathbf{S}(t) \rangle(\mathbf{r}, \omega) \, dx dy = \pm \sum_{k_{tn} < k(\omega)} P_{n\nu}^E |a_{n\nu}^{\pm}|^2$$

Losses in walls

$$P(z) = P(0)e^{-\alpha z}$$

$$\alpha = \begin{cases} \frac{\omega \epsilon \epsilon_0}{\sigma \delta k_{tn}^2 k_{zn}} \oint_{\Gamma} |\hat{\mathbf{z}} \times \nabla_T v_n|^2 \, dl, & \text{TM} \\ \frac{k_{tn}^2}{\omega \sigma \delta \mu_0 \mu k_{zn}} \oint_{\Gamma} \left(\frac{k_{zn}^2}{k_{tn}^4} |\nabla_T w_n|^2 + |w_n|^2 \right) \, dl, & \text{TE} \end{cases}$$

Sources and determination of expansion coefficients

$$a_{n\nu}^{\pm} = \frac{1}{4U_{n\nu}} \iiint_V \mathbf{J} \cdot \mathbf{E}_{n\nu}^{\mp} \, dv$$

Chapter 7

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.

7.1 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 7.1 and 7.2.

7.1.1 The resonances in a lossless cavity

Assume a volume V with a homogeneous lossless material bounded by a perfectly conducting surface S with inward pointing normal $\hat{\mathbf{n}}$. 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 $\hat{\mathbf{n}} \times \mathbf{E}(\omega, \mathbf{r}) = \mathbf{0}$. The eigenvalue problem for the electric field is

$$\begin{cases} \nabla^2 \mathbf{E}(\mathbf{r}) + k^2 \mathbf{E}(\mathbf{r}) = \mathbf{0} & \mathbf{r} \in V \\ \hat{\mathbf{n}} \times \mathbf{E}(\mathbf{r}) = \mathbf{0} & \mathbf{r} \in S \end{cases}$$

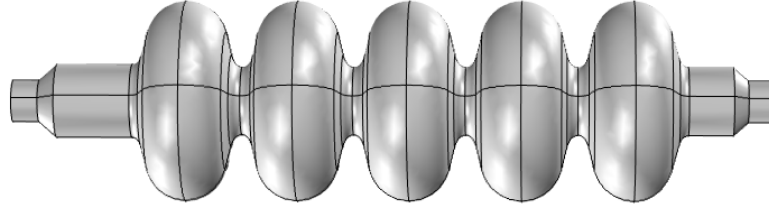


Figure 7.1: The elliptic cavity to be used in the LINAC of ESS

A general vector field $\mathbf{E}$ can according to Helmholtz theorem be written as a sum of a solenoidal field $\nabla \times \mathbf{N}$ and an irrotational field $-\nabla\Phi$ such that

$$\mathbf{E}(\mathbf{r}) = \nabla \times \mathbf{N}(\mathbf{r}) - \nabla\Phi(\mathbf{r})$$

In order to express the electric field from a source in a cavity both solenoidal and irrotational vector functions are needed. The solenoidal functions satisfy the eigenvalue problem

$$\begin{cases} \nabla^2 \mathbf{E}_n + k_n^2 \mathbf{E}_n = \mathbf{0} & \text{in } V \\ \hat{\mathbf{n}} \times \mathbf{E}_n = \mathbf{0} & \text{on } S \\ \nabla \cdot \mathbf{E}_n = 0 & \text{in } V \text{ and on } S \end{cases} \quad (7.1)$$

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 are orthogonal and can be chosen to be real. Assuming real eigenfunctions, we normalize them according to

$$\iiint_V \mathbf{E}_m \cdot \mathbf{E}_n \, dV = \delta_{n,m} \quad (7.2)$$

for all m and n . Applying Gauss's theorem to the vector identity

$$\nabla \cdot (\mathbf{E}_n \times (\nabla \times \mathbf{E}_m)) = (\nabla \times \mathbf{E}_m) \cdot (\nabla \times \mathbf{E}_n) - k_m^2 \mathbf{E}_m \cdot \mathbf{E}_n$$

yields

$$\iiint_V (\nabla \times \mathbf{E}_m) \cdot (\nabla \times \mathbf{E}_n) \, dV = k_m^2 \iiint_V \mathbf{E}_m \cdot \mathbf{E}_n \, dV = k_m^2 \delta_{n,m} \quad (7.3)$$

The set of solenoidal vector eigenfunctions constitute a complete set of vector functions for the volume V . It means that any solenoidal vector field $\mathbf{A}(\mathbf{r}, t)$ that satisfies $\hat{\mathbf{n}} \times \mathbf{A}(\mathbf{r}, t) = \mathbf{0}$ on S can be expanded in a series of the solenoidal vector eigenfunctions. In a source free cavity $\nabla \cdot \mathbf{E} = 0$ and hence the irrotational part of $\mathbf{E}$ is zero and we expand the electric field in a series of solenoidal eigenfunctions.

$$\mathbf{E}(\mathbf{r}, t) = \sum_{n=0}^{\infty} e_n(t) \mathbf{E}_n(\mathbf{r}) \quad (7.4)$$

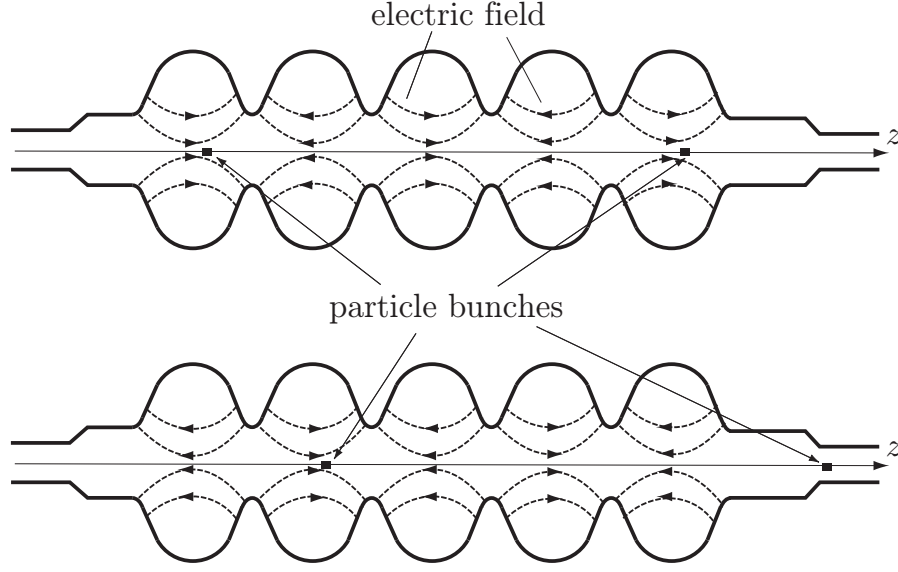


Figure 7.2: The cross section of the elliptic cavity and the electric field. The bunches are traveling to the right. The frequency of the fundamental mode is 704 MHz and the bunches come with a frequency of 352MHz. The lower figure is at a time half a period later than the upper one.

From the Maxwell equations it is seen that the electric field satisfies the source free wave equation

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

We insert the series expansion into this equation

$$\sum_{m=0}^{\infty} k_m^2 e_m(t) \mathbf{E}_m(\mathbf{r}) + \sum_{m=0}^{\infty} \frac{1}{c^2} \frac{\partial^2 e_m(t)}{\partial t^2} \mathbf{E}_m(\mathbf{r}) = \mathbf{0}$$

We take the scalar product of the wave equation with $\mathbf{E}_n$, integrate over the volume of the cavity and use the orthogonality of the eigenfunctions, see (7.2). This leads to

$$\frac{\partial^2 e_n(t)}{\partial t^2} + \omega_n^2 e_n(t) = 0 \quad (7.5)$$

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.

7.1.2 The resonances in a cavity with losses

There are two main phenomena that cause losses in a cavity: lossy material inside the cavity and walls that are not perfectly conducting. We first discuss a cavity with PEC walls filled with a homogenous material with a small conductivity. The wavenumber is then related to the frequency by, see also (3.9),

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

To find the angular frequency ω we first solve the eigenvalue problem (7.1). The eigenvalues k_n^2 and the eigenfunctions $\mathbf{E}_n$ are not affected by the material inside the cavity. It means that

$$k_n = \tilde{\omega}_n \sqrt{\epsilon_0 \mu_0 (\epsilon + i\sigma/(\tilde{\omega}_n \epsilon_0))} \quad (7.6)$$

where $\tilde{\omega}_n = \omega_n - i\alpha_n$ is a complex angular frequency. If the losses are small then

$$\begin{aligned} \alpha_n &\ll \omega_n \\ \sigma &\ll \omega_n \epsilon_0 \epsilon \end{aligned}$$

and

$$\sqrt{\epsilon_0 \mu_0 (\epsilon + i\sigma/(\tilde{\omega}_n \epsilon_0))} \approx \sqrt{\epsilon_0 \mu_0 \epsilon} \left(1 + i \frac{\sigma}{2\omega_n \epsilon_0 \epsilon} \right)$$

From (7.6) we then obtain

$$\begin{aligned} \omega_n &\approx \frac{k_n}{\sqrt{\epsilon_0 \mu_0 \epsilon}} \\ \alpha_n &\approx \frac{\sigma}{2\epsilon_0 \epsilon} \end{aligned} \quad (7.7)$$

For a cavity with vacuum inside and walls that are not PEC, the losses are due to the currents in the walls. One can then use the impedance boundary condition in (2.18) to solve the problem

$$\begin{cases} \nabla^2 \mathbf{E}_n + k_n^2 \mathbf{E}_n = \mathbf{0} & \text{in } V \\ \mathbf{E}_n - \hat{\mathbf{n}}(\mathbf{E}_n \cdot \hat{\mathbf{n}}) = -\eta_s \hat{\mathbf{n}} \times \mathbf{H}_n & \text{on } S \\ \nabla \cdot \mathbf{E}_n = 0 & \text{in } V \text{ and on } S \end{cases} \quad (7.8)$$

That leads to complex wavenumbers k_n , and since $k_n = \omega/c_0$ the angular frequencies will be complex with real part $\omega_n = c_0 \text{Re}\{k_n\}$ and imaginary part $\alpha_n = -c_0 \text{Im}\{k_n\}$. The other option is to say that the losses are small enough so that they do not affect ω_n . We solve (7.1) and get $\omega_n = c_0 k_n$. The quality factor Q is now calculated by the method described in the next subsection and obtained $\alpha_n = \frac{\omega_n}{2Q}$. For most metallic cavities both of these methods give approximately the same values for ω_n and α_n .

7.1.3 The Quality factor (Q-factor) for a mode in a cavity

The losses add a dissipative term in (7.5)

$$\frac{d^2 e_n(t)}{dt^2} + 2\alpha_n \frac{de_n(t)}{dt} + \omega_n^2 e_n(t) = 0 \quad (7.9)$$

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 quality factor Q measures energy losses within the cavity for a given mode n . In this section, we derive an expression for Q_n and demonstrate its relationship to the cavity bandwidth.

We define the Q factor for cavity mode n as:

$$Q_n = 2\pi \frac{\text{Stored energy in cavity (time-averaged over one cycle)}}{\text{Dissipated energy during one cycle at resonance}} \quad (7.10)$$

Both the stored and dissipated energies are evaluated over the same time interval $[t, t+T_n]$, where $T_n = 2\pi/\omega_n$ represents the period of mode n . Due to spatial orthogonality, the time-averaged energy $U_n(t)$ for mode n is:

$$U_n(t) = \frac{\epsilon_0 \epsilon}{2T_n} \int_t^{t+T_n} |A_n|^2 \cos^2(\omega_n t' + \phi) e^{-2\alpha_n t'} dt' \int_V |\mathbf{E}_n(\mathbf{r})|^2 dV \quad (7.11)$$

We use the normalization in (7.2) and assume that the attenuation is negligible during a cycle such that $\alpha_n T_n \ll 1$. Then (7.11) simplifies to

$$U_n(t) = \frac{\epsilon_0 \epsilon}{4} |A_n|^2 e^{-2\alpha_n t}$$

Consequently, the time-averaged energy decays exponentially as $U_n(t) = U_n(0)e^{-2\alpha_n t}$, where $U_n(0)$ is the time-averaged energy at $t = 0$. The total energy dissipated during the first cycle $[0, T_n]$ is given by:

$$P_n(0) = U_n(0) - U_n(T_n) = U_n(0) (1 - e^{-2\alpha_n T_n})$$

Since $\alpha_n \ll \omega_n$, we have $2\alpha_n T_n \ll 1$ and $e^{-2\alpha_n T_n} \approx 1 - 2\alpha_n T_n$. Substituting this back into the definition (7.10) yields:

$$Q_n = \frac{2\pi U_n(0)}{2\alpha_n T_n U_n(0)} = \frac{2\pi}{2\alpha_n T_n} = \frac{\omega_n}{2\alpha_n}$$

As known from circuit theory, Q is directly related to the half-power bandwidth (BW):

$$\text{BW} = f_+ - f_- = \frac{\alpha_n}{\pi} = \frac{f_n}{Q_n}$$

The dissipated energy during the cycle $[t, t+T_n]$ is given by

$$U_d(t) = R_S \int_t^{t+T_n} \iint_S |\mathbf{J}_{Sn}(\mathbf{r}, t)|^2 dS = R_S \int_t^{t+T_n} \iint_S |\hat{\mathbf{n}} \times \mathbf{H}_n(\mathbf{r}, t)|^2 dS$$

where $R_S = \frac{1}{\sigma_c \delta} = \sqrt{\frac{\omega_n \mu_c \mu_0}{2\sigma_c}}$ is the surface resistance, c.f., (4.44) and (6.48). The magnetic field is related to the electric field by the induction law, i.e.,

$$\mu_0 \frac{\partial \mathbf{H}_n(\mathbf{r}, t)}{\partial t} = -e_n(t) \nabla \times \mathbf{E}_n(\mathbf{r})$$

Example 7.1

Each mode in the cavity can be modeled by the parallel circuit in Figure 7.3. From Kirchhoff's current law we get the equation for the voltage

$$\frac{d^2 v(t)}{dt^2} + 2\alpha \frac{dv(t)}{dt} + \omega_0^2 v(t) = \frac{1}{C} \frac{di(t)}{dt} \quad (7.12)$$

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

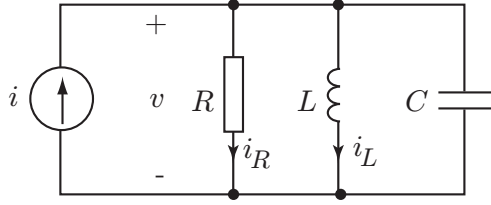


Figure 7.3: The circuit analogues of a cavity mode. The values of R , L , and C are different for different modes.

a cavity, i.e., with the electric field $\mathbf{E}(\mathbf{r}, t) = e_n(t)\mathbf{E}_n(\mathbf{r})$. 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 parallel resonance circuit can be found in most textbooks on basic circuit theory. It is then seen that the Q-factor 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

$$\begin{aligned} W_e(t) &= \frac{\epsilon_0 \epsilon}{2} \iiint_V |\mathbf{E}_n(\mathbf{r}, t)|^2 dV = \frac{1}{2} C (v(t))^2 \\ W_m(t) &= \frac{\mu_0}{2} \iiint_V |\mathbf{H}_n(\mathbf{r}, t)|^2 dV = \frac{1}{2} L (i_L(t))^2 \\ P_d(t) &= R_S \iint_S |\mathbf{J}_{Sn}(\mathbf{r}, t)|^2 dS = R (i_R(t))^2 = \frac{(v(t))^2}{R} \end{aligned} \quad (7.13)$$

■

Example 7.2

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

$$\mathbf{E}(\mathbf{r}, t) = \sum_n e_n(t) \mathbf{E}_n(\mathbf{r}) \quad (7.14)$$

The dissipated power is expressed in terms of the surface current densities $\mathbf{J}_{Sn}(\mathbf{r}, t)$ as

$$P_d(t) = R_S \iint_S \left| \sum_n \mathbf{J}_{Sn}(\mathbf{r}, t) \right|^2 dS \quad (7.15)$$

A relevant question is if this power equals the sum of the dissipated powers for the individual modes?

The answer is yes. To prove it we use Poynting's theorem in section 2.2

$$\begin{aligned} \iint_S \mathbf{E} \times \mathbf{H} \cdot \hat{\mathbf{n}} \, dS &= \iiint_V \nabla \cdot \mathbf{S} \, dv \\ &= - \iiint_V \left[\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \right] dv - \iiint_V \mathbf{E} \cdot \mathbf{J} \, dv \end{aligned} \quad (7.16)$$

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. Thus

$$\iint_S \mathbf{E} \times \mathbf{H} \cdot \hat{\mathbf{n}} \, dS = 0$$

The dissipated power in the metal is the term $P_d(t) = \iiint_V \mathbf{E} \cdot \mathbf{J} \, dv$, where only the volume in the metal gives a contribution. Thus

$$P_d(t) = - \iiint_V \left[\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \right] dv \quad (7.17)$$

Since $e_n(t) = A_n \cos \omega_n t e^{-\alpha_n t}$ we get

$$P_d(t) = - \iiint_V \left[\sum_n \mathbf{H}_n(\mathbf{r}) e_n(t) \cdot \sum_m \mathbf{B}_m(\mathbf{r}) \frac{\partial e_m(t)}{\partial t} + \sum_n \mathbf{E}_n(\mathbf{r}) e_n(t) \cdot \sum_m \mathbf{D}_m(\mathbf{r}) \frac{\partial e_m(t)}{\partial t} \right] dv \quad (7.18)$$

We use the orthogonality and get

$$P_d(t) = -2\epsilon_0 \epsilon \sum_n e_n(t) \frac{\partial e_n(t)}{\partial t} = 2\epsilon_0 \epsilon \sum_n |A_n|^2 (\alpha_n \cos^2(\omega_n t) + \cos(\omega_n t) \sin(\omega_n t)) e^{-2\alpha_n t} \quad (7.19)$$

The time average over one cycle from t to $t + T$ is

$$P_d = \epsilon_0 \sum_n |A_n|^2 \alpha_n e^{-2\alpha_n t} \quad (7.20)$$

Thus the time average of the total dissipated power is the sum of the dissipated powers for the modes. ■

7.1.4 Slater's theorem

Slater's theorem is due to J. C. Slater [21] 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)

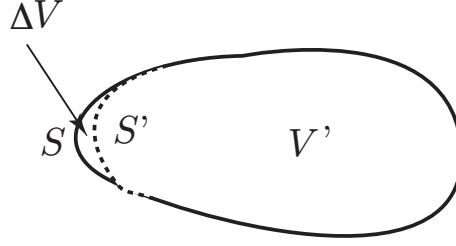


Figure 7.4: The volume V eclipsed by the surface S and the perturbed surface S' . The volume between S and S' is denoted ΔV .

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

$$\frac{f^2 - f_n^2}{f_n^2} = \iiint_{\Delta V} \left(\frac{\mu_0}{\epsilon_0} |\mathbf{H}_n(\mathbf{r})|^2 - |\mathbf{E}_n(\mathbf{r})|^2 \right) dV \quad (7.21)$$

The integration is over $V - V'$, see Figure 7.4. Here $\mathbf{H}_n = (i\omega_n\mu_0)^{-1}\nabla \times \mathbf{E}_n$ is the magnetic field of mode n . Notice that $\frac{1}{4}\epsilon_0|\mathbf{E}_n|^2$ is the electric energy density and $\frac{1}{4}\mu_0|\mathbf{H}_n|^2$ the magnetic energy density of mode n . If we remove a volume where the electric energy is larger than the magnetic energy the frequency decreases, otherwise it increases.

To prove Slater's theorem we start with the Helmholtz equations for the electric and magnetic fields $\mathbf{E}$ and $\mathbf{H}$ in the source free volume V'

$$\begin{aligned} \nabla \times (\nabla \times \mathbf{E}) - k^2 \mathbf{E} &= \mathbf{0} \\ \nabla \times (\nabla \times \mathbf{H}) - k^2 \mathbf{H} &= \mathbf{0} \end{aligned}$$

We consider mode number n and let $\mathbf{E}$ and $\mathbf{H}$ be the perturbed versions of $\mathbf{E}_n$ and $\mathbf{H}_n$. 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 $\mathbf{E}_n$ and use the relation $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$ two times to get

$$(\nabla \times (\nabla \times \mathbf{E})) \cdot \mathbf{E}_n = \nabla \cdot ((\nabla \times \mathbf{E}) \times \mathbf{E}_n + \mathbf{E} \times (\nabla \times \mathbf{E}_n)) + k_n^2 \mathbf{E} \cdot \mathbf{E}_n$$

We integrate over V' and use Gauss theorem

$$(k^2 - k_n^2) \iiint_{V'} \mathbf{E} \cdot \mathbf{E}_n dV = \iint_{S'} \hat{\mathbf{n}} \cdot ((\nabla \times \mathbf{E}) \times \mathbf{E}_n + \mathbf{E} \times (\nabla \times \mathbf{E}_n)) dS$$

The last integral is zero since $\hat{\mathbf{n}} \times \mathbf{E} = \mathbf{0}$ on S' and then

$$(k^2 - k_n^2) \iiint_{V'} \mathbf{E} \cdot \mathbf{E}_n dV = \iint_{S'} \hat{\mathbf{n}} \cdot ((\nabla \times \mathbf{E}) \times \mathbf{E}_n) dS$$

We make use of the completeness of the system $\mathbf{E}_n$. Since V' is inside V we can expand $\mathbf{E}$ in a series of $\mathbf{E}_n$ as

$$\mathbf{E} = \sum_{n=0}^{\infty} \mathbf{E}_n \iiint_{V'} \mathbf{E} \cdot \mathbf{E}_n dV$$



Figure 7.5: 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.

The eigenfunction $\mathbf{E}$ is very close to $\mathbf{E}_n$ in V' , and we approximate $\mathbf{E}$ as

$$\mathbf{E} \approx \mathbf{E}_n \int \int \int_{V'} \mathbf{E} \cdot \mathbf{E}_n dV$$

That means that

$$(k^2 - k_n^2) \int \int \int_{V'} \mathbf{E}_n \cdot \mathbf{E}_n dV \approx \int \int_{S'} \hat{\mathbf{n}} \cdot ((\nabla \times \mathbf{E}_n) \times \mathbf{E}_n) dS$$

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 $\hat{\mathbf{n}} \times \mathbf{E}_n = 0$ on S . That gives

$$\begin{aligned} (k^2 - k_n^2) \int \int \int_{V'} \mathbf{E}_n \cdot \mathbf{E}_n dV &\approx \int \int_{S'} \hat{\mathbf{n}} \cdot ((\nabla \times \mathbf{E}_n) \times \mathbf{E}_n) dS = - \int \int \int_{\Delta V} \nabla \cdot ((\nabla \times \mathbf{E}_n) \times \mathbf{E}_n) dV \\ &= - \int \int \int_{\Delta V} \mathbf{E}_n \cdot (\nabla \times (\nabla \times \mathbf{E}_n)) - (\nabla \times \mathbf{E}_n) \cdot (\nabla \times \mathbf{E}_n) dV \end{aligned}$$

We use $\nabla \times (\nabla \times \mathbf{E}_n) = k_n^2 \mathbf{E}_n$, $\mathbf{H}_n = (i\omega_n \mu_0)^{-1} \nabla \times \mathbf{E}_n$, $k = 2\pi f/c$, $k_n = 2\pi f_n/c$, and the normalization of $\mathbf{E}_n$ over V .

$$f^2 - f_n^2 = f_n^2 \int \int \int_{\Delta V} \left(\frac{\mu_0}{\epsilon_0} |\mathbf{H}_n|^2 - |\mathbf{E}_n|^2 \right) dV \quad (7.22)$$

which is the same as (7.21).

