Contents

Preface

This is the second version of the lecture notes of *Applied Electromagnetics*. The first author (Daniël De Zutter) has been teaching Electromagnetics for many years now: "Electromagnetics I, II and III" in the electrical engineering curriculum and "Electromagnetics and Special Relativity" in the applied physics engineering curriculum. The original version of those courses was based on the extensive lecture notes of prof. em. Jean Van Bladel and on the research experience of its author. In 1964 J. Van Bladel founded the "Laboratory of Electromagnetism and Acustics", the present "Department of Information Technology" (INTEC) headed by P. Lagasse. When Frank Olyslager became a staff member of the Electromagnetics Group of INTEC, he thoroughly reworked and enlarged the "Electromagnetics and Special Relativity" course notes and updated the Electromagnetics III course.

With the transition to the Bachelor-Master structure, came the need for a new introductory course on electromagnetics for the third bachelor year of the Electrical Engineering bachelor. The title of this course, "Applied Electromagnetics", emphasises the fact that the course is intended to provide the fundamentals of electromagnetics at the same time introducing the electrical engineering student to applications in wave and signal propagation, multiconductor transmission lines, lumped element representation of interconnections, waveguides, antennas, ... This does not imply that the course is an encyclopedia of applications of electromagnetics in electrical engineering. On the contrary. The authors are convinced that students benefit most from a thorough understanding of the basic principles of Electromagnetics. However, in comparison to the electromagnetic courses for the applied physics engineering students, less prominence is given to a number of theoretical concepts and to the physics of materials.

The first author assumes the full responsibility for writing these notes in English. In this way the student will quickly familiarise himself with the typical terminology of Electromagnetics and will all the much easier find his way to the engineering literature. The authors also want to contribute to the much desired internationalisation of higher education curricula. However, the course itself will be taught in Dutch and due care will be taken to explore the rich possibilities of our own language for science teaching.

We hope the students will appreciate this course. There invaluable comments and input when teaching the course for the first time, provided the basis for the second version of these notes. Students interested in the long-standing

and internationally recognised research of the Electromagnetics Group are invited to visit the web pages of the group at
http://www.intec.rug.ac.be/em/.

Daniël De Zutter,
Frank Olyslager,

September 2004.

Chapter 1

Introduction

1.1 Electromagnetics

The basic equations of Electromagnetics are Maxwell's equations first put forward by James Clark Maxwell in 1864. They describe the coupling between electric and magnetic fields and lead to the concept of electromagnetic waves. Maxwell's equations constitute the first "modern" physics theory, soon to be followed by the Special Theory of Relativity and by Quantum Physics. Maxwell's equations predict the speed of light for wave propagation in vacuum and remain invariant under the Lorentz transformation, i.e. they are compatible with the Special Theory of Relativity.

From last year's physics course, the student is already familiar with electric and magnetic fields. As is customary in many physics courses, the basic concepts of Electromagnetism are introduced retracing the historical developments. The laws of Coulomb, Gauss, Ampère and Faraday are formulated starting from the experiments leading to their discovery. Maxwell's equations then provide the final unifying framework. In this course, Maxwell's equations form the starting point to study the propagation and scattering of waves, the principles of antennas and antenna communication, transmission line theory, signal propagation along multiconductor lines and waveguides but also to study electrostatics, magnetostatics and the so-called quasi-static behaviour of circuits and their description in terms of lumped elements. Advantage is taken of the extensive mathematical background of the third year's bachelor student to provide a mathematically rigourous treatment of the various topics.

Although Maxwell's equations have been around for about one-an-a-half century, Electromagnetism still constitutes a flourishing research domain. The main reason for this is the prominent role played by electromagnetic fields in signal propagation, both in modern wireless communication systems such as e.g. GSM, UMTS and Hyperlan as in electronic circuits operating at clock speeds above 1GHz. Even for chip level interconnect, full-wave effects start to play a role as clock harmonics up to 10GHz have to be taken into account. In nowadays opto-electronics, scalar approximations to Maxwell's equations no longer suffice and the full vectorial nature of wave propagation must be considered. An area

of great interest is that of photonic bandgap materials or more in general that of electromagnetic crystals: artificial materials intended to "mold the flow of light". Another increasingly important domain is that of remote sensing and inverse scattering using electromagnetic waves for the non-destructive testing of materials, for the detection of buried objects such as landmines or ore deposits and for medical imaging.

All these applications require the development of specialised and powerful CAD-tools for the solution of Maxwell's equations. In this introductory course only very limited attention is devoted to numerical techniques such as finite elements, finite differences, integral equations in conjunction with the method of moments, the finite difference time domain method and many others. These numerical techniques are the subject matter of advanced courses in Computational Electromagnetics.

During the past 15 years the Electromagnetics Group of the Department of Information Technology has been on the forefront of electromagnetic simulation techniques with applications in a large number of areas including microwave and RF circuits, scattering, waveguides, packaging for digital systems, Electromagnetic Compatibility (EMC) and indoor propagation. Much attention has been devoted both to the development of powerful simulation tools and to industrially relevant applications. Research results were reported in more than 160 international journal papers and resulted in 25 Ph.Ds.

1.2 Notations

In this course italic symbols such as a or α stand for scalar quantities. Their bold counterparts $\mathbf{a}$ and $\boldsymbol{\alpha}$ are used for vectors. The symbol $\mathbf{u}$ stands for a unit vector; $\mathbf{u}_x$, $\mathbf{u}_y$ and $\mathbf{u}_z$ are the unit vectors along the co-ordinate axes in a cartesian co-ordinate system, while $\mathbf{u}_r$, $\mathbf{u}_\phi$ and $\mathbf{u}_\theta$ are the unit vectors for the co-ordinates in a spherical co-ordinate system. The notation $\mathbf{u}_n$ stands for a unit normal vector to a curved line or to a surface. The notation for the place vector is $\mathbf{r} = x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z = r\mathbf{u}_r$.

$\overline{\overline{A}}$ and $\overline{\overline{\alpha}}$ stand for tensors. In a particular co-ordinate system, tensors can be represented by a matrix, e.g. tensor $\overline{\overline{A}}$ in cartesian co-ordinates becomes

$$\overline{\overline{A}} = \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix}. \tag{1.1}$$

The scalar product of a tensor and a vector in cartesian co-ordinates is

$$\begin{aligned} \overline{\overline{A}} \cdot \mathbf{a} &= (A_{xx}a_x + A_{xy}a_y + A_{xz}a_z)\mathbf{u}_x + (A_{yx}a_x + A_{yy}a_y + A_{yz}a_z)\mathbf{u}_y \\ &\quad +(A_{zx}a_x + A_{zy}a_y + A_{zz}a_z)\mathbf{u}_z. \end{aligned} \tag{1.2}$$

1.3 Frequency domain, sinusoidal regime, phasors

In the major part of this course, as in much of the engineering literature on electromagnetics, frequency domain phenomena are studied. The main reason for this is that in passing from the time domain to the frequency domain, derivatives with respect to time are replaced by a multiplicative $j\omega$-factor, resulting, as we will soon see, in a considerable simplification of Maxwell's differential equations. If $f(t)$ represents a time domain signal, the corresponding frequency domain signal is found by the Fourier transformation

$$f(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} \mathrm{d}t, \tag{1.3}$$

with ω the angular frequency. No separate notation is used to distinguish between the time domain version of $f(t)$ and its Fourier transform $f(\omega)$. The distinction between time and frequency domain will always be made clear to the reader, either through the context of a particular reasoning, proof or calculation or by explicitly mentioning the argument t or ω. The inverse Fourier transformation is defined as

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{j\omega t} \mathrm{d}\omega. \tag{1.4}$$

Fourier transformation of a first order time domain differential equation of the form

$$a \frac{d}{dt} f(t) + b f(t) = g(t), \tag{1.5}$$

gives

$$j\omega a f(\omega) + b f(\omega) = g(\omega). \tag{1.6}$$

To obtain (1.6) partial integration was used together with the fact that $f(t \to \infty) = 0$ and $f(t \to -\infty) = 0$. From (1.6) it is clear that in the Fourier domain the time derivative $\frac{d}{dt}$ is replaced by $j\omega$.

A pure (and real) sinusoidal signal with angular frequency ω_0 can in general be written as

$$f(t) = |f| \cos(\omega_0 t + \phi) = \Re[|f| e^{j\phi} e^{j\omega_0 t}] = \Re[f e^{j\omega_0 t}], \tag{1.7}$$

with $f = |f| e^{j\phi}$ a complex number known as the *phasor*. $|f|$ is the amplitude of f and ϕ is its phase. We again use the same letter f to represent the phasor. The Fourier transform of the sinusoidal signal (1.7) is

$$f(\omega) = \pi[f \delta(\omega - \omega_0) + f^* \delta(\omega + \omega_0)]. \tag{1.8}$$

with $\delta(\tau)$ the Dirac delta-function of argument τ. We now again turn to the first order differential equation (1.5) and its Fourier transform (1.6). Substituting $f(\omega)$ from (1.8) in (1.6) and a similar substitution for $g(\omega)$, shows that

$$j\omega_0 a f + b f = g, \qquad -j\omega_0 a f^* + b f^* = g^*, \tag{1.9}$$

with f and g the phasors of $f(t)$ and $g(t)$. The above equations result from the identification of the contributions to the right-hand and left-hand side of (1.6) at $\omega = \omega_0$ and $\omega = -\omega_0$. These equations are identical and one of them suffices. We observe that the time derivative $\frac{d}{dt}$ is now replaced by a multiplication of the phasor$f$ with $j\omega_0$, i.e the obtained result is formally identical with the rule for the Fourier transform of $\frac{d}{dt}f(t)$. Hence, in the sequel no distinction will be made between the frequency domain and the pure sinusoidal regime. To emphasise this point, the inverse Fourier transformation (1.4) is rewritten as

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(\omega)e^{j\omega t}\mathrm{d}\omega \\ &= \frac{1}{2\pi}\int_{0}^{+\infty} f(\omega)e^{j\omega t}\mathrm{d}\omega + \frac{1}{2\pi}\int_{-\infty}^{0} f(\omega)e^{j\omega t}\mathrm{d}\omega \\ &= \frac{1}{2\pi}\int_{0}^{+\infty} f(\omega)e^{j\omega t}\mathrm{d}\omega + \frac{1}{2\pi}\int_{0}^{+\infty} f^*(\omega)e^{-j\omega t}\mathrm{d}\omega \\ &= \frac{1}{\pi}\int_{\infty}^{0} \Re[f(\omega)e^{j\omega t}]\mathrm{d}\omega, \end{aligned} \tag{1.10}$$

where the fact that for a real signal $f(t)$, $f(-\omega) = f^*(\omega)$ holds, was taken into account. Not withstanding the factor $1/\pi$, (1.10) is the superposition of signals of the form (1.7), which again shows the intimate relationship between the use of Fourier transformed signals and the use of phasors in sinusoidal regime.

To conclude this section, suppose that $f(t)$ and $g(t)$ are two sinusoidal signals i.e.

$$f(t) = |f|\cos(\omega t+\phi) = \Re[fe^{j\omega t}], \qquad g(t) = |g|\cos(\omega t+\psi) = \Re[ge^{j\omega t}]. \tag{1.11}$$

Their product $f(t)g(t)$ is no longer sinusoidal but has two frequency components, a DC component and a component at twice the original frequency ω,

$$f(t)g(t) = \frac{1}{2}|f||g|[\cos(2\omega t + \phi + \psi) + \cos(\phi - \psi)]. \tag{1.12}$$

The time average value $\overline{f(t)g(t)}$, i.e. the average over a single period, of this signal is

$$\overline{f(t)g(t)} = \frac{1}{2}|f||g|\cos(\phi - \psi) = \frac{1}{2}\Re[fg^*]. \tag{1.13}$$

The product of a phasor f with the complex conjugate g^* of the phasor g, divided by 2, equals the time average over one period of the signal $f(t)g(t)$. In particular, we have that

$$\overline{f^2(t)} = \frac{1}{2}\Re[ff^*] = \frac{1}{2}|f|^2. \tag{1.14}$$

The student is invited to examine how (1.13) translates to the Fourier transform domain (hint: use Parseval's theorem).

The phasor definition (1.7) adopts the convention of the engineering literature i.e. a $e^{j\omega t}$-dependence. In the physics literature a $e^{-i\omega t}$-dependence is most often used. It suffices to replace j by $-i$ and vice versa in each equation to switch between conventions.

1.4 Polarisation

Polarisation is a general property of vectors with a sinusoidal time dependence. Hence, electric and magnetic fields in the frequency domain will also be polarised. To conclude this chapter, we take a closer look at the physical meaning of polarisation. To this end consider the vector $\mathbf{a}(t)$ which varies sinusoidally in time

$$\begin{aligned}\mathbf{a}(t) &= |a_x|\cos(\omega t+\phi_x)\mathbf{u}_x + |a_y|\cos(\omega t+\phi_y)\mathbf{u}_y + |a_z|\cos(\omega t+\phi_z)\mathbf{u}_z \\ &= \Re[(a_x\mathbf{u}_x + a_y\mathbf{u}_y + a_z\mathbf{u}_z)e^{j\omega t}]. \end{aligned} \tag{1.15}$$

This can be compactly rewritten as

$$\mathbf{a}(t) = \Re[\mathbf{a}e^{j\omega t}], \tag{1.16}$$

with $\mathbf{a}$ a complex vector representing the phasor of $\mathbf{a}(t)$. This complex phasor can be written as the sum of its real and imaginary part

$$\mathbf{a} = \mathbf{a}_r + j\mathbf{a}_i, \tag{1.17}$$

with the real vectors $\mathbf{a}_r$ and $\mathbf{a}_i$ given by

$$\mathbf{a}_r = |a_x|\cos\phi_x\mathbf{u}_x + |a_y|\cos\phi_y\mathbf{u}_y + |a_z|\cos\phi_z\mathbf{u}_z, \tag{1.18}$$

$$\mathbf{a}_i = |a_x|\sin\phi_x\mathbf{u}_x + |a_y|\sin\phi_y\mathbf{u}_y + |a_z|\sin\phi_z\mathbf{u}_z. \tag{1.19}$$

Consequently, the time domain vector $\mathbf{a}(t)$ becomes

$$\mathbf{a}(t) = \Re[(\mathbf{a}_r + j\mathbf{a}_i)e^{j\omega t}] = \mathbf{a}_r\cos\omega t - \mathbf{a}_i\sin\omega t. \tag{1.20}$$

Expression (1.20) shows that, as a function of time, $\mathbf{a}(t)$ moves in a plane defined by the real vectors $\mathbf{a}_r$ and $\mathbf{a}_i$. This plane is the *polarisation plane*. Let us now introduce two new vectors $\tilde{\mathbf{a}}_r$ and $\tilde{\mathbf{a}}_i$ defined as

$$\tilde{\mathbf{a}}_r = \mathbf{a}_r\cos\phi + \mathbf{a}_i\sin\phi, \tag{1.21}$$

$$\tilde{\mathbf{a}}_i = -\mathbf{a}_r\sin\phi + \mathbf{a}_i\cos\phi, \tag{1.22}$$

with the angle ϕ given by

$$\mathrm{tg}2\phi = \frac{2\mathbf{a}_r\cdot\mathbf{a}_i}{\mathbf{a}_r\cdot\mathbf{a}_r - \mathbf{a}_i\cdot\mathbf{a}_i}. \tag{1.23}$$

It can readily be verified that $\tilde{\mathbf{a}}_r\cdot\tilde{\mathbf{a}}_i = 0$, meaning that $\tilde{\mathbf{a}}_r$ and $\tilde{\mathbf{a}}_i$ are perpendicular to each other. Remark that this is not the case for $\mathbf{a}_r$ and $\mathbf{a}_i$ implying that (1.21) and (1.22) do not represent a simple rotation. The inverse transformation is

$$\mathbf{a}_r = \tilde{\mathbf{a}}_r\cos\phi - \tilde{\mathbf{a}}_i\sin\phi, \tag{1.24}$$

$$\mathbf{a}_i = \tilde{\mathbf{a}}_r\sin\phi + \tilde{\mathbf{a}}_i\cos\phi. \tag{1.25}$$

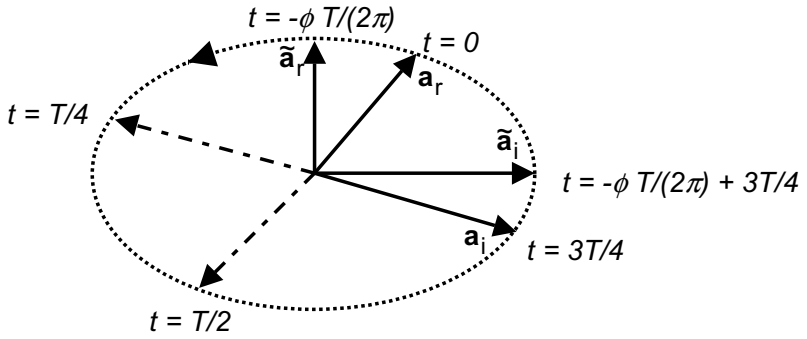


Figure 1.1: Polarisation ellipse.