7.1.5 Measuring electric and magnetic fields in cavities

Slater's theorem can be used 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 7.5. The sphere gives a shift of the resonance frequencies of the cavity, in concordance with the Slater's theorem. 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 $\mathbf{E}_0 = E_0 \hat{\mathbf{z}} = E_0 \cos \theta \hat{\mathbf{r}} - E_0 \sin \theta \hat{\boldsymbol{\theta}}$ 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

$$\mathbf{E}_{\text{sphere}}(r, \theta) = \mathbf{E}_0 + \mathbf{E}_{\text{dipole}} = E_0 \hat{\mathbf{z}} + \frac{p}{4\pi\epsilon_0 r^3} (2\hat{\mathbf{r}} \cos \theta + \hat{\boldsymbol{\theta}} \sin \theta)$$

where p is the induced dipole moment. The dipole moment is obtained from the fact that the tangential component $\mathbf{E}_{\text{sphere}}$ is zero on the surface of the sphere. The tangential component is the θ -component and since $\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{z}} = -\sin \theta$ we get

$$\hat{\boldsymbol{\theta}} \cdot \mathbf{E}_{\text{sphere}}(R_0, \theta) = -E_0 \sin \theta + \frac{p}{4\pi\epsilon_0 R^3} \sin \theta = 0$$

Thus $p = 4\pi\epsilon_0 R^3 E_0$ and

$$\mathbf{E}_{\text{sphere}}(r, \theta) = E_0 \cos \theta \left(1 + \frac{2R^3}{r^3}\right) \hat{\mathbf{r}} - E_0 \sin \theta \left(1 - \frac{R^3}{r^3}\right) \hat{\boldsymbol{\theta}}$$

We first assume that the magnetic field is zero. We have to be careful when we determine the integral

$$\iiint_{\Delta V} |\mathbf{E}_n|^2 dV$$

since the field $\mathbf{E}_n$ 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

$$\begin{aligned} \frac{(f + \delta f)^2 - f^2}{f^2} &= -2\pi \int_0^\pi \int_R^{R+\delta R} |\mathbf{E}_{\text{sphere}}(r, \theta)|^2 r^2 \sin \theta dr d\theta \\ &\approx -2\pi \int_0^\pi |\mathbf{E}_{\text{sphere}}(R, \theta)|^2 R^2 \sin \theta d\theta \delta R \\ &= -12\pi R^2 E_0^2 \delta R \end{aligned}$$

We let $\delta f = df$ be infinitesimal and get $((f + df)^2 - f^2)/f^2 \approx 2df/f$. 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

$$\Delta f \approx -2\pi R_0^3 E_0^2 f_n \quad (7.23)$$

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. (7.23).

7.2 Example: Cylindrical cavities

We analyze a common resonant cavity consisting of a hollow waveguide with simply connected cross section Ω and bounded by perfectly conducting metallic planes at $z = 0$ and $z = d$ (see Figure 7.6).

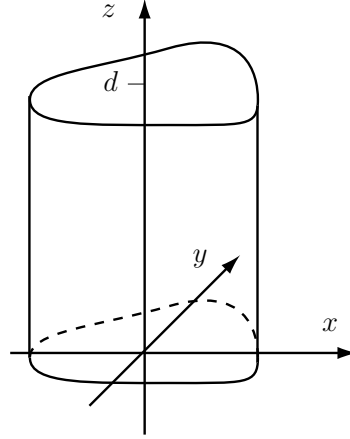


Figure 7.6: Geometry for cylindric resonance cavity. The cavity is a hollow waveguide with cross section Ω and terminated by perfectly conducting plates at $z = 0$ and $z = d$.

In order to determine the modes 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 $z = d$. Because $\mathbf{E}_T(\boldsymbol{\rho}, 0) = \mathbf{E}_T(\boldsymbol{\rho}, d) = \mathbf{0}$ for all $\boldsymbol{\rho}$, the transverse divergence vanishes at these surfaces:

$$\nabla_T \cdot \mathbf{E}_T(\boldsymbol{\rho}, 0) = \nabla_T \cdot \mathbf{E}_T(\boldsymbol{\rho}, d) = 0$$

Since the interior of the cavity is charge-free ($\nabla \cdot \mathbf{E} = 0$), Gauss's law expands to:

$$\nabla_T \cdot \mathbf{E}_T(\mathbf{r}) + \partial E_z(\mathbf{r}) / \partial z = 0$$

Consequently, the normal derivative of E_z must vanish at $z = 0$ and $z = d$.

For the magnetic field, the normal component of magnetic flux density $\mathbf{B}$ is everywhere continuous across boundaries. Because the fields vanish inside the perfect conductor, the normal component H_z must be identically zero at the metal plates.

Thus, the boundary conditions at $z = 0$ and $z = d$ are summarized as:

$$\begin{cases} \frac{\partial E_z(\boldsymbol{\rho}, 0)}{\partial z} = \frac{\partial E_z(\boldsymbol{\rho}, d)}{\partial z} = 0 \\ H_z(\boldsymbol{\rho}, 0) = H_z(\boldsymbol{\rho}, d) = 0 \end{cases} \quad (7.24)$$

The fields in the cavity is a superposition of waveguide modes propagating in the positive and negative z -directions. From the expansions in (6.21) and (6.22) the z -components of the fields of mode n are expressed as

$$\begin{cases} E_z(\mathbf{r}) = (a_{n\nu}^+ e^{ik_z z} - a_{n\nu}^- e^{-ik_z z}) v_n(\boldsymbol{\rho}) & \nu = \text{TM} \\ H_z(\mathbf{r}) = (a_{n\nu}^+ e^{ik_z z} + a_{n\nu}^- e^{-ik_z z}) w_n(\boldsymbol{\rho}) & \nu = \text{TE} \end{cases}$$

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

$$k_{z\ell} = \frac{\ell\pi}{d} \quad \begin{cases} \ell = 0, 1, 2, \dots & \nu = \text{TM} \\ \ell = 1, 2, \dots & \nu = \text{TE} \end{cases}$$

There exists no TE mode with value $\ell = 0$ since $H_z(\mathbf{r}) = 0$ when $\ell = 0$. However, we saw in Subsection 6.3 that if the region Ω of the waveguide is not simply connected then TEM modes with $\ell = 0$ can exist.

The frequencies that can exist in the cavity are determined by $k^2 = k_{tn}^2 + k_{z\ell}^2$ and thus

$$f_{n\ell} = \frac{c}{2\pi} \sqrt{k_{tn}^2 + \left(\frac{\ell\pi}{d}\right)^2} \quad (7.25)$$

The fields for the corresponding resonances follow from (6.21) and (6.22)

$$\begin{cases} \mathbf{E}_{n\ell}(\mathbf{r}) = A_{n\ell\nu} \sqrt{\frac{\varepsilon_\ell}{d}} \left(i\mathbf{E}_{Tn\nu}(\boldsymbol{\rho}) \sin \frac{\ell\pi z}{d} + v_n(\boldsymbol{\rho}) \cos \frac{\ell\pi z}{d} \hat{\mathbf{z}} \right) & \nu = \text{TM} \\ \mathbf{H}_{n\ell}(\mathbf{r}) = A_{n\ell\nu} \sqrt{\frac{\varepsilon_\ell}{d}} \mathbf{H}_{Tn\nu}(\boldsymbol{\rho}) \cos \frac{\ell\pi z}{d} & \ell = 0, 1, 2, \dots \end{cases} \quad (7.26)$$

where $\varepsilon_\ell = 2 - \delta_{\ell,0}$ and

$$\begin{cases} \mathbf{E}_{n\ell}(\mathbf{r}) = A_{n\ell\nu} i \sqrt{\frac{2}{d}} \mathbf{E}_{Tn\nu}(\boldsymbol{\rho}) \sin \frac{\ell\pi z}{d} & \nu = \text{TE} \\ \mathbf{H}_{n\ell}(\mathbf{r}) = A_{n\ell\nu} \sqrt{\frac{2}{d}} \left(\mathbf{H}_{Tn\nu}(\boldsymbol{\rho}) \cos \frac{\ell\pi z}{d} + iw_n(\boldsymbol{\rho}) \sin \frac{\ell\pi z}{d} \hat{\mathbf{z}} \right) & \ell = 1, 2, \dots \end{cases} \quad (7.27)$$

Notice that $\mathbf{H}_{nl}$ and $\mathbf{E}_{nl}$ are related via the induction law $\mathbf{H}_{nl}(\mathbf{r}) = (i\omega\mu_0)^{-1} \nabla \times \mathbf{E}_{nl}(\mathbf{r})$. Both $\mathbf{E}_{Tn\nu}(\boldsymbol{\rho})$ and $\mathbf{H}_{Tn\nu}(\boldsymbol{\rho})$ are imaginary, as seen from Eq. (6.24). Thus $\mathbf{E}_{nl}$ is real and $\mathbf{H}_{nl}$ imaginary. The amplitudes $A_{n\ell\nu}$ are determined from the normalization

$$\iiint_V |\mathbf{E}_{nl}|^2 dV = \frac{\mu_0}{\epsilon_0\epsilon} \iiint_V |\mathbf{H}_{nl}|^2 dV = 1$$

Then

$$\begin{aligned} A_{nl\text{TM}} &= \frac{k_{tn}}{k_{nl}} \\ A_{nl\text{TE}} &= \frac{k_{tn}}{\omega_{nl}\mu_0} \end{aligned} \quad (7.28)$$

We are now prepared to determine the Q-factor and bandwidth of a cylindrical cavity with metal walls. The conductivity of the walls is σ_c and the relative permeability μ .

With suitable normalization, the electric field of cavity mode n is given by

$$\mathbf{E}(\mathbf{r}, t) = \cos(\omega_n t + \phi) e^{-\alpha t} \mathbf{E}_n(\mathbf{r})$$

where $\mathbf{E}_n(\mathbf{r})$ is given by (7.26) for TM modes and (7.27) 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 of $\mathbf{E}_n(\mathbf{r})$,

$$U(t) = 2U_e(t) = \frac{\epsilon_0\epsilon}{2} e^{-2\alpha_n t} \quad (7.29)$$

The time average of the losses at the end surfaces are

$$P_e(t) = \frac{R_S}{2} \iint_{\Omega} |\mathbf{H}_T(\boldsymbol{\rho}, 0)|^2 + |\mathbf{H}_T(\boldsymbol{\rho}, d)|^2 dx dy e^{-2\alpha t} = R_S \iint_{\Omega} |\mathbf{H}_T(\boldsymbol{\rho}, 0)|^2 dx dy e^{-2\alpha t}$$

where R_S is the surface resistance given by (4.44) and (6.48). From (6.24) and the orthogonality relation in (6.33) we get

$$P_e(t) = \begin{cases} R_S \frac{\varepsilon_\ell \epsilon_0 \epsilon}{\mu_0 d} e^{-2\alpha_n t} & \text{TM} \\ R_S \frac{k_{z\ell}^2 \epsilon_0 \epsilon}{\mu_0^2 k_{tn}^2 d} e^{-2\alpha_n t} & \text{TE} \end{cases}$$

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

$$\frac{P_e(t)}{U(t)} = \begin{cases} R_S \frac{2\varepsilon_\ell}{\mu_0 d} & \text{TM} \\ R_S \frac{2k_{z\ell}^2}{\mu_0 k_{nl}^2 d} & \text{TE} \end{cases}$$

The time average of the losses in the lateral surface is

$$P_m(t) = \frac{1}{2} R_S \int_0^d \int_\Gamma |\hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r})|^2 d\Gamma dz e^{-2\alpha t}$$

where $\hat{\mathbf{n}}$ is the normal to the lateral surface and Γ is the boundary curve. We now utilize equations (6.21)–(6.24) and (7.28) to get

$$P_m(t) = \begin{cases} \frac{R_S \epsilon_0 \epsilon}{2\mu_0 k_{tn}^2} e^{-2\alpha_n t} \int_\Gamma |\hat{\mathbf{n}} \cdot \nabla_T v_n(\boldsymbol{\rho})|^2 d\Gamma & \text{TM} \\ \frac{R_S \epsilon_0 \epsilon}{2\mu_0 k_{nl}^2} e^{-2\alpha_n t} \int_\Gamma \frac{k_{z\ell}^2}{k_{tn}^4} |\hat{\mathbf{n}} \times \nabla_T w_n(\boldsymbol{\rho})|^2 + |w_n(\boldsymbol{\rho})|^2 d\Gamma & \text{TE} \end{cases} \quad (7.30)$$

We conclude that the Q-factor is related to the attenuation coefficient α via $Q = \omega_{nl}/2\alpha$ and that α is given by

$$\alpha = \frac{P_e(t)}{2U(t)} = \begin{cases} \frac{R_S}{\mu_0} \left(\frac{\varepsilon_\ell}{d} + \frac{1}{2k_{tn}^2} \int_\Gamma |\hat{\mathbf{n}} \cdot \nabla_T v_n(\boldsymbol{\rho})|^2 d\Gamma \right) & \text{TM} \\ \frac{R_S}{\mu_0 k_{nl}^2} \left(\frac{k_{z\ell}^2}{d} + \frac{1}{2} \int_\Gamma \frac{k_{z\ell}^2}{k_{tn}^2} |\hat{\mathbf{n}} \times \nabla_T w_n(\boldsymbol{\rho})|^2 + k_{tn}^2 |w_n(\boldsymbol{\rho})|^2 d\Gamma \right) & \text{TE} \end{cases} \quad (7.31)$$

Example 7.3

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 6.4 on

page 90. By inserting those expressions into (7.30) the power losses in the lateral surface are obtained

$$P_m(t) = R_S e^{-2\alpha t} \begin{cases} (k_{mnl}a)^2 \xi_{mn}^{-2} \eta_0^{-2} a^{-1} & \text{TM} \\ \frac{\eta_{mn}^2}{(\eta_{mn}^2 - m^2)a} \left(m^2 \frac{(k_{z\ell}a)^2}{\eta_{mn}^4} + 1 \right) & \text{TE} \end{cases}$$

Hence the Q-factors are

$$\begin{cases} Q_{TM} = \frac{\eta_0 k_{mnl}a}{2R_S(1 + \varepsilon_\ell a/d)} \\ Q_{TE} = \frac{(k_{mnl}a)^3 \eta_0 (1 - m^2 \eta_{mn}^{-2})}{2R_S \eta_{mn}^2 (1 + (k_{z\ell}a)^2 m^2 \eta_{mn}^{-4} + 2(1 - m^2 \eta_{mn}^{-2}) k_{z\ell}^2 a^3 d^{-1} \eta_{mn}^{-2})} \end{cases}$$

Since $k_{mnl}a = \sqrt{\eta_{mn}^2 + (\ell\pi a/d)^2}$ and $k_{mnl}a = \sqrt{\xi_{mn}^2 + (\ell\pi a/d)^2}$ the Q-factors Q_{TM} and Q_{TE} can both be written in the form $f(a/d)/R_S$. ■

7.3 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 the Maxwells equations in spherical coordinates.

7.3.1 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]

$$\begin{aligned} \mathbf{A}_{1\sigma ml}(\theta, \phi) &= \frac{1}{\sqrt{l(l+1)}} \nabla \times (\mathbf{r} Y_{\sigma ml}(\theta, \phi)) \\ \mathbf{A}_{2\sigma ml}(\theta, \phi) &= \hat{\mathbf{r}} \times \mathbf{A}_{1\sigma ml}(\theta, \phi) = \frac{1}{\sqrt{l(l+1)}} r \nabla Y_{\sigma ml}(\theta, \phi) \\ \mathbf{A}_{3\sigma ml}(\theta, \phi) &= \hat{\mathbf{r}} Y_{\sigma ml}(\theta, \phi). \end{aligned} \tag{7.32}$$

where $Y_{\sigma ml}$ are the spherical harmonics:

$$Y_{\sigma ml}(\theta, \phi) = \sqrt{\frac{\varepsilon_m}{2\pi}} \sqrt{\frac{2l+1}{2} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix} \tag{7.33}$$

where $\varepsilon_m = 2 - \delta_{m0}$ and τ, m, l take the values

$$\sigma = \begin{pmatrix} e \\ o \end{pmatrix}, \quad m = 0, 1, 2, \dots, l, \quad l = 0, 1, \dots \tag{7.34}$$

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

$$\int_{\Omega} \mathbf{A}_{\tau n}(\theta, \phi) \cdot \mathbf{A}_{\tau' n'}(\theta, \phi) d\Omega = \delta_{\tau\tau'} \delta_{nn'}, \tag{7.35}$$

where the integration is over the unit sphere and where n is the multi index $n = \sigma ml$.

7.3.2 Regular spherical vector waves

The regular divergence-free spherical vector waves are defined by

$$\begin{cases} \mathbf{u}_{1n}(\mathbf{r}) = j_l(kr) \mathbf{A}_{1\sigma ml}(\theta, \phi) \\ \mathbf{u}_{2n}(\mathbf{r}) = \frac{1}{k} \nabla \times (j_l(kr) \mathbf{A}_{1\sigma ml}(\theta, \phi)) \\ \quad = j'_l(kr) \mathbf{A}_{2\sigma ml}(\theta, \phi) + \frac{1}{kr} j_l(kr) (\mathbf{A}_{2\sigma ml}(\theta, \phi) + \sqrt{l(l+1)} \mathbf{A}_{3\sigma ml}(\theta, \phi)), \end{cases} \quad (7.36)$$

where $j_l(kr)$ is the spherical Bessel function and $n = \sigma ml$ is a multiindex. They satisfy the vector Helmholtz equation

$$\nabla \times (\nabla \times \mathbf{u}_{\tau nn}(\mathbf{r})) - k^2 \mathbf{u}_{\tau nn}(\mathbf{r}) = \mathbf{0}$$

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.

7.3.3 Resonance frequencies in a spherical cavity

Consider a spherical cavity with radius a and filled with air, or vacuum. The boundary condition for the complex electric field $\mathbf{E}(\mathbf{r})$ is

$$\hat{\mathbf{r}} \times \mathbf{E}(a, \theta, \phi) = \mathbf{0}$$

The electric field $\mathbf{E}(\mathbf{r})$ can be expanded in the system of regular divergence-free spherical vector waves. The modes with $\mathbf{E}(\mathbf{r}) = \mathbf{u}_{1n}(\mathbf{r})$ are referred to as TE modes since the electric field lacks radial component. Modes with $\mathbf{E}(\mathbf{r}) = \mathbf{u}_{2n}(\mathbf{r})$ have a magnetic field given by

$$\mathbf{H}(\mathbf{r}) = \frac{1}{i\omega\mu_0} \nabla \times \mathbf{u}_2(\mathbf{r}) = \frac{k}{i\omega\mu_0} \mathbf{u}_{1n}(\mathbf{r})$$

Since $\mathbf{H}(\mathbf{r})$ lacks radial component these modes are referred to as TM modes. It follows that the resonances are obtained from

$$\begin{cases} \hat{\mathbf{r}} \times \mathbf{u}_{1n}(a\hat{\mathbf{r}}) = \mathbf{0} & \text{TE} \\ \hat{\mathbf{r}} \times \mathbf{u}_{2n}(a\hat{\mathbf{r}}) = \mathbf{0} & \text{TM} \end{cases} \quad (7.37)$$

That leads to the equations for the wavenumbers

$$j_l(ka) = 0, \text{ TE} \quad (7.38)$$

and

$$j'_l(ka) + (ka)^{-1} j_l(ka) = 0, \text{ TM} \quad (7.39)$$

Table (7.3.3) 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 $m = 0, \sigma = \text{e}$ and $m = 1, 2, \dots, l, \sigma = \begin{pmatrix} \text{e} \\ \text{o} \end{pmatrix}$. The number of

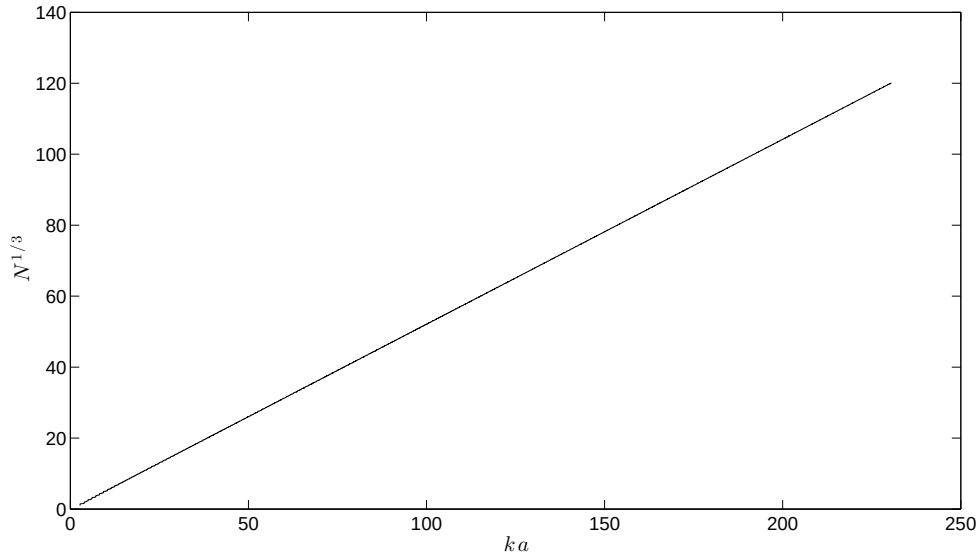


Figure 7.7: The cube root, $N^{1/3}$, of the total number of modes for a sphere with radius a with wavenumber less than k , as a function of ka . The relation $N^{1/3} \sim ka$ is valid for all cavities.

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 7.7 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 k in the graph. The number of resonances with a wavelength larger 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 $\mathbf{E} = \mathbf{u}_{1n}$ and then the time average of the electric field is

$$U_e = \frac{\epsilon_0}{4} \iiint_V |\mathbf{u}_1|^2 dv$$

where V is the volume of the sphere. Due to the orthogonality in (7.35)

$$\iiint_V |\mathbf{u}_1(\mathbf{r})|^2 dv = \int_0^a (j_l(kr))^2 r^2 dr \quad (7.40)$$

From appendix A.3 we get $j_l(x) = \sqrt{\frac{2}{\pi x}} J_{l+0.5}(x)$ and the integral

$$\int x (J_\nu(x))^2 dx = \frac{x^2}{2} (J'_\nu(x))^2 + \frac{1}{2} (x^2 - \nu^2) (J_\nu(x))^2 + \text{constant} \quad (7.41)$$

and then

$$U_e = \frac{\epsilon_0}{4} \iiint_V |\mathbf{u}_{1n}(\mathbf{r})|^2 dv = \frac{a^3 \epsilon_0}{8} (j'_l(ka))^2 \quad (7.42)$$

Table 7.1: 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 (7.38) and (7.39). 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

For $\tau = 2$ we have $\mathbf{E}(\mathbf{r}) = \mathbf{u}_2(\mathbf{r})$. Then

$$\mathbf{H}(\mathbf{r}) = \frac{1}{i\omega\mu_0} \nabla \times \mathbf{u}_2(\mathbf{r}) = \frac{1}{i\omega\mu_0 k} \nabla \times (\nabla \times \mathbf{u}_1(\mathbf{r})) = \frac{k}{i\omega\mu_0} \mathbf{u}_1(\mathbf{r})$$

It is seen that

$$\begin{aligned} \int_V |\mathbf{u}_1(\mathbf{r})|^2 dv &= \int_0^a (j_l(kr))^2 r^2 dr = k^{-3} \frac{ka}{8} (1 + 4(ka)^2 - (2l+1)^2) (j_l(ka))^2 \\ &= k^{-3} \frac{ka}{2} ((ka)^2 - l(l+1)) (j_l(ka))^2 \end{aligned} \quad (7.43)$$

The electric and magnetic energies are

$$U_e = U_m = \frac{\mu_0}{4} \left(\frac{k}{\omega\mu_0} \right)^2 \int_V |\mathbf{u}_1(\mathbf{r})|^2 dv = \frac{a^3 \epsilon_0}{8(ka)^2} ((ka)^4 - (ka)^2 l(l+1)) \left(\frac{j_l(ka)}{ka} \right)^2 \quad (7.44)$$

7.3.4 Q-factors for spherical cavity

The Q-factor for a mode in cavity was defined in (7.10). Consider first the modes with $\tau = 1$, i.e., $\mathbf{E}_n(\mathbf{r}) = \mathbf{u}_{1n}(\mathbf{r})$. According to (7.42) the time averaged stored energy in the cavity is

$$U = 2U_e = \frac{\epsilon_0}{2} \iiint_V |\mathbf{u}_{1n}(\mathbf{r})|^2 dv = a^3 \frac{\epsilon_0}{4} (j'_l(ka))^2$$

The dissipated power during one period, T , is

$$P = \frac{T}{2} R_S \iint_S |\mathbf{J}_S(\mathbf{r})|^2 dS = \frac{T}{2} R_S \iint_S |\hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r})|^2 dS$$

where $\mathbf{H}(\mathbf{r}) = -\frac{i}{\omega\mu_0}\nabla \times \mathbf{u}_1(\mathbf{r}) = -\frac{ik}{\omega\mu_0}\mathbf{u}_2(\mathbf{r})$. The surface resistance is R_S is given by (4.44) and (6.48).

The orthogonality of the vector spherical harmonics in (7.36) and (7.38) give

$$P = \frac{T}{2}R_S \left(\frac{k}{\omega\mu_0}\right)^2 \iint_S |\mathbf{u}_2(\mathbf{r})|^2 dS = a^2 \frac{T}{2}R_S \left(\frac{k}{\omega\mu_0}\right)^2 (j'_l(ka))^2 = a^2 \frac{T}{2}R_S \eta_0^{-2} (j'_l(ka))^2$$

where $\eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$. The Q-factor is

$$Q = 2\pi \frac{U}{P} = a \frac{1}{2R_S} \omega \mu_c \mu_0 = \frac{a}{\delta} \quad (7.45)$$

where δ is the skin depth.

For $\tau = 2$ the stored energy is given by (7.44)

$$U = \frac{\epsilon_0}{2} \iiint_V |\mathbf{u}_{2n}(\mathbf{r})|^2 dv = \frac{a^3 \epsilon_0}{4(ka)^2} ((ka)^4 - (ka)^2 l(l+1)) \left(\frac{j_l(ka)}{ka}\right)^2 \quad (7.46)$$

The dissipated power integrated over one cycle is

$$P = \frac{T}{2}R_S \iint |\mathbf{J}_S(\mathbf{r})|^2 dS = \frac{T}{2}R_S \iint |\hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r})|^2 dS$$

where $\mathbf{H}(\mathbf{r}) = -\frac{i}{\omega\mu_0}\nabla \times \mathbf{u}_2(\mathbf{r}) = -\frac{ik}{\omega\mu_0}\mathbf{u}_1(\mathbf{r}) = -i\eta_0^{-1}\mathbf{u}_1(\mathbf{r})$. The orthogonality of the vector spherical harmonics gives

$$P = \frac{T}{2}R_S (\eta_0^{-1})^2 \iint_S |\hat{\mathbf{n}} \times \mathbf{u}_1(\mathbf{r})|^2 dS = a^2 \frac{T}{2\eta_0^2} R_S (j_l(ka))^2 \quad (7.47)$$

Thus

$$Q = 2\pi \frac{U}{P} = a\omega \frac{\mu_0}{2R_S(ka)^4} ((ka)^4 - (ka)^2 l(l+1)) = \frac{a}{\delta} \left(1 - \frac{l(l+1)}{(ka)^2}\right) \quad (7.48)$$

7.3.5 Two concentric spheres

Consider two concentric conducting spheres with radius a and $b = a + h$. The general solutions are

$$\mathbf{E}(\mathbf{r}) = \alpha \mathbf{u}_{1l}(\mathbf{r}) + \beta \mathbf{w}_{1l}(\mathbf{r}) \quad (7.49)$$

for TE modes and

$$\mathbf{E}(\mathbf{r}) = \gamma_l \mathbf{u}_{2l}(\mathbf{r}) + \kappa_l \mathbf{w}_{2l}(\mathbf{r}) \quad (7.50)$$

for TM modes. The vector functions $\mathbf{w}_{\tau l}(\mathbf{r})$ are defined as

$$\begin{cases} \mathbf{w}_{1n}(\mathbf{r}) = n_l(kr) \mathbf{A}_{1\sigma ml}(\theta, \phi) \\ \mathbf{w}_{2n}(\mathbf{r}) = \frac{1}{k} \nabla \times (n_l(kr) \mathbf{A}_{1\sigma ml}(\theta, \phi)) \\ \quad = n'_l(kr) \mathbf{A}_{2\sigma ml}(\theta, \phi) + \frac{1}{kr} n_l(kr) (\mathbf{A}_{2\sigma ml}(\theta, \phi) + \sqrt{l(l+1)} \mathbf{A}_{3\sigma ml}(\theta, \phi)), \end{cases} \quad (7.51)$$

where $n_l(kr)$ is the spherical Neumann function.

The boundary conditions are that $\hat{\mathbf{n}} \times \mathbf{E} = 0$ for $r = a$ and $r = b$. By utilizing the orthogonality of the vector spherical harmonics two systems of equations are obtained

$$\begin{aligned}\alpha j_l(ka) + \beta n_l(ka) &= 0 \\ \alpha j_l(kb) + \beta n_l(kb) &= 0\end{aligned}\tag{7.52}$$

for TE modes and

$$\begin{aligned}\gamma \left(j'_l(ka) + \frac{j_l(ka)}{ka} \right) + \beta \left(n'_l(ka) + \frac{n_l(ka)}{ka} \right) &= 0 \\ \gamma \left(j'_l(kb) + \frac{j_l(kb)}{kb} \right) + \beta \left(n'_l(kb) + \frac{n_l(kb)}{kb} \right) &= 0\end{aligned}\tag{7.53}$$

for TM modes. The resonance wavenumbers k are determined from the determinants

$$\begin{cases} j_l(ka)n_l(kb) - j_l(kb)n_l(ka) = 0 & \text{TE} \\ \left(j'_l(ka) + \frac{j_l(ka)}{ka} \right) \left(n'_l(kb) + \frac{n_l(kb)}{kb} \right) - \left(j'_l(kb) + \frac{j_l(kb)}{kb} \right) \left(n'_l(ka) + \frac{n_l(ka)}{ka} \right) = 0 & \text{TM} \end{cases}\tag{7.54}$$

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

$$n'_l(kb) + \frac{n_l(kb)}{kb} = n'_l(ka) + \frac{n_l(ka)}{ka} + kh \left(n''_l(ka) + \frac{n'_l(ka)}{ka} - \frac{n_l(ka)}{(ka)^2} \right) + \mathcal{O}((kh)^2)\tag{7.55}$$

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

$$ka \approx \sqrt{l(l+1)}\tag{7.56}$$

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 (7.56). The lowest resonances are $f_1 = 10.6$ Hz, $f_2 = 18.4$ Hz, $f_3 = 26.0$ Hz, $f_4 = 33.5$ 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 (7.56) 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 (2.18) together with the expansions in (7.49). 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.

7.4 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 $e^{in\phi}$. Then the problem is reduced to a two-dimensional problem in the cylindrical coordinates ρ 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 in Subsection 6.6. Then the resonance frequencies of the cavity are obtained from (7.25).
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 case.

Example 7.4

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

$$f(\rho, \phi, z) = \sum_{m=-\infty}^{\infty} f_m(\rho, z)e^{im\phi}$$

Fields with different m -values do not couple to each other and for 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.

Determining the Q-factor using FEM

The Q-factor for a cavity mode can be determined in at least two different ways:

Method 1: Explicit Calculation of Stored Energy and Power Dissipation

The first approach uses the definitions of stored energy and dissipated power from (7.13) alongside the Q -factor relation in (7.10).

The absorbed power in the cavity walls is given by:

$$P_c = \frac{1}{2}R_s \iint_S |\mathbf{J}_s(\mathbf{r})|^2 dS = \frac{1}{2}R_s \iint_S |\mathbf{H}(\mathbf{r})|^2 dS$$

where S represents the internal boundary surface of the cavity and R_s is the surface resistance.

The Power dissipation and stored energy are computed by COMSOL Multiphysics in the following manner:

- Power Dissipation (P_c): Evaluate the surface integral using Results → Derived Values → Line Integration (or Surface Integration, depending on your model dimension). Ensure that Integration Settings → Compute surface integral (or volume integral for axisymmetric models) is selected. In axisymmetric configurations, the magnitude of the magnetic field $|\mathbf{H}(\mathbf{r})|$ is a pre-defined variable suitable for direct integration.
- Stored Energy (U): Integrate the pre-defined Electric energy density via Results → Derived Values → Surface Integration (checking Compute volume integral in 2D axisymmetric mode). Multiply this result by 2, using the peak equality $U = 2U_e$.

Method 2: Lossy Boundaries via the Impedance Boundary Condition Alternatively, cavity losses can be modeled directly using the Impedance Boundary Condition:

Table 7.2: 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 7.1

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

- Assign the appropriate conductor properties to the boundary metal under Materials (e.g., along the circular boundary arc).
- Select Impedance Boundary Condition under the Electromagnetic Fields interface for those boundaries.
- Because the boundary condition depends on the unknown eigenfrequency, the underlying eigenvalue problem becomes non-linear. To resolve this, specify a linearization point: Right-click Eigenvalue Solver (under Study Settings $\rightarrow$ Solver Configurations). Enter an initial frequency estimate (e.g., 10^9 for 1 GHz) in the Transform point box. For this cavity geometry, the precise value of the linearization point is not critical.
- Solve the eigenfrequency study.
- Extract the Q-factor directly as a pre-defined variable via Results $\rightarrow$ Derived Values $\rightarrow$ Global Evaluation by selecting the desired mode and choosing Quality factor.

In Figure 7.8 we see the electric and magnetic fields for the three lowest TE and TM resonances.



Example 7.5

In order to obtain also the non-axisymmetric modes we choose $m > 0$ in **Electromagnetic waves**.



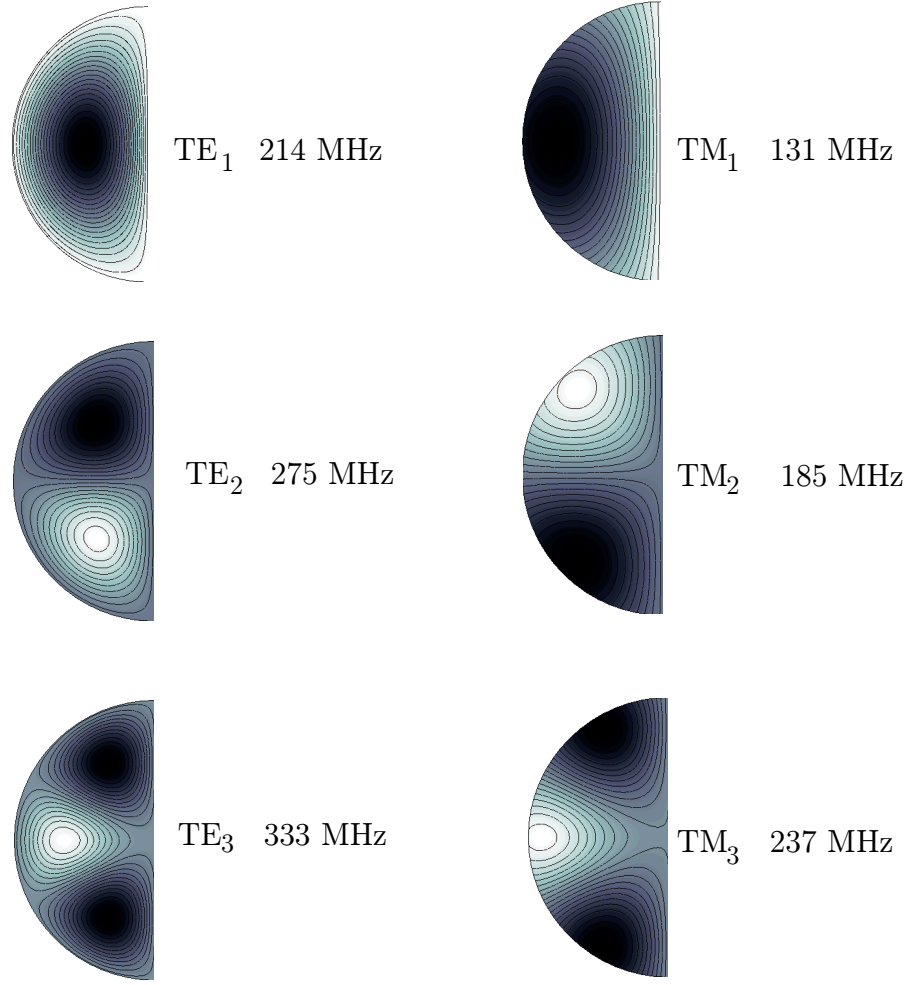


Figure 7.8: The electric field for the three lowest axisymmetric TE modes and the magnetic field for the three lowest axisymmetric TM modes in a spherical cavity with radius 1 m. The electric field of the TE modes and the magnetic field of the TM modes are directed in the $\hat{\phi}$ -direction.

7.5 Excitation of modes in a cavity

Electromagnetic fields in a cavity are excited by electrical currents, which are typically fed through a waveguide or coaxial cable (see Figure 7.9). To excite a single cavity mode, the driving current is usually tuned to a specific target frequency. In particle accelerators, however, the charged particle beam itself acts as a current source. Because the beam carries a broad frequency spectrum, it excites a wide array of cavity modes simultaneously.

We first analyze excitation in the frequency domain, defining the complex current density $\mathbf{J}(\mathbf{r}, \omega)$ as the Fourier transform of the time-dependent current density. Next, we derive the governing equations in the time domain. Note that irrotational modes persist only while the source current is active; once the source is turned off, only the solenoidal modes remain.

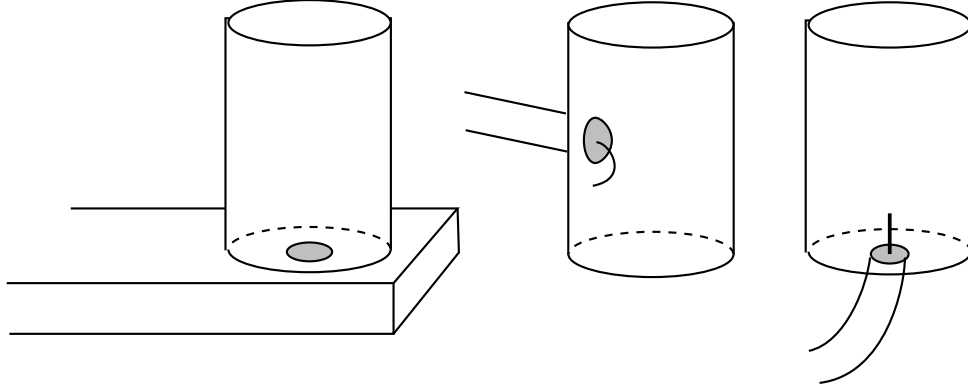


Figure 7.9: Three examples of feeding systems for microwave cavities. The left cavity is fed by a waveguide through a hole in the bottom surface. The waveguide in the middle is fed by a coaxial cable that is attached to the wall of the cavity. The cavity to the right is fed by a coaxial cable where the inner conductor continues vertically into the cavity.

Sources in a lossless cavity

The Maxwell equations read

$$\begin{aligned}\nabla \times \mathbf{E}(\mathbf{r}) &= i\omega\mu_0\mathbf{H}(\mathbf{r}) \\ \nabla \times \mathbf{H}(\mathbf{r}) &= \mathbf{J}(\mathbf{r}, \omega) - i\omega\epsilon_0\mathbf{E}(\mathbf{r})\end{aligned}$$

The following identity is used

$$0 = \iiint_V \nabla \cdot (\mathbf{E}_n(\mathbf{r}) \times \mathbf{H}(\mathbf{r})) dV = \iiint_V \mathbf{H}(\mathbf{r}) \cdot \nabla \times \mathbf{E}_n(\mathbf{r}) - \mathbf{E}_n(\mathbf{r}) \cdot \nabla \times \mathbf{H}(\mathbf{r}) dV \quad (7.57)$$

We utilize the frequency domain version of the expansion in Eq. (7.4) and expand the electric and magnetic fields

$$\mathbf{E}(\mathbf{r}) = \sum_{n=0}^{\infty} e_n(\omega) \mathbf{E}_n(\mathbf{r}) + \text{irrotational electric field} \quad (7.58)$$

$$\mathbf{H}(\mathbf{r}) = (i\omega\mu_0)^{-1} \sum_{n=0}^{\infty} e_n(\omega) \nabla \times \mathbf{E}_n(\mathbf{r}) + \text{irrotational magnetic field}$$

where $e_n(\omega)$ is the Fourier transform of $e_n(t)$. The rotation of the irrotational magnetic field is zero and thus

$$\nabla \times \mathbf{H}(\mathbf{r}) = (i\omega\mu_0)^{-1} \sum_{n=0}^{\infty} e_n(\omega) \nabla \times (\nabla \times \mathbf{E}_n(\mathbf{r})) = (i\omega\mu_0)^{-1} \sum_{n=0}^{\infty} e_n(\omega) k_n^2 \mathbf{E}_n(\mathbf{r}) \quad (7.59)$$

where we used $\nabla \times (\nabla \times \mathbf{E}_n(\mathbf{r})) = -\nabla^2 \mathbf{E}_n(\mathbf{r}) = k_n^2 \mathbf{E}_n(\mathbf{r})$.

The Ampère law gives

$$\nabla \times \mathbf{H}(\mathbf{r}) = \mathbf{J}(\mathbf{r}, \omega) - i\omega\epsilon_0 \left(\sum_{n=0}^{\infty} e_n(\omega) \mathbf{E}_n(\mathbf{r}) + \text{irrotational electric field} \right) \quad (7.60)$$

We now determine the equation for the coefficients $e_n(\omega)$. We insert the expansions in (7.58) and (7.59) into (7.60), take the scalar product with $\mathbf{E}_n$, integrate over V , and use the orthogonality of the vector functions (7.2). Thus

$$0 = -i \frac{k_n^2}{\omega \mu_0} e_n(\omega) - \iiint_V \mathbf{J}(\mathbf{r}, \omega) \cdot \mathbf{E}_n(\mathbf{r}) dV + i\omega \epsilon_0 e_n(\omega)$$

Thus the amplitudes of the electric solenoidal modes are

$$e_n(\omega) = -i\omega \mu_0 \frac{\iiint_V \mathbf{J}(\mathbf{r}, \omega) \cdot \mathbf{E}_n(\mathbf{r}) dV}{k^2 - k_n^2} \quad (7.61)$$

The singularity at $k = k_n$ implies that $e_n(t)$ has a part that oscillates with angular frequency ω_n also after the source is turned off. The amplitude of the oscillating part can be obtained by Fourier transformation of (7.61).

Time-domain equations

We obtain the time domain relation for $e_n(t)$ by Fourier transformation of (7.61). Since multiplication with $-i\omega$ 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 (7.61)

$$\frac{\partial^2 e_n(t)}{\partial t^2} + \omega_n^2 e_n(t) = -\epsilon_0^{-1} \frac{\partial}{\partial t} \iiint_V \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{E}_n(\mathbf{r}) dV \quad (7.62)$$

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 (7.62)

$$\frac{\partial^2 e_n(t)}{\partial t^2} + 2\alpha_n \frac{\partial e_n(t)}{\partial t} + \omega_n^2 e_n(t) = -\epsilon_0^{-1} \frac{\partial}{\partial t} \iiint_V \mathbf{J}(\mathbf{r}, t) \cdot \mathbf{E}_n(\mathbf{r}) dV \quad (7.63)$$

This leads us to the relation in the frequency domain

$$e_n(\omega) = -i\omega \mu_0 \frac{\iiint_V \mathbf{J}(\mathbf{r}, \omega) \cdot \mathbf{E}_n(\mathbf{r}) dV}{k^2 + 2ik\alpha_n c^{-1} - k_n^2} \quad (7.64)$$

where $k_n = \omega_n/c$.

Problems in Chapter 7

- 7.1** Show that the time averages of the stored electric and magnetic energies in a resonance cavity are equal.

- 7.2** 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.
- 7.3** a) Determine Q for the TE_{101} mode in a rectangular parallelepiped 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).
- 7.4** 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 7

Resonance cavities

$\alpha = \begin{cases} \frac{R_S}{\mu_0} \left(\frac{\varepsilon_\ell}{d} + \frac{1}{2k_{tn}^2} \int_{\Gamma} \hat{\mathbf{n}} \cdot \nabla_T v_n(\boldsymbol{\rho}) ^2 d\Gamma \right) & \text{TM} \\ \frac{R_S}{\mu_0 k_{n\ell}^2} \left(\frac{2k_{z\ell}^2}{d} + \frac{1}{2} \int_{\Gamma} \frac{k_{z\ell}^2}{k_{tn}^2} \hat{\mathbf{n}} \times \nabla_T w_n(\boldsymbol{\rho}) ^2 + k_{tn}^2 w_n(\boldsymbol{\rho}) ^2 d\Gamma \right) & \text{TE} \end{cases}$
$Q = \frac{\omega_{n\ell}}{2\alpha} \quad B = 2\alpha$

Chapter 8

Transients in waveguides

In Chapter 6 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

$$\begin{aligned} E_z(\mathbf{r}, \omega) &= v(\boldsymbol{\rho})a(z, \omega) = v(\boldsymbol{\rho})A(\omega)e^{ik_z(\omega)z} \quad (\text{TM}) \\ H_z(\mathbf{r}, \omega) &= w(\boldsymbol{\rho})a(z, \omega) = w(\boldsymbol{\rho})A(\omega)e^{ik_z(\omega)z} \quad (\text{TE}) \end{aligned} \quad (8.1)$$

where the longitudinal wavenumber is

$$k_z(\omega) = \{k^2(\omega) - k_t^2\}^{1/2} = \left\{ \frac{\omega^2}{c_0^2} - k_t^2 \right\}^{1/2}$$

Note that $v(\boldsymbol{\rho})$, $w(\boldsymbol{\rho})$, k_t^2 and c_0 are independent of the frequency¹. Even the case with material dispersion can be handled (ϵ frequency dependent). We transform the expressions in equation (8.1) to the time domain in order to study the transient wave phenomena in the waveguide. The inverse Fourier transform of equation (8.1) gives the z -component of the fields in a point z

$$\begin{aligned} E_z(\mathbf{r}, t) &= v(\boldsymbol{\rho})a(z, t) \\ H_z(\mathbf{r}, t) &= w(\boldsymbol{\rho})a(z, t) \end{aligned}$$

where

$$a(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} a(z, \omega)e^{-i\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(\omega)e^{ik_z(\omega)z}e^{-i\omega t} d\omega$$

The amplitude $A(\omega)$ is the Fourier transform of $a(0, t)$

$$A(\omega) = \int_{-\infty}^{\infty} a(0, t)e^{i\omega t} dt$$

¹With 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 $\exp\{i\omega z/c_0\}$, that is determined by the asymptotic behavior of the function $k_z(\omega)$ for high frequencies, i.e.,

$$a(z, \omega) = A(\omega)e^{i\omega z/c_0} + A(\omega)e^{i\omega z/c_0} \left(e^{ik_z(\omega)z - i\omega z/c_0} - 1 \right)$$

Since a product of two Fourier transforms gives a convolution in the time domain, and multiplication with a factor $\exp(-i\omega t_0)$ gives a time delay t_0 , the z -component of the electric field at a point z reads

$$a(z, t) = a(0, t - z/c_0) + \int_{-\infty}^{\infty} P(z, t - z/c_0 - t') a(0, t') dt' \quad (8.2)$$

where

$$\begin{cases} P(z, t - z/c_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(e^{ik_z(\omega)z - i\omega z/c_0} - 1 \right) e^{-i\omega(t - z/c_0)} d\omega \\ a(0, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(\omega) e^{-i\omega t} d\omega \end{cases}$$

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

$$\begin{cases} \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(1 - e^{-ib\omega + ib(\omega^2 - a^2)^{1/2}} \right) e^{-i\omega t} d\omega = H(t) ab \frac{J_1 \left(a\sqrt{t^2 + 2bt} \right)}{\sqrt{t^2 + 2bt}} \\ ab \int_0^{\infty} \frac{J_1 \left(a\sqrt{t^2 + 2bt} \right)}{\sqrt{t^2 + 2bt}} e^{i\omega t} dt = 1 - e^{-ib\omega + ib(\omega^2 - a^2)^{1/2}} \end{cases}$$

where $H(t)$ is the Heaviside function ($H(t) = 1, t \geq 0$, otherwise zero) and furthermore $\text{Im}(\omega^2 - a^2)^{1/2} > 0$, $\arg w \in [0, 2\pi]$. This gives us an explicit expression for $P(z, t)$

$$P(z, t) = -c_0 k_t z \frac{J_1 \left(k_t \sqrt{c_0^2 t^2 + 2z c_0 t} \right)}{\sqrt{c_0^2 t^2 + 2z c_0 t}} H(t)$$

and an expression for the amplitude of the waveguide mode

$$a(z, t + z/c_0) = a(0, t) - c_0 k_t z \int_{-\infty}^t \frac{J_1 \left(k_t \sqrt{c_0^2 (t - t')^2 + 2c_0 z (t - t')} \right)}{\sqrt{c_0^2 (t - t')^2 + 2c_0 z (t - t')}} a(0, t') dt' \quad (8.3)$$

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

$$\begin{cases} \zeta = k_t z \\ s = k_t c_0 t \end{cases} \quad (8.4)$$

are suitable dimensionless parameters to describe the wave propagation in the waveguide. Let u be H_z for a TE mode or E_z for a TM mode, but expressed in the dimensionless

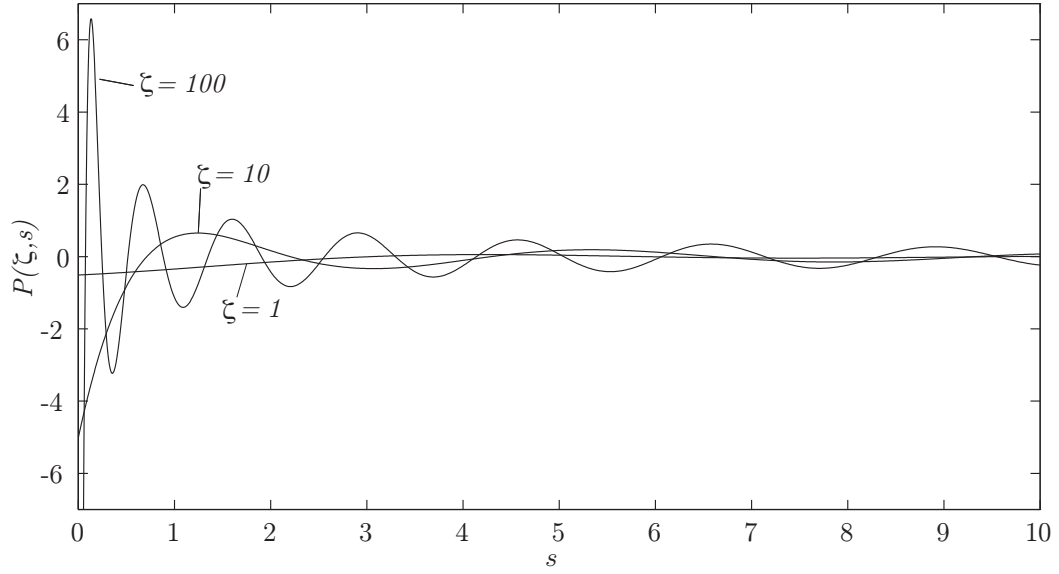


Figure 8.1: The propagator kernel $P(\zeta, s)$ at three different positions ζ . Note that when ζ increases the propagator kernel is compressed and increases in amplitude. One can prove that in the limit $\zeta \rightarrow \infty$ it holds that $P(\zeta, s) \rightarrow -\delta(s)$.

coordinates s and ζ . Then

$$u(\zeta, s + \zeta) = u(0, s) + \int_{-\infty}^s P(\zeta, s - s') u(0, s') ds'$$

$$P(\zeta, s) = -\zeta \frac{J_1\left(\sqrt{s^2 + 2\zeta s}\right)}{\sqrt{s^2 + 2\zeta s}} H(s)$$

The different modes only differ by a scaling in time and space given by equation (8.4). In figure 8.1 the propagator kernel $P(\zeta, s)$ is depicted at three different positions ζ .