Substituting (1.24) and (1.25) into (1.20) yields

$$\mathbf{a}(t) = \tilde{\mathbf{a}}_r \cos(\omega t + \phi) - \tilde{\mathbf{a}}_i \sin(\omega t + \phi). \tag{1.26}$$

As $\tilde{\mathbf{a}}_r$ and $\tilde{\mathbf{a}}_i$ are perpendicular, $\mathbf{a}(t)$ describes an ellipse, the *polarisation ellipse*. The ellipticity κ is defined as

$$\kappa = \frac{|\tilde{\mathbf{a}}_i|}{|\tilde{\mathbf{a}}_r|} \tag{1.27}$$

and $\tilde{\mathbf{a}}_r$ and $\tilde{\mathbf{a}}_i$ are the main axes. The table below gives some particular values of $\mathbf{a}(t)$ as a function of time

t	$\mathbf{a}(t)$
0	$\mathbf{a}_r$
$T/4$	$-\mathbf{a}_i$
$T/2$	$-\mathbf{a}_r$
$3T/4$	$\mathbf{a}_i$
$-\phi T/(2\pi)$	$\tilde{\mathbf{a}}_r$
$-\phi T/(2\pi) + 3T/4$	$\tilde{\mathbf{a}}_i$

The quantity T is the period of the sinusoidal signal, i.e. $2\pi/\omega$. From the above table it is seen that $\mathbf{a}(t)$ rotates from $\mathbf{a}_i$ towards $\mathbf{a}_r$. This is also shown in Fig. 1.1.

1.4.1 Linear polarisation

Suppose that $\mathbf{a}_r$ and $\mathbf{a}_i$ are parallel to each other

$$\mathbf{a}_r = a_r \mathbf{u}, \qquad \mathbf{a}_i = a_i \mathbf{u}, \tag{1.28}$$

with $\mathbf{u}$ a unit vector along the common direction. This allows to rewrite $\mathbf{a}(t)$ as

$$\mathbf{a}(t) = \Re[(a_r + j a_i) e^{j\omega t}]\mathbf{u}, \tag{1.29}$$

or, with $a_r + j a_i = A e^{j\phi}$, as

$$\mathbf{a}(t) = A \cos(\omega t + \phi)\mathbf{u}. \tag{1.30}$$

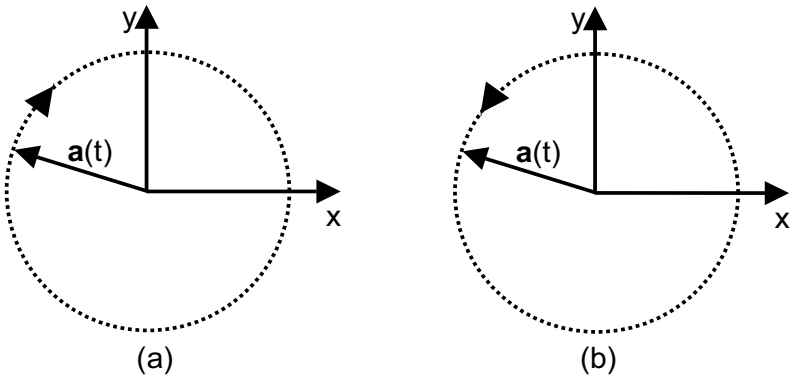


Figure 1.2: Left hand circular (LHC) or clockwise circular polarisation (a); right hand circular (RHC) or counter clockwise circular polarisation (b).

Hence, $\mathbf{a}(t)$ moves sinusoidally up and down a line defined by the unit vector $\mathbf{u}$ and this with amplitude A. The vector is *linearly* polarised and all field components are in phase.

From (1.26) we see that any elliptical polarisation can be written as the superposition of two linear polarisations

$$\mathbf{a}(t) = \tilde{\mathbf{a}}_r \cos(\omega t + \phi) - \tilde{\mathbf{a}}_i \cos(\omega t + \phi - \pi/2). \tag{1.31}$$

Verify that $\mathbf{a} \times \mathbf{a}^* = 0$ is a necessary and sufficient condition for a vector to be linearly polarised.

1.4.2 Circular polarisation

Another special case is obtained for main axes of equal length

$$|\tilde{\mathbf{a}}_r| = |\tilde{\mathbf{a}}_i| = \tilde{A}. \tag{1.32}$$

The polarisation ellipse is now a circle ($\kappa = 1$) and the polarisation is *circular*. Expressions (1.24) en (1.25) show that

$$|\mathbf{a}_r|^2 = |\tilde{\mathbf{a}}_r|^2 \cos^2\phi + |\tilde{\mathbf{a}}_i|^2 \sin^2\phi = \tilde{A}^2, \tag{1.33}$$
$$|\mathbf{a}_i|^2 = |\tilde{\mathbf{a}}_r|^2 \sin^2\phi + |\tilde{\mathbf{a}}_i|^2 \cos^2\phi = \tilde{A}^2, \tag{1.34}$$
$$\mathbf{a}_r \cdot \mathbf{a}_i = (|\tilde{\mathbf{a}}_r|^2 - |\tilde{\mathbf{a}}_i|^2) \cos\phi \sin\phi = 0, \tag{1.35}$$

hence, $\mathbf{a}_r$ and $\mathbf{a}_i$ are also perpendicular and can play the role of main axes. Now verify that $\mathbf{a} \cdot \mathbf{a} = 0$ is a necessary and sufficient condition for a vector to be circularly polarised (i.e. the vector is perpendicular to itself!). Remark that the condition $|\mathbf{a}_r| = |\mathbf{a}_i|$ does not suffice: $\mathbf{a}_r$ and $\mathbf{a}_i$ must be main axes, i.e. perpendicular to each other and of equal length.

Let us select the xy-plane as the polarisation plane such that

$$\mathbf{a}_r = A\mathbf{u}_x, \qquad \mathbf{a}_i = \pm A\mathbf{u}_y. \tag{1.36}$$

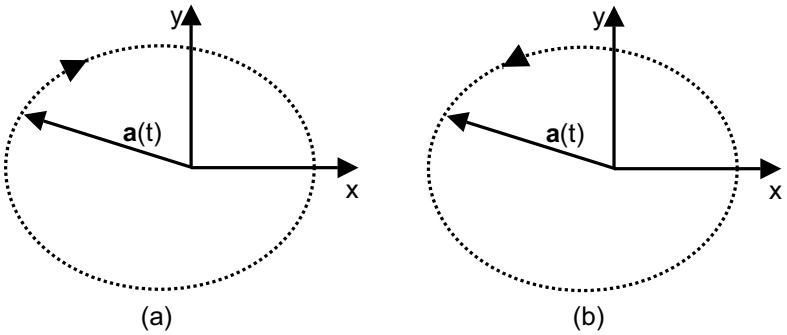


Figure 1.3: Left hand elliptical polarisation (a); right hand elliptical polarisation (b).

This allows to write $\mathbf{a}(t)$ as

$$\mathbf{a}(t) = A\Re[(\mathbf{u}_x \pm j\mathbf{u}_y)e^{j\omega t}] = A(\cos\omega t\mathbf{u}_x \mp \sin\omega t\mathbf{u}_y). \tag{1.37}$$

For $\mathbf{a}_i = A\mathbf{u}_y$ and as depicted in Fig. 1.2a, $\mathbf{a}(t)$ describes the circle in a clockwise way (from y to x) and the polarisation is therefore called clockwise circular polarisation or also left hand circular (LHC) polarisation. Fig. 1.2b, for $\mathbf{a}_i = -A\mathbf{u}_y$, shows the counter clockwise or right hand circular (RHC) polarisation. The phase difference between the two perpendicular components is $\pi/2$ for LHC polarisation and $-\pi/2$ for RHC polarisation. Similarly, elliptical polarisation can either be left or right hand elliptical as shown in Fig. 1.3.

We leave it to the exercises to prove that any elliptical polarisation can be written as the superposition of a LHC and a RHC polarisation.

Chapter 2

Maxwell's Equations

2.1 Differential and integral formulation

The essence of Electromagnetism is captured by Maxwell's equations formulated by James Clark Maxwell in 1864. These equations unify the experimental results obtained earlier by such eminent scientists as Coulomb, Gauss, Faraday and Ampère, amongst others. Maxwell's equations are a set of partial differential equations, originally formulated using quaternions and scalars. Heaviside's version of 1880, using vectors, is the one used today.
The basic equations are the curl equations

$$\nabla \times \mathbf{e}(\mathbf{r},t) = -\frac{\partial}{\partial t}\mathbf{b}(\mathbf{r},t), \tag{2.1}$$

$$\nabla \times \mathbf{h}(\mathbf{r},t) = \frac{\partial}{\partial t}\mathbf{d}(\mathbf{r},t) + \mathbf{j}(\mathbf{r},t), \tag{2.2}$$

with $\mathbf{e}(\mathbf{r},t)$ the electric field in V/m, $\mathbf{h}(\mathbf{r},t)$ the magnetic field in A/m, $\mathbf{d}(\mathbf{r},t)$ the electric induction or dielectric displacement in C/m^2 and $\mathbf{b}(\mathbf{r},t)$ the magnetic induction in Wb/m^2. $\mathbf{j}(\mathbf{r},t)$ is the electrical current density in A/m^2. The curl equations are complemented by two divergence equations

$$\nabla \cdot \mathbf{d}(\mathbf{r},t) = \rho(\mathbf{r},t), \tag{2.3}$$

$$\nabla \cdot \mathbf{b}(\mathbf{r},t) = 0, \tag{2.4}$$

with $\rho(\mathbf{r},t)$ the electrical charge density in C/m^3. Curl and divergence equations together form Maxwell's equations. Combination of (2.2) with the time-derivative of (2.3) yields the important law of charge conservation

$$\nabla \cdot \mathbf{j}(\mathbf{r},t) = -\frac{\partial}{\partial t}\rho(\mathbf{r},t). \tag{2.5}$$

The divergence equation (2.4) can be derived from (2.1) by taking the divergence of both sides of the equation (remember that $\nabla \cdot \nabla \times \mathbf{a} = 0$ for any vector $\mathbf{a}$). However, we have to assume that $\nabla \cdot \mathbf{b}(\mathbf{r},t_0) = 0$ at one particular moment t_0. Similarly, (2.3) can be derived from (2.2) when accepting the law of charge conservation and again assuming that $\nabla \cdot \mathbf{d}(\mathbf{r},t_0) = \rho(\mathbf{r},t_0)$ at one particular moment t_0. Hence, the divergence equations are not superfluous. They act as

initial conditions for the curl equations.
In the frequency domain the curl equations become

$$\nabla \times \mathbf{e}(\mathbf{r}) = -j\omega \mathbf{b}(\mathbf{r}), \tag{2.6}$$
$$\nabla \times \mathbf{h}(\mathbf{r}) = j\omega \mathbf{d}(\mathbf{r}) + \mathbf{j}(\mathbf{r}), \tag{2.7}$$

while the law of charge conservation now reads

$$\nabla \cdot \mathbf{j}(\mathbf{r}) = -j\omega \rho(\mathbf{r}). \tag{2.8}$$

The divergence equations

$$\nabla \cdot \mathbf{d}(\mathbf{r}) = \rho(\mathbf{r}), \tag{2.9}$$
$$\nabla \cdot \mathbf{b}(\mathbf{r}) = 0, \tag{2.10}$$

now follow directly from the curl equations and the charge conservation law. In sinusoidal regime, they are not required for completeness. Although no magnetic charges exist, it turns out to be useful to introduce a magnetic charge density $\pi(\mathbf{r})$ in Wb/m^3 and the corresponding magnetic current density $\mathbf{k}(\mathbf{r})$ in V/m^2. In this introductory course however, magnetic charges and currents will not be used. In practice, they allow to elegantly formulate and solve a number of problems of great practical interest, e.g. the penetration of waves through an opening in a metallic enclosure. To accommodate magnetic currents and charges, (2.6) and (2.10) must be modified as follows

$$\nabla \times \mathbf{e}(\mathbf{r}) = -j\omega \mathbf{b}(\mathbf{r}) - \mathbf{k}(\mathbf{r}), \tag{2.11}$$
$$\nabla \cdot \mathbf{b}(\mathbf{r}) = \pi(\mathbf{r}), \tag{2.12}$$

while the law of conservation of magnetic charge is

$$\nabla \cdot \mathbf{k}(\mathbf{r}) = -j\omega \pi(\mathbf{r}). \tag{2.13}$$

Remark that the introduction of magnetic currents and charges makes Maxwell's equations much more symmetric.

Equations (2.1)–(2.4) are Maxwell's equations in differential form. The corresponding integral form provides additional physical insight and it was also the preferred formulation when introducing electrostatic, magnetostatic and electromagnetic phenomena in last year's physics course. Integration of (2.1) over a surface S with boundary curve c (see Fig. 2.1) and application of Stokes's theorem yield

$$\oint_c \mathbf{e}(\mathbf{r},t) \cdot \mathbf{dc} = -\frac{\partial}{\partial t} \int_S \mathbf{b}(\mathbf{r},t) \cdot \mathbf{u}_n dS = -\frac{\partial}{\partial t} \Phi_b. \tag{2.14}$$

The direction of $\mathbf{u}_n$ is linked to the integration sense along c. Eqn. (2.14) is Faraday's law: a changing magnetic field or more precisely, the change in the flux Φ_b of the magnetic induction through a closed loop, induces an electromotoric force or emf. This emf is proportional to the amplitude of $\mathbf{b}(\mathbf{r},t)$, to the surface of the loop and to the rate of change of the flux. The flux change can be brought about by a change of $\mathbf{b}(\mathbf{r},t)$ but equally so by a change in the surface, position or orientation of the loop. Faraday's law further implies that a

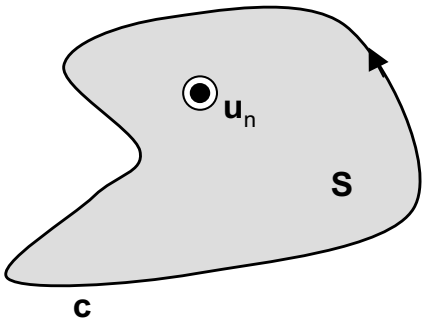


Figure 2.1: Faraday's law.

unique definition of the potential ϕ from which the electric field can be derived through $\mathbf{e} = -\nabla\phi$, as in the electrostatic case, is no longer possible as this uniqueness depends on the conservative nature of the static electric field, i.e. $\oint_c \mathbf{e}(\mathbf{r},t)\cdot\mathbf{dc} = 0$. The integral of $\mathbf{e}$ over a closed loop, $\oint_c \mathbf{e}(\mathbf{r},t)\cdot\mathbf{dc}$, is known as the circulation of $\mathbf{e}$.

Stokes's theorem can similarly be applied to the second curl equation (2.2)

$$\oint_c \mathbf{h}(\mathbf{r},t)\cdot\mathbf{dc} = \int_S \mathbf{j}(\mathbf{r},t)\cdot\mathbf{u}_n dS + \frac{\partial}{\partial t}\int_S \mathbf{d}(\mathbf{r},t)\cdot\mathbf{u}_n dS = i(\mathbf{r},t) + \frac{\partial}{\partial t}\Phi_d. \quad (2.15)$$

In the absence of the contribution of the flux Φ_d of the electric induction, (2.15) reduces to Ampère's law

$$\oint_c \mathbf{h}(\mathbf{r},t)\cdot\mathbf{dc} = \int_S \mathbf{j}(\mathbf{r},t)\cdot\mathbf{u}_n dS = i(\mathbf{r},t). \quad (2.16)$$

This law states that the magnetic field, even in the magnetostatic case, is not conservative. The integral of the magnetic field over a closed loop, i.e. the circulation of $\mathbf{h}$, yields the total current $i(\mathbf{r},t)$ passing through the surface defined by that loop. The important step taken by Maxwell was to recognise that Ampère's law is incomplete. The curl of the magnetic field and hence its integral over a closed loop is not only influenced by currents but also by the rate of change of the flux of the electric induction. Only in the magnetostatic case, the magnetic field can be derived from the sole knowledge of the current distribution.

As will be shown later in this chapter, the existence of wave solutions depends on the presence of the electric induction term. Whereas (2.14) shows that a changing magnetic fields induces an electric field, (2.15) shows that a changing electric field induces a magnetic field (even in the absence of current!). It was Hertz who could first experimentally show the existence of electromagnetic waves in 1887.

Gauss's law is obtained by integrating (2.3) over a volume V with boundary surface S (see Fig. 2.2)

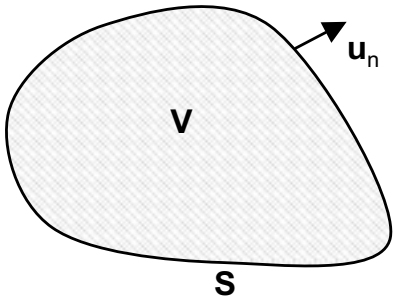


Figure 2.2: Gauss's law

$$\int_S \mathbf{d}(\mathbf{r},t)\cdot\mathbf{u}_n dS = \int_V \rho(\mathbf{r},t)dV = q_V(t). \tag{2.17}$$

The flux of the electric induction or dielectric displacement through any closed surface equals the total charge $q_V(t)$ inside the volume enclosed by that surface. Similarly, (2.4) leads to

$$\int_S \mathbf{b}(\mathbf{r},t)\cdot\mathbf{u}_n dS = 0. \tag{2.18}$$

In the absence of magnetic charges, the flux of the magnetic induction through any closed surface is zero. Finally, the law of charge conservation in integral form becomes

$$\int_S \mathbf{j}(\mathbf{r},t)\cdot\mathbf{u}_n dS = -\frac{\partial}{\partial t}q_V(t), \tag{2.19}$$

stating that the total current flowing out of a closed surface entails a decrease of the total charge inside the volume enclosed by that surface.

The reader can easily derive the frequency domain counterparts of the above integral formulations of Maxwell's equations. It must be remarked that the integral formulation is completely equivalent to the differential formulation. This follows from the fact that the volumes, surfaces and boundary contours used to derive the integral formulation are completely arbitrary.

2.2 Constitutive equations

The vectorial curl equations enforce six scalar relationships on the twelve scalar components of the fields $\mathbf{e}(\mathbf{r})$, $\mathbf{h}(\mathbf{r})$, $\mathbf{d}(\mathbf{r})$ and $\mathbf{b}(\mathbf{r})$. We will soon discuss the role of $\mathbf{j}(\mathbf{r})$ in detail. However, for the moment, simple suppose the current density $\mathbf{j}(\mathbf{r})$ is known. Hence, to solve Maxwell's equations, additional relationships between the field components are necessary. These additional relationships are the constitutive equations. In vacuum, they are given by

$$\mathbf{d}(\mathbf{r}) = \epsilon_0\mathbf{e}(\mathbf{r}), \tag{2.20}$$

$$\mathbf{b}(\mathbf{r}) = \mu_0\mathbf{h}(\mathbf{r}), \tag{2.21}$$

with μ_0 the permeability of vacuum given by $4\pi 10^{-7} H/m$ and ϵ_0 the permittivity of vacuum defined as

$$\epsilon_0 = \frac{1}{c^2 \mu_0} = 8.854187818\ 10^{-12} F/m, \tag{2.22}$$

with c the speed of light in vacuum, by convention fixed at $2.99792458\ 10^8 m/s$. In this course, the term "free space" will often be used to refer to vacuum, as in much of the technical literature. As will be proved in the chapter on plane waves, the speed of electromagnetic waves in free space is

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}. \tag{2.23}$$

The constitutive equations for material media can be obtained by passing from a microscopic field description to a macroscopic one. This problem will not be treated here. In sinusoidal regime, the most simple macroscopic constitutive equations for materials are

$$\mathbf{d}(\mathbf{r}, \omega) = \epsilon(\mathbf{r}, \omega)\mathbf{e}(\mathbf{r}, \omega), \tag{2.24}$$

$$\mathbf{b}(\mathbf{r}, \omega) = \mu(\mathbf{r}, \omega)\mathbf{h}(\mathbf{r}, \omega), \tag{2.25}$$

with $\epsilon(\mathbf{r}, \omega)$ the permittivity and $\mu(\mathbf{r}, \omega)$ the permeability of the material. A medium in which $\epsilon(\mathbf{r})$ and $\mu(\mathbf{r})$ do not depend on the fields, as in (2.24) and (2.25), is a linear medium. Furthermore, at any point $\mathbf{r}$, $\mathbf{d}(\mathbf{r})$ and $\mathbf{b}(\mathbf{r})$ only depend on $\mathbf{e}(\mathbf{r})$ and $\mathbf{h}(\mathbf{r})$ at that point: the medium is local. Although in this course the ω-dependence is suppressed when writing down sinusoidal regime equations, we have explicitly shown this dependence in the above equations to make the reader aware of the fact that all quantities, including the permittivity and permeability, in principle depend on frequency. If $\epsilon(\mathbf{r}, \omega)$ and $\mu(\mathbf{r}, \omega)$ are frequency independent, the medium is said to be non-dispersive. If not so, the medium is dispersive. Suppose e.g. that the time-domain constitutive equation for $\mathbf{d}$ takes the following form

$$\mathbf{d}(\mathbf{r}, t) = \epsilon_1(\mathbf{r})\mathbf{e}(\mathbf{r}, t) + \epsilon_2(\mathbf{r})\frac{\partial}{\partial t}\mathbf{e}(\mathbf{r}, t). \tag{2.26}$$

The corresponding frequency domain expression is

$$\epsilon(\mathbf{r}) = \epsilon_1(\mathbf{r}) + j\omega\epsilon_2(\mathbf{r}), \tag{2.27}$$

showing that $\epsilon(\mathbf{r})$ can be complex-valued (and this of course also applies to $\mu(\mathbf{r})$). To explicitly introduce the real and imaginary part of $\epsilon(\mathbf{r})$ and $\mu(\mathbf{r})$, the notation $\epsilon = \epsilon_R + j\epsilon_I$ and $\mu = \mu_R + j\mu_I$ will be used. For a lossy dielectric it is customary to define the loss tangent as

$$\mathrm{tg}\delta = -\frac{\epsilon_I}{\epsilon_R}. \tag{2.28}$$

In a good dielectric $\mathrm{tg}\delta$ and hence δ are very small, i.e. of the order of 10^{-3} or lower such that $\epsilon_R \gg -\epsilon_I$. In a good conductor this is just the other way around. Here $\epsilon_R \ll -\epsilon_I = \sigma/\omega$. For example for copper at $100GHz$ with $\sigma = 5.8\ 10^7 S/m$ we have that $\epsilon_R = \epsilon_0 \approx 8.85\ 10^{-12} F/m \ll -\epsilon_I = 9.2\ 10^{-5} F/m$

such that ϵ_R can be neglected in all calculations.
When ϵ en μ are place independent, the medium is said to be homogeneous. The relative permittivity ϵ_r and permeability μ_r are defined as

$$\epsilon_r = \epsilon/\epsilon_0, \qquad \mu_r = \mu/\mu_0, \tag{2.29}$$

while the refractive index n of a material is given by $n = \sqrt{\epsilon_r \mu_r}$.
Eqn. (2.24) also implies that $\mathbf{d}(\mathbf{r})$ is independent of the direction or polarisation of $\mathbf{e}(\mathbf{r})$ and from (2.25) this also applies to $\mathbf{b}(\mathbf{r})$ and $\mathbf{h}(\mathbf{r})$. The medium has no preferential directions: it is isotropic. In forthcoming chapters, free space, homogeneous or piecewise homogeneous linear, isotropic and non-dispersive materials will be used. However, many solids exhibit preferential directions due to their crystalline structure. This translates into the constitutive equations as

$$\mathbf{d}(\mathbf{r}) = \overline{\overline{\epsilon}}(\mathbf{r}) \cdot \mathbf{e}(\mathbf{r}), \tag{2.30}$$

$$\mathbf{b}(\mathbf{r}) = \overline{\overline{\mu}}(\mathbf{r}) \cdot \mathbf{h}(\mathbf{r}), \tag{2.31}$$

with $\overline{\overline{\epsilon}}(\mathbf{r})$ and $\overline{\overline{\mu}}(\mathbf{r})$ the permittivity and permeability tensors. The medium is still linear but anisotropic.

We now turn our attention to the current density $\mathbf{j}(\mathbf{r})$. Up to now we simply assumed this current density to be known, i.e. externally enforced. These currents are the sources of the fields. However, charge displacements and hence currents can also be caused by the fields themselves. In that case we talk about induced currents. The current $\mathbf{j}(\mathbf{r})$ should hence be written as the sum of two contributions

$$\mathbf{j}(\mathbf{r}) = \mathbf{j}_i(\mathbf{r}) + \mathbf{j}_e(\mathbf{r}), \tag{2.32}$$

with $\mathbf{j}_i(\mathbf{r})$ the induced current density and $\mathbf{j}_e(\mathbf{r})$ the externally enforced source current density. We now need an additional constitutive equation to specify the relationship of the induced current to the fields. The simplest one is that for the current in a conductor

$$\mathbf{j}_i(\mathbf{r}) = \sigma(\mathbf{r})\mathbf{e}(\mathbf{r}). \tag{2.33}$$

This is Ohm's law with $\sigma(\mathbf{r})$ the electrical conductivity (or simply conductivity) in S/m.

In an isotropic medium with constitutive equations (2.24), (2.25) and (2.33), the curl equations can be written as

$$\nabla \times \mathbf{e}(\mathbf{r}) = -j\omega\mu(\mathbf{r})\mathbf{h}(\mathbf{r}), \tag{2.34}$$

$$\nabla \times \mathbf{h}(\mathbf{r}) = j\omega\epsilon(\mathbf{r})\mathbf{e}(\mathbf{r}) + \sigma(\mathbf{r})\mathbf{e}(\mathbf{r}) + \mathbf{j}_e(\mathbf{r}). \tag{2.35}$$

The second equation can be more concisely written as

$$\nabla \times \mathbf{h}(\mathbf{r}) = j\omega\tilde{\epsilon}(\mathbf{r})\mathbf{e}(\mathbf{r}) + \mathbf{j}_e(\mathbf{r}), \tag{2.36}$$

with the generalised permittivity $\tilde{\epsilon}(\mathbf{r})$ defined as

$$\tilde{\epsilon}(\mathbf{r}) = \epsilon(\mathbf{r}) + \frac{\sigma(\mathbf{r})}{j\omega}. \tag{2.37}$$

Sometimes $\tilde{\epsilon}(\mathbf{r})$ is indicated as the "complex" permittivity, showing that in the frequency domain the conductivity can be merged with the permittivity. The

term "complex" permittivity is quite misleading as $\epsilon(\mathbf{r})$ itself can already be complex (see (2.27)). In a similar way the generalised electric induction is defined as

$$\tilde{\mathbf{d}}(\mathbf{r}) = \tilde{\epsilon}(\mathbf{r})\mathbf{e}(\mathbf{r}) = \mathbf{d}(\mathbf{r}) + \frac{1}{j\omega}\mathbf{j}_i(\mathbf{r}), \tag{2.38}$$

such that the divergence equation

$$\nabla \cdot \tilde{\mathbf{d}}(\mathbf{r}) = \rho_e(\mathbf{r}), \tag{2.39}$$

only involves externally enforced charge densities or source charge densities. In the sequel generalised permittivities and permeabilities will be used, unless explicitly mentioned otherwise and for notational simplicity the tildes will be left out. Maxwell's curl equations in a linear isotropic medium then become

$$\nabla \times \mathbf{e}(\mathbf{r}) = -j\omega\mu(\mathbf{r})\mathbf{h}(\mathbf{r}), \tag{2.40}$$

$$\nabla \times \mathbf{h}(\mathbf{r}) = j\omega\epsilon(\mathbf{r})\mathbf{e}(\mathbf{r}) + \mathbf{j}_e(\mathbf{r}). \tag{2.41}$$

A special idealised medium of considerable interest to Electromagnetics is the perfect electric conductor or PEC. In a perfect electric conductor all fields are zero and no current or charge densities can exist in its volume. However, its surface supports electric surface current and charge densities. Magnetic current densities or charges can neither exist in its volume nor on its surface. As will later become clear, a perfect electric conductor is a mathematical limiting case obtained when $\sigma \to \infty$. A perfect conductor is a good model for many good conductors (such as e.g. copper) at least over a certain frequency range.
In a PMC or perfect magnetic conductor the role of electric charges and currents in a PEC is taken by magnetic charges and currents: its surface admits magnetic charge and current densities but no electric ones.

2.3 Conservation of energy - Poynting's vector

In this section conservation of energy will be discussed in relationship to Maxwell's equations. This will force us to recognise that radiating fields carry a certain amount of energy and leads to the introduction of Poynting's vector.

Scalar multiplication of (2.6) with $\mathbf{h}^*$ and with $\mathbf{j} = \mathbf{j}_e + \mathbf{j}_i$, of the complex conjugate of (2.7) with $\mathbf{e}$ and subtracting the results, yields

$$\mathbf{h}^* \cdot (\nabla \times \mathbf{e}) - \mathbf{e} \cdot (\nabla \times \mathbf{h}^*) = -j\omega\mathbf{h}^* \cdot \mathbf{b} - \mathbf{e} \cdot (\mathbf{j}_e^* + \mathbf{j}_i^*) + j\omega\mathbf{e} \cdot \mathbf{d}^*, \tag{2.42}$$

or

$$\nabla \cdot (\mathbf{e} \times \mathbf{h}^*) = -j\omega\mathbf{h}^* \cdot \mathbf{b} - \mathbf{e} \cdot (\mathbf{j}_e^* + \mathbf{j}_i^*) + j\omega\mathbf{e} \cdot \mathbf{d}^*. \tag{2.43}$$

Now consider a volume V with boundary surface S and outward directed unit normal $\mathbf{u}_n$. Fig. 2.3 shows two examples of such volumes, either with V bounded by a single surface or by two surfaces ($S = S_1 \cup S_2$). Integration of (2.43) over V and application of the divergence theorem shows that

$$\int_S (\mathbf{e} \times \mathbf{h}^*) \cdot \mathbf{u}_n dS = \int_V (-j\omega\mathbf{h}^* \cdot \mathbf{b} - \mathbf{e} \cdot (\mathbf{j}_e^* + \mathbf{j}_i^*) + j\omega\mathbf{e} \cdot \mathbf{d}^*) dV. \tag{2.44}$$

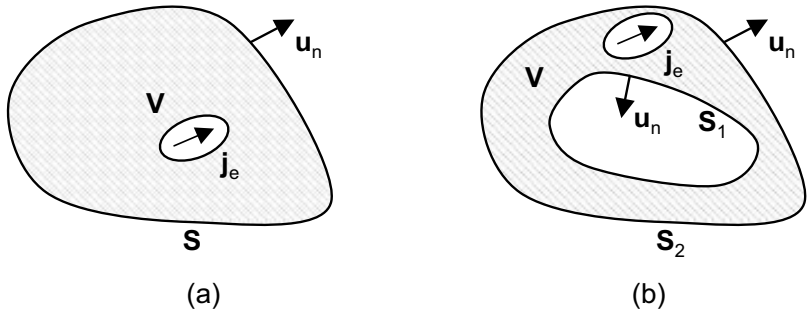


Figure 2.3: Integration volume for the energy calculations: bounded by a single surface (a), or by two surfaces (b).

At this point the complex *Poynting vector* $\mathbf{p}(\mathbf{r})$, defined as

$$\mathbf{p}(\mathbf{r}) = \frac{1}{2}\mathbf{e}(\mathbf{r}) \times \mathbf{h}^*(\mathbf{r}), \tag{2.45}$$

is introduced, allowing to rewrite (2.44) as

$$-\frac{1}{2}\int_V \mathbf{e}\cdot\mathbf{j}_e^* dV = \int_S \mathbf{p}\cdot\mathbf{u}_n dS + \frac{1}{2}\int_V \mathbf{e}.\mathbf{j}_i^* dV + \frac{1}{2}\int_V (j\omega\mathbf{h}^*\cdot\mathbf{b} - j\omega\mathbf{e}\cdot\mathbf{d}^*)dV. \tag{2.46}$$

Let us first take a closer look at this result for *free space*. Taking the real and imaginary part of (2.46) then yields

$$-\frac{1}{2}\Re\int_V \mathbf{e}\cdot\mathbf{j}_e^* dV = \Re\int_S \mathbf{p}\cdot\mathbf{u}_n dS, \tag{2.47}$$

$$-\frac{1}{2}\Im\int_V \mathbf{e}\cdot\mathbf{j}_e^* dV = \Im\int_S \mathbf{p}\cdot\mathbf{u}_n dS + \frac{1}{2}\omega\mu_0\int_V |\mathbf{h}|^2 dV - \frac{1}{2}\omega\epsilon_0\int_V |\mathbf{e}|^2 dV. \tag{2.48}$$

From (1.13) we conclude that the left hand side of (2.47) is the average over a single period of $-\mathbf{e}(\mathbf{r},t)\cdot\mathbf{j}_e(\mathbf{r},t)$ integrated over the volume V. Physically, $-\mathbf{e}(\mathbf{r},t)\cdot\mathbf{j}_e(\mathbf{r},t)$ is the power (or energy per unit of time) per unit of volume delivered by the source $\mathbf{j}_e(\mathbf{r},t)$ at each point $\mathbf{r}$ in order to sustain the electromagnetic fields. If V encompasses all sources, the left hand member is the total power delivered by these sources. As free space is lossless, all energy generated inside V must be radiated through its surface S. This leads to the interpretation of Poynting's vector: the integral over a closed surface S of the normal component of the real part of Poynting's vector, or the total flux of $\Re\mathbf{p}$ through S, corresponds to the total power flowing through that surface (or more precisely the time-average of that power over a single period). In the right-hand member of (2.48) we can identify the average values over a period of

$$u_e(t) = \frac{1}{2}\int_V \epsilon_0[\mathbf{e}(t)]^2 dV, \tag{2.49}$$

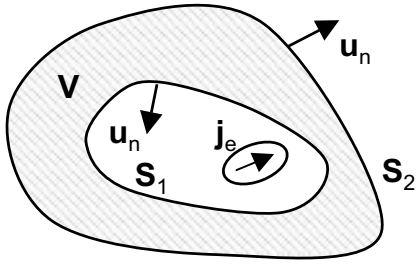


Figure 2.4: Energy considerations for a sourceless volume.

and

$$u_m(t) = \frac{1}{2}\int_V \mu_0[\mathbf{h}(t)]^2 dV. \tag{2.50}$$

From electrostatics and magnetostatics we know that $u_e(t)$ and $u_m(t)$ represent the total electric and magnetic energy stored inside volume V at time t. This allows us to rewrite (2.48) as

$$-\frac{1}{2}\Im\int_V \mathbf{e}\cdot\mathbf{j}_e^* dV = \Im\int_S \mathbf{p}\cdot\mathbf{u}_n dS + 2\omega(\overline{u_m(t)} - \overline{u_e(t)}). \tag{2.51}$$

The left-hand side is the reactive power generated by the sources. Part of this reactive power also flows through the boundary surface(s) S as the total flux of $\Im\mathbf{p}$, the remaining part is 2ω times the difference between the time average value of the magnetic and the electric energy stored inside V.
An interesting special case if obtained for the situation depicted in Fig. 2.4 with all sources located inside surface S_1. Consequently, no sources are present in the volume bounded by S_1 and S_2. According to (2.47) (the left-handed member of which is zero) we find

$$\Re\int_{S_1\cup S_2} \mathbf{p}\cdot\mathbf{u}_n dS = 0, \tag{2.52}$$

or

$$\Re\int_{S_1} \mathbf{p}\cdot(-\mathbf{u}_n) dS = \Re\int_{S_2} \mathbf{p}\cdot\mathbf{u}_n dS. \tag{2.53}$$

The total power entering S_1 equals the total power leaving S_2. A similar conclusion applies to the reactive power.