Example 8.1

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 8.1 we see that the oscillations of the tail increases when the wave propagates. ■

Example 8.2

The expression (8.3) 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 $\sqrt{c_0^2 t^2 + 2c_0 z t} = i\sqrt{-2c_0 z t - c_0^2 t^2}$. One can show that in this case

$$P(z, t) = -c_0 k_t z \frac{I_1(k_t \sqrt{-2c_0 z t - c_0^2 t^2})}{\sqrt{-2c_0 z t - c_0^2 t^2}} H(t)$$

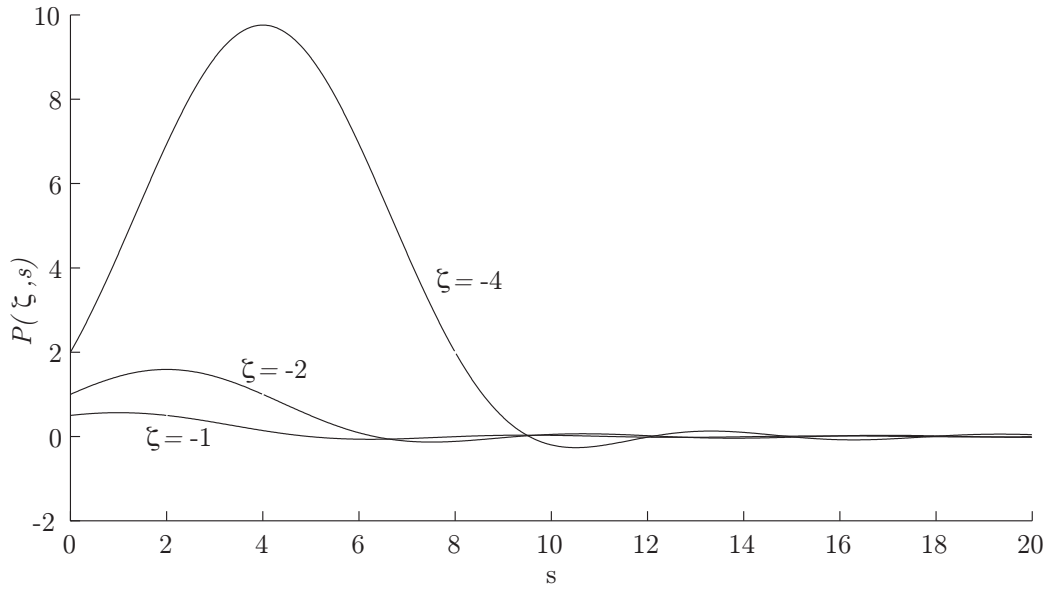


Figure 8.2: The propagator kernel $P(\zeta, s)$ for three values of ζ .

where $I_1(y) = -iJ_1(iy)$ is the modified Bessel function. In figure 8.2 we see the graph of $P(\zeta, s)$ for some negative values of ζ . ■

Problems in Chapter 8

- 8.1** 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 susceptibility 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 the 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 $\Delta t \ll 1/\omega_0$ directly after the wave front, then show that the precursor in a dispersive material satisfies the same equation as the transient field in the waveguide.

- 8.2** 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.

- 8.3** 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 8

The propagator

$$\begin{aligned} E_z(\mathbf{r}, t) &= v(\boldsymbol{\rho}) a(z, t) \\ a(z, t) &= a(0, t - z/c_0) + \int_{-\infty}^{\infty} P(z, t - z/c_0 - t') a(0, t') dt' \\ P(z, t) &= -c_0 k_t z \frac{J_1 \left(k_t \sqrt{c_0^2 t^2 + 2z c_0 t} \right)}{\sqrt{c_0^2 t^2 + 2z c_0 t}} H(t) \end{aligned}$$

Chapter 9

Dielectric waveguides

In this chapter we analyze dielectric waveguides. In these waveguides, electromagnetic waves are guided through a dielectric core that has a lower wave speed, and hence higher index of refraction, than the surrounding medium.

According to Snell's law, total internal reflection occurs when a plane wave impinges on an interface between two dielectric media under two specific conditions: The first is that the refractive index of the incident medium (n_1) is greater than that of the surrounding medium (n_2). The second is that the angle of incidence exceeds the critical angle, given by:

$$\theta_c = \arcsin(n_2/n_1)$$

It is precisely this total internal reflection that allows electromagnetic waves to be trapped and guided within a dielectric slab or cylinder.

A major portion of this chapter is devoted to optical fibers. The development of optical fibers began in 1966 when Charles Kao¹ realized that purifying glass could reduce signal attenuation to below 20 dB/km, making long-distance optical communication feasible. Reaching this threshold took several years, but in 1970, Robert Maurer, Donald Keck, and Peter Schultz at Corning Glass Works successfully produced ultra-pure silica fibers with an attenuation of just 17 dB/km. By 1975, they had further reduced losses to 4 dB/km.

The optical fibers developed at Corning Glass Works share the basic design of modern fiber-optic systems. They consist of a circular silica cylinder with a central core doped with materials such as germanium dioxide (GeO_2) to elevate its refractive index, thereby trapping and guiding electromagnetic waves within the core.

In standard silica fibers, material attenuation reaches a minimum near an infrared vacuum wavelength of $\lambda = 1.55 \mu\text{m}$, making it the standard operating wavelength for most communications networks.

For long-distance transmission, single-mode fibers are preferred. By restricting the core radius to a few micrometers (typically $r \approx 5\text{--}10 \mu\text{m}$), only a single fundamental mode is supported, which virtually eliminates intermodal dispersion. Conversely, short-distance links often employ multimode fibers. These feature a much larger core radius, allowing hundreds of modes to propagate simultaneously. While multimode systems exhibit higher modal dispersion, they are significantly less expensive to manufacture and couple to light sources.

¹Kao was awarded the 2009 Nobel Prize in Physics for his pioneering work.

Optical fibers offer vastly greater bandwidth than conventional copper wires, with standard single-fiber systems routinely achieving transmission rates of 40 Gbit/s. Beyond their data capacity, optical fibers occupy significantly less space in cable ducts, are entirely immune to electromagnetic interference, and provide high security due to the difficulty of wiretapping. Consequently, optical fibers are far superior to electrical wiring for long-distance communication.

9.1 Planar dielectric waveguides

We begin by examining waves guided by planar dielectric slabs. Consider an infinitely extended plate in the y - and z -directions that occupies the region $0 \leq x \leq d$. The plate is lossless with a relative permittivity ϵ_1 , surrounded by vacuum ($x < 0$ and $x > d$). We restrict our analysis to time-harmonic waves propagating in the z -direction.

Guided waves in the slab decouple into TE modes, where the electric field is oriented along the y -axis, and TM modes, where the electric field lies in the xz -plane. In this section, we focus exclusively on TM modes; the derivation for TE modes is left as an exercise at the end of the chapter.

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

$$v(x) = \begin{cases} Ae^{ik_{0x}(x-d)}, & x > d \\ B \sin(k_{1x}(x-d/2)), & 0 \leq x \leq d \\ -Ae^{-ik_{0x}x}, & x < 0 \end{cases}$$

where $k_0 = \omega/c_0$ is the wave number in vacuum, $k_1 = k_0\sqrt{\epsilon_1}$ 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

$$A = B \sin(k_{1x}d/2)$$

$$ik_{0x}A = \frac{\epsilon_1 k_{0x}^2}{k_{1x}^2} k_{1x} B \cos(k_{1x}d/2)$$

By combining the two equations the equation for non-trivial solutions is obtained

$$\tan(k_{1x}d/2) = -i \frac{\epsilon_1 k_{0x}}{k_{1x}}$$

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 $k_{0x} = iq$. Then

$$k_{1x} = \sqrt{k_1^2 - k_0^2 - q^2}$$

and 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 equation². 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 $\omega \rightarrow 0$ then $k_{1x} \rightarrow 0$ and $\tan(k_{1x}d/2) \rightarrow k_{1x}d/2$. 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 $\omega \rightarrow 0$ 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 $k_0 d \rightarrow 0$ then $k_z \rightarrow k_0$. 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 the wavenumber of the surrounding medium.

²A transcendental function is a function that is not a polynomial

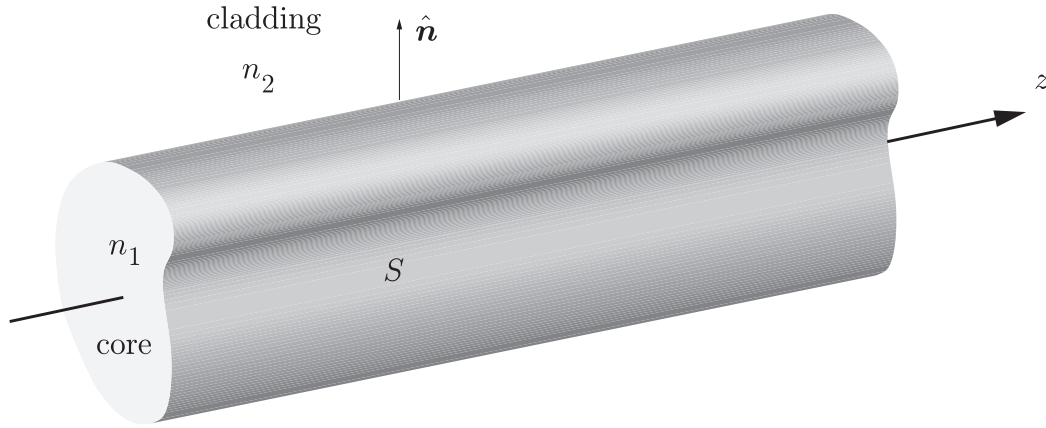


Figure 9.1: The dielectric waveguide with envelope surface S .

9.2 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 , see Figure 9.1. 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

$$\begin{aligned} n_1 &= c_0/c_1 = \sqrt{\epsilon_1\mu_1} \\ n_2 &= c_0/c_2 = \sqrt{\epsilon_2\mu_2} \end{aligned}$$

The wavenumber in vacuum is $k_0 = \omega/c_0$, and the wavenumber in the core and cladding are

$$\begin{aligned} k_1 &= \omega/c_1 = k_0 n_1 \\ k_2 &= \omega/c_2 = k_0 n_2 \end{aligned}$$

9.2.1 The electromagnetic fields

In Chapter 5 we introduced a general decomposition of the electromagnetic fields, that is useful for the dielectric waveguides. Since the fields propagate along the z -direction, we make the following ansatz:

$$\begin{cases} E_z(\mathbf{r}, \omega) = v(\boldsymbol{\rho}, \omega) e^{ik_z(\omega)z} \\ H_z(\mathbf{r}, \omega) = w(\boldsymbol{\rho}, \omega) e^{ik_z(\omega)z} \end{cases}$$

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 satisfy

$$\begin{cases} \nabla_T^2 v(\boldsymbol{\rho}) + k_{1t}^2 v(\boldsymbol{\rho}) = 0 \\ \nabla_T^2 w_p(\boldsymbol{\rho}) + k_{1t}^2 w(\boldsymbol{\rho}) = 0 \end{cases} \quad \text{when } \boldsymbol{\rho} \text{ is inside } \Gamma, \quad (9.1)$$

$$\begin{cases} \nabla_T^2 v(\boldsymbol{\rho}) + k_{2t}^2 v(\boldsymbol{\rho}) = 0 \\ \nabla_T^2 w_p(\boldsymbol{\rho}) + k_{2t}^2 w(\boldsymbol{\rho}) = 0 \end{cases} \quad \text{when } \boldsymbol{\rho} \text{ is outside } \Gamma \quad (9.2)$$

The transverse wavenumbers k_{it} are related to the wavenumbers k_1 , k_2 , and k_z as

$$k_{it}^2 = k_i^2 - k_z^2, \quad i = 1, 2$$

9.2.2 Boundary conditions

We need the boundary conditions at the surface S in order to determine $v(\rho)$, $w(\rho)$ 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 (2.16) on page 6)

$$\begin{cases} \hat{\mathbf{n}} \times \mathbf{E}_1 = \hat{\mathbf{n}} \times \mathbf{E}_2 \\ \hat{\mathbf{n}} \times \mathbf{H}_1 = \hat{\mathbf{n}} \times \mathbf{H}_2 \end{cases} \quad \mathbf{r} \text{ on } S$$

In analogy with the analysis of hollow waveguides we introduce the decomposition of $\mathbf{E}$ and $\mathbf{H}$ fields in a longitudinal and transverse part. In the decomposed fields the boundary conditions are

$$\begin{cases} \hat{\mathbf{n}} \times (\mathbf{E}_T + \hat{\mathbf{z}}E_z) \\ \hat{\mathbf{n}} \times (\mathbf{H}_T + \hat{\mathbf{z}}H_z) \end{cases} \quad \text{are continuous over } S$$

It follows that

$$\begin{cases} E_z \\ H_z \end{cases} \quad \text{and} \quad \begin{cases} \hat{\mathbf{n}} \times \mathbf{E}_T \\ \hat{\mathbf{n}} \times \mathbf{H}_T \end{cases} \quad \text{are continuous over } S$$

We use (5.7) to express the transverse field components in the longitudinal components, and find that the following quantities are continuous over S :

$$\begin{cases} E_z \\ H_z \end{cases} \quad \text{and} \quad \begin{cases} \frac{1}{k_t^2} \left[k_z \hat{\mathbf{z}} \cdot (\hat{\mathbf{n}} \times \nabla_T E_z) - \omega \mu_0 \mu \frac{\partial H_z}{\partial n} \right] \\ \frac{1}{k_t^2} \left[k_z \hat{\mathbf{z}} \cdot (\hat{\mathbf{n}} \times \nabla_T H_z) + \omega \epsilon_0 \epsilon \frac{\partial E_z}{\partial n} \right] \end{cases}$$

Since $e^{ik_z z}$ is a common factor for all of the terms, the continuity is also valid for:

$$\begin{cases} v(\rho) \\ w(\rho) \end{cases} \quad \text{and} \quad \begin{cases} \frac{1}{k_t^2} \left[k_z \hat{\mathbf{z}} \cdot (\hat{\mathbf{n}} \times \nabla_T v(\rho)) - \omega \mu_0 \mu \frac{\partial w(\rho)}{\partial n} \right] \\ \frac{1}{k_t^2} \left[k_z \hat{\mathbf{z}} \cdot (\hat{\mathbf{n}} \times \nabla_T w(\rho)) + \omega \epsilon_0 \epsilon \frac{\partial v(\rho)}{\partial n} \right] \end{cases} \quad (9.3)$$

where k_t, ϵ, μ are $k_{1t}, \epsilon_1, \mu_1$ in the core and $k_{2t}, \epsilon_2, \mu_2$ in the cladding. 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.

9.3 Circular dielectric waveguide

We now focus on a circular dielectric waveguide. As shown in Figure 9.2, the structure consists of a cylinder core of radius $\rho = a$ with material parameters ϵ_1 and μ_1 , surrounded by a cladding region ($\rho > a$) with parameters ϵ_2 and μ_2 . Both regions are initially assumed to be lossless (ϵ and μ are real), losses can be incorporated into the results later.

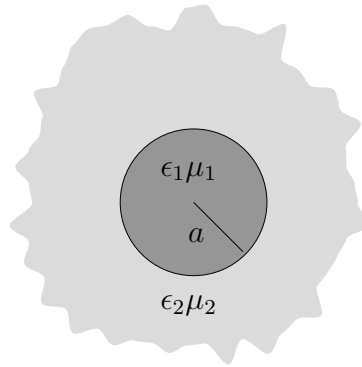


Figure 9.2: Geometry for the circular dielectric waveguide.

Similar to a hollow waveguide, a circular dielectric waveguide supports a finite number of propagating modes at any given frequency. Crucially, one fundamental mode has a zero cutoff frequency, allowing it to propagate at all frequencies. If the core radius is sufficiently small, this becomes the only supported propagating mode—a regime known as single-mode operation.

9.3.1 Waveguide modes

In this section, we derive the field expressions, dispersion relations, cutoff frequencies, and power flow for the modes that can exist in a circular dielectric waveguide.

As established, $v(\rho, \phi)$ and $w(\rho, \phi)$ satisfy the Helmholtz equations (9.1) and (9.2). In cylindrical coordinates the equations read

$$\nabla_T^2 = \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \quad (9.4)$$

A separation of Helmholtz equation in cylindrical coordinates gives the eigenfunction system in the azimuthal angle $\{1, \sin m\phi, \cos m\phi\}$, where $m = 1, 2, \dots$, c.f., section 6.5.3 on page 89. 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 first consider modes with $m > 0$, for which the two choices are physically equivalent. Choosing E_z to be an odd function of ϕ (and consequently H_z as an even function), and restricting attention to waves propagating in the positive z -direction, we obtain:

$$\begin{aligned} v(\rho, \phi) &= \psi_E(\rho) \sin m\phi \\ w(\rho, \phi) &= \psi_H(\rho) \cos m\phi \end{aligned}$$

where the functions ψ_E and ψ_H satisfy Bessel's differential equation

$$\left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} - \frac{m^2}{\rho^2} + k_t^2 \right) \psi(\rho) = 0 \quad (9.5)$$

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.

- In the core $\rho < a$ we use the combination

$$\psi(\rho) = AJ_m(k_{1t}\rho) + \tilde{A}N_m(k_{1t}\rho) \quad (9.6)$$

where J_m is the Bessel function and N_m is the Neumann function of order m .

- If k_{2t} is real we use the combination

$$\psi(\rho) = AH_m^1(k_{2t}\rho) + \tilde{A}H_m^2(k_{2t}\rho) \quad (9.7)$$

in the cladding $\rho > a$, where H_m^1 and H_m^2 are the Hankel functions of the first and second kind, respectively, see Appendix A.

- If k_{2t} is imaginary we use

$$\psi(\rho) = AK_m(q\rho) + \tilde{A}I_m(q\rho) \quad (9.8)$$

where $q = -ik_{2t}$ and where K_m and I_m are modified Bessel functions, see Appendix A.

We now find the 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 $k_{2t}^2 < 0$. We require that $v_m(\rho, \phi)$ and $w_m(\rho, \phi)$ are finite at $\rho = 0$ and $\rho = \infty$. Since $N_m(k_{1t}\rho)$ is infinite at $\rho = 0$ and $I_m(q\rho)$ is infinite at $\rho = \infty$ the proper ansatz when $m > 0$ is

$$\begin{cases} v_m(\rho, \phi) = AJ_m(h\rho) \sin m\phi \\ w_m(\rho, \phi) = BJ_m(h\rho) \cos m\phi \end{cases} \quad \rho < a \quad (9.9)$$

$$\begin{cases} v_m(\rho, \phi) = CK_m(q\rho) \sin m\phi \\ w_m(\rho, \phi) = DK_m(q\rho) \cos m\phi \end{cases} \quad \rho > a \quad (9.10)$$

where

$$\begin{cases} h = k_{1t} \\ q = -ik_{2t} \end{cases}$$

If we instead let $v_m(\rho, \phi)$ be an even function of ϕ we make the ansatz

$$\begin{cases} v_m(\rho, \phi) = AJ_m(h\rho) \cos m\phi \\ w_m(\rho, \phi) = -BJ_m(h\rho) \sin m\phi \end{cases} \quad \rho < a \quad (9.11)$$

$$\begin{cases} v_m(\rho, \phi) = CK_m(q\rho) \cos m\phi \\ w_m(\rho, \phi) = -DK_m(q\rho) \sin m\phi \end{cases} \quad \rho > a \quad (9.12)$$

The ansatz (9.9) and (9.10) and the ansatz (9.11) and (9.12) give identical characteristic equations for the relations between A , B , C , and D .