Next, the general result (2.46) is applied to an isotropic medium characterised by a complex-valued $\epsilon(\mathbf{r})$ and $\mu(\mathbf{r})$ and by a real-valued conductivity $\sigma(\mathbf{r})$

$$\begin{aligned}-\frac{1}{2}\Re\int_V \mathbf{e}\cdot\mathbf{j}_e^* dV = \Re\int_S \mathbf{p}\cdot\mathbf{u}_n dS + \frac{1}{2}\int_V \sigma|\mathbf{e}|^2 dV \\ -\frac{1}{2}\omega\int_V \mu_I|\mathbf{h}|^2 dV + \frac{1}{2}\omega\int_V \epsilon_I|\mathbf{e}|^2 dV,\end{aligned} \tag{2.54}$$

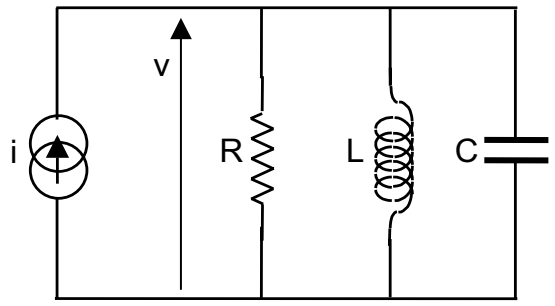


Figure 2.5: Conservation of energy in a simple circuit

$$-\frac{1}{2}\Im\int_V \mathbf{e}\cdot\mathbf{j}_e^* dV = \Im\int_S \mathbf{p}\cdot\mathbf{u}_n dS + \frac{1}{2}\omega\int_V \mu_R|\mathbf{h}|^2 dV - \frac{1}{2}\omega\int_V \epsilon_R|\mathbf{e}|^2 dV. \tag{2.55}$$

When comparing (2.54) with the corresponding free space result (2.47), three additional terms are found in the right-hand member. The first term corresponds to the Joule losses, also called Ohmic losses. The second and third term correspond to the time average electric and magnetic power losses in V. This part of the electromagnetic energy is transformed into other forms of energy (typically heat). Result (2.55) for the reactive power differs little from the corresponding result in free space.

Energy considerations have led to the introduction of the complex Poynting vector, the total flux of which through a closed surface is the total complex power flowing through that surface. However, from this one should *not* draw the conclusion that $\mathbf{p}(\mathbf{r})$ can locally be interpreted as the power flux through the point $\mathbf{r}$! The only valid conclusion is that $\nabla\cdot\mathbf{p}(\mathbf{r})$ can be interpreted as the total power flux emerging from the point $\mathbf{r}$. This follows from the application of (2.46) to an infinitesimally small volume.

To conclude this section, we urge the reader to reflect on the analogy between the conservation of energy results derived from Maxwell's equations and those we are familiar with from circuit theory. To this end consider the simple circuit depicted in Fig. 2.5 . For this circuit the energy balance in sinusoidal regime reads

$$\frac{1}{2}vi^* = \frac{1}{2}\frac{|v|^2}{R} + \frac{1}{2}j\omega(L|i|^2 - C|v|^2), \tag{2.56}$$

with i the current delivered by the current source and v the voltage over the parallel RLC-circuit.

2.4 Boundary conditions

Maxwell's equations in differential form (2.6)–(2.10) tacitly assume that the fields and their spatial derivatives are continuous. This is certainly not always

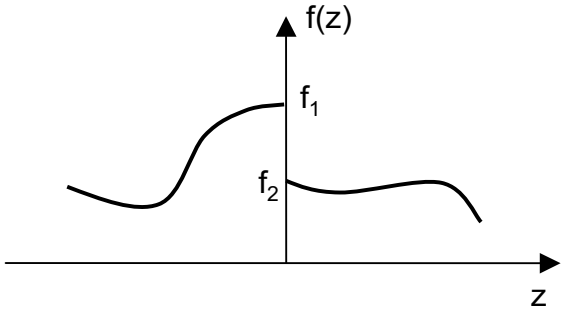


Figure 2.6: Discontinuous function.

the case. Fields can exhibit a step discontinuity or even become singular. By interpreting Maxwell's equations in the sense of the distributions, it is possible to examine which type of jump discontinuities occur. This then leads to the boundary conditions satisfied by the fields at the interface between two different media or materials or to the boundary conditions at an interface supporting surface currents or surface charges. It is also possible (and this is the preferred way in many textbooks) to derive these boundary conditions from the integral form of Maxwell's equations by carefully selecting the volumes or surfaces occurring in these equations. Here we prefer the elegant derivation based on distributions.

Consider a plane interface between two regions and let this interface coincide with the xy-plane. Suppose that $f(z)$ is a function exhibiting a jump discontinuity at $z = 0$ with $f(0+) = f_2$ and $f(0-) = f_1$ as depicted in Fig. 2.6. The derivative of $f(z)$ can be written as

$$\frac{\mathrm{d}}{\mathrm{d}z} f(z) = \{\frac{\mathrm{d}}{\mathrm{d}z} f(z)\} + (f_2 - f_1)\delta(z), \tag{2.57}$$

with $\{\frac{\mathrm{d}}{\mathrm{d}z} f(z)\}$ the regular part of the derivative and with the singular contribution expressed in terms of the delta function. Similarly, the divergence of a vector function $\mathbf{f}(\mathbf{r})$ exhibiting a jump discontinuity across the xy-plane can be written as the sum of a regular and a singular part

$$\nabla \cdot \mathbf{f}(\mathbf{r}) = \{\nabla \cdot \mathbf{f}(\mathbf{r})\} + \mathbf{u}_z \cdot (\mathbf{f}_2 - \mathbf{f}_1)\delta(z). \tag{2.58}$$

For the curl of a vector function the result is

$$\nabla \times \mathbf{f}(\mathbf{r}) = \{\nabla \times \mathbf{f}(\mathbf{r})\} + \mathbf{u}_z \times (\mathbf{f}_2 - \mathbf{f}_1)\delta(z). \tag{2.59}$$

In (2.58) and (2.59), $\mathbf{f}_1$ and $\mathbf{f}_2$ of course do not depend on z but still depend on x and y. It will be clear to the reader that the above results can easily be generalised to any plane interface by replacing $\mathbf{u}_z$ by the normal $\mathbf{u}_n$ to that plane, $\frac{\mathrm{d}}{\mathrm{d}z}$ by $\frac{\mathrm{d}}{\mathrm{d}n}$ and $\delta(z)$ by $\delta(n)$. The result can further be generalised to a curved interface as in Fig. 2.7 . The normal will then of course differ from point to point.

With the above knowledge, Maxwell's equations can now be written in the distributional sense as

$$\{\nabla \times \mathbf{e}(\mathbf{r})\} + \mathbf{u}_n \times (\mathbf{e}_2 - \mathbf{e}_1)\delta(n) = -j\omega\mathbf{b}(\mathbf{r}), \tag{2.60}$$

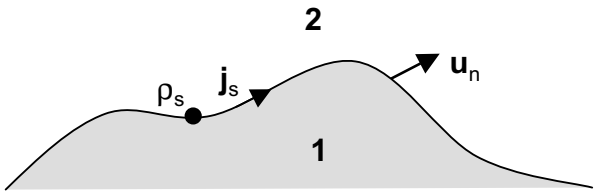


Figure 2.7: Interface between two media.

$$\{\nabla \times \mathbf{h}(\mathbf{r})\} + \mathbf{u}_n \times (\mathbf{h}_2 - \mathbf{h}_1)\delta(n) = j\omega\mathbf{d}(\mathbf{r}) + \{\mathbf{j}(\mathbf{r})\} + \mathbf{j}_s\delta(n), \quad (2.61)$$

$$\{\nabla \cdot \mathbf{d}(\mathbf{r})\} + \mathbf{u}_n \cdot (\mathbf{d}_2 - \mathbf{d}_1)\delta(n) = \{\rho(\mathbf{r})\} + \rho_s\delta(n), \quad (2.62)$$

$$\{\nabla \cdot \mathbf{b}(\mathbf{r})\} + \mathbf{u}_n \cdot (\mathbf{b}_2 - \mathbf{b}_1)\delta(n) = 0. \quad (2.63)$$

We have assumed that the current density and the charge density can contribute to the singular part of the equations. Physically this means that a surface current density $\mathbf{j}_s$ and/or a surface charge density ρ_s can be present on the considered interface.
The elegance of the distributional approach results from the fact that singular and regular contributions must now be equal at both sides of equations (2.60)–(2.63), leading to

$$\mathbf{u}_n \times (\mathbf{e}_2 - \mathbf{e}_1) = 0, \quad (2.64)$$

$$\mathbf{u}_n \times (\mathbf{h}_2 - \mathbf{h}_1) = \mathbf{j}_s, \quad (2.65)$$

$$\mathbf{u}_n \cdot (\mathbf{d}_2 - \mathbf{d}_1) = \rho_s, \quad (2.66)$$

$$\mathbf{u}_n \cdot (\mathbf{b}_2 - \mathbf{b}_1) = 0. \quad (2.67)$$

These equations state the jump conditions that must be satisfied by the fields at an interface. This interface can be the interface between two different materials or it can be a surface carrying a surface current density and/or a surface charge density.

We immediately conclude that the tangential component of the electric field and the normal component of the magnetic field are always continuous. The tangential component of the magnetic field exhibits a jump discontinuity determined by the possible presence of surface currents. If no surface currents are present, the tangential magnetic field also remains continuous. A possible jump discontinuity of the normal component of the dielectric displacement is determined by the presence of surface charges. This normal component remains continuous when no surface charges are present.

When also considering magnetic surface currents and charges, (2.64) and (2.67) are replaced by

$$\mathbf{u}_n \times (\mathbf{e}_2 - \mathbf{e}_1) = -\mathbf{k}_s, \quad (2.68)$$

$$\mathbf{u}_n \cdot (\mathbf{b}_2 - \mathbf{b}_1) = \pi_s, \quad (2.69)$$

With $\mathbf{j} = \{\mathbf{j}\} + \mathbf{j}_s\delta(n)$ and $\rho = \{\rho\} + \rho_s\delta(n)$, the law of charge conservation can be written as

$$\{\nabla \cdot \mathbf{j}\} + \mathbf{u}_n \cdot (\mathbf{j}_2 - \mathbf{j}_1)\delta(n) + \nabla_s \cdot \mathbf{j}_s\delta(n) = -j\omega\{\rho\} - j\omega\rho_s\delta(n), \quad (2.70)$$

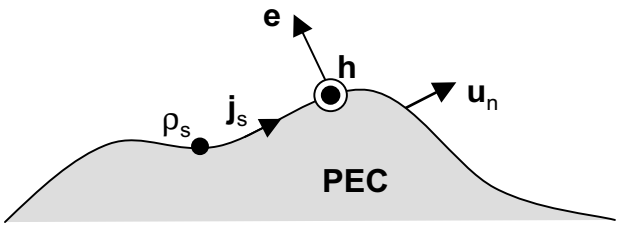


Figure 2.8: Boundary conditions at a PEC.

with ∇_s the surface divergence operator. From the above equation we easily derive the surface version of the charge conservation law

$$\mathbf{u}_n \cdot (\mathbf{j}_2 - \mathbf{j}_1) + \nabla_s \cdot \mathbf{j}_s = -j\omega\rho_s. \tag{2.71}$$

The first terms expresses the net current flowing onto the surface.

Now suppose that the interface coincides with the surface of a perfect electric conductor (PEC) and that $\mathbf{u}_n$ is the outward pointing unit normal vector to the PEC surface as depicted in Fig. 2.8 . From (2.64)–(2.67) and taking into account that the fields inside the PEC vanish, the following boundary conditions for the fields just outside the perfect conductor can be derived

$$\mathbf{u}_n \times \mathbf{e} = 0, \tag{2.72}$$
$$\mathbf{u}_n \times \mathbf{h} = \mathbf{j}_s, \tag{2.73}$$
$$\mathbf{u}_n \cdot \mathbf{d} = \rho_s, \tag{2.74}$$
$$\mathbf{u}_n \cdot \mathbf{b} = 0. \tag{2.75}$$

The tangential component of the electric field and the normal component of the magnetic induction vanish at a PEC surface. The tangential component of the magnetic field is equal to the surface current density rotated over 90°, while the normal component of the dielectric displacement is equal to the surface charge density. The surface current density can either be externally enforced or it can be induced. In the first case the surface current is the source of the magnetic field near the PEC surface while in the second case the surface current is induced by the magnetic field near the PEC surface.

2.5 Elementary dipole sources

In this course analysis is restricted to linear media. In that case the solution of Maxwell's equations for an arbitrary source can be derived (by superposition) from the solution for an "elementary" source.

To investigate the meaning of an elementary electric source current density, consider a small cylinder with radius a and length l, as depicted in Fig. 2.9. Suppose that the cylinder is placed in the origin of the (ρ, ϕ, z) cylindrical co-ordinate system with its axis along the z-axis. The current density in the cylinder is completely homogeneous and the total vertical (i.e. z-directed)

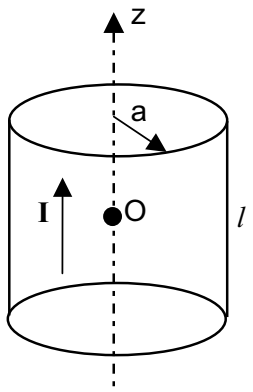


Figure 2.9: Small cylinder carrying a vertical current I.

current is I. In cylindrical co-ordinates (ρ, ϕ, z) the source current density is expressed as

$$\mathbf{j}(\mathbf{r}) = \frac{I}{\pi a^2}[u(z + \frac{l}{2}) - u(z - \frac{l}{2})]u(a - \rho)\mathbf{u}_z, \tag{2.76}$$

with $u(x)$ the Heaviside step function ($u(x) = 1$ for $x > 0$, $u(x) = 0$ for $x < 0$ and $u(x) = 1/2$ for $x = 0$). To simplify the notations we drop the subscript "e" on the source current and simply use $\mathbf{j}$. The expression for the elementary current density is obtained by taking the limit for $a \to 0$ and $l \to 0$ and for I becoming infinite but in such a way that Il remains finite. The result is

$$\mathbf{j}(\mathbf{r}) = \lim_{a,l\to 0;I\to\infty} \frac{I}{\pi a^2}[u(z + \frac{l}{2}) - u(z - \frac{l}{2})]u(a - \rho)\mathbf{u}_z = Il\delta(\mathbf{r})\mathbf{u}_z. \tag{2.77}$$

It is interesting to determine the corresponding charge density. The charge conservation law (2.8) dictates that

$$\begin{aligned} \rho(\mathbf{r}) &= -\frac{1}{j\omega}\nabla \cdot \mathbf{j}(\mathbf{r}) = -\frac{1}{j\omega} \lim_{a,l\to 0;I\to\infty} \frac{I}{\pi a^2}[\delta(z + \frac{l}{2}) - \delta(z - \frac{l}{2})]u(a - \rho), \\ &= \frac{1}{j\omega} \lim_{l\to 0;I\to\infty} I[\delta(\mathbf{r} - \frac{l}{2}\mathbf{u}_z) - \delta(\mathbf{r} + \frac{l}{2}\mathbf{u}_z)]. \end{aligned} \tag{2.78}$$

This result can be interpreted as two opposite charges $q = \pm I/(j\omega)$ separated by a distance l, i.e. an electric dipole with dipole moment $p_e = Il/(j\omega)$. Introducing the vectorial dipole moment $\mathbf{p}_e = p_e\mathbf{u}_z$ allows us to express the current density in terms of this dipole moment as $\mathbf{j}(\mathbf{r}) = j\omega\mathbf{p}_e\delta(\mathbf{r})$.

Hence, we conclude that the elementary electric source current density or in short, elementary electric current, is the time-derivative of a vectorial dipole moment. This elementary electric dipole is known as the electric Hertz dipole.