To relate the Fourier coefficients A , B , C , and D , and to determine the longitudinal wavenumber k_z (and consequently h and q), we apply the continuity of the fields in (9.3) at $\rho = a$. Since $\hat{n} = \hat{\rho}$ for a circle it follows from (9.3) that

$$\begin{cases} v(a^-, \phi) = v(a^+, \phi) \\ w(a^-, \phi) = w(a^+, \phi) \\ \frac{1}{h^2} \left[\frac{k_z}{a} \frac{\partial v(a^-, \phi)}{\partial \phi} - \omega\mu_0\mu_1 \frac{\partial w(a^-, \phi)}{\partial \rho} \right] = -\frac{1}{q^2} \left[\frac{k_z}{a} \frac{\partial v(a^+, \phi)}{\partial \phi} - \omega\mu_0\mu_2 \frac{\partial w(a^+, \phi)}{\partial \rho} \right] \\ \frac{1}{h^2} \left[\frac{k_z}{a} \frac{\partial w(a^-, \phi)}{\partial \phi} + \omega\epsilon_0\epsilon_1 \frac{\partial v(a^-, \phi)}{\partial \rho} \right] = -\frac{1}{q^2} \left[\frac{k_z}{a} \frac{\partial w(a^+, \phi)}{\partial \phi} + \omega\epsilon_0\epsilon_2 \frac{\partial v(a^+, \phi)}{\partial \rho} \right] \end{cases}$$

By inserting (9.11) and (9.12) we obtain

$$\begin{pmatrix} J_m(ha) & 0 & -K_m(qa) & 0 \\ 0 & J_m(ha) & 0 & -K_m(qa) \\ \frac{mk_z}{h^2a} J_m(ha) & -\frac{\omega\mu_0\mu_1}{h} J'_m(ha) & \frac{mk_z}{q^2a} K_m(qa) & -\frac{\omega\mu_0\mu_2}{q} K'_m(qa) \\ \frac{\omega\epsilon_0\epsilon_1}{h} J'_m(ha) & -\frac{mk_z}{h^2a} J_m(ha) & \frac{\omega\epsilon_0\epsilon_2}{q} K'_m(qa) & -\frac{mk_z}{q^2a} K_m(qa) \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (9.13)$$

Non-trivial solutions exist if and only if the coefficient matrix has a vanishing determinant:

$$\begin{vmatrix} J_m(ha) & 0 & -K_m(qa) & 0 \\ 0 & J_m(ha) & 0 & -K_m(qa) \\ \frac{mk_z}{h^2a} J_m(ha) & -\frac{\omega\mu_0\mu_1}{h} J'_m(ha) & \frac{mk_z}{q^2a} K_m(qa) & -\frac{\omega\mu_0\mu_2}{q} K'_m(qa) \\ \frac{\omega\epsilon_0\epsilon_1}{h} J'_m(ha) & -\frac{mk_z}{h^2a} J_m(ha) & \frac{\omega\epsilon_0\epsilon_2}{q} K'_m(qa) & -\frac{mk_z}{q^2a} K_m(qa) \end{vmatrix} = 0$$

This condition forms the characteristic equation determining the longitudinal wavenumber k_z . Expanding the determinant along the top row yields

$$\begin{aligned} J_m(ha) & \begin{vmatrix} 0 & -K_m(qa) & 0 \\ -\frac{\omega\mu_1}{c_0h} J'_m(ha) & \frac{mk_z}{q^2a} K_m(qa) & -\frac{\omega\mu_2}{c_0q} K'_m(qa) \\ -\frac{mk_z}{h^2a} J_m(ha) & \frac{\omega\epsilon_2}{c_0q} K'_m(qa) & -\frac{mk_z}{q^2a} K_m(qa) \end{vmatrix} \\ -K_m(qa) & \begin{vmatrix} 0 & J_m(ha) & -K_m(qa) \\ \frac{mk_z}{h^2a} J_m(ha) & -\frac{\omega\mu_1}{c_0h} J'_m(ha) & -\frac{\omega\mu_2}{c_0q} K'_m(qa) \\ \frac{\omega\epsilon_1}{c_0h} J'_m(ha) & -\frac{mk_z}{h^2a} J_m(ha) & -\frac{mk_z}{q^2a} K_m(qa) \end{vmatrix} = 0 \end{aligned}$$

Evaluating these subdeterminants and grouping terms gives the characteristic equation

$$\begin{aligned} m^2 \left(\frac{1}{q^2a^2} + \frac{1}{h^2a^2} \right)^2 \left(\frac{k_z}{k_0} \right)^2 \\ = \left(\frac{\mu_1 J'_m(ha)}{haJ_m(ha)} + \frac{\mu_2 K'_m(qa)}{qaK_m(qa)} \right) \left(\frac{\epsilon_1 J'_m(ha)}{haJ_m(ha)} + \frac{\epsilon_2 K'_m(qa)}{qaK_m(qa)} \right) \end{aligned} \quad (9.14)$$

where $k_0 = \omega/c_0 = \omega\sqrt{\epsilon_0\mu_0}$ is the free space wave number. This equation is quadratic in $J'_m(ha)/(haJ_m(ha))$ and has two solutions:

$$\begin{aligned} \frac{J'_m(ha)}{haJ_m(ha)} &= -\frac{K'_m(qa)}{2qaK_m(qa)} \frac{\mu_1\epsilon_2 + \mu_2\epsilon_1}{\mu_1\epsilon_1} \\ &\pm \left[\left(\frac{K'_m(qa)}{2qaK_m(qa)} \frac{\mu_1\epsilon_2 - \mu_2\epsilon_1}{\mu_1\epsilon_1} \right)^2 + \left(\frac{mk_z}{k_1} \right)^2 \left(\frac{1}{q^2a^2} + \frac{1}{h^2a^2} \right)^2 \right]^{1/2} \end{aligned} \quad (9.15)$$

Equation (9.14) yields two distinct sets of solutions—corresponding to the plus and minus signs—which define two separate families of waveguide modes. The solutions associated with the plus sign are designated as *EH modes*, while those corresponding to the

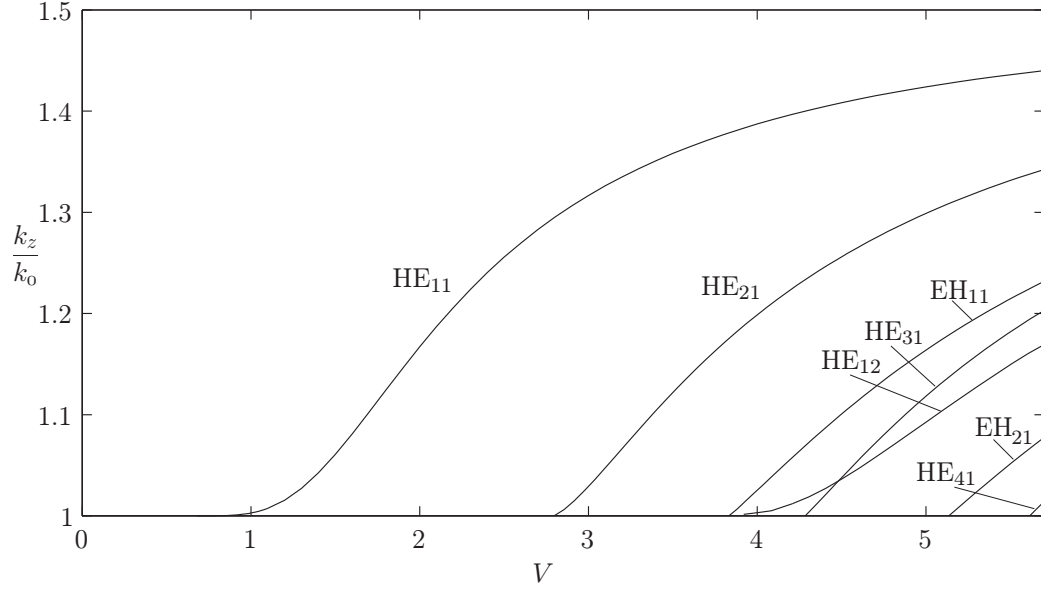


Figure 9.3: The normalized wavenumber k_z/k_0 as a function of the normalized frequency $V = k_0 a \sqrt{(n_1^2 - n_2^2)}$ for the lowest HE and EH modes when $n_1 = 1.5$ and $n_2 = 1$.

minus sign are designated as *HE modes*. These characteristic equations determine the longitudinal wavenumber k_z . The corresponding transverse wavenumbers in the core and cladding are $h = \sqrt{k_1^2 - k_z^2}$ and $k_{2t} = iq$.

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 (9.10) then is real and positive. No active power is then 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 $q = 0$. At the cut-off frequency we have

$$h_{\text{cut-off}} = \sqrt{k_1^2 - k_z^2} = k_0 \sqrt{n_1^2 - n_2^2}$$

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. This fourth coefficient sets the overall amplitude of the mode, which requires knowing the field's source to determine. From the equations in (9.13) we get the following relations:

$$\frac{C}{A} = \frac{D}{B} = \frac{J_m(ha)}{K_m(qa)} \quad (9.16)$$

$$\frac{B}{A} = \frac{D}{C} = \frac{mk_z}{k_0} \left(\frac{1}{h^2 a^2} + \frac{1}{q^2 a^2} \right) \left(\frac{\mu_1 J'_m(ha)}{ha J_m(ha)} + \frac{\mu_2 K'_m(qa)}{qa K_m(qa)} \right)^{-1} \quad (9.17)$$

9.3.2 HE modes

For $m \neq 0$ only modes with non-zero E_z and H_z components can exist. These modes are called *hybrid modes* and are distinct from pure TE or TM modes. The characteristic

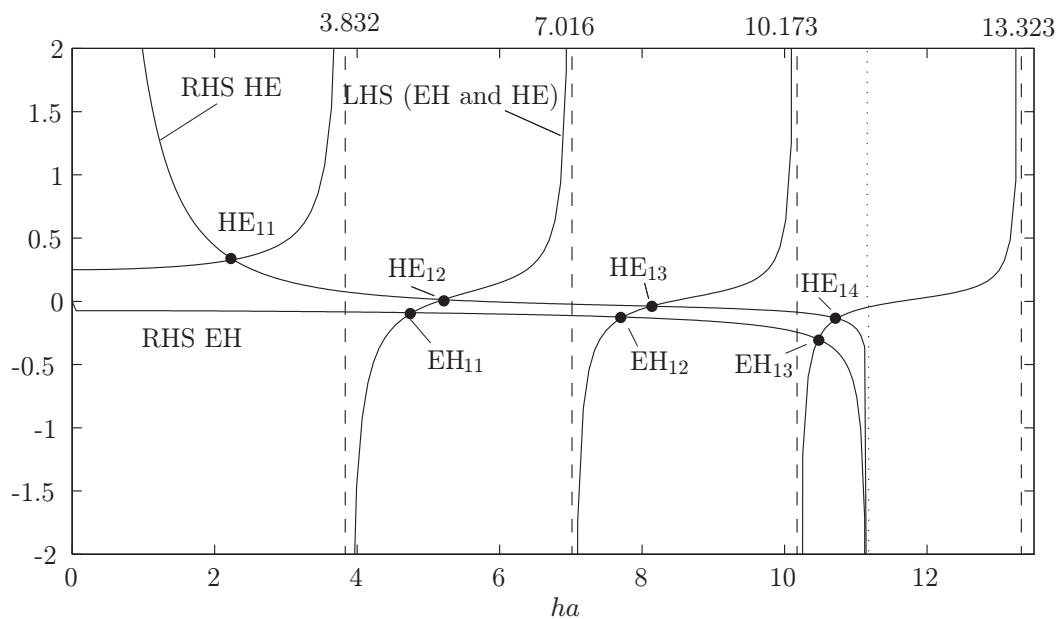


Figure 9.4: The left and right hand sides for the characteristic equation for the EH- and HE-modes with $m = 1$, $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.5$, $\sqrt{\epsilon_2} = n_2 = 1$ and $k_0a = 10$. The points of interaction give the transverse wavenumber h for the propagating modes. The vertical dotted line $ha = k_0a\sqrt{n_1^2 - n_2^2}$ is the limit for the right hand sides of the characteristic equations.

equation that determines the longitudinal wavenumber for the HE modes can be written as, see (9.15),

$$\begin{aligned} \frac{J_{m+1}(ha)}{haJ_m(ha)} &= \frac{K'_m(qa)}{2qaK_m(qa)} \frac{\mu_1\epsilon_2 + \mu_2\epsilon_1}{\mu_1\epsilon_1} + \frac{m}{(ha)^2} \\ &+ \left[\left(\frac{K'_m(qa)}{2qaK_m(qa)} \frac{\mu_1\epsilon_2 - \mu_2\epsilon_1}{\mu_1\epsilon_1} \right)^2 + \left(\frac{mk_z}{k_1} \right)^2 \left(\frac{1}{q^2a^2} + \frac{1}{h^2a^2} \right)^2 \right]^{1/2} \end{aligned} \quad (9.18)$$

Figure 9.4 plots both sides of this equation as a function of ha for $m = 1$. The parameter values used are $m = 1$, $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.5$, $\sqrt{\epsilon_2} = n_2 = 1$ and $k_0a = 10$. 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 $q \rightarrow 0$, i.e., where the right hand sides approach $-\infty$.

As the frequency decreases, the dotted line at $ha = k_0a\sqrt{n_1^2 - n_2^2}$ shifts to the left, reducing the number of intersections and, consequently, the number of propagating modes. However, regardless of how low the frequency drops, at least one intersection always remains. This guarantees that the fundamental $m = 1$ HE mode propagates across all frequencies—a critical property that forms the basis of single-mode optical fibers.

Determining the cutoff frequencies for HE_{mn} modes requires careful asymptotic analysis. For $m = 1$, taking the first two terms in the power series expansion of $K_1(qa)$ and $K_2(qa)$ reveals that the right-hand side of equation (9.18) diverges as $qa \rightarrow 0$. Conse-

quently, the cutoff frequencies are given by

$$f_{1n} = \frac{c_0}{2\pi a} \frac{\xi_{1n}}{\sqrt{n_1^2 - n_2^2}} \quad (9.19)$$

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 $qa \rightarrow 0$ one can prove that the cut-off frequencies for the modes HE_{mn} are given by

$$f_{mn} = \frac{c_0}{2\pi a} \frac{x_{mn}}{\sqrt{n_1^2 - n_2^2}} \quad (9.20)$$

where

$$x_{mn} \frac{J_m(x_{mn})}{J_{m-1}(x_{mn})} = (m-1) \frac{n_1^2 + n_2^2}{n_2^2}, \quad m > 1$$

The dispersion curves for the lowest HE modes are depicted in Figure 9.3. 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 $V J_m(V)/J_{m-1}(V) = (m-1)(n_1^2 + n_2^2)/n_2^2$.

9.3.3 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 (9.15), can be written as

$$\begin{aligned} \frac{J_{m+1}(ha)}{ha J_m(ha)} &= \frac{K'_m(qa)}{2qa K_m(qa)} \frac{\mu_1 \epsilon_2 + \mu_2 \epsilon_1}{\mu_1 \epsilon_1} + \frac{m}{(ha)^2} \\ &- \left[\left(\frac{K'_m(qa)}{2qa K_m(qa)} \frac{\mu_1 \epsilon_2 - \mu_2 \epsilon_1}{\mu_1 \epsilon_1} \right)^2 + \left(\frac{mk_z}{k_1} \right)^2 \left(\frac{1}{q^2 a^2} + \frac{1}{h^2 a^2} \right)^2 \right]^{1/2} \end{aligned}$$

The left-hand side remains identical to that of the HE modes, whereas the right-hand side is altered. Figure 9.4 plots both sides of the characteristic equation for $m = 1$ as functions of ha , using parameters $\mu_1 = \mu_2 = 1$, $n_1 = \sqrt{\epsilon_1} = 1.5$, $n_2 = \sqrt{\epsilon_2} = 1$, and $k_0 a = 10$. The intersection points yield the dispersion solutions; for this particular frequency, exactly three propagating EH modes exist for $m = 1$.

Vertical dashed lines indicate where the left-hand side diverges to infinity, whereas the vertical dotted line marks the cutoff limit $q \rightarrow 0$, where the right-hand side approaches $-\infty$. As the frequency decreases below $k_0 a \sqrt{n_1^2 - n_2^2} < 3.832$, no intersection points remain, and thus no propagating EH modes exist.

The dispersion curves for the lowest EH modes are depicted in Figure 9.3. The graphs give the normalized longitudinal wavenumber, k_z/k_0 , plotted as a function of the normalized frequency $V = k_0 a \sqrt{n_1^2 - n_2^2}$ when $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.5$ and $\sqrt{\epsilon_2} = n_2 = 1$.

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. 9.25 we see that the right hand side goes to infinity when $q \rightarrow 0$ 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 9.3 the dispersion curves start at the cut-off frequency V , where $J_m(V) = 0$ and $k_z = k_2$.

9.3.4 TE and TM modes

The axially symmetric solutions have $m = 0$, where the odd modes are TE modes with $E_z = 0$ and $H_z \neq 0$. The even modes have $E_z \neq 0$ and $H_z = 0$ and are TM modes. When $m = 0$, the system of equations (9.13) is split up into two independent systems

$$\begin{pmatrix} J_0(ha) & -K_0(qa) \\ \frac{\epsilon_1}{ha} J'_0(ha) & \frac{\epsilon_2}{qa} K'_0(qa) \end{pmatrix} \begin{pmatrix} A_0 \\ C_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (9.21)$$

for TM modes and

$$\begin{pmatrix} J_0(ha) & -K_0(qa) \\ \frac{\mu_1}{ha} J'_0(ha) & \frac{\mu_2}{qa} K'_0(qa) \end{pmatrix} \begin{pmatrix} B_0 \\ D_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (9.22)$$

for TE modes. If we utilize that $J'_0 = -J_1$ and $K'_0 = -K_1$ the determinant condition gives the characteristic equation

$$\frac{J_1(ha)}{haJ_0(ha)} = -\frac{\epsilon_2}{\epsilon_1} \frac{K_1(qa)}{qaK_0(qa)} \quad (9.23)$$

that determines k_z for the TM modes and

$$\frac{J_1(ha)}{haJ_0(ha)} = -\frac{\mu_2}{\mu_1} \frac{K_1(qa)}{qaK_0(qa)} \quad (9.24)$$

that determines k_z for the TE modes. The left and right hand sides are plotted as functions of ha in Figure 9.5. 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 $k_0a = 10$. 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 (A.6) on page 194)

$$\begin{aligned} K_0(qa) &\sim -\ln(qa/2) - \gamma + O((qa)^2) \\ K_m(qa) &\sim 2^{m-1}(m-1)!(qa)^{-m} \end{aligned} \quad (9.25)$$

The right hand side goes to $-\infty$ when qa goes to zero. This point corresponds to $ha = k_0a\sqrt{n_1^2 - n_2^2}$ and is indicated by a dotted vertical line in Figure 9.5. 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

$$f_{0n} = \frac{c_0}{2\pi a} \frac{\xi_{0n}}{\sqrt{n_1^2 - n_2^2}}$$

where $J_0(\xi_{0n}) = 0$

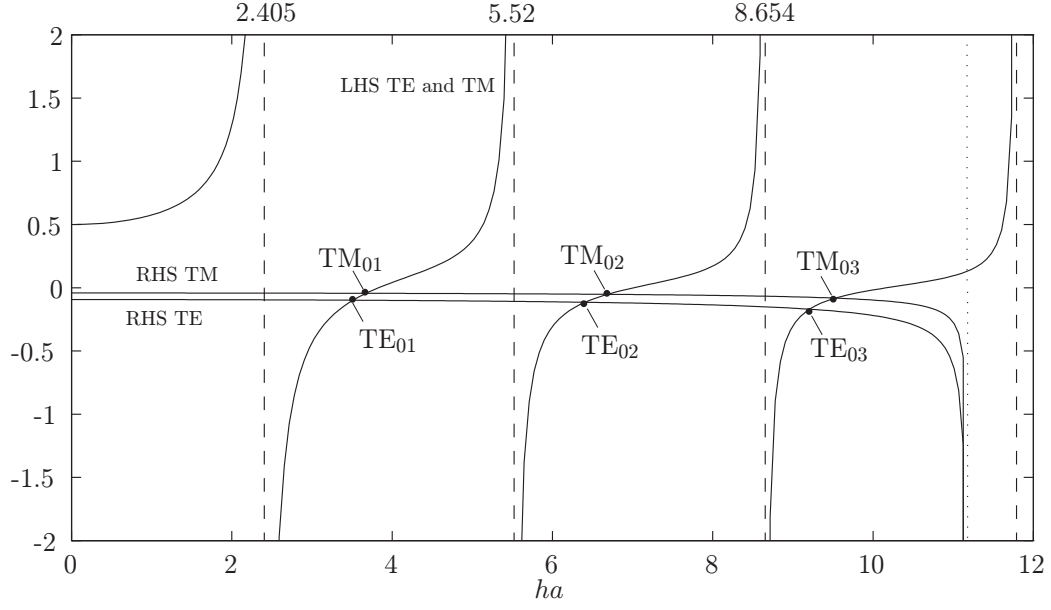


Figure 9.5: Left and right hand sides of the characteristic equations (9.24) and (9.23) for the TE and TM modes when $n_1 = 1.5$, $n_2 = 1$ and $k_0 a = 10$. The points of intersection give the transverse wavenumber $h = k_{1t}$ for the propagating modes. The dotted vertical line indicates the highest 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 190. The smallest zero is $\xi_{01} = 2.405$ which means that for frequencies $f < 2.405c_0/2\pi a\sqrt{n_1^2 - n_2^2}$ there are no propagating TE- and TM modes.

9.4 Optical fibers

An optical fiber is a dielectric waveguide designed for extremely low signal attenuation at its operating wavelength. The most common configuration is a circular silica (SiO_2) fiber with a doped core. This doping slightly raises the core's refractive index n_1 above that of the surrounding cladding n_2 , see Figure 9.6.

Fiber-optic communication systems typically operate at wavelengths where transmission losses are minimized. In Figure 9.6 we see the losses in a 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 $\lambda = 1.55 \mu\text{m}$, 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 $k_0 a \sqrt{n_1^2 - n_2^2} < 2.405$.

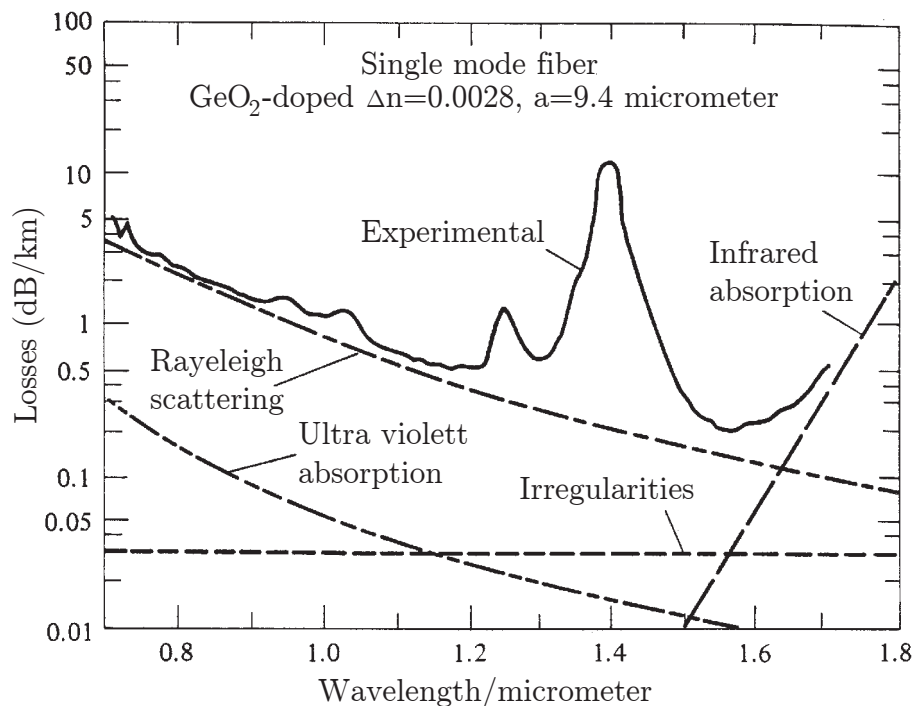


Figure 9.6: Spectrum for the losses for a single mode fiber made out of silica and with the core doped with GeO_2 .

Since 1980 communication systems based on optical fibers are invaluable for both long 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 ³:

- 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.

³A 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.

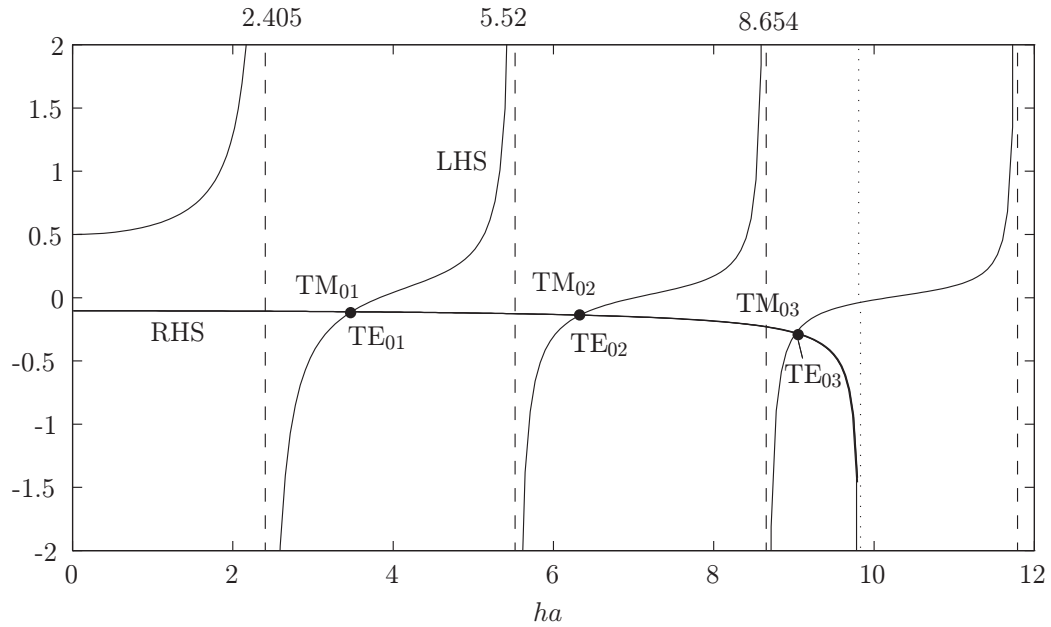


Figure 9.7: The left and right hand sides of the characteristic equations for the TE and TM modes when $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.52$, $\sqrt{\epsilon_2} = n_2 = 1.50$, and $k_0 a = 40$. The points of intersection give the transverse wavenumber $h = k_{1t}$ for the propagating 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.

- 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.
- 2010 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.