It is also possible to define an elementary magnetic source current density

$$\mathbf{k}(\mathbf{r}) = Kl\delta(\mathbf{r})\mathbf{u}_z. \tag{2.79}$$

This current density can again be interpreted as an infinitesimally small cylinder of length l carrying a magnetic current K. However, another interpretation is

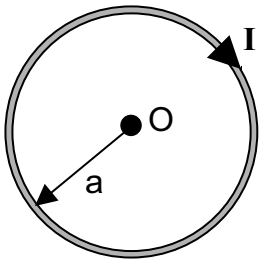


Figure 2.10: Elementary electric loop current

also possible. It can be shown (the proof is left out here) that the fields generated by $\mathbf{k}(\mathbf{r})$ are identical (or rather almost identical, i.e. except at the source point) to the fields generated by an elementary current $\mathbf{j}(\mathbf{r})$ given by

$$\mathbf{j}(\mathbf{r}) = \frac{1}{j\omega\mu}\nabla \times \mathbf{k}(\mathbf{r}). \tag{2.80}$$

The reader might wonder how to physically interpret (2.80). To this end inserting (2.79) into (2.80) yields

$$\begin{aligned} \mathbf{j}(\mathbf{r}) &= \frac{1}{j\omega\mu} \lim_{a,l\to 0;K\to\infty} \frac{K}{\pi a^2}[u(z+\frac{l}{2}) - u(z-\frac{l}{2})]\delta(\rho - a)\mathbf{u}_\phi \\ &= \frac{1}{j\omega\mu} \lim_{a\to 0} \frac{Kl}{\pi a^2}\delta(z)\delta(\rho - a)\mathbf{u}_\phi. \end{aligned} \tag{2.81}$$

This expression corresponds to the loop current density in a circular loop with radius a (see Fig. 2.10) carrying a total current

$$I = \frac{Kl}{j\omega\mu\pi a^2}. \tag{2.82}$$

A magnetic dipole moment is defined as $p_m = I\pi a^2$ or $p_m = Kl/(j\omega\mu)$. The corresponding vectorial moment $\mathbf{p}_m = p_m\mathbf{u}_z$ allows to write the elementary magnetic current density as $\mathbf{k}(\mathbf{r}) = j\omega\mu\mathbf{p}_m\delta(\mathbf{r})$. Remark that the direction of the magnetic dipole, i.e. $\mathbf{u}_z$, is the unit normal to the surface of the loop current (and with the orientation of the unit normal defined by the circulation sense of the current in the loop). The elementary magnetic current density or equivalently, the elementary electric loop current is known as the magnetic Hertz dipole.

It will be clear to the reader that it suffices to replace $\mathbf{u}_z$ by an arbitrary unit vector $\mathbf{u}$, to obtain the expressions for the electric and magnetic Hertz dipole oriented along $\mathbf{u}$.

2.6 Potentials and Green's functions

To solve an electromagnetic field problem it is obvious to start from Maxwell's equations. However, sometimes it is advantageous to use potentials. Several

types of potentials are known in electromagnetic literature, e.g. the scalar and vector Lorenz potentials, the vector Hertz potential, the scalar Hertz potentials and the Debye potentials. In this section the Lorenz potentials will be derived. They will play a prominent role in some of the forthcoming chapters.

Vector analysis dictates that a divergenceless vector can be derived from the curl of another vector called the vector potential. This theorem holds provided some very general conditions (not discussed here) are satisfied. As $\nabla \cdot \mathbf{b}(\mathbf{r}) = 0$, this can be invoked to derive $\mathbf{b}(\mathbf{r})$ from a vector potential $\mathbf{a}(\mathbf{r})$ in Wb/m as

$$\mathbf{b}(\mathbf{r}) = \nabla \times \mathbf{a}(\mathbf{r}). \tag{2.83}$$

Combined with $\nabla \times \mathbf{e}(\mathbf{r}) = -j\omega\mathbf{b}(\mathbf{r})$, this shows that

$$\nabla \times [\mathbf{e}(\mathbf{r}) + j\omega\mathbf{a}(\mathbf{r})] = 0. \tag{2.84}$$

Vector analysis can now again be invoked as a curl-free field can always be derived from a scalar potential. In our case $\mathbf{e}(\mathbf{r}) + j\omega\mathbf{a}(\mathbf{r})$ can be derived from a scalar potential $\phi(\mathbf{r})$ such that $\mathbf{e}(\mathbf{r})$ can be expressed as

$$\mathbf{e}(\mathbf{r}) = -j\omega\mathbf{a}(\mathbf{r}) - \nabla\phi(\mathbf{r}). \tag{2.85}$$

Suppose now that the medium that we are working with is homogeneous and isotropic and characterised by ϵ and μ. For such a medium $\mathbf{h}(\mathbf{r})$ and $\mathbf{d}(\mathbf{r})$ can be expressed in terms of the potentials $\mathbf{a}$ and ϕ as

$$\mathbf{h}(\mathbf{r}) = \frac{1}{\mu}\nabla \times \mathbf{a}(\mathbf{r}), \tag{2.86}$$

$$\mathbf{d}(\mathbf{r}) = -j\omega\epsilon\mathbf{a}(\mathbf{r}) - \epsilon\nabla\phi(\mathbf{r}). \tag{2.87}$$

Combination of the divergence equation (2.9) and (2.87) yields

$$\nabla^2\phi(\mathbf{r}) + j\omega\nabla \cdot \mathbf{a}(\mathbf{r}) = -\frac{\rho_e(\mathbf{r})}{\epsilon}, \tag{2.88}$$

while the combination of the curl equation (2.7) and (2.86) shows that

$$\nabla^2\mathbf{a}(\mathbf{r}) + k^2\mathbf{a}(\mathbf{r}) = \nabla[\nabla \cdot \mathbf{a}(\mathbf{r}) + j\omega\epsilon\mu\phi(\mathbf{r})] - \mu\mathbf{j}_e(\mathbf{r}). \tag{2.89}$$

We have introduced the wave number k in $1/m$, defined as $k = \omega\sqrt{\epsilon\mu}$.

The potentials that we have introduced above are not unique. Indeed, consider a new set of potentials given by

$$\hat{\mathbf{a}}(\mathbf{r}) = \mathbf{a}(\mathbf{r}) - \nabla\theta(\mathbf{r}), \tag{2.90}$$

$$\hat{\phi}(\mathbf{r}) = \phi(\mathbf{r}) + j\omega\theta(\mathbf{r}), \tag{2.91}$$

with $\theta(\mathbf{r})$ an arbitrary scalar function. The reader can easily verify that the fields derived from this new set of potentials remain unchanged (remember that $\nabla \times \nabla f(\mathbf{r})$ is always zero). This degree of freedom can be used to enforce an additional condition on $\mathbf{a}(\mathbf{r})$ and $\phi(\mathbf{r})$. This condition is known as the gauge condition. In this course the familiar Lorenz gauge is used

$$\nabla \cdot \mathbf{a}(\mathbf{r}) + j\omega\epsilon\mu\phi(\mathbf{r}) = 0. \tag{2.92}$$

It derives its name from Ludwig Lorenz who was the first to introduce this particular gauge. With this Lorenz gauge, (2.88) and (2.89) simplify considerably, yielding

$$\nabla^2\mathbf{a}(\mathbf{r}) + k^2\mathbf{a}(\mathbf{r}) = -\mu\mathbf{j}_e(\mathbf{r}), \tag{2.93}$$

$$\nabla^2\phi(\mathbf{r}) + k^2\phi(\mathbf{r}) = -\frac{\rho_e(\mathbf{r})}{\epsilon}, \tag{2.94}$$

with $\mathbf{a}(\mathbf{r})$ and $\phi(\mathbf{r})$ the Lorenz potentials.

A differential equation of the form

$$\nabla^2 f(\mathbf{r}) + k^2 f(\mathbf{r}) = source(\mathbf{r}) \tag{2.95}$$

is called a scalar Helmholtz equation. If $f(\mathbf{r})$ and the source are replaced by their vector counterparts, the differential equation becomes a vector Helmholtz equation. We conclude that $\mathbf{a}(\mathbf{r})$ satisfies a vector Helmholtz equation with $-\mu\mathbf{j}_e(\mathbf{r})$ as its source, while $\phi(\mathbf{r})$ satisfies a scalar Helmholtz equation with $-\frac{\rho_e(\mathbf{r})}{\epsilon}$ as its source.

A general analytical solution to (2.93) and (2.94) is not available. We consider the very important special case of a homogeneous and isotropic medium of infinite extent. To this end first consider the following scalar Helmholtz equation for the function $G(\mathbf{r})$

$$\nabla^2 G(\mathbf{r}) + k^2 G(\mathbf{r}) = -\delta(\mathbf{r}). \tag{2.96}$$

The source term is a point source at $\mathbf{r}$, mathematically represented by minus the delta function $\delta(\mathbf{r})$. The function $G(\mathbf{r})$, i.e. the solution for a point source, is known as the Green's function of the problem. This Green's function can be determined analytically. In spherical co-ordinates, due to the symmetry of the problem, $G(\mathbf{r})$ only depends on the radial co-ordinate r, i.e. $G(\mathbf{r}) = G(r)$. Hence, in a spherical co-ordinate system (2.96) becomes

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}r^2\frac{\mathrm{d}}{\mathrm{d}r}G(r) + k^2 G(r) = 0, \tag{2.97}$$

for $r \neq 0$. It is easily verified that the general solution of the above equation is

$$G(r) = A\frac{e^{-jkr}}{r} + B\frac{e^{jkr}}{r}, \tag{2.98}$$

with A and B as yet unknown but constant coefficients. When defining $k = \omega\sqrt{\epsilon\mu}$ nothing was said about the sign of the square root. By convention and in line with the electromagnetic literature, we choose $\Re k \geq 0$ and $\Im k \leq 0$. For a lossless medium k is a real number. The time domain signal corresponding to (2.98) is

$$G(r,t) = \Re[A\frac{e^{-jkr+j\omega t}}{r}] + \Re[B\frac{e^{jkr+j\omega t}}{r}]. \tag{2.99}$$

The first term represents a spherical wave propagating from the origin ($r = 0$) to infinity. In a lossy medium and in addition to the typical $1/r$-dependence for spherical waves, the amplitude of that wave decreases exponentially as $\Im k \leq 0$. The second term also represents a spherical wave, this time coming from infinity

and propagating inwards towards the origin. As the point source is located in the origin, the wave coming from infinity is non-physical and hence we are forced to make the choice $B = 0$. Remark that the speed of the spherical waves is given by $v = \omega/k = 1/\sqrt{\epsilon\mu}$ which for free space corresponds to (2.23), proving that Maxwell leads to wave solutions with waves propagating at the speed of light. Coefficient A is determined by integrating (2.96) over a sphere V_α with radius α

$$\int_{V_\alpha} \nabla^2 G(r) dV + k^2 \int_{V_\alpha} G(r) dV = -1. \tag{2.100}$$

The first term can be transformed to a surface integral using Gauss's divergence theorem, leading to

$$\int_{S_\alpha} \nabla G(r) \cdot \mathbf{u}_r dS + k^2 \int_{V_\alpha} G(r) dV = -1, \tag{2.101}$$

with S_α the surface of V_α. Substituting the value of G and integration yields

$$4\pi A[-(1 + jk\alpha)e^{-jk\alpha}] + 4\pi A[(1 + jk\alpha)e^{-jk\alpha} - 1] = -1, \tag{2.102}$$

or $A = 1/(4\pi)$. The final result for the Green's function is

$$G(r) = \frac{e^{-jkr}}{4\pi r}. \tag{2.103}$$

The importance of the Green's function solution now shows from the fact that the solution of (2.93) and (2.94) can immediately be obtained by superposition. To see this, it e.g. suffices to rewrite the source term $\frac{\rho_e(\mathbf{r})}{\epsilon}$ of (2.94) as

$$\frac{\rho_e(\mathbf{r})}{\epsilon} = \int_V \delta(\mathbf{r} - \mathbf{r}') \frac{\rho_e(\mathbf{r}')}{\epsilon} dV'. \tag{2.104}$$

The integration in the left hand member is with respect to $\mathbf{r}'$. A similar expression can of course be written down for $\mu\mathbf{j}_e(\mathbf{r})$. The final expressions for $\mathbf{a}(\mathbf{r})$ and $\phi(\mathbf{r})$ as functions of the sources $\rho_e(\mathbf{r})$ and $\mathbf{j}_e(\mathbf{r})$ are

$$\mathbf{a}(\mathbf{r}) = \mu \int_V G(|\mathbf{r} - \mathbf{r}'|)\mathbf{j}_e(\mathbf{r}')dV' = \frac{\mu}{4\pi} \int_V \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \mathbf{j}_e(\mathbf{r}')dV', \tag{2.105}$$

$$\phi(\mathbf{r}) = \frac{1}{\epsilon} \int_V G(|\mathbf{r} - \mathbf{r}'|)\rho_e(\mathbf{r}')dV' = \frac{1}{4\pi\epsilon} \int_V \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \rho_e(\mathbf{r}')dV'. \tag{2.106}$$

Once the potentials are known, application of (2.83) and (2.85) immediately leads to the fields. The sources in (2.105) and (2.106) are not independent as the law of charge conservation dictates that $\nabla \cdot \mathbf{j}_e(\mathbf{r}) + j\omega\rho_e = 0$. This leads to the expectation that $\mathbf{a}(\mathbf{r})$ and $\phi(\mathbf{r})$ and hence $\mathbf{e}(\mathbf{r})$ and $\mathbf{b}(\mathbf{r})$ can be derived from the sole knowledge of $\mathbf{j}_e(\mathbf{r})$. We will not go into detail at this point and we only present the final result for the electric field, i.e.

$$\mathbf{e}(\mathbf{r}) = \int_V \overline{\overline{G}}(\mathbf{r} - \mathbf{r}') \cdot \mathbf{j}_e(\mathbf{r}')dV', \tag{2.107}$$

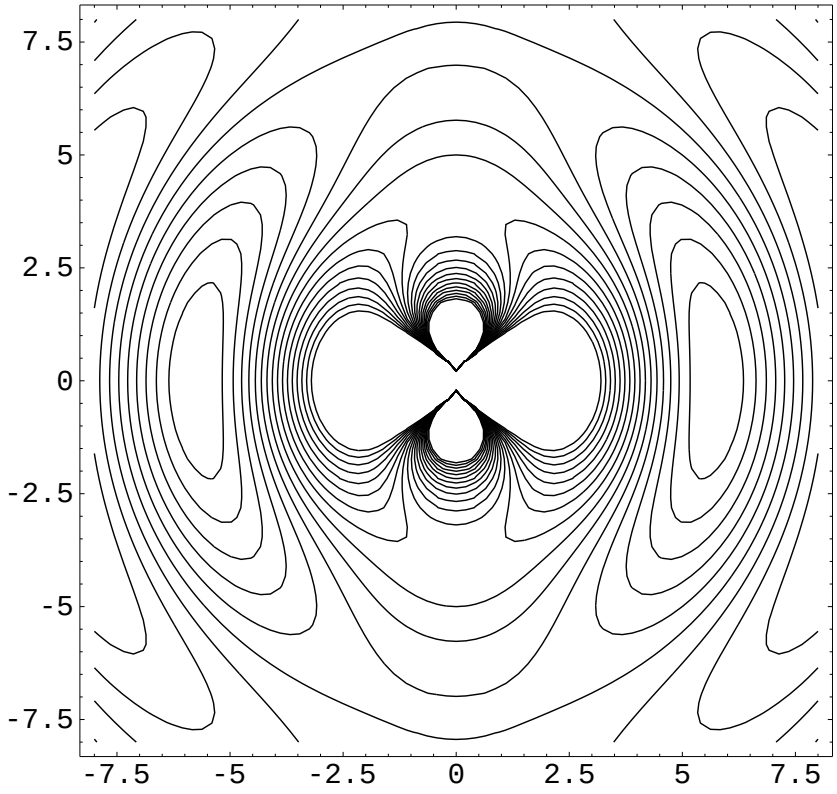


Figure 2.11: Magnitude of the electric field in the xy-plane of an elementary x-oriented electric dipole located at the origin.

with $\overline{\overline{G}}(\mathbf{r}-\mathbf{r}')$ the Green's tensor. In cartesian co-ordinates this tensor is given by

$$\overline{\overline{G}}(\mathbf{r}) = -j\omega\mu \begin{pmatrix} 1+\frac{1}{k^2}\frac{\partial^2}{\partial x^2} & \frac{1}{k^2}\frac{\partial^2}{\partial x \partial y} & \frac{1}{k^2}\frac{\partial^2}{\partial x \partial z} \\ \frac{1}{k^2}\frac{\partial^2}{\partial x \partial y} & 1+\frac{1}{k^2}\frac{\partial^2}{\partial y^2} & \frac{1}{k^2}\frac{\partial^2}{\partial y \partial z} \\ \frac{1}{k^2}\frac{\partial^2}{\partial x \partial z} & \frac{1}{k^2}\frac{\partial^2}{\partial y \partial z} & 1+\frac{1}{k^2}\frac{\partial^2}{\partial z^2} \end{pmatrix} G(r). \qquad (2.108)$$

From (2.108) the electric field of an elementary electrical dipole $\mathbf{j}_e(\mathbf{r}) = Il\delta(\mathbf{r})\mathbf{u}$ is given by $\mathbf{e}(\mathbf{r}) = Il\overline{\overline{G}}(\mathbf{r}) \cdot \mathbf{u}$ with $r = |\mathbf{r}|$. Fig. 2.11 shows the amplitude in the xy-plane of the electric field of an elementary electric dipole $\mathbf{j}_e(\mathbf{r}) = Il\delta(\mathbf{r})\mathbf{u}_x$. The wavenumber k was taken to be one.

Extending the above results to piecewise homogeneous media is not straightforward as $\mathbf{a}$ and ϕ couple due to the boundary conditions at the interfaces between materials. No general analytical solution is available and numerical methods must be invoked.

2.7 Wave equations

Consider a homogeneous and isotropic medium characterised by ϵ and μ and with sources $\mathbf{j}_e(\mathbf{r})$. The fields in this medium satisfy

$$\nabla \times \mathbf{e}(\mathbf{r}) = -j\omega\mu\mathbf{h}(\mathbf{r}), \tag{2.109}$$

$$\nabla \times \mathbf{h}(\mathbf{r}) = j\omega\epsilon\mathbf{e}(\mathbf{r}) + \mathbf{j}_e(\mathbf{r}), \tag{2.110}$$

$$\nabla \cdot \mathbf{e}(\mathbf{r}) = \frac{\rho_e}{\epsilon} = -\frac{\nabla \cdot \mathbf{j}_e(\mathbf{r})}{j\omega\epsilon}, \tag{2.111}$$

$$\nabla \cdot \mathbf{h}(\mathbf{r}) = 0. \tag{2.112}$$

Taking the curl of the first equation and substituting the result in the second equation gives

$$\nabla \times (\nabla \times \mathbf{e}(\mathbf{r})) = -j\omega\mu[j\omega\epsilon\mathbf{e}(\mathbf{r}) + \mathbf{j}_e(\mathbf{r})]. \tag{2.113}$$

Using $\nabla \times (\nabla \times \mathbf{e}(\mathbf{r})) = \nabla(\nabla \cdot \mathbf{e}(\mathbf{r})) - \nabla^2\mathbf{e}(\mathbf{r})$ and (2.111) immediately leads to

$$\nabla^2\mathbf{e}(\mathbf{r}) + k^2\mathbf{e}(\mathbf{r}) = j\omega\mu[\mathbf{j}_e(\mathbf{r}) + \frac{1}{k^2}\nabla(\nabla \cdot \mathbf{j}_e(\mathbf{r}))]. \tag{2.114}$$

This shows that not only the Lorenz potentials but also the electric field satisfies a Helmholtz equation also called wave equation for the electric field. The source term is much more complex as compared to the source term of the Helmholtz equation for the vector potential. It can be shown that the solution of (2.114) is identical to (2.107), obtained through the Lorenz potentials. Another way to obtain a wave equation is to take the curl of the second equation (2.110) and to eliminate $\mathbf{e}(\mathbf{r})$ by using the first one (2.109). By finally invoking $\nabla \cdot \mathbf{h}(\mathbf{r}) = 0$, the following Helmholtz equation or wave equation for the magnetic field is found

$$\nabla^2\mathbf{h}(\mathbf{r}) + k^2\mathbf{h}(\mathbf{r}) = -\nabla \times \mathbf{j}_e(\mathbf{r}). \tag{2.115}$$

Equations (2.114) and (2.115) are of particular importance in source-free regions as will become clear in the sequel. From these equations the reader might incorrectly deduce that $\mathbf{e}(\mathbf{r})$ and $\mathbf{h}(\mathbf{r})$ are independent, in particular in source-free regions. This is *not* the case as $\mathbf{e}(\mathbf{r})$ and $\mathbf{h}(\mathbf{r})$ remain linked through (2.109) or (2.110). Finally, remark that (2.114) and (2.115) are very useful in Cartesian co-ordinates as each of the Cartesian components of $\mathbf{e}(\mathbf{r})$ and $\mathbf{h}(\mathbf{r})$ satisfy a separate scalar Helmholtz equation. This is however not the case in other co-ordinate systems! In cylindrical or spherical co-ordinates e.g. e_ϕ does not satisfy a scalar Helmholtz equation.

2.8 Image theory and image sources

As in the previous section, a homogeneous and isotropic medium is considered but now in the presence of electric source current densities with the following symmetry properties

$$\mathbf{j}_t(\mathbf{r}) = -\mathbf{j}_t(\mathbf{r}'), \qquad j_z(\mathbf{r}) = j_z(\mathbf{r}'), \tag{2.116}$$

with $\mathbf{r} = x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z$ and $\mathbf{r}' = x\mathbf{u}_x + y\mathbf{u}_y - z\mathbf{u}_z$. This means that $\mathbf{r}'$ is the image of $\mathbf{r}$ with respect to the xy-plane. To simplify the notations, we have

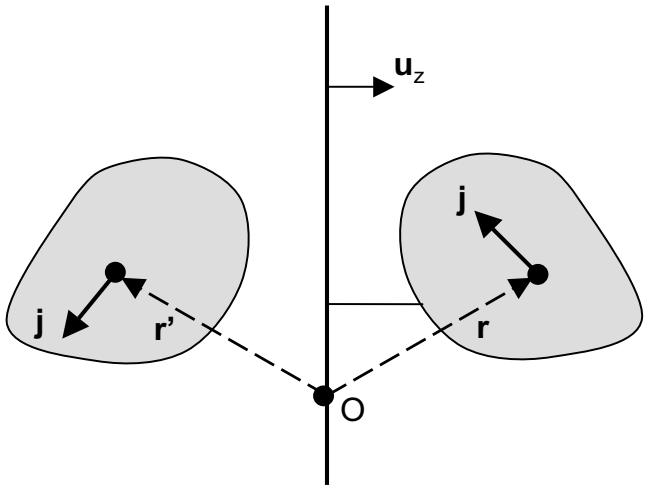


Figure 2.12: Current density with symmetric longitudinal component and anti-symmetrical transversal component with respect to the xy-plane.

dropped the subscript "e" of the source terms. The component of $\mathbf{j}(\mathbf{r})$ in the xy-plane is given by $\mathbf{j}_t(\mathbf{r})$ and is the so-called transversal component while $j_z(\mathbf{r})$ is the longitudinal component. The above symmetry properties state that the transversal component of the current is anti-symmetric with respect to the xy-plane while the longitudinal component is symmetric, as depicted in Fig. 2.12.
It can now be proved that the transversal electric field $\mathbf{e}_t(x, y, z = 0)$ and the normal magnetic field $h_z(x, y, z = 0)$ in the symmetry plane, generated by the sources (2.116), are zero. The conditions $\mathbf{e}_t(x, y, z = 0) = 0$ and $h_z(x, y, z = 0) = 0$, i.e. zero tangential electric field and zero normal magnetic field component, are the boundary conditions at a perfect electric conductor. This means that the symmetry plane can be replaced by a PEC without influencing the fields. The reader who is interested in the proof of the above assertion can read through the following paragraph.
The proof starts by rewriting Maxwell's curl equations in terms of the transversal and longitudinal field components. The result is

$$\frac{\partial}{\partial z}\mathbf{u}_z \times \mathbf{e}_t(\mathbf{r}) + \nabla_t \times e_z(\mathbf{r})\mathbf{u}_z = -j\omega\mu\mathbf{h}_t(\mathbf{r}), \tag{2.117}$$

$$\nabla_t \times \mathbf{e}_t(\mathbf{r}) = -j\omega\mu h_z(\mathbf{r})\mathbf{u}_z, \tag{2.118}$$

$$\frac{\partial}{\partial z}\mathbf{u}_z \times \mathbf{h}_t(\mathbf{r}) + \nabla_t \times h_z(\mathbf{r})\mathbf{u}_z = j\omega\epsilon\mathbf{e}_t(\mathbf{r}) + \mathbf{j}_t(\mathbf{r}), \tag{2.119}$$

$$\nabla_t \times \mathbf{h}_t(\mathbf{r}) = j\omega\epsilon e_z(\mathbf{r})\mathbf{u}_z + j_z(\mathbf{r})\mathbf{u}_z, \tag{2.120}$$

with $\nabla_t = \frac{\partial}{\partial x}\mathbf{u}_x + \frac{\partial}{\partial y}\mathbf{u}_y$. Application of the co-ordinate transformation $\mathbf{r} \to \mathbf{r}'$ and taking into account (2.116) yields

$$-\frac{\partial}{\partial z}\mathbf{u}_z \times \mathbf{e}_t(\mathbf{r}') + \nabla_t \times e_z(\mathbf{r}')\mathbf{u}_z = -j\omega\mu\mathbf{h}_t(\mathbf{r}'), \tag{2.121}$$

$$\nabla_t \times \mathbf{e}_t(\mathbf{r}') = -j\omega\mu h_z(\mathbf{r}')\mathbf{u}_z, \tag{2.122}$$

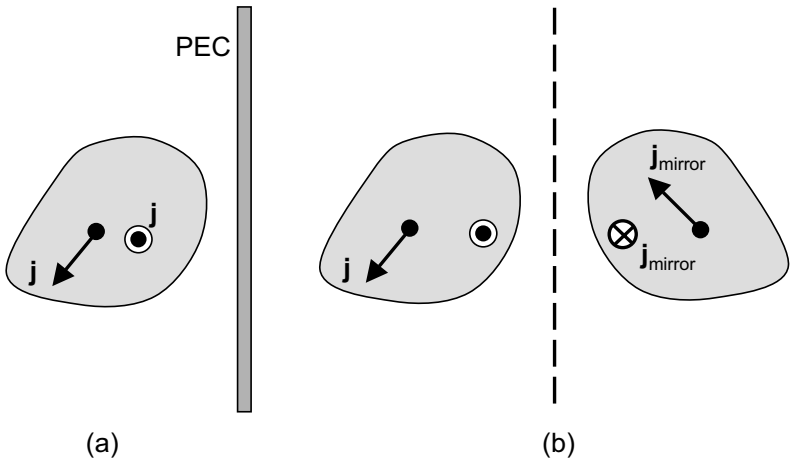


Figure 2.13: Current source in front of a PEC plane (a) and its equivalent based on image theory (b).

$$-\frac{\partial}{\partial z}\mathbf{u}_z \times \mathbf{h}_t(\mathbf{r}') + \nabla_t \times h_z(\mathbf{r}')\mathbf{u}_z = j\omega\epsilon\mathbf{e}_t(\mathbf{r}') - \mathbf{j}_t(\mathbf{r}), \tag{2.123}$$

$$\nabla_t \times \mathbf{h}_t(\mathbf{r}') = j\omega\epsilon e_z(\mathbf{r}')\mathbf{u}_z + j_z(\mathbf{r})\mathbf{u}_z. \tag{2.124}$$

Comparing (2.117)–(2.120) to (2.121)–(2.124) shows that these equations will be identical (as they should due to the symmetry of the situation) provided

$$\mathbf{e}_t(\mathbf{r}') = -\mathbf{e}_t(\mathbf{r}), \qquad \mathbf{h}_t(\mathbf{r}') = \mathbf{h}_t(\mathbf{r}), \tag{2.125}$$

$$e_z(\mathbf{r}') = e_z(\mathbf{r}), \qquad h_z(\mathbf{r}') = -h_z(\mathbf{r}). \tag{2.126}$$

This immediately shows that for $z = 0$, $\mathbf{e}_t = 0$ and $h_z = 0$.
In practice, the result obtained above will be used to solve field problems involving an infinite PEC plane as in Fig. 2.13a. To solve such a field problem, it suffices to remove the PEC plane while introducing image sources satisfying (2.116) as depicted in Fig. 2.13b.
The rules for imaging magnetic current densities are

$$\mathbf{k}_t(\mathbf{r}) = \mathbf{k}_t(\mathbf{r}'), \qquad k_z(\mathbf{r}) = -k_z(\mathbf{r}'). \tag{2.127}$$

Suppose now that the medium in front of the PEC-plane is *inhomogeneous*. In that case it suffices to also mirror the medium in the PEC-plane such that $\epsilon(\mathbf{r}') = \epsilon(\mathbf{r})$ and $\mu(\mathbf{r}') = \mu(\mathbf{r})$. This also implies that perfectly conducting structures in front of a PEC-mirror must also be replaced by their image. If the considered media are anisotropic the situation becomes more complex and will not be treated here.

To conclude this section, we remind the reader of the image theory for charges. For charges in front of a PEC-plane, the rule is quite simple: the PEC-plane must be replaced by mirror charges which take the opposite sign, as illustrated in Fig. 2.14. This immediately follows from the charge conservation law, written as

$$\nabla_t \cdot \mathbf{j}_t(\mathbf{r}) + \frac{\partial}{\partial z} j_z(\mathbf{r}) + j\omega\rho(\mathbf{r}) = 0. \tag{2.128}$$

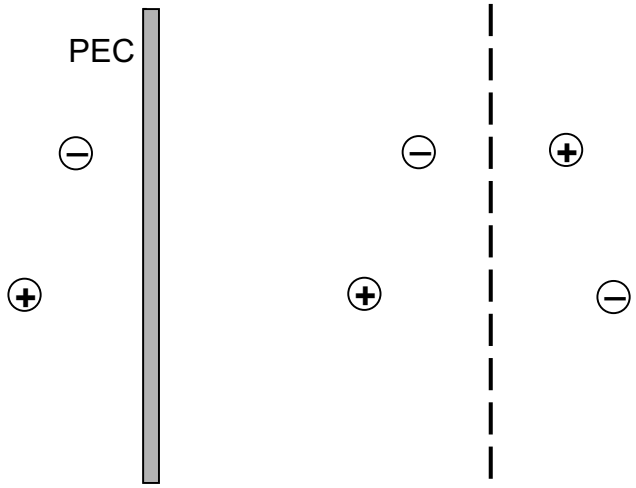


Figure 2.14: Image theory for charges.

Application of the co-ordinate transformation $\mathbf{r} \rightarrow \mathbf{r}'$ and taking into account (2.116) transform (2.128) into

$$-\nabla_t \cdot \mathbf{j}_t(\mathbf{r}) - \frac{\partial}{\partial z} j_z(\mathbf{r}) + j\omega\rho(\mathbf{r}') = 0. \tag{2.129}$$

Comparing this with (2.128) shows that $\rho(\mathbf{r}') = -\rho(\mathbf{r})$.

Chapter 3

Electrostatics

3.1 Maxwell's equations in the static case

In this and the next chapter, we turn our attention to static phenomena. By definition, static phenomena do not depend on time and hence this will also be the case for all quantities in Maxwell's equations. Putting the derivatives with respect to time equal to zero in (2.1)–(2.4) reduces Maxwell's equations to

$$\nabla \times \mathbf{e}(\mathbf{r}) = 0, \tag{3.1}$$
$$\nabla \cdot \mathbf{d}(\mathbf{r}) = \rho(\mathbf{r}) \tag{3.2}$$

and

$$\nabla \times \mathbf{h}(\mathbf{r}) = \mathbf{j}(\mathbf{r}), \tag{3.3}$$
$$\nabla \cdot \mathbf{b}(\mathbf{r}) = 0. \tag{3.4}$$

The first set of equations are the equations of Electrostatics, the second set those of Magnetostatics. If we only consider an externally enforced current density $\mathbf{j}_e(\mathbf{r})$, the electric and magnetic field are no longer coupled in the static case. In the presence of conductors however, the coupling between the two set of equations is still present and is solely due to the conduction current $\mathbf{j}(\mathbf{r}) = \sigma\mathbf{e}(\mathbf{r})$. In the rest of this chapter, as in the next one, we will simply note the current, either source current or conduction current, as $\mathbf{j}$. The difference between the two will be clear from the context.

The rest of this chapter is devoted to Electrostatics. In last year's physics course, the student was already introduced to this topic. We recapitulate the basic notions such as Coulomb force, electric field, potential, resistance and capacitance and discuss the difference between a dielectric and a conductor. Furthermore, Laplace's equation is introduced together with the appropriate boundary conditions to determine the resistance of an arbitrary conductor or the capacitance matrix of a set of conductors. The chapter concludes with a brief introduction to the numerical solution of Laplace's equation.

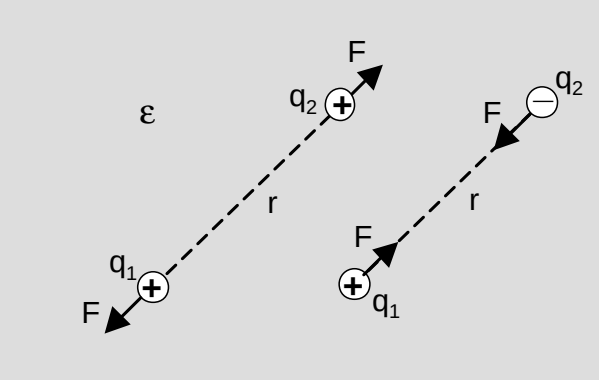


Figure 3.1: Coulomb forces between two positive point charges (a), between a positive and a negative point charge (b).

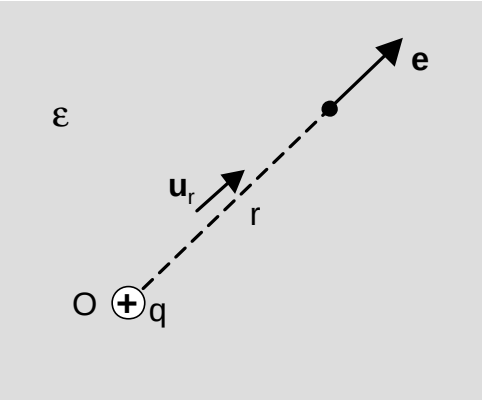


Figure 3.2: Electric field of a positive point charge.