Example 9.1

If we assume a fiber with $n_1 = 1.5028$ and $n_2 = 1.5$ where infrared light with vacuum wavelength $\lambda = 1.55 \mu\text{m}$ 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 $a < 6.5 \mu\text{m}$. ■

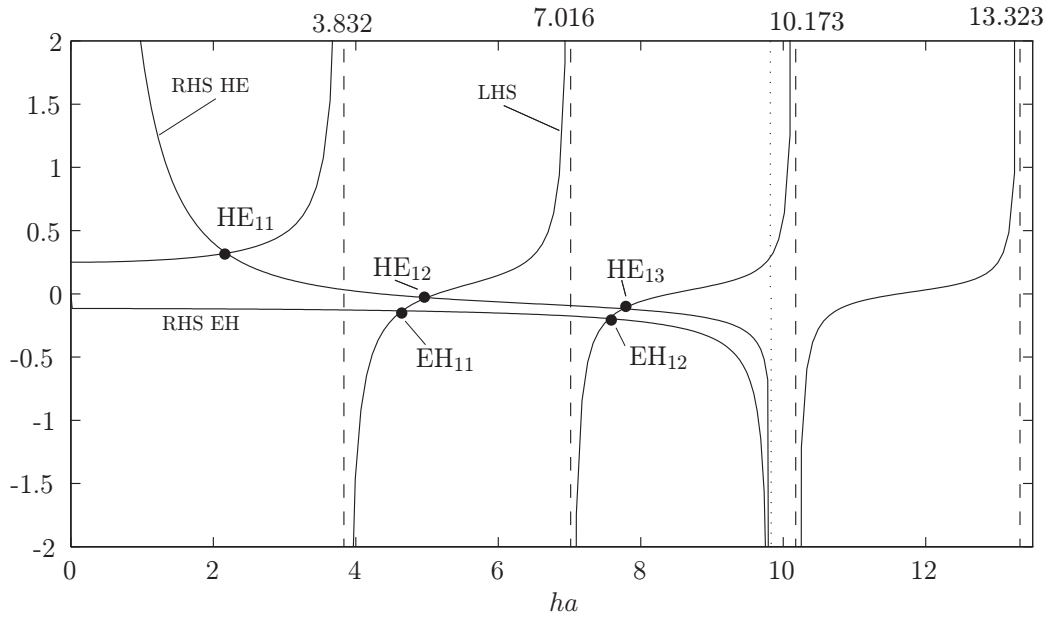


Figure 9.8: The left and right hand sides of the characteristic equations for the EH and HE modes with $m = 1$. The optical fiber has $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.52$, $\sqrt{\epsilon_2} = n_2 = 1.50$ and $k_0a = 40$. The points of intersection give the transverse wavenumber h for the propagating modes.

Figure 9.7 shows the left and right hand sides of the characteristic equations (9.24) and (9.23) for the TE and TM modes, plotted as functions of ha . The fiber has $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.52$, $\sqrt{\epsilon_2} = n_2 = 1.50$ and $k_0a = 40$. 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 $q \rightarrow 0$ the right hand side goes to $-\infty$. This point corresponds to $ha = k_0a\sqrt{n_1^2 - n_2^2}$ 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 9.8. At $k_0a = 40$ we see that the modes HE_{11} , HE_{12} , HE_{13} , EH_{11} and EH_{12} propagate. In Figure 9.9 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 $V = k_0a\sqrt{n_1^2 - n_2^2}$.

9.4.1 Effective index of refraction and phase velocity

For a propagating mode we can write the longitudinal wavenumber k_z as:

$$k_z = n_e k_0 \quad (9.26)$$

where $k_0 = \omega/c_0$ is the vacuum wavenumber and n_e is the *effective index of refraction*. The *phase velocity* for a propagating mode is defined by (c.f., the definition on page 77)

$$v_f = \frac{\omega}{k_z} = \frac{c_0}{n_e} \quad (9.27)$$

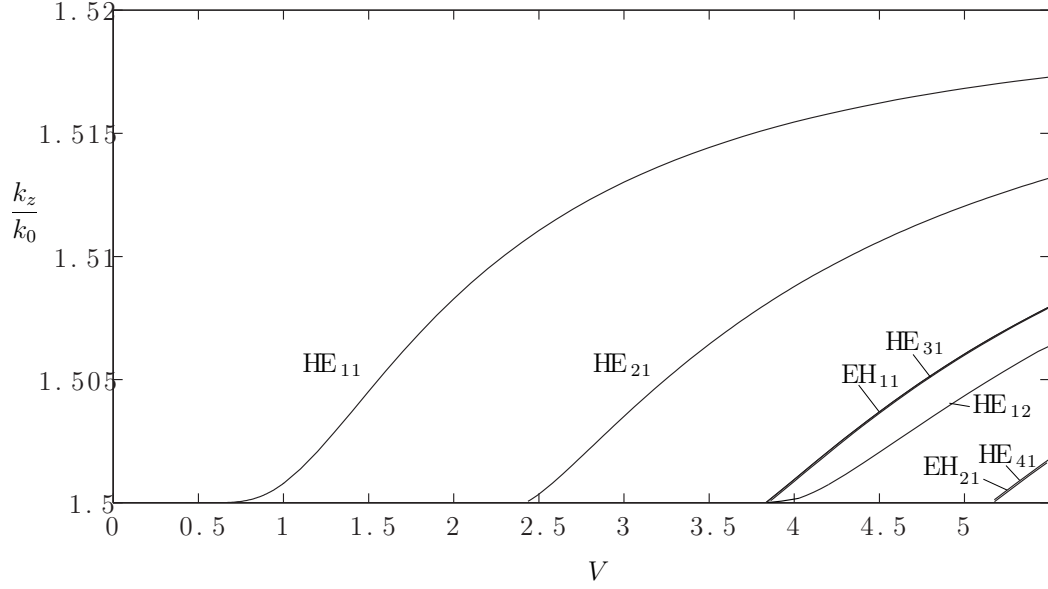


Figure 9.9: The normalized longitudinal wavenumber k_z/k_0 as a function of the normalized frequency $V = k_0 a \sqrt{(n_1^2 - n_2^2)}$ for the lowest HE and EH modes when $\mu_1 = \mu_2 = 1$, $\sqrt{\epsilon_1} = n_1 = 1.52$ and $\sqrt{\epsilon_2} = n_2 = 1.50$. We see that the dispersion curves for HE_{31} and EH_{11} are almost identical. These form together a linearly polarized mode.

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 $q^2 > 0$. Thus $k_2 < k_z < k_1$, $n_2 < n_e < n_1$ and $c_1 < v_f < c_2$.

9.4.2 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 our primary focus is single-mode fibers, multimode dispersion is excluded from this section. The sum of waveguide dispersion and material dispersion is known as *chromatic dispersion*, characterized by the dispersion parameter

$$D = \frac{d\tau_g}{d\lambda} \quad (9.28)$$

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, also called normal, dispersion. If D is greater than zero, the waveguide has negative (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 6, with vacuum or air, always have anomalous dispersion since the group velocity increases with frequency and the phase velocity decreases with frequency. That means that high frequencies in a pulse travels faster than low frequencies. This is seen from the analysis in Chapter 8.

Without chromatic dispersion, the group velocity, and consequently the propagation delay τ , is independent of frequency, meaning the dispersion parameter $D = 0$. A pulse propagating in a waveguide with zero chromatic dispersion does not change its shape, but might be attenuated.

Conversely, if D is large then τ_g and the group velocity vary rapidly with λ . The large difference in group velocity between different frequencies leads to significant pulse broadening over distance.

The chromatic dispersion can be decomposed in the contribution from the material dispersion D_m and the waveguide dispersion D_v as

$$D = D_m + D_v \quad (9.29)$$

When $n_1 - n_2 \ll 1$ we can determine the material dispersion by utilizing the approximation $n_e \approx n_1$ which gives

$$D_m = -\frac{\lambda}{c_0} \frac{d^2 n_1}{d\lambda^2} \quad (9.30)$$

When we consider the waveguide dispersion we keep n_1 and n_2 constant, i.e.,

$$D_v = -\frac{\lambda}{c_0} \frac{d^2 n_e}{d\lambda^2} \Big|_{n_1, n_2 \text{ konst.}}$$

Example 9.2

Figure 9.10 shows the material dispersion as a function of the wavelength. We have already seen from Figure 9.6 that the most suitable vacuum wavelength for a single mode fiber made out of silica is $\lambda = 1.55 \mu\text{m}$. At this wavelength the material dispersion is positive. In order to obtain a dispersion free fiber at the wavelength $\lambda = 1.55 \mu\text{m}$ 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

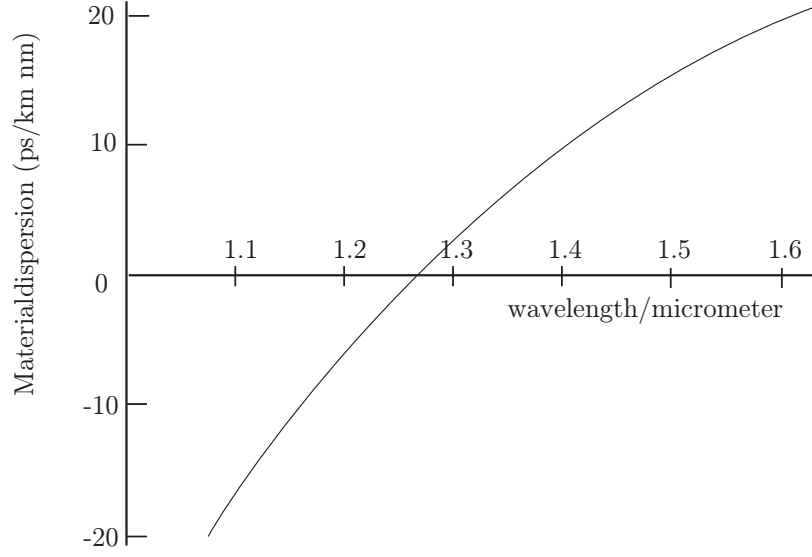


Figure 9.10: Material dispersion for silica.

choosing a suitable 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.

■

9.4.3 Attenuation in optical fibers

Up to this point, the fiber material has been treated as lossless. In practice, all optical fibers experience small propagation losses arising primarily from two mechanisms: scattering from material impurities and intrinsic attenuation in the silica glass itself, see Figure 9.6.

One can purify the material such that the scattering from the impurities are negligible, but intrinsic material losses set a fundamental lower bound on attenuation. The absorption in silica is, as we have seen, very dependent on the wavelength and reaches its minima at the vacuum wavelength $\lambda = 1.55 \mu\text{m}$. 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 $e^{ik_z z}$ with $e^{i\text{Re}\{k_z\}z} e^{-\alpha z}$. The attenuation α is approximately the same as for plane wave propagation in a dielectric material with small losses

$$\alpha = \text{Im}\{k_z\} = k_0 \text{Im}\{n_e\} \simeq k_0 \text{Im}\{\sqrt{\epsilon_1}\} \simeq \frac{k_0 \text{Im}\{\epsilon_1\}}{2\sqrt{\text{Re}\{\epsilon_1\}}} \quad (9.31)$$

where $\epsilon_1 = \text{Re}\{\epsilon_1\} + i\text{Im}\{\epsilon_1\}$ is the complex permittivity.

9.4.4 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

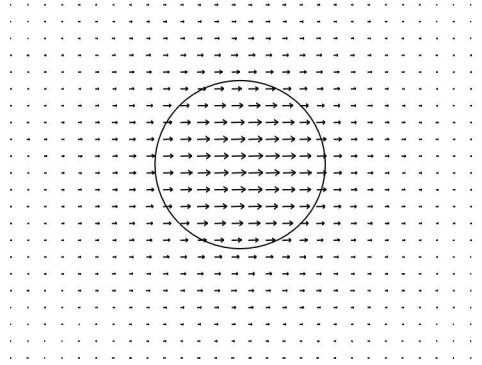


Figure 9.11: Electric field for the HE_{11} mode of a circular optical fiber.

of problems. We give three examples here, the circular dielectric waveguide, a dielectric waveguide with elliptic cross section and finally a dielectric resonator.

Example 9.3

Consider a circular dielectric waveguide with radius $a = 4 \mu\text{m}$ and with $n = 1.503$ in the core and $n = 1.5$ in the cladding. The wavelength in vacuum is $\lambda = 1.55 \mu\text{m}$. Find k_z and the electric field distribution. We have called n_e the effective index of refraction but COMSOL refers to it as the effective mode index. We do the following steps:

- Choose 2D and **2D > Electromagnetic waves > Mode analysis**.
- Draw a circle with radius around $40 \mu\text{m}$ (cladding) and a circle with radius $a = 4 \mu\text{m}$ (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 $n_e = 1.5$ and 1.503.

■

In Figure 9.11 the electric field for the HE_{11} mode of circular waveguide is depicted. Note that there exists also a HE_{11} mode with the electric field rotated 90° compared to the one in the figure. In Figure 9.12 we have exchanged the circle for an ellipse with half axes $a = 5 \mu\text{m}$ and $b = 3 \mu\text{m}$. 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.

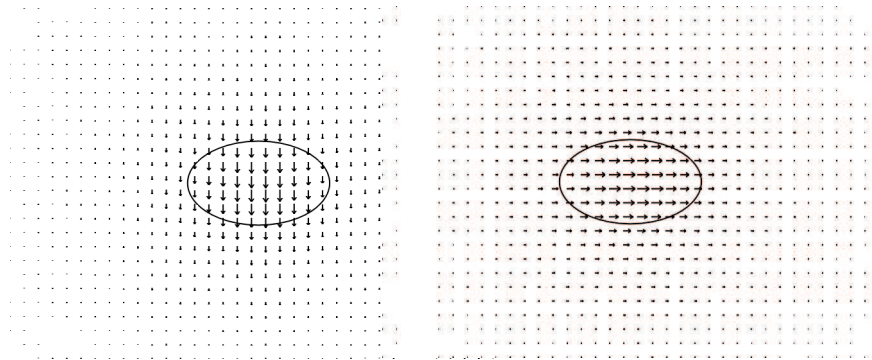


Figure 9.12: The electric fields the two fundamental modes for an elliptic optical fiber with half axes $a = 5 \mu\text{m}$ and $b = 3 \mu\text{m}$. COMSOL gives the mode index 1.500543 for the mode to the left, i.e., with the electric field along the minor axis and 1.500544 for the mode to the right, i.e., with the electric field along the major axis. The two modes correspond to the two HE_{11} modes in a circular optical fiber.

9.4.5 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 resonators are used as bandpassfilters with very high Q-factors, i.e., they are very narrow banded. Their primary applications include:

- **High-Q Bandpass Filters:** Their ability to achieve exceptionally high Q-factors makes them perfect for creating extremely narrow-band filters that precisely isolate specific frequencies.
- **Resonating Antennas:** Dielectric resonator antennas (DRAs) have recently surged in popularity as highly efficient, compact alternatives to traditional metallic antennas.

Dielectric resonators can be analyzed with FEM. The shape of the resonator and the dielectric material's permittivity determine its resonant frequencies, with the primary resonance typically occurring when the resonator's diameter roughly equals half the wavelength inside 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 spurious resonances with electric fields that are not confined in the resonator and it is important that we disregard these. One way to filter out all of the spurious resonances— 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**.

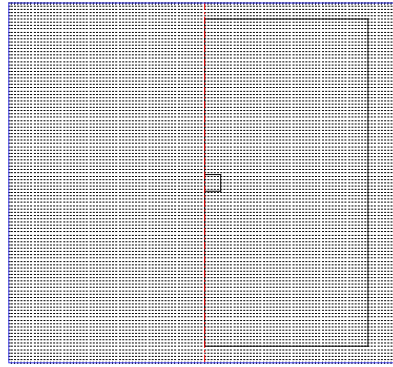


Figure 9.13: The geometry in COMSOL for the cylindric dielectric resonator. Notice that we only draw that half of the cylinders that are to the right of the symmetry line.

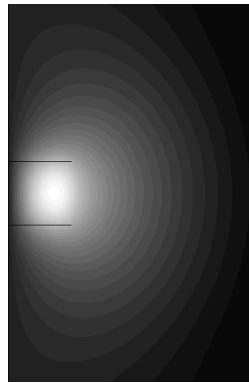


Figure 9.14: The electric field of the lowest resonant frequency $f_r = 46.2$ GHz for the cylindric resonator with $\epsilon = 10$, radius $a = 1$ mm and height $h = 1$ mm. The mode is a TE mode, i.e., the electric field is directed along the azimuthal direction.

- 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 9.13. 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 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.

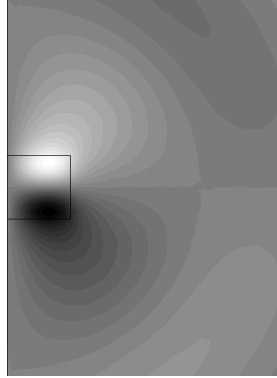


Figure 9.15: The electric field of the second lowest resonant TE mode. The frequency is $f_r = 75.7$ GHz for the cylindric resonator with $\epsilon = 10$, radius $a = 1$ mm and height $h = 1$ mm.

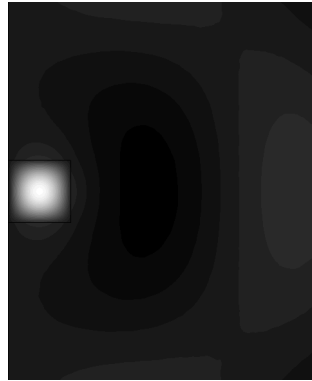


Figure 9.16: The magnetic field of the lowest resonant TM mode. The frequency is $f_r = 73.4$ GHz for the cylindric resonator with $\epsilon = 10$, radius $a = 1$ mm and height $h = 1$ mm.

- 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 Figure 9.14 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., $\mathbf{E} = E(r_c, z)\hat{\phi}$. In Figure 9.15 the next resonance is depicted. The frequency is 75.7 GHz. The lowest TM mode is depicted in 9.16.

Problems in Chapter 9

9.1 Prove that the boundary conditions

$$\begin{cases} \mu_1 \hat{\mathbf{n}} \cdot \mathbf{H}_1 = \mu_2 \hat{\mathbf{n}} \cdot \mathbf{H}_2 \\ \epsilon_1 \hat{\mathbf{n}} \cdot \mathbf{E}_1 = \epsilon_2 \hat{\mathbf{n}} \cdot \mathbf{E}_2 \end{cases} \quad \mathbf{r} \text{ on } S$$



Figure 9.17: Geometry for problem 9.5.

follow from the boundary conditions in section 9.2.2.

- *9.2** Determine the TE modes in a planar waveguide $0 < y < b$ that for $0 < y < a$ is filled with a dielectric material with permittivity ϵ and for $a < y < b$ is filled with air. Determine also the transcendental equation that determines the longitudinal wavenumber.
- 9.3** Determine $v_m(\rho)$ and $w_m(\rho)$ in the cladding for a circular dielectric waveguide at the cut-off frequency and show that these functions satisfy the Laplace equation $\nabla_T^2 \psi(\rho) = 0$.
- *9.4** Show that there are no propagating modes in a circular waveguide when $n_2 > n_1$.
- *9.5** A circular hollow waveguide with radius b , is for $0 < \rho < a$ filled with a dielectric material with relative permittivity ϵ , see Figure 9.17. Determine all axially symmetric modes.

Summary of chapter 9

Expressions for the fields in a circular dielectric waveguide

$$\begin{aligned}
 &\begin{cases} v_m(\rho, \phi) = AJ_m(h\rho) \sin m\phi \\ w_m(\rho, \phi) = BJ_m(h\rho) \cos m\phi \end{cases} & \rho < a \\
 &\begin{cases} v_m(\rho, \phi) = CK_m(q\rho) \sin m\phi \\ w_m(\rho, \phi) = DK_m(q\rho) \cos m\phi \end{cases} & \rho > a \\
 &\frac{C}{A} = \frac{D}{B} = \frac{J_m(ha)}{K_m(qa)} \\
 &\frac{B}{A} = \frac{D}{C} = \frac{mk_z}{k_0} \left(\frac{1}{h^2a^2} + \frac{1}{q^2a^2} \right) \left(\frac{\mu_1 J'_m(ha)}{haJ_m(ha)} + \frac{\mu_2 K'_m(qa)}{qaK_m(qa)} \right)^{-1} \\
 &h = k_{1t} \\
 &q = -ik_{2t} = |k_{2t}|
 \end{aligned}$$

TE and TM modes (m=0)

Cut-off frequencies (both cases)

$$\begin{aligned}
 f_{0n} &= \frac{c_0}{2\pi a} \frac{x_{0n}}{\sqrt{n_1^2 - n_2^2}} \\
 &\text{where } J_0(x_{0n}) = 0
 \end{aligned}$$

Characteristic equation for the TE modes

$$\begin{array}{l}
 A = 0, B \neq 0 \\
 \frac{J_1(ha)}{haJ_0(ha)} = -\frac{\mu_2}{\mu_1} \frac{K_1(qa)}{qaK_0(qa)}
 \end{array}$$

Characteristic equation for the TM modes

$$\begin{array}{l}
 B = 0, A \neq 0 \\
 \frac{J_1(ha)}{haJ_0(ha)} = -\frac{\epsilon_2}{\epsilon_1} \frac{K_1(qa)}{qaK_0(qa)}
 \end{array}$$

EH modes ($m > 0$)**Cut-off frequencies:**

$$f_{mn}^{EH} = \frac{c_0}{2\pi a} \frac{x_{mn}}{\sqrt{n_1^2 - n_2^2}}, \quad J_m(x_{mn}) = 0$$

Characteristic equation:

$$\begin{array}{l}
 \frac{J_{m+1}(ha)}{haJ_m(ha)} = \frac{K'_m(qa)}{2qaK_m(qa)} \frac{\mu_1\epsilon_2 + \mu_2\epsilon_1}{\mu_1\epsilon_1} + \frac{m}{(ha)^2} \\
 - \left[\left(\frac{K'_m(qa)}{2qaK_m(qa)} \frac{\mu_1\epsilon_2 - \mu_2\epsilon_1}{\mu_1\epsilon_1} \right)^2 + \left(\frac{mk_z}{k_1} \right)^2 \left(\frac{1}{q^2a^2} + \frac{1}{h^2a^2} \right)^2 \right]^{1/2}
 \end{array}$$

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

$$D = D_m + D_v$$

$$D_m = -\frac{\lambda}{c_0} \frac{d^2 n_1}{d\lambda^2}$$

$$D_v = -\frac{\lambda}{c_0} \frac{d^2 n_e}{d\lambda^2} \Big|_{n_1, n_2 \text{ konst.}}$$

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

$$z^2 \frac{d^2}{dz^2} Z_n(z) + z \frac{d}{dz} Z_n(z) + (z^2 - n^2) Z_n(z) = 0 \quad (\text{A.1})$$

where n is assumed integer¹.

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

$$J_n(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{z}{2}\right)^{n+2k} \quad (\text{A.2})$$

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

$$J_n(z) = \frac{1}{\pi} \int_0^\pi \cos(z \sin t - nt) dt = \frac{1}{2\pi} \int_0^{2\pi} e^{iz \cos t} e^{in(t - \frac{1}{2}\pi)} dt \quad (\text{A.3})$$

¹A more general definition with eg., complex-valued n is also possible, but the expressions and the results often differ.

Order n	Root j			
	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 ($\nu = 4n^2$)

$$\begin{cases} P_n(z) \sim 1 - \frac{(\nu-1)(\nu-9)}{2!(8z)^2} + \frac{(\nu-1)(\nu-9)(\nu-25)(\nu-49)}{4!(8z)^4} - \dots \\ Q_n(z) \sim \frac{\nu-1}{8z} - \frac{(\nu-1)(\nu-9)(\nu-25)}{3!(8z)^3} + \dots \end{cases} \quad (\text{A.4})$$

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

$$\xi_{nj} = j\pi + \left(n - \frac{1}{2}\right) \frac{\pi}{2}, \quad \eta_{nj} = j\pi + \left(n - \frac{3}{2}\right) \frac{\pi}{2}$$

Another, linearly independent solution to the Bessel differential equation, which is real-valued for real arguments, is the Neumann function² $N_n(z)$. The power series expansion is

$$\begin{aligned} N_n(z) = & \frac{2}{\pi} \left(\ln\left(\frac{z}{2}\right) + \gamma - \frac{1}{2} \sum_{k=1}^n \frac{1}{k} \right) J_n(z) \\ & - \frac{1}{\pi} \sum_{k=0}^{\infty} (-1)^k \frac{\left(\frac{z}{2}\right)^{n+2k}}{k!(n+k)!} \sum_{l=1}^k \left(\frac{1}{l} + \frac{1}{l+n} \right) \\ & - \frac{1}{\pi} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!} \left(\frac{z}{2}\right)^{-n+2k} \end{aligned}$$

²These solutions are also called Bessel functions of the second kind.

Order n	Root j			
	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 $\gamma = 0.577\,215\,66\dots$, 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

$$N_0(z) = \frac{2}{\pi} \left(\ln \left(\frac{z}{2} \right) + \gamma \right) + O(z^2)$$

$$N_n(z) = -\frac{(n-1)!}{\pi} \left(\frac{z}{2} \right)^{-n} + \dots$$

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 natural³. These are defined as

$$H_n^{(1)}(z) = J_n(z) + iN_n(z)$$

$$H_n^{(2)}(z) = J_n(z) - iN_n(z)$$

The Hankel functions of the first and second kind have integral representations

$$H_n^{(1)}(z) = \frac{2}{i\pi} e^{-in\frac{\pi}{2}} \int_0^\infty e^{iz \cosh s} \cosh ns \, ds, \quad 0 < \arg z < \pi$$

$$H_n^{(2)}(z) = \frac{2i}{\pi} e^{in\frac{\pi}{2}} \int_0^\infty e^{-iz \cosh s} \cosh ns \, ds, \quad -\pi < \arg z < 0$$

For large argumens, the Hankel functions have asymptotic expansions

$$H_n^{(1)}(z) = \left(\frac{2}{\pi z} \right)^{1/2} e^{i(z - \frac{n\pi}{2} - \frac{\pi}{4})} (P_n(z) + iQ_n(z)), \quad -\pi < \arg z < 2\pi$$

$$H_n^{(2)}(z) = \left(\frac{2}{\pi z} \right)^{1/2} e^{-i(z - \frac{n\pi}{2} - \frac{\pi}{4})} (P_n(z) - iQ_n(z)), \quad -2\pi < \arg z < \pi$$
(A.5)

³These 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 ($n = 0, 1, 2, \dots, m = 0, 1, 2, \dots$)⁴

$$\begin{aligned} Z_{n-1}(z) - Z_{n+1}(z) &= 2Z'_n(z) \\ Z_{n-1}(z) + Z_{n+1}(z) &= \frac{2n}{z}Z_n(z) \\ Z_{n+1}(z) &= \frac{n}{z}Z_n(z) - Z'_n(z) \\ Z'_n(z) &= Z_{n-1}(z) - \frac{n}{z}Z_n(z) \\ \left(\frac{d}{zdz}\right)^m [z^n Z_n(z)] &= z^{n-m}Z_{n-m}(z) \\ \left(\frac{d}{zdz}\right)^m [z^{-n}Z_n(z)] &= (-1)^m z^{-n-m}Z_{n+m}(z) \end{aligned}$$

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 ($n = 0, 1, 2, \dots$)

$$\begin{aligned} \int x^{n+1} Z_n(x) dx &= x^{n+1} Z_{n+1}(x) = -x^{n+1} \left(Z'_n(x) - \frac{n}{x} Z_n(x) \right) \\ \int x^{-n+1} Z_n(x) dx &= -x^{-n+1} Z_{n-1}(x) = -x^{-n+1} \left(Z'_n(x) + \frac{n}{x} Z_n(x) \right) \\ \int x (Z_n(x))^2 dx &= \frac{x^2}{2} \left[(Z_n(x))^2 - Z_{n-1}(x) Z_{n+1}(x) \right] \\ &= \frac{x^2}{2} (Z'_n(x))^2 + \frac{1}{2} (x^2 - n^2) (Z_n(x))^2 \end{aligned}$$

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 ($n = 0, 1, 2, \dots, m = 0, 1, 2, \dots$)

$$\begin{aligned} \int \left[(\alpha^2 - \beta^2)x - \frac{m^2 - n^2}{x} \right] Z_m(\alpha x) Y_n(\beta x) dx &= \beta x Z_m(\alpha x) Y_{n-1}(\beta x) \\ &\quad - \alpha x Z_{m-1}(\alpha x) Y_n(\beta x) + (m - n) Z_m(\alpha x) Y_m(\beta x) \\ \int x Z_m(\alpha x) Y_m(\beta x) dx &= \frac{\beta x Z_m(\alpha x) Y_{m-1}(\beta x) - \alpha x Z_{m-1}(\alpha x) Y_m(\beta x)}{\alpha^2 - \beta^2} \\ \int \frac{Z_m(\alpha x) Y_n(\alpha x)}{x} dx &= \alpha x \frac{Z_{m-1}(\alpha x) Y_n(\alpha x) - Z_m(\alpha x) Y_{n-1}(\alpha x)}{m^2 - n^2} - \frac{Z_m(\alpha x) Y_n(\alpha x)}{m + n} \end{aligned}$$

⁴These 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

$$\begin{cases} J_n(z^*) = (J_n(z))^* \\ N_n(z^*) = (N_n(z))^* \end{cases} \quad \begin{cases} H_n^{(1)}(z^*) = (H_n^{(2)}(z))^* \\ H_n^{(2)}(z^*) = (H_n^{(1)}(z))^* \end{cases}$$

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

$$Z_n(w) \begin{pmatrix} \cos n\phi \\ \sin n\phi \end{pmatrix} = \sum_{k=-\infty}^{\infty} Z_{n+k}(u) J_k(v) \begin{pmatrix} \cos k\alpha \\ \sin k\alpha \end{pmatrix}, \quad |ve^{\pm i\alpha}| < |u|$$

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)

$$\int_0^{2\pi} e^{iz \cos \phi} e^{in\phi} d\phi = 2\pi i^n J_n(z)$$

From this we easily conclude by a simple change of variables that

$$\int_0^{2\pi} e^{iz \cos(\phi-\alpha)} e^{in\phi} d\phi = e^{in\alpha} \int_{-\alpha}^{2\pi-\alpha} e^{iz \cos \psi} e^{in\psi} d\psi = 2\pi i^n J_n(z) e^{in\alpha}$$

This integral is a function of the variables z and α .

In scattering problems the following integral is often encountered:

$$I(\boldsymbol{\tau}, \rho) = \int_0^{2\pi} \hat{\boldsymbol{\phi}} e^{i\boldsymbol{\rho} \cdot \boldsymbol{\tau}} e^{in\phi} d\phi = \int_0^{2\pi} \hat{\boldsymbol{\phi}} e^{i\rho\tau \cos(\phi-\alpha)} e^{in\phi} d\phi$$

where the position vector in the x - y -plane is denoted $\boldsymbol{\rho} = \hat{\mathbf{x}}\rho \cos \phi + \hat{\mathbf{y}}\rho \sin \phi$ and $\boldsymbol{\tau} = \hat{\mathbf{x}}\tau \cos \alpha + \hat{\mathbf{y}}\tau \sin \alpha$. This integral can be rewritten as

$$I(\boldsymbol{\tau}, \rho) = \frac{1}{i\rho} \hat{\mathbf{z}} \times \nabla_{\boldsymbol{\tau}} \int_0^{2\pi} e^{i\rho\tau \cos(\phi-\alpha)} e^{in\phi} d\phi$$

where the nabla-operator in the τ variable is denoted $\nabla_{\boldsymbol{\tau}}$, and where we used

$$\hat{\mathbf{z}} \times \nabla_{\boldsymbol{\tau}} e^{i\boldsymbol{\rho} \cdot \boldsymbol{\tau}} = i\hat{\mathbf{z}} \times \boldsymbol{\rho} e^{i\boldsymbol{\rho} \cdot \boldsymbol{\tau}} = i\rho \hat{\mathbf{z}} \times \hat{\boldsymbol{\rho}} e^{i\boldsymbol{\rho} \cdot \boldsymbol{\tau}} = i\rho \hat{\boldsymbol{\phi}} e^{i\boldsymbol{\rho} \cdot \boldsymbol{\tau}}$$

From this we infer

$$I(\boldsymbol{\tau}, \rho) = \frac{2\pi i^{n-1}}{\rho} \hat{\mathbf{z}} \times \nabla_{\boldsymbol{\tau}} J_n(\rho\tau) e^{in\alpha} = 2\pi i^{n-1} \hat{\mathbf{z}} \times \left(\hat{\boldsymbol{\tau}} J'_n(\rho\tau) e^{in\alpha} + in\hat{\boldsymbol{\alpha}} \frac{J_n(\rho\tau)}{\rho\tau} e^{in\alpha} \right)$$

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 section 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:

$$\begin{aligned} I_n(z) &= i^{-n} J_n(iz) \\ K_n(z) &= \frac{i^{n+1} \pi}{2} H_n^{(1)}(iz) \end{aligned}$$

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 $z \rightarrow 0$.

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 $\gamma = 0.57721566\dots$, 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

$$\begin{cases} K_0(z) = -\left(\ln\left(\frac{z}{2}\right) + \gamma\right) + O(z^2) \\ K_n(z) = \frac{(n-1)!}{2} \left(\frac{z}{2}\right)^{-n} + \dots, \quad n \neq 0 \end{cases} \quad (\text{A.6})$$

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

$$K_n(z) = \sqrt{\frac{\pi}{2z}} e^{-z} Q_n(z), \quad -\frac{3\pi}{2} < \arg z < \frac{3\pi}{2}$$

where the functions $P_n(z)$ and $Q_n(z)$ have the following asymptotic expansions ($\nu = 4n^2$):

$$\begin{cases} P_n(z) \sim 1 - \frac{\nu-1}{8z} + \frac{(\nu-1)(\nu-9)}{2!(8z)^2} - \frac{(\nu-1)(\nu-9)(\nu-25)}{3!(8z)^3} + \dots \\ Q_n(z) \sim 1 + \frac{\nu-1}{8z} + \frac{(\nu-1)(\nu-9)}{2!(8z)^2} + \frac{(\nu-1)(\nu-9)(\nu-25)}{3!(8z)^3} + \dots \end{cases}$$

There are recursion relations that relate modified Bessel functions of different order. For $I_n(z)$, $n = 0, 1, 2, \dots$, the most important relations are ($m = 0, 1, 2, \dots$)⁵

$$\begin{aligned} I_{n-1}(z) + I_{n+1}(z) &= 2I'_n(z) \\ I_{n-1}(z) - I_{n+1}(z) &= \frac{2n}{z} I_n(z) \\ I_{n+1}(z) &= I'_n(z) - \frac{n}{z} I_n(z) \\ I'_n(z) &= I_{n-1}(z) - \frac{n}{z} I_n(z) \\ \left(\frac{d}{z dz}\right)^m [z^n I_n(z)] &= z^{n-m} I_{n-m}(z) \\ \left(\frac{d}{z dz}\right)^m [z^{-n} I_n(z)] &= z^{-n-m} I_{n+m}(z) \end{aligned}$$

och för $K_n(z)$ motsvarande ($n = 0, 1, 2, \dots, m = 0, 1, 2, \dots$)⁵

$$\begin{aligned} K_{n-1}(z) + K_{n+1}(z) &= -2K'_n(z) \\ K_{n+1}(z) - K_{n-1}(z) &= \frac{2n}{z} K_n(z) \\ K_{n+1}(z) &= \frac{n}{z} K_n(z) - K'_n(z) \\ K'_n(z) &= -K_{n-1}(z) - \frac{n}{z} K_n(z) \\ \left(\frac{d}{z dz}\right)^m [z^n K_n(z)] &= (-1)^m z^{n-m} K_{n-m}(z) \\ \left(\frac{d}{z dz}\right)^m [z^{-n} K_n(z)] &= (-1)^m z^{-n-m} K_{n+m}(z) \end{aligned}$$

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) .

⁵These 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 expansion⁶

$$j_n(z) = 2^n z^n \sum_{k=0}^{\infty} \frac{(-1)^k (k+n)!}{k! (2k+2n+1)!} z^{2k} \quad (\text{A.7})$$

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

$$\begin{aligned} h_n^{(1)}(z) &= j_n(z) + i n_n(z) \\ h_n^{(2)}(z) &= j_n(z) - i n_n(z) \end{aligned}$$

The spherical Bessel, Neumann, and Hankel functions for $l = 0$ are

$$\begin{cases} j_0(z) = \frac{\sin z}{z} \\ n_0(z) = -\frac{\cos z}{z} \end{cases} \quad \begin{cases} h_0^{(1)}(z) = -\frac{i}{z} e^{iz} \\ h_0^{(2)}(z) = \frac{i}{z} e^{-iz} \end{cases}$$

and for $l = 1$ they are

$$\begin{cases} j_1(z) = \frac{\sin z}{z^2} - \frac{\cos z}{z} \\ n_1(z) = -\frac{\cos z}{z^2} - \frac{\sin z}{z} \end{cases} \quad \begin{cases} h_1^{(1)}(z) = e^{iz} \left(-\frac{1}{z} - \frac{i}{z^2} \right) \\ h_1^{(2)}(z) = e^{-iz} \left(-\frac{1}{z} + \frac{i}{z^2} \right) \end{cases}$$

⁶Between 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 $z = -i$, and the roots in the finite complex plane of $(zh_1^{(1)}(z))'$ are $z = (\pm\sqrt{3} - i)/2$.

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 $l = 1$

$$\begin{cases} j_1(z) \approx \frac{z}{3}, \text{ as } z \rightarrow 0 \\ h_1^{(1)}(z) \approx \frac{1}{iz^2}, \text{ as } z \rightarrow 0 \end{cases} \quad \begin{cases} \frac{(zj_1(z))'}{z} \approx \frac{2}{3}, \text{ as } z \rightarrow 0 \\ \frac{(zh_1^{(1)}(z))'}{z} \approx -\frac{1}{iz^3}, \text{ as } z \rightarrow 0 \end{cases}$$

An alternative expansion of the spherical Bessel and Neumann functions in finite trigonometric series are:

$$\begin{cases} j_n(z) = \frac{1}{z} \left\{ p_n(z) \sin\left(z - \frac{n\pi}{2}\right) + q_n(z) \cos\left(z - \frac{n\pi}{2}\right) \right\} \\ n_n(z) = \frac{1}{z} \left\{ q_n(z) \sin\left(z - \frac{n\pi}{2}\right) - p_n(z) \cos\left(z - \frac{n\pi}{2}\right) \right\} \end{cases} \quad (\text{A.8})$$

The explicit for of the finite sums are

$$\begin{cases} p_n(z) = \sum_{k=0}^{[n/2]} (-1)^k \frac{(n+2k)!}{(2k)!(n-2k)!} \frac{1}{(2z)^{2k}} = 1 - \frac{b_n}{z^2} + O(z^{-4}) \\ q_n(z) = \sum_{k=0}^{[(n-1)/2]} (-1)^k \frac{(n+2k+1)!}{(2k+1)!(n-2k-1)!} \frac{1}{(2z)^{2k+1}} = \frac{a_n}{z} + O(z^{-3}) \end{cases} \quad (\text{A.9})$$

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

$$\begin{cases} h_n^{(1)}(z) = \frac{e^{iz-i(n+1)\pi/2}}{z} \sum_{k=0}^n \frac{(n+k)!}{k!(n-k)!} \frac{1}{(-2iz)^k} = \frac{e^{iz-i(n+1)\pi/2}}{z} (p_n(z) + iq_n(z)) \\ h_n^{(2)}(z) = \frac{e^{-iz+i(n+1)\pi/2}}{z} \sum_{k=0}^n \frac{(n+k)!}{k!(n-k)!} \frac{1}{(2iz)^k} = \frac{e^{-iz+i(n+1)\pi/2}}{z} (p_n(z) - iq_n(z)) \end{cases} \quad (\text{A.10})$$

In scattering problems another combination of spherical Bessel and Hankel functions occurs, i.e.,

$$\begin{cases} \frac{(zj_n(z))'}{z} = \frac{1}{z} \left\{ P_n(z) \sin \left(z - \frac{n\pi}{2} \right) + Q_n(z) \cos \left(z - \frac{n\pi}{2} \right) \right\} \\ \frac{(zh_n^{(1)}(z))'}{z} = \frac{e^{iz-i(n+1)\pi/2}}{z} (P_n(z) + iQ_n(z)) \\ \frac{(zh_n^{(2)}(z))'}{z} = \frac{e^{-iz+i(n+1)\pi/2}}{z} (P_n(z) - iQ_n(z)) \end{cases} \quad (\text{A.11})$$

where

$$\begin{cases} P_n(z) = p'_n(z) - q_n(z) = -\frac{a_n}{z} + O(z^{-3}) \\ Q_n(z) = q'_n(z) + p_n(z) = 1 - \frac{a_n + b_n}{z^2} + O(z^{-4}) \end{cases} \quad (\text{A.12})$$

Between different solutions of spherical Bessel functions there are recursion relations. The most important ones are ($n = 0, 1, 2, \dots, m = 0, 1, 2, \dots$)

$$\begin{cases} f_{n-1}(z) + f_{n+1}(z) = \frac{2n+1}{z} f_n(z) \\ nf_{n-1}(z) - (n+1)f_{n+1}(z) = (2n+1)f'_n(z) \\ nf_n(z) - zf_{n+1}(z) = zf'_n(z) \\ \left(\frac{d}{z dz} \right)^m [z^{n+1} f_n(z)] = z^{n-m+1} f_{n-m}(z) \\ \left(\frac{d}{z dz} \right)^m [z^{-n} f_n(z)] = (-1)^m z^{-n-m} f_{n+m}(z) \end{cases} \quad (\text{A.13})$$

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:

$$\frac{(zf_n(z))'}{z} = \frac{f_n(z)}{z} + f'_n(z) = \frac{(n+1)f_{n-1}(z) - nf_{n+1}(z)}{2n+1} = \frac{(n+1)f_n(z)}{z} - f_{n+1}(z) \quad (\text{A.14})$$

The Wronskian relation for $j_n(z)$ and $n_n(z)$ is

$$j_n(z)n'_n(z) - j'_n(z)n_n(z) = \frac{1}{z^2} \quad (\text{A.15})$$

and for $j_n(z)$ and $h_n^{(1)}(z)$ it is

$$j_n(z)h_n^{(1)'}(z) - j'_n(z)h_n^{(1)}(z) = \frac{i}{z^2} \quad (\text{A.16})$$

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

$$\begin{aligned}\nabla\psi &= \hat{\mathbf{x}} \frac{\partial\psi}{\partial x} + \hat{\mathbf{y}} \frac{\partial\psi}{\partial y} + \hat{\mathbf{z}} \frac{\partial\psi}{\partial z} \\ \nabla^2\psi &= \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\end{aligned}$$

The divergence, the curl, and the Laplace-operator of a vector field $\mathbf{A}(x, y, z) = \hat{\mathbf{x}}A_x(x, y, z) + \hat{\mathbf{y}}A_y(x, y, z) + \hat{\mathbf{z}}A_z(x, y, z)$ are

$$\begin{aligned}\nabla \cdot \mathbf{A} &= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \\ \nabla \times \mathbf{A} &= \hat{\mathbf{x}} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \hat{\mathbf{y}} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \hat{\mathbf{z}} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \\ \nabla^2 \mathbf{A} &= \hat{\mathbf{x}} \nabla^2 A_x + \hat{\mathbf{y}} \nabla^2 A_y + \hat{\mathbf{z}} \nabla^2 A_z\end{aligned}$$

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

$$\begin{aligned}\nabla\psi &= \hat{\boldsymbol{\rho}}\frac{\partial\psi}{\partial\rho} + \hat{\boldsymbol{\phi}}\frac{1}{\rho}\frac{\partial\psi}{\partial\phi} + \hat{\mathbf{z}}\frac{\partial\psi}{\partial z} \\ \nabla^2\psi &= \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\psi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\psi}{\partial\phi^2} + \frac{\partial^2\psi}{\partial z^2}\end{aligned}$$

The divergence, the curl, and the Laplace-operator of a vector field $\mathbf{A}(\rho, \phi, z) = \hat{\boldsymbol{\rho}}A_\rho(\rho, \phi, z) + \hat{\boldsymbol{\phi}}A_\phi(\rho, \phi, z) + \hat{\mathbf{z}}A_z(\rho, \phi, z)$ are

$$\begin{aligned}\nabla \cdot \mathbf{A} &= \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho A_\rho) + \frac{1}{\rho}\frac{\partial A_\phi}{\partial\phi} + \frac{\partial A_z}{\partial z} \\ \nabla \times \mathbf{A} &= \hat{\boldsymbol{\rho}}\left(\frac{1}{\rho}\frac{\partial A_z}{\partial\phi} - \frac{\partial A_\phi}{\partial z}\right) + \hat{\boldsymbol{\phi}}\left(\frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial\rho}\right) + \hat{\mathbf{z}}\frac{1}{\rho}\left(\frac{\partial}{\partial\rho}(\rho A_\phi) - \frac{\partial A_\rho}{\partial\phi}\right) \\ \nabla^2\mathbf{A} &= \hat{\boldsymbol{\rho}}\left(\nabla^2 A_\rho - \frac{A_\rho}{\rho^2} - \frac{2}{\rho^2}\frac{\partial A_\phi}{\partial\phi}\right) + \hat{\boldsymbol{\phi}}\left(\nabla^2 A_\phi - \frac{A_\phi}{\rho^2} + \frac{2}{\rho^2}\frac{\partial A_\rho}{\partial\phi}\right) + \hat{\mathbf{z}}\nabla^2 A_z\end{aligned}$$

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

$$\begin{aligned}\nabla\psi &= \hat{\mathbf{r}}\frac{\partial\psi}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial\psi}{\partial\theta} + \hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\phi} \\ \nabla^2\psi &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} \\ &= \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\end{aligned}$$

and the divergence, the curl, and the Laplace-operator of a vector field $\mathbf{A}(r, \theta, \phi) = \hat{\mathbf{r}}A_r(r, \theta, \phi) + \hat{\boldsymbol{\theta}}A_\theta(r, \theta, \phi) + \hat{\boldsymbol{\phi}}A_\phi(r, \theta, \phi)$ are

$$\begin{aligned}\nabla \cdot \mathbf{A} &= \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 A_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta A_\theta) + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi} \\ \nabla \times \mathbf{A} &= \hat{\mathbf{r}} \frac{1}{r \sin \theta} \left(\frac{\partial}{\partial \theta} (\sin \theta A_\phi) - \frac{\partial A_\theta}{\partial \phi} \right) \\ &\quad + \hat{\boldsymbol{\theta}} \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{\partial}{\partial r} (r A_\phi) \right) + \hat{\boldsymbol{\phi}} \frac{1}{r} \left(\frac{\partial}{\partial r} (r A_\theta) - \frac{\partial A_r}{\partial \theta} \right) \\ \nabla^2 \mathbf{A} &= \hat{\mathbf{r}} \left(\nabla^2 A_r - \frac{2A_r}{r^2} - \frac{2}{r^2} \frac{\partial A_\theta}{\partial \theta} - \frac{2 \cot \theta}{r^2} A_\theta - \frac{2}{r^2 \sin \theta} \frac{\partial A_\phi}{\partial \phi} \right) \\ &\quad + \hat{\boldsymbol{\theta}} \left(\nabla^2 A_\theta - \frac{A_\theta}{r^2 \sin^2 \theta} + \frac{2}{r^2} \frac{\partial A_r}{\partial \theta} - \frac{2 \cos \theta}{r^2 \sin^2 \theta} \frac{\partial A_\phi}{\partial \phi} \right) \\ &\quad + \hat{\boldsymbol{\phi}} \left(\nabla^2 A_\phi - \frac{A_\phi}{r^2 \sin^2 \theta} + \frac{2}{r^2 \sin \theta} \frac{\partial A_r}{\partial \phi} + \frac{2 \cos \theta}{r^2 \sin^2 \theta} \frac{\partial A_\theta}{\partial \phi} \right)\end{aligned}$$

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

$$\begin{aligned}\mu_0 &= 4\pi \cdot 10^{-7} \text{ N/A}^2 \\ \epsilon_0 &= \frac{1}{c_0^2 \mu_0} \text{ F/m}\end{aligned}$$

Approximative values of these constants are

$$\begin{aligned}\mu_0 &\approx 12.566\,370\,614 \cdot 10^{-7} \text{ N/A}^2 \\ \epsilon_0 &\approx 8.854\,187\,817 \cdot 10^{-12} \text{ F/m}\end{aligned}$$

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

$$\begin{aligned}e &\approx 1.602\,177\,33 \cdot 10^{-19} \text{ C} \\ m &\approx 9.109\,389\,8 \cdot 10^{-31} \text{ kg} \\ e/m &\approx 1.758\,819\,63 \cdot 10^{11} \text{ C/kg}\end{aligned}$$

Appendix D

Notation

Appropriate notation leads to a more easily 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 where they are introduced, but some more general notation that is often used is collected in this appendix.