3.2 Coulomb force and the electric field

Electrostatics often takes Coulomb's force law as its starting point. Consider two point charges q_1 and q_2 placed in a homogeneous dielectric with dielectric constant ϵ as depicted in Fig. 3.1 . The magnitude of the forces between these charges is

$$F = \frac{1}{4\pi\epsilon}\frac{q_1 q_2}{r^2}, \tag{3.5}$$

with r the distance between the charges. The direction of the forces depends on the sign of the charges: repulsive forces for charges of equal sign, attractive forces for charges of opposite sign. The electric field $\mathbf{e}(\mathbf{r})$ derives from Coulomb's force law by defining the electric field of a charge q in a point $\mathbf{r}$ as the force exerted by that charge on a unit positive charge placed at $\mathbf{r}$ (see Fig. 3.2). For a point charge q placed at the origin, the electric field in V/m is

$$\mathbf{e}(\mathbf{r}) = \mathbf{e}(r) = \frac{1}{4\pi\epsilon}\frac{q}{r^2}\mathbf{u}_r, \tag{3.6}$$

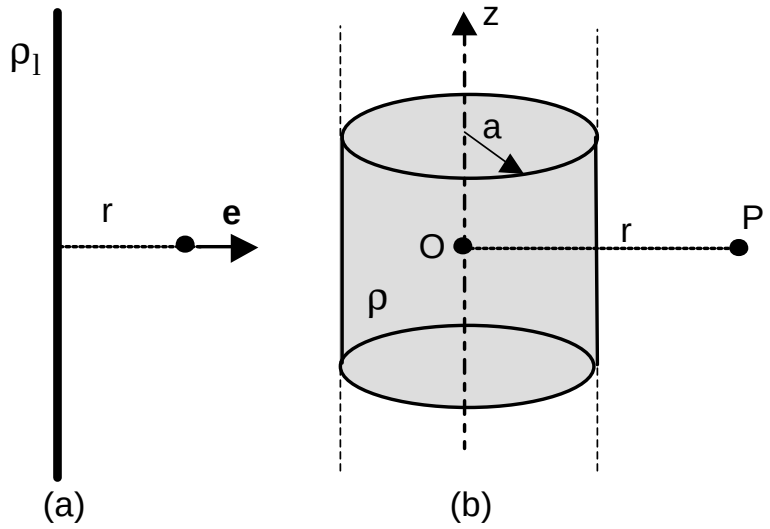


Figure 3.3: Electric field of a uniform line charge: original line charge (a); cylindrical volume charge (b).

with r the distance to the origin and with $\mathbf{u}_r$ the unit vector along the radial direction. This field exhibits spherical symmetry. By superposition, the above result is easily generalised to a volume charge density $\rho(\mathbf{r})$ in C/m^3 as

$$\begin{aligned} \mathbf{e}(\mathbf{r}) &= \frac{1}{4\pi\epsilon}\int_V \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^2}\mathbf{u}\, dV' \\ &= \frac{1}{4\pi\epsilon}\int_V \frac{\rho(\mathbf{r}')(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}\, dV', \end{aligned} \tag{3.7}$$

with $\mathbf{u} = \frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}$ the unit vector in the direction joining $\mathbf{r}'$ to $\mathbf{r}$. The integration in (3.7) extends over the source volume V. The special case for a surface charge distribution $\rho_s(\mathbf{r})$ in C/m^2 residing on a surface S will be of particular importance when considering conductors. In that case $\rho(\mathbf{r}')$ in (3.7) is replaced by $\rho_s(\mathbf{r}')$ with the integration extending over the surface S. Another special case is a line charge density $\rho_l(\mathbf{r}')$ in C/m with the charge distribution concentrated on a line segment.

By way of example, consider an infinitely long and uniform line charge density ρ_l as shown in Fig. 3.3a . To determine the electric field of such a line charge, first consider a uniform volume charge density ρ in an infinitely long cylinder of radius a (Fig. 3.3b). Due to the symmetry of this configuration, the electric field will only depend on the distance r to the axis. Hence, let us calculate the field in a point P with co-ordinates $x = r$, $y = 0$ and $z = 0$. With each source point in the cylinder described in terms of its polar co-ordinates

(r', ϕ', z'), (3.7) becomes

$$\mathbf{e}(x = r) = \frac{1}{4\pi\epsilon}\rho \int_{-\infty}^{\infty}\int_{0}^{2\pi}\int_{0}^{a} \frac{(r - r'\cos\phi')\mathbf{u}_x - r'\sin\phi'\mathbf{u}_y - z'\mathbf{u}_z}{[(r - r'\cos\phi')^2 + (r'\sin\phi')^2 + z'^2]^{3/2}} r' dr' d\phi' dz'. \tag{3.8}$$

The terms in $\mathbf{u}_z$ and $\mathbf{u}_y$ immediately drop out (why?). The solution for the line charge is obtained when the radius of the cylinder becomes infinitely small. In that case (3.8) can be approximated by

$$\begin{aligned} \mathbf{e}(x = r) &= \frac{1}{4\pi\epsilon}\rho \int_{-\infty}^{\infty}\int_{0}^{2\pi}\int_{0}^{a} \frac{r\mathbf{u}_x}{(r^2 + z'^2)^{3/2}} r' dr' d\phi' dz' \qquad (3.9) \\ &= \frac{1}{4\pi\epsilon}\rho r \frac{a^2}{2}(2\pi)\left[\frac{|z'|}{r^2\sqrt{r^2 + z'^2}}\right]_{-\infty}^{\infty}, \end{aligned}$$

from which we conclude that the electric field of the line charge ρ_l is

$$\mathbf{e} = \frac{1}{2\pi\epsilon}\frac{\rho_l}{r}\mathbf{u}_r. \tag{3.10}$$

To obtain this result, the following relationship between the constant charge density ρ_l of the line charge and the constant charge density ρ of the cylinder

$$\rho_l = \lim_{a\to 0} \pi a^2 \rho, \tag{3.11}$$

was invoked, i.e. the volume charge density becomes infinite but in such a way that its integral over the cross section of the cylinder remains finite.

As in the general dynamic case, (3.2) is the differential form of Gauss's law (2.17)

$$\int_S \mathbf{d}(\mathbf{r}) \cdot \mathbf{u}_n dS = \int_V \rho(\mathbf{r}) dV = q_V. \tag{3.12}$$

The flux of the electric induction or dielectric displacement through any closed surface equals the total charge q_V inside the volume enclosed by that surface. The above calculation of the electric field of a line charge can be much simplified by applying Gauss's law to the cylindrical volume with radius r and length L shown in Fig. 3.4 . As for symmetry reasons the field must be radial and z-independent, (3.12) immediately shows that $2\pi r L \epsilon e_r = \rho_l L$, confirming (3.10).

3.3 The electric potential

When the curl of a vector field is zero, this field is said to be conservative or irrotational and can be expressed as the gradient of a scalar potential. As in the static case, $\nabla \times \mathbf{e}(\mathbf{r}) = 0$, we introduce the electric potential $\phi(\mathbf{r})$ with

$$\mathbf{e}(\mathbf{r}) = -\nabla\phi(\mathbf{r}). \tag{3.13}$$

Remember that in the previous chapter the electromagnetic fields were already expressed in terms of the Lorenz potentials $\mathbf{a}(\mathbf{r})$ and $\phi(\mathbf{r})$. For $\omega = 0$ (2.85)

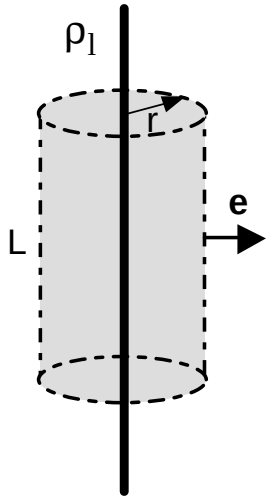


Figure 3.4: Electric field of a uniform line charge and Gauss's law.

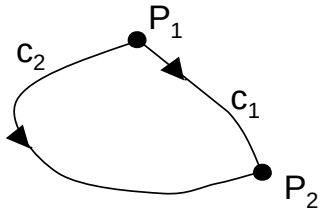


Figure 3.5: Two integration paths between P_1 and P_2.

indeed reduces to (3.13).
The potential difference ϕ_{12} between two points P_1 and P_2 is given by

$$\phi_{12} = \phi_1 - \phi_2 = -\int_{P_1}^{P_2} \nabla\phi \cdot \mathbf{dc} = \int_{P_1}^{P_2} \mathbf{e} \cdot \mathbf{dc}. \tag{3.14}$$

It is not necessary to specify the integration path between P_1 and P_2. This is a consequence of the irrotational character of the electric field. Indeed, consider two different integration paths c_1 and c_2 between P_1 and P_2 as depicted in Fig. 3.5 . With a self-explanatory notation, we have that

$$[\phi_1 - \phi_2]_{c_1} = \int_{c_1} \mathbf{e} \cdot \mathbf{dc},$$

$$[\phi_1 - \phi_2]_{c_2} = \int_{c_2} \mathbf{e} \cdot \mathbf{dc}.$$

As the circulation of $\mathbf{e}$ is zero, i.e. $\oint_c \mathbf{e}(\mathbf{r}) \cdot \mathbf{dc} = 0$, we immediately conclude that

$$\int_{c_1} \mathbf{e}(\mathbf{r}) \cdot \mathbf{dc} - \int_{c_2} \mathbf{e}(\mathbf{r}) \cdot \mathbf{dc} = 0 \tag{3.15}$$

and hence $[\phi_1 - \phi_2]_{c_1} = [\phi_1 - \phi_2]_{c_2}$.
For two points P_1 and P_2 on a perfect conductor, the potential difference ϕ_{12} is zero. This follows from the fact that the electric field is normal to the perfect conductor and hence $\mathbf{e} \cdot \mathbf{dc}$ is zero in (3.14). In the static case, a perfect conductor is an equipotential surface.

As adding or subtracting a constant from $\phi(\mathbf{r})$ does not change the corresponding electric field nor the potential difference between two points, absolute potential values are meaningless. Consequently, if we want to define the potential or in circuit terms, the voltage, at a particular point, it is necessary to define a reference point P_{ref} to which we assign a zero reference potential. The potential or voltage at a point P now becomes

$$\phi(P) = -\int_{P_{ref}}^{P} \nabla\phi \cdot \mathbf{dc}. \tag{3.16}$$

In many textbooks this reference point is placed at infinity. In our experience and in relation to topics such as capacitance and inductance matrices and multiconductor transmission lines or more generally, when considering the relationship between fields and lumped or distributed circuit equivalents, it is advantageous to select a particular point or most often a PEC as the reference. Sometimes, this reference is also called "ground". This is bad practice. Ground and shielding are concepts that arise in the context of Electromagnetic Compatibility or EMC (EMC is by definition "the extent to which a piece of hardware will tolerate electrical interference from other equipment, and will interfere with other equipment"). A "ground" is not automatically an equipotential surface.

Multiplying $\mathbf{e}(\mathbf{r})$ with $\epsilon(\mathbf{r})$ to obtain $\mathbf{d}(\mathbf{r})$ and application of (3.2), shows that

$$\nabla \cdot [\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -\rho(\mathbf{r}). \tag{3.17}$$

This is *Poisson's equation*, the solution of which allows to determine ϕ for a given charge distribution $\rho(\mathbf{r})$. Remark that in the above expression the permittivity $\epsilon(\mathbf{r})$ can still differ from point to point. In a piecewise homogeneous medium (3.17) reduces to

$$\nabla^2\phi(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon}. \tag{3.18}$$

For a charge free region, (3.18) yields *Laplace's equation*

$$\nabla^2\phi(\mathbf{r}) = 0. \tag{3.19}$$

A function satisfying Laplace's equation is said to be *harmonic*.

For an infinite homogeneous medium the solution of (3.18) can be written down explicitly. The final result immediately follows from the expression (2.85) for ϕ derived in the previous chapter, by putting ω to zero

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon}\int_V \frac{1}{|\mathbf{r} - \mathbf{r}'|}\rho(\mathbf{r}')dV'. \tag{3.20}$$

This result expresses that the potential is the superposition of the contributions from individual elementary charges $\rho(\mathbf{r}')dV'$. The potential due to a discrete charge q located at $\mathbf{r}'$ is

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon}\frac{q}{|\mathbf{r}-\mathbf{r}'|}. \tag{3.21}$$

Two-dimensional potential problems are an important class of potential problems, more in particular for the study of multiconductor transmission lines in the low frequency limit. In the two-dimensional case, the geometry and the charge distribution are independent of one co-ordinate, say z. We already encountered one such example when determining the field of a uniform line charge. Poisson's equation in a piecewise homogeneous medium becomes

$$\nabla^2_{xy}\phi(\boldsymbol{\rho}) = \frac{\partial^2}{\partial x^2}\phi(\boldsymbol{\rho}) + \frac{\partial^2}{\partial y^2}\phi(\boldsymbol{\rho}) = -\frac{\rho_l(\boldsymbol{\rho})}{\epsilon}, \tag{3.22}$$

with $\boldsymbol{\rho} = x\mathbf{u}_x + y\mathbf{u}_y$, the position vector in the xy-plane. The potential due to a discrete line charge ρ_l in a homogeneous dielectric, located at $\boldsymbol{\rho}'$, is

$$\phi(\boldsymbol{\rho}) = -\frac{\rho_l}{2\pi\epsilon}\ln\frac{|\boldsymbol{\rho}-\boldsymbol{\rho}'|}{|\boldsymbol{\rho}_{ref}-\boldsymbol{\rho}'|}, \tag{3.23}$$

with $\boldsymbol{\rho}_{ref}$ the reference point (or rather reference line) a zero potential is assigned to. The corresponding result for a line charge density $\rho_l(\boldsymbol{\rho})$ thus becomes

$$\phi(\boldsymbol{\rho}) = -\int_S \frac{\rho_l(\boldsymbol{\rho}')}{2\pi\epsilon}\ln\frac{|\boldsymbol{\rho}-\boldsymbol{\rho}'|}{|\boldsymbol{\rho}_{ref}-\boldsymbol{\rho}'|}dS', \tag{3.24}$$

with the integration extending over the cross section S of two-dimensional charge density.

3.4 Dielectrics - Electric dipole - Polarisation

Consider a pair of opposite charges in a homogeneous medium separated by a small distance d. Applying (3.21) to find the potential of these two charges gives

$$\phi(\mathbf{r}) = \frac{q}{4\pi\epsilon}\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{1}{|\mathbf{r}-\mathbf{r}''|}\right), \tag{3.25}$$

with $\mathbf{r}'$ and $\mathbf{r}''$ the position vectors of the charges. With the origin as in Fig. 3.6a we have that

$$|\mathbf{r}-\mathbf{r}'| = \sqrt{r^2 - rd\cos\theta + d^2/4} \approx r - \frac{d}{2}\cos\theta \tag{3.26}$$

and

$$|\mathbf{r}-\mathbf{r}''| = \sqrt{r^2 + rd\cos\theta + d^2/4} \approx r + \frac{d}{2}\cos\theta, \tag{3.27}$$

and for $d \ll r$, (3.25) becomes

$$\phi(\mathbf{r}) = \frac{q}{4\pi\epsilon}\frac{d\cos\theta}{r^2}. \tag{3.28}$$

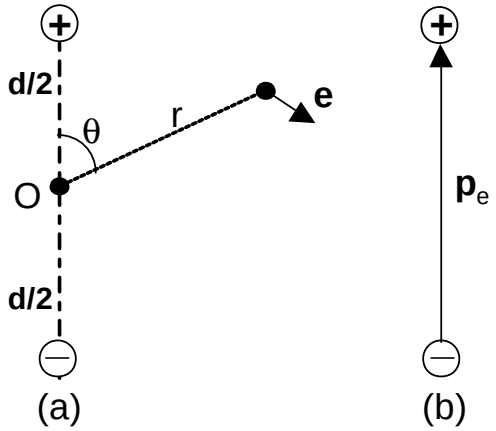


Figure 3.6: Electric field of a dipole (a) and resulting electric dipole moment $\mathbf{p}_e$ (b)

The final result is obtained by the limiting process $d \to 0$ while keeping the product qd finite. This product is the amplitude $p_e = qd$ of the electric dipole. A co-ordinate system independent expression is derived by introducing the dipole moment $\mathbf{p}_e$ as $\mathbf{p}_e = q\mathbf{d}$, with $\mathbf{d}$ the vector joining the negative charge to the positive one (see Fig. 3.6b). The potential (3.28) of the dipole becomes

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon}\frac{\mathbf{p}_e \cdot \mathbf{u}_r}{r^2} = \frac{1}{4\pi\epsilon}\frac{\mathbf{p}_e \cdot \mathbf{r}}{r^3}. \tag{3.29}$$

This potential exhibits rotational symmetry with respect to the line joining the charges. Prove that the electric field of the dipole is

$$\mathbf{e}(\mathbf{r}) = \frac{1}{4\pi\epsilon}\frac{qd}{r^3}(2\cos\theta\mathbf{u}_r + \sin\theta\mathbf{u}_\theta). \tag{3.30}$$

In this course, little will be said about the properties of materials. The macroscopic properties of many dielectric materials can be qualitatively explained in terms of electric dipoles. In nonpolar materials and in the absence of an external electric field $\mathbf{e}_{ext}$, the centre of the cloud of electrons of each atom coincides with the centre of the nucleus (Fig. 3.7a). The presence of an external electric field distorts the electron cloud such that its centre no longer coincides with the centre of the nucleus (Fig. 3.7b). This process is called *polarisation*. Fig. 3.7c shows the resulting dipole. Its direction coincides with that of the external field. The field of the dipole is induced by the external field and is the polarisation field. When a dielectric is placed in an external field, all the induced dipoles align themselves as shown in Fig. 3.8a . The net result (see Fig. 3.8b) is that the upper surface of the dielectric is positively charged, while the lower surface is negatively charged. The field of these surface charges points in the opposite direction of the external field. For a polar dielectric such as water, the molecules already exhibit a permanent dipole moment. In the absence of an external field, these moments are randomly oriented and no net dipole moment is created, but this changes under the influence of an external field. Both

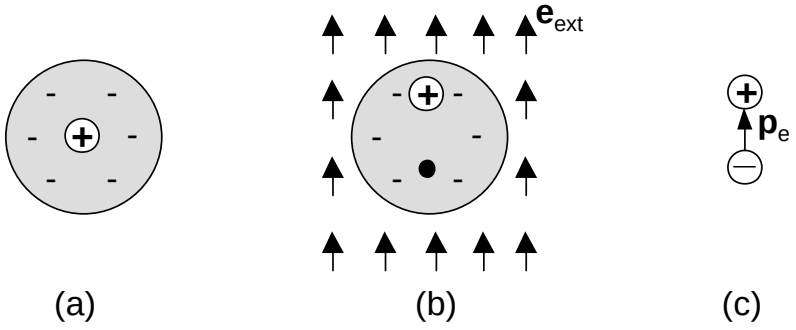


Figure 3.7: Polarisation due to an external electric field.

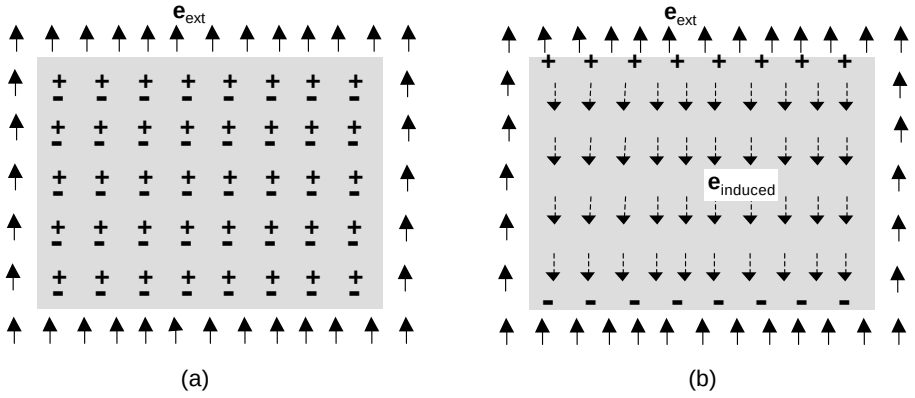


Figure 3.8: Nonpolar dielectric in an external electric field (a) and equivalent field problem (b).

for nonpolar and polar dielectrics, the dielectric displacement vector $\mathbf{d}(\mathbf{r})$ is no longer given by $\mathbf{d} = \epsilon_0 \mathbf{e}$ as in free space, but by

$$\mathbf{d} = \epsilon_0 \mathbf{e} + \mathbf{p}_e, \tag{3.31}$$

with $\mathbf{p}_e$ the electric polarisation field due to the polarisation process. For isotropic and linear materials, $\mathbf{p}_e$ and $\mathbf{e}$ are proportional

$$\mathbf{p}_e(\mathbf{r}) = \epsilon_0 \chi_e \mathbf{e}(\mathbf{r}), \tag{3.32}$$

with χ_e the electric susceptibility. Hence, (3.31) becomes

$$\mathbf{d}(\mathbf{r}) = \epsilon_0(1 + \chi_e)\mathbf{e}(\mathbf{r}) = \epsilon_0 \epsilon_r \mathbf{e}(\mathbf{r}) = \epsilon \mathbf{e}(\mathbf{r}). \tag{3.33}$$

We conclude this section with two remarks.

1. The relative dielectric constant ϵ_r of air is about 1.0006 at sea level and decreases to 1 with increasing height. Hence in most calculations (except when considering bending of waves in the atmosphere) no distinction is made between air and free space.

2. When the electric field strength of $\mathbf{e}_{ext}$ exceeds a certain value, the electrons of the dielectric will break free and give rise to a conduction current

that will often damage the material. This phenomenon is known as *dielectric breakdown* and the *dielectric strength* of the material is the field strength above which breakdown occurs. This breakdown strength is material dependent but also depends on temperature and humidity. For air the breakdown strength is of the order of $3MV/m$ while for polystyrene ($\epsilon_r = 2.6$) it amounts to $20MV/m$.

3.5 Boundary conditions

An important remark should be made at this point. Although we have introduced the concept of polarisation charges to explain the behaviour of dielectrics, the dielectric remains completely neutral and the polarisation charges do *not* appear in the right-hand member of (3.2). Moreover, as already shown in the previous chapter (2.66), the boundary condition for the dielectric displacement at the interface between two non-conducting dielectrics (or between a dielectric and free space) is

$$\mathbf{u}_n \cdot (\mathbf{d}_2 - \mathbf{d}_1) = 0, \tag{3.34}$$

with $\mathbf{u}_n$ the unit normal pointing in the direction of medium 2. We emphasise that the surface charge density ρ_s in (2.66) either stands for free surface charges (deliberately placed between the two dielectrics) or for surface charges at the interface between two conductors or at the interface between a conductor and a non-conducting dielectric. In none of these cases the polarisation charges enter into the boundary conditions. As in the general dynamic case, the tangential electric field remains continuous across the interface

$$\mathbf{u}_n \times (\mathbf{e}_2 - \mathbf{e}_1) = 0. \tag{3.35}$$

In the static case, it is advantageous to express the boundary conditions in terms of the potential ϕ. First, observe that this potential must be continuous everywhere as a jump discontinuity in the potential results in a non-physical delta function type of contribution to the electric field. Consequently, at a dielectric interface we have that $\phi_1 = \phi_2$. If this is the case everywhere on the boundary interface, the tangential derivatives of ϕ will also be continuous and hence, (3.35) is automatically satisfied. This implies that in the static case boundary conditions (3.34) and (3.35) can be replaced by

$$\phi_1 = \phi_2, \tag{3.36}$$

$$\epsilon_1 \left(\frac{\partial \phi}{\partial n}\right)_1 = \epsilon_2 \left(\frac{\partial \phi}{\partial n}\right)_2. \tag{3.37}$$

3.6 Conductors - Resistance - Joule's law

In contrast to a dielectric, a conductor has a number of electrons in the outermost shell of its atoms that are only loosely attached. In the presence of an external electric field, these electrons move from one atom to the next and this in the opposite direction of the external field, giving rise to the conduction current. The average velocity of the electrons is the electron drift velocity $\mathbf{m}_e$ in m/s, given by

$$\mathbf{m}_e = -\mu_e \mathbf{e}, \tag{3.38}$$

with μ_e the electron mobility of the conductor in m^2/Vs. Although semiconductors are not considered in this course, it is interesting to mention here that the current in semiconductors is not only due to moving electrons but also to moving holes. These holes are positively charged and the associated average hole drift velocity $\mathbf{m}_h$ is

$$\mathbf{m}_h = \mu_h \mathbf{e}, \tag{3.39}$$

with μ_h the hole mobility of the semiconductor. The conduction current density due to electrons and holes becomes

$$\begin{aligned} \mathbf{j} &= \mathbf{j}_e + \mathbf{j}_h \\ &= \rho_{ve}\mathbf{m}_e + \rho_{vh}\mathbf{m}_h \\ &= (-\rho_{ve}\mu_e + \rho_{vh}\mu_h)\mathbf{e} \\ &= e(N_e\mu_e + N_h\mu_h)\mathbf{e}, \end{aligned} \tag{3.40}$$

with $\rho_{ve} = -N_e e$ and $\rho_{vh} = N_h e$, the volume charge densities of the electrons and the holes, $e = 1.6 \times 10^{-19} C$ the absolute charge of a single electron or of a single hole and with N_e and N_h the number of free electrons and the number of free holes per unit of volume. From (3.40) we derive Ohm's law

$$\begin{aligned} \mathbf{j} &= \sigma \mathbf{e} \\ &= eN_e\mu_e\mathbf{e} \qquad \text{for a conductor,} \\ &= e(N_e\mu_e + N_h\mu_h)\mathbf{e} \qquad \text{for a semiconductor.} \end{aligned} \tag{3.41}$$

Let us now take a closer look at what happens to free charges in a conductor. As our starting point we take the time-domain charge conservation law (2.5)

$$\nabla \cdot \mathbf{j}(\mathbf{r}, t) + \frac{\partial}{\partial t}\rho(\mathbf{r}, t) = 0. \tag{3.42}$$

With $\mathbf{j} = \sigma\mathbf{e} = (\sigma/\epsilon)\mathbf{d}$ and $\nabla \cdot \mathbf{d} = \rho$, (3.42) becomes

$$\frac{\sigma}{\epsilon}\rho(\mathbf{r}, t) + \frac{\partial}{\partial t}\rho(\mathbf{r}, t) = 0, \tag{3.43}$$

from which we immediately conclude that

$$\rho(\mathbf{r}, t) = \rho(\mathbf{r}, t = 0)e^{-\frac{\sigma}{\epsilon}t} = \rho(\mathbf{r}, t = 0)e^{-\frac{t}{\tau}}, \tag{3.44}$$

with $\tau = \epsilon/\sigma$ the relaxation constant of the conductor. For a good conductor such as copper, $\sigma = 5.8 \times 10^7 S/m$, $\epsilon = \epsilon_0 = 8.85 \times 10^{-12} F/m$ and $\tau = 1.53 \times 10^{-19} s$. This shows that any free charge concentration in a conductor very quickly diffuses (except at extremely high frequencies). All the free charges end up at the boundary surface of the conductor. We will come back to this point in the sequel.

Next, we turn our attention to the resistance of a conductor. Fig. 3.9 shows a cylindrical piece of resistor connected to a DC source. If we suppose that the contact surfaces S_A and S_B are equipotential surfaces and taking into account the circular symmetric nature of the problem, the current density $\mathbf{j}(\mathbf{r})$ in the conductor is

$$\mathbf{j}(\mathbf{r}) = (I/S)\mathbf{u}_z, \tag{3.45}$$

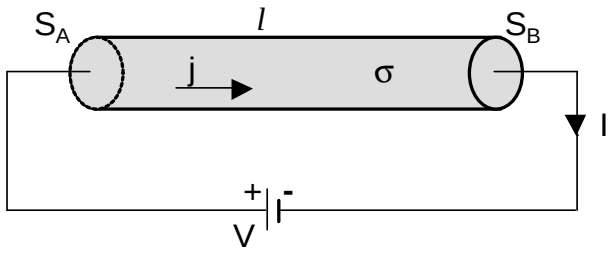


Figure 3.9: Cylindrical resistor.

with I the total current running through the circuit and S the cross section of the conductor. The corresponding electric field is $(I/\sigma S)\mathbf{u}_z$ and from (3.14) the potential difference $\phi_A - \phi_B = V$ between the contact surfaces is

$$V = \int_0^l \mathbf{e} \cdot \mathbf{u}_z dz = \frac{l}{\sigma S} I, \tag{3.46}$$

with l the length of the resistor and with the integration along the axis of the cylinder. Hence, the resistance R is

$$R = \frac{l}{\sigma S}, \tag{3.47}$$

a formula the reader is already familiar with. The resistance is proportional to the length and inversely proportional to the cross section. For an arbitrary resistor, (3.46) can be generalised by expressing $\phi_A - \phi_B$ and I as

$$\begin{aligned} \phi_A - \phi_B = V &= \int_A^B \mathbf{e}(\mathbf{r}) \cdot \mathbf{dc}, \\ I &= \int_S \sigma \mathbf{e}(\mathbf{r}) \cdot \mathbf{dS}. \end{aligned} \tag{3.48}$$

The line integral is taken along an arbitrary line between surface S_A and surface S_B. The surface integral is over any cross section of the resistor, with $\mathbf{dS} = \mathbf{u}_n dS$ and $\mathbf{u}_n$ the unit normal to S pointing in the direction of contact surface S_B. The resistance is thus given by

$$R = \frac{V}{I} = \frac{\int_A^B \mathbf{e}(\mathbf{r}) \cdot \mathbf{dc}}{\int_S \sigma \mathbf{e}(\mathbf{r}) \cdot \mathbf{dS}}. \tag{3.49}$$

Let us take a closer look at the field problem that has to be solved inside an arbitrary conductor to determine the electric field and hence to determine its resistance. As no free charges can build up inside the conductor, the potential ϕ satisfies Laplace's equation (3.19). This differential equation must be complemented with appropriate boundary conditions. Suppose the potential of contact surface S_A is unity (see Fig. 3.10) and that of S_B is zero (this choice is warranted

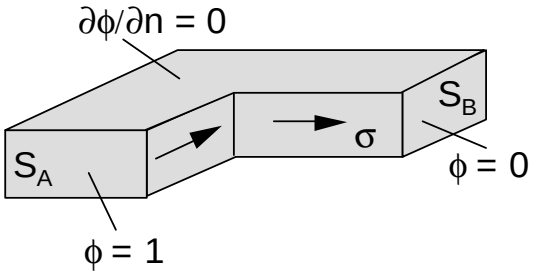


Figure 3.10: Laplace's problem for a piece of conductor.

as the total current I is proportional to the potential difference). Because the contact surfaces are equipotential surfaces, the current density is perpendicular to them (why?). Now consider the surface version of the charge conservation law (2.71) which, in the static case, in the absence of surface currents and for a conductor in free space or embedded in a non-conducting dielectric, reduces to

$$\mathbf{u}_n \cdot \mathbf{j} = \frac{\partial \phi}{\partial n} = 0 \tag{3.50}$$

at the surface of the conductor. This is easily understood. In the static case the normal component of the current at the surface of the conductor must vanish, otherwise an infinite amount of charge would accumulate there. Starting from (3.50), the reader can easily proof that in the resistance definition (3.49) any cross section S yields the same total current I. From the above considerations, the potential problem for the conductor can be formulated as

$$\nabla^2 \phi(\mathbf{r}) = 0 \qquad \text{inside the conductor,} \tag{3.51}$$

$$\phi = 1 \qquad \text{on contact surface } S_A, \tag{3.52}$$

$$\phi = 0 \qquad \text{on contact surface } S_B, \tag{3.53}$$

$$\frac{\partial \phi}{\partial n} = 0 \qquad \text{on the outer conductor surface.} \tag{3.54}$$

To better understand the impact of the contact surface boundary conditions on the solution of Laplace's equation, consider the simple conductor of length l with rectangular cross section depicted in Fig. 3.11 . To solve (3.51) the potential ϕ is represented by the following double Fourier cosine series

$$\phi(x, y, z) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} A_{nm}(z) \cos \frac{n\pi x}{a} \cos \frac{m\pi y}{b}. \tag{3.55}$$

By selecting the cosine series, boundary condition (3.54) is automatically satisfied. The z-variation of the potential is taken into account by the z-dependent Fourier coefficients $A_{nm}(z)$. To find these coefficients, (3.55) is inserted in (3.51) yielding

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2} A_{nm}(z) - \left(\left(\frac{n\pi}{a} \right)^2 + \left(\frac{m\pi}{b} \right)^2 \right) A_{nm}(z) = 0. \tag{3.56}$$

The solution of (3.56) is quite straightforward:

$$A_{nm}(z) = P_{nm} e^{\alpha_{nm} z} + Q_{nm} e^{-\alpha_{nm} z} \qquad \forall m, n \text{ except } n = m = 0, \tag{3.57}$$

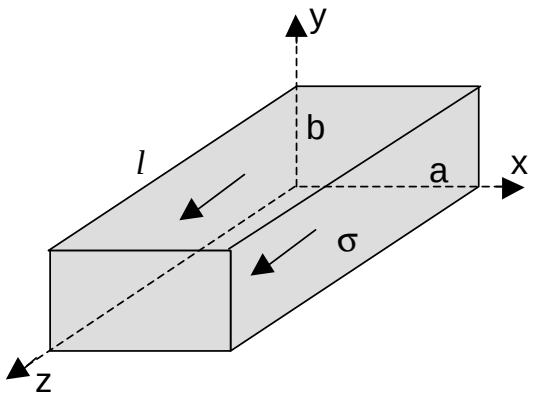


Figure 3.11: Conductor with rectangular cross section.

$$\alpha_{nm} = \sqrt{\left(\frac{n\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2},$$
$$A_{00} = Cz + D, \tag{3.58}$$

with P_{nm}, Q_{nm}, C and D constant coefficients. It is clear that Laplace's equation only combined with boundary condition (3.54) allows a complex solution for ϕ and hence for the current. However, we still have to invoke the contact surface conditions (3.52) and (3.53). As the potential must be constant over the whole cross section, both at $z = 0$ and at $z = l$, we