- Vector-valued quantities (mostly in $\mathbb{R}^3$) is denoted in slanted bold face, eg., $\mathbf{a}$ and $\mathbf{b}$, and vectors of unit length have a “hat” or caret ($\hat{\cdot}$) over a symbol, eg., $\hat{\mathbf{x}}$ and $\hat{\boldsymbol{\rho}}$.
- The (Euclidean) scalar product between two vectors, $\mathbf{a}$ and $\mathbf{b}$, is denoted in the usual standard way by a dot ($\cdot$), i.e., $\mathbf{a} \cdot \mathbf{b}$. If the vectors are complex-valued the appropriate scalar product is $\mathbf{a}^* \cdot \mathbf{b}$, 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

$$(f, g) = \iiint_{\Omega} f(\mathbf{r}) g^*(\mathbf{r}) \, dv$$

- We make a distinction between a vector $\mathbf{a}$ 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.,

$$[\mathbf{a}] = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}$$

where

$$\mathbf{a} = \hat{\mathbf{x}}a_x + \hat{\mathbf{y}}a_y + \hat{\mathbf{z}}a_z$$

- Linear vector-valued transformations (dyadics) are denoted in bold roman fonts, eg., $\mathbf{A}$. A linear transformation $\mathbf{A}$ acting on a vector field $\mathbf{a}$ gives a new vector field $\mathbf{b}$ and we use the notation

$$\mathbf{b} = \mathbf{A} \cdot \mathbf{a}$$

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 $\mathbf{b}$ is then

$$[\mathbf{b}] = [\mathbf{A}] \cdot [\mathbf{a}]$$

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

$$[\mathbf{I}]_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad [\mathbf{0}]_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

or in two dimensions

$$[\mathbf{I}]_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad [\mathbf{0}]_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

- 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.,

$$\begin{aligned} A_{ij}^t &= A_{ji} \\ A_{ij}^\dagger &= A_{ji}^* \end{aligned}$$

- We use the symbols o and O defined by

$$\begin{cases} f(x) = o(g(x)), & x \rightarrow a & \Leftrightarrow & \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0 \\ f(x) = O(g(x)), & x \rightarrow a & \Leftrightarrow & \frac{f(x)}{g(x)} \text{ bounded in a neighborhood of } a \end{cases}$$

- The symbol $\blacksquare$ denotes the end of an example.
- The real and the imaginary part of a complex number $z = x + iy$ are denoted $\operatorname{Re} z$ and $\operatorname{Im} z$, respectively, dvs.

$$\begin{aligned} \operatorname{Re} z &= x = \frac{1}{2}(z + z^*) \\ \operatorname{Im} z &= y = \frac{1}{2i}(z - z^*) \end{aligned}$$

A star $(^*)$ is used to denote the complex conjugate of a complex number, i.e., $z^* = x - iy$.

- The real and imaginary part of a dyadic can also be defined in terms of its Hermitian transpose. We adopt

$$\begin{aligned}\operatorname{Re} \mathbf{A} &= \frac{1}{2}(\mathbf{A} + \mathbf{A}^\dagger) \\ \operatorname{Im} \mathbf{A} &= \frac{1}{2i}(\mathbf{A} - \mathbf{A}^\dagger)\end{aligned}$$

where the dagger (†) denotes the Hermitian transpose. We see that the real and imaginary parts satisfy

$$\begin{aligned}\operatorname{Re} \mathbf{A} &= \operatorname{Re} \mathbf{A}^\dagger \\ \operatorname{Im} \mathbf{A}^\dagger &= -\operatorname{Im} \mathbf{A}\end{aligned}$$

We see that $\operatorname{Re} \mathbf{A}$ and $\operatorname{Im} \mathbf{A}$ are both Hermitian dyadics and that

$$\mathbf{A} = \operatorname{Re} \mathbf{A} + i \operatorname{Im} \mathbf{A}$$

- 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

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \theta = \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2 + y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2 + y^2}} & y < 0 \end{cases} \end{cases}$$

The 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] G. Kirchhoff, On the motion of electricity in conductors. *Annalen der Physic*, vol 152, 1857.
- [11] J.A. Kong. *Electromagnetic Wave Theory*. John Wiley & Sons, New York, 1986.
- [12] J.D. Kraus. *Electromagnetics*. McGraw-Hill, New York, fourth edition, 1992.
- [13] G. Kristensson. *Elektromagnetisk vågutbredning*. Studentlitteratur, Lund, 1999.
- [14] S.Y. Liao. *Engineering Applications of Electromagnetic Theory*. West Publishing Company, St. Paul, 1988.
- [15] A Dynamical Theory of the Electromagnetic Field. *The Philosophical Transactions of the Royal Society of London*, 1865.
- [16] B.D. Popović. *Introductory Engineering Electromagnetics*. Addison-Wesley, Reading, 1971.
- [17] D.M. Pozar. *Microwave Engineering*. Addison-Wesley, Reading, 1990.

- [18] G. Prescott, *History, theory, and practice of electrical telegraph*. Ticknor and fields, Boston, 1866. (available online)
 - [19] G. Russakoff. A Derivation of the Macroscopic Maxwell Equations. American Journal of Physics, vol. 38, 1188-1195, 1970
 - [20] SIS-Standardiseringskommissionen i Sverige. *SI måttenheter*, September 1988.
 - [21] J.C. Slater Microwave Electronics. Reviews of Modern Physics, Vol 18, 441-514, 1946
 - [22] T. Standage, *The Victorian Internet*. Bloomsbury (2014).
 - [23] J.A. Stratton. *Electromagnetic Theory*. McGraw-Hill, New York, 1941.
 - [24] J. Van Bladel. *Electromagnetic Fields*. Hemisphere Publishing Corporation, New York, 1985.
 - [25] A. Yariv. *Optical Electronics*. Saunders College Publishing, Fort Worth, 1991.
-

Answers

2.1

2.2

2.3

3.1

4.1 a) $V = V_0 e^{j\pi/4}$ b) $v(t) = V_0 \cos(\omega t - \pi/4)$

4.2

4.3

4.6 $L = 0,24 \mu\text{H/m}$ $C = 95 \text{ pF/m}$ $R = 50 \text{ m}\Omega/\text{m}$ $G = 20 \cdot 10^{-6} \text{ S/m}$

4.7 a) V_0 b) 0 c) $\frac{2}{5} V_0$ d) $\frac{2}{5} V_0$ e) $\frac{16}{25} V_0$

4.8 a) $\alpha = 0.014 \text{ m}^{-1}$ b) 125 dB/km

4.9 $Z_0 = 60 \Omega$ $\beta = 5,5 \cdot 10^{-3} \text{ rad/m}$

4.10 $\ell = 0.265 \text{ m}$ $Z_0 = 134 \Omega$

4.11 $\Gamma = 0,5$ $\text{SWR} = 3$

4.12 $R_b = 150 \Omega$ or $R_b = 17 \Omega$

4.13 $Z(0) = (120 - j60) \Omega$

4.14 a) $v(t) = \begin{cases} 0 & t < 2L/v \\ \frac{V_0}{3} & t \geq 2L/v \end{cases}$ b) $Z_{\text{in}} = 2R$

4.15 $x = 3\lambda/8$ or $x = 7\lambda/8$

4.16 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 $Z(0) = 0.38R$

5.1

$$\begin{aligned} \mathbf{A}_T &= \nabla_T \times (\nabla_T \times \mathbf{F}_T) + \nabla_T \frac{\partial}{\partial z} F_z - \frac{\partial^2}{\partial z^2} \mathbf{F}_T \\ &= \nabla_T (\nabla_T \cdot \mathbf{F}_T) - \nabla_T^2 \mathbf{F}_T + \nabla_T \frac{\partial}{\partial z} F_z - \frac{\partial^2}{\partial z^2} \mathbf{F}_T \\ A_z &= \frac{\partial}{\partial z} (\nabla_T \cdot \mathbf{F}_T) - \nabla_T^2 F_z \end{aligned}$$

6.3 TM modes:

$$\begin{cases} v_{mn}(\rho, \phi) = C_{mn} \underbrace{\{J_m(\zeta_{mn}\rho/a)N_m(\zeta_{mn}) - J_m(\zeta_{mn})N_m(\zeta_{mn}\rho/a)\}}_{Z_{mn}(\zeta_{mn}\rho/a)} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix} \\ k_{tmn} = \zeta_{mn}/a \end{cases}$$

where $\{\zeta_{mn}\}_{n=1}^{\infty}$ are the solutions to, $m = 0, 1, 2, 3, \dots$

$$Z_{mn}(\zeta_{mn}b/a) = J_m(\zeta_{mn}b/a)N_m(\zeta_{mn}) - J_m(\zeta_{mn})N_m(\zeta_{mn}b/a) = 0$$

and the normalization constant C_{mn} is determined by ($\epsilon_m = 2 - \delta_{m,0}$)

$$C_{mn}^{-2} = \frac{2\pi}{\epsilon_m} \int_a^b Z_{mn}^2(\zeta_{mn}\rho/a) \rho \, d\rho = \frac{\pi}{\epsilon_m} \left\{ b^2 (Z'_{mn}(\zeta_{mn}b/a))^2 - a^2 (Z'_{mn}(\zeta_{mn}))^2 \right\}$$

TE modes:

$$\begin{cases} w_{mn}(\rho, \phi) = D_{mn} \underbrace{\{J_m(\gamma_{mn}\rho/a)N'_m(\gamma_{mn}) - J'_m(\gamma_{mn})N_m(\gamma_{mn}\rho/a)\}}_{Y_{mn}(\gamma_{mn}\rho/a)} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix} \\ k_{tmn} = \gamma_{mn}/a \end{cases}$$

where $\{\gamma_{mn}\}_{n=1}^{\infty}$ are the solutions to, $m = 0, 1, 2, 3, \dots$

$$Y'_{mn}(\gamma_{mn}b/a) = J'_m(\gamma_{mn}b/a)N'_m(\gamma_{mn}) - J'_m(\gamma_{mn})N'_m(\gamma_{mn}b/a) = 0$$

and the normalization constant D_{mn} is determined by

$$\begin{aligned} D_{mn}^{-2} &= \frac{2\pi}{\epsilon_m} \int_a^b Y_{mn}^2(\gamma_{mn}\rho/a) \rho \, d\rho \\ &= \frac{\pi a^2}{\epsilon_m \gamma_{mn}^2} \left\{ (\gamma_{mn}^2 b^2/a^2 - m^2) (Y_{mn}(\gamma_{mn}b/a))^2 - (\gamma_{mn}^2 - m^2) (Y_{mn}(\gamma_{mn}))^2 \right\} \end{aligned}$$

TEM mode:

$$\nabla_T \psi(\rho, \phi) = \hat{\rho} \frac{1}{\rho} \frac{1}{2\pi \ln \frac{a}{b}}$$

6.4 TM modes:

$$\begin{cases} v_{mn}(\rho, \phi) = C_{mn} J_{m/2}(\zeta_{mn}\rho/a) \sin \frac{m}{2}\phi, & m = 1, 2, 3, \dots \\ k_{tmn} = \zeta_{mn}/a \end{cases}$$

where $\{\zeta_{mn}\}_{n=1}^{\infty}$ are the solutions to, $m = 1, 2, 3, \dots$

$$J_{m/2}(\zeta_{mn}) = 0$$

and the normalization constant C_{mn} is determined by

$$C_{mn}^{-2} = \frac{\pi a^2}{\zeta_{mn}^2} \int_0^{\zeta_{mn}} (J_{m/2}(x))^2 x \, dx$$

TE modes:

$$\begin{cases} w_{mn}(\rho, \phi) = D_{mn} J_{m/2}(\gamma_{mn} \rho/a) \cos \frac{m}{2} \phi, & m = 0, 1, 2, \dots \\ k_{tmn} = \gamma_{mn}/a \end{cases}$$

where $\{\gamma_{mn}\}_{n=1}^{\infty}$ are the solution to, $m = 0, 1, 2, \dots$

$$J'_{m/2}(\gamma_{mn}) = 0$$

and the normalization constant D_{mn} is determined by

$$D_{mn}^{-2} = \frac{\pi a^2}{\gamma_{mn}^2} \int_0^{\gamma_{mn}} (J_{m/2}(x))^2 x \, dx$$

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%.

6.5 a) $P_r/P_i = 6.1 \cdot 10^{-3}$

b) $P_r = 0$ for the frequency $f = \frac{c_0 \sqrt{b^2 + a^2}}{2ab} \sqrt{\frac{\epsilon + 1}{\epsilon}} = 7.65 \text{ GHz}$

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.

6.6 a) TM_{01} , TE_{11} and TE_{21}

b) $\alpha = \frac{\omega \mu_0 \sigma}{2(\omega^2 \mu_0 \epsilon_0 \epsilon - k_{tnj}^2)^{1/2}} = \frac{\sigma \eta}{2(1 - (f_c/f)^2)^{1/2}}$

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 $f_c = ck_{tnj}/(2\pi) = 1.7 \text{ GHz}$ we get

$$\alpha = \frac{1.1 \cdot 10^{-9}}{(1 - (1.7 \cdot 10^9/f)^2)^{1/2}} \text{ Np/m}$$

The attenuation in dB/km is $20 \log(E(0\text{km})/E(1\text{km})) = \alpha \cdot 8.7 \cdot 10^3 \text{ dB/km}$ with α in Np/m.

6.7 a) $x_0 = 4 \text{ cm}$

b) $P_r/P_i = 0$

c) $P_r/P_i = 0$

6.8 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:

$$E_z = B_{nj} J_{2n}(k_{t2n,j} \rho) \sin(2n\phi) \exp(ik_{znj} z)$$

where $k_{znj}^2 = (\omega/c_0)^2 - (k_{t2nj})^2$ and $J_{2n}(k_{t2nj} R) = 0$.

6.10 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.

7.3 a)

$$Q = \frac{\omega_{101} \mu_0 a b d (a^2 + d^2)}{2R_S(2b(a^3 + d^3) + ad(a^2 + d^2))}$$

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 $Q=31300$.

7.4 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 $1.34 \cdot 10^{20}$.

8.2

$$a(0, t - L/c_0) = H(t) - H(t - t_0) + \int_{t - \min(t, t_0)}^t P(-L, t') dt'$$

$$\text{where } P(-L, t) = \begin{cases} c_0 k_t L \frac{I_1(k_t \sqrt{2c_0 L t - c_0^2 t^2})}{\sqrt{2c_0 L t - c_0^2 t^2}}, & \text{when } t < 2L/c_0 \\ c_0 k_t L \frac{J_1(k_t \sqrt{-2c_0 L t + c_0^2 t^2})}{\sqrt{-2c_0 L t + c_0^2 t^2}}, & \text{when } t > 2L/c_0 \end{cases}$$

9.2

$$H_z = \begin{cases} H_n \cos(k_{tn} y) e^{ik_{zn} z}, & \text{då } 0 < y < a \\ H_n \frac{\cos(k_{tn} a) \cos(k'_{tn}(b - y))}{\cos(k'_{tn}(b - a))} e^{ik_{zn} z}, & \text{då } a < y < b \end{cases}$$

where k_{tn} , k'_{tn} and k_{zn} are given by

$$\begin{aligned} k_{tn}^2 &= (\omega/c_0)^2 \epsilon - k_{zn}^2 \\ k'_{tn}{}^2 &= (\omega/c_0)^2 - k_{zn}^2 \\ k_{tn} \tan(k'_{tn}(b - a)) &= -k'_{tn} \tan(k_{tn} a) \end{aligned}$$

9.3

$$\begin{aligned} v_m &= A J_m \left(k_0 \sqrt{n_1^2 - n_2^2} \right) \left(\frac{a}{\rho} \right)^m \sin m\phi \\ w_m &= -n_2 A J_m \left(k_0 \sqrt{n_1^2 - n_2^2} \right) \left(\frac{a}{\rho} \right)^m \cos m\phi \end{aligned}$$

9.5

$$\mathbf{H} = A_n (\hat{\mathbf{z}} J_0(k_{tn}\rho) + \hat{\boldsymbol{\rho}} i \frac{k_{zn}}{k_{tn}} J'_0(k_{tn}\rho)) \exp(ik_{zn}z), \quad 0 < \rho < a$$

$$\begin{aligned} \mathbf{H} = & B_n (\hat{\mathbf{z}} (J_0(k'_{tn}\rho) - \frac{J'_0(k'_{tn}b)}{Y'_0(k'_{tn}b)} Y_0(k'_{tn}\rho)) \\ & + i \hat{\boldsymbol{\rho}} \frac{k_{zn}}{k'_{tn}} (J'_0(k'_{tn}\rho) - \frac{J'_0(k'_{tn}b)}{Y'_0(k'_{tn}b)} Y'_0(k'_{tn}\rho))) \exp(ik_{zn}z), \quad a < \rho < b \end{aligned}$$

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

$$\begin{aligned} & k_{tn} J_0(k_{tn}a) (J'_0(k'_{tn}a)Y'_0(k'_{tn}b) - J'_0(k'_{tn}b)Y'_0(k'_{tn}a)) \\ &= k'_{tn} J'_0(k'_{tn}a) (J_0(k'_{tn}a)Y'_0(k'_{tn}b) - J'_0(k'_{tn}b)Y_0(k'_{tn}a)) \end{aligned}$$

Index

- Q -value for cavity, 136–140
- Addition theorem
 - Bessel functions, 197
- Ampère’s (generalized) law, 1
- Attenuation, 183
- Attenuation coefficient, 118
- Average
 - temporal, 35
- Bessel function
 - modified, 198–199
 - spherical, 199
- Bessel functions, 166, 193–197
- Boundary conditions, 7
 - dielectric waveguide, 165
 - hollow wave guide, 93
 - hollow waveguide, 91
 - perfectly conducting medium, 8
- Cascaded waveguides, 134
- Causality, 14
- Cavity
 - circular cylindric, 139
- Characteristic equation, 168
- Charge density
 - surface charge density, 6
- Circular cylindrical cavity, 139
- Clausius-Mossotti’s law, 23
- Cole-Cole plot, 33
- Conductivity, 17–20, 25, 32
- Conservation of charge, 2
 - time harmonic fields, 31
- Constitutive relations, 3, 14–20, 26
 - classification, 26
 - conductivity model, 20
 - different formulation, 13
 - dispersion model, 19
 - permeability, 32
 - permittivity, 32
 - time harmonic field, 32–33
- Current density, 1, 2
 - surface current density, 6
- Cut-off frequencies
 - dielectric waveguides, 171
- Cut-off frequency
 - dielectric waveguide, 169, 170, 172, 179
- Debye’s model
 - frequency domain, 33
 - time domain, 20–22
- Dielectric waveguide, 187
 - circular, 165–171
 - optical fiber, 173–187
- Dielectric waveguides, 161
- Dispersion, 15, 181
 - chromatic, 181
 - dispersion parameter, 181
 - material, 181
 - multi mode, 181
 - waveguide, 181
- Dispersion curves, 178
- Drude’s model, 25, 34
- Effective index of refraction, 181
- Electric field, 1
- Electric flux density, 1
- Faraday’s law of induction, 1
- Gauss law, 2
- Graf’s addition theorem, 197
- Hankel function
 - spherical, 199
- Hankel functions, 193–197
- Heaviside function, 156
- Hybrid modes, 169
- Index of refraction, 164
- Invariance under time translation, 14
- Isotropic media, 14–20

- Kaskadkoppling, 134
- Linear material, 14
- Linearly polarized modes, 175–180
- Longitudinal wave number, 87, 99
- Lorentz' model
 - frequency domain, 33
 - time domain, 22–26
- Lorentz's force, 2
- Lorenz-Lorentz' law, 23
- Magnetic field, 1
- Magnetic flux density, 1
- Magnetization, 3
 - induced, 3
 - permanent, 3
- Material dispersion, 100, 181
- Maxwell equations, 1, 2
 - time harmonic fields, 31
 - weak solutions, 6
- Mode, 99
- Mode matching method, 129–134
- Mode power, 116
- Modiferaade Bessel function, 198–199
- Multi mode dipersion, 181
- Neumann function
 - spherical, 199
- Neumann functions, 193–197
- Ohm's law, 17, 32
- Optical fiber, 173–181
 - cladding, 164
 - core, 164
- Optical response, 16
- Permeability, 32
- Permittivity, 32
- Phase velocity, 181
- Plasma frequency, 24
- Polarization, 3
 - induced, 3
 - permanent, 3
- Power
 - modes, 116
- Power flow, 116–120
 - LP-modes, 179–180
- Poynting's theorem, 9, 35
 - time harmonic fields, 35
- Poynting's vector, 9
- Propagatorerna, 156
- Relaxation model
 - frequency domain, 33
 - time domain, 20–22
- Relaxation time, 21
- Resonance cavities, 134–140
- Resonance model
 - frequency domain, 33
 - time domain, 22–26
- SEM-poles, 201
- Singel mode fiber, 161
- Singelmodfiber, 170
- Single mode fiber, 173
- Skin depth, 121
- Spherical Bessel function, 199
- Spherical Hankel function, 199
- Spherical Neumann function, 199
- Surface charge density, 6
- Surface current density, 6
- Susceptibility function, 16
- The conductivity model, 20
- The dispersion model, 19
- Time average, 35
- Time harmonic fields, 29–30
 - condition for real field, 30
- Transient fields, 155–157
- Transmission-lines, 37
- Transverse wave number, 96
- Transverse wavenumber, 88
- Wave impedance
 - relative, 103
 - vacuum, 86
- Wave number
 - longitudinal, 87, 99
 - transverse, 88, 96
- Wavefront, 156
- Waveguide dispersion, 100, 181
- Waveguide modes
 - EH-mode, 168
 - EH-modes, 172
 - HE-mode, 168
 - HE-modes, 171
 - LP-modes, 175–180
 - TE- and TM-modes, 172–173

- TEM-modes, 101–103
- Waveguide modes
 - HE-modes, 169
- Waveguidemodes
 - EH-modes, 171
- Waveguides
 - closed, 91
 - hollow, 91
 - open, 91

Relations between basis vectors

Cylindrical coordinates (ρ, ϕ, z)

Spherical coordinates (r, θ, ϕ)

$$\begin{cases} \rho = \sqrt{x^2 + y^2} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2 + y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2 + y^2}} & y < 0 \end{cases} \\ z = z \end{cases}$$

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \theta = \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \\ \phi = \begin{cases} \arccos \frac{x}{\sqrt{x^2 + y^2}} & y \geq 0 \\ 2\pi - \arccos \frac{x}{\sqrt{x^2 + y^2}} & y < 0 \end{cases} \end{cases}$$

$$(r, \theta, \phi) \longrightarrow (x, y, z)$$

$$\begin{cases} \hat{r} = \hat{x} \sin \theta \cos \phi + \hat{y} \sin \theta \sin \phi + \hat{z} \cos \theta \\ \hat{\theta} = \hat{x} \cos \theta \cos \phi + \hat{y} \cos \theta \sin \phi - \hat{z} \sin \theta \\ \hat{\phi} = -\hat{x} \sin \phi + \hat{y} \cos \phi \end{cases}$$

$$(x, y, z) \longrightarrow (r, \theta, \phi)$$

$$\begin{cases} \hat{x} = \hat{r} \sin \theta \cos \phi + \hat{\theta} \cos \theta \cos \phi - \hat{\phi} \sin \phi \\ \hat{y} = \hat{r} \sin \theta \sin \phi + \hat{\theta} \cos \theta \sin \phi + \hat{\phi} \cos \phi \\ \hat{z} = \hat{r} \cos \theta - \hat{\theta} \sin \theta \end{cases}$$

$$(\rho, \phi, z) \longrightarrow (x, y, z)$$

$$\begin{cases} \hat{\rho} = \hat{x} \cos \phi + \hat{y} \sin \phi = (\hat{x}x + \hat{y}y)/\sqrt{x^2 + y^2} \\ \hat{\phi} = -\hat{x} \sin \phi + \hat{y} \cos \phi = (-\hat{x}y + \hat{y}x)/\sqrt{x^2 + y^2} \\ \hat{z} = \hat{z} \end{cases}$$

$$(x, y, z) \longrightarrow (\rho, \phi, z)$$

$$\begin{cases} \hat{x} = \hat{\rho} \cos \phi - \hat{\phi} \sin \phi \\ \hat{y} = \hat{\rho} \sin \phi + \hat{\phi} \cos \phi \\ \hat{z} = \hat{z} \end{cases}$$

$$(r, \theta, \phi) \longrightarrow (\rho, \phi, z)$$

$$\begin{cases} \hat{r} = \hat{\rho} \sin \theta + \hat{z} \cos \theta \\ \hat{\theta} = \hat{\rho} \cos \theta - \hat{z} \sin \theta \\ \hat{\phi} = \hat{\phi} \end{cases}$$

$$(\rho, \phi, z) \longrightarrow (r, \theta, \phi)$$

$$\begin{cases} \hat{\rho} = \hat{r} \sin \theta + \hat{\theta} \cos \theta \\ \hat{\phi} = \hat{\phi} \\ \hat{z} = \hat{r} \cos \theta - \hat{\theta} \sin \theta \end{cases}$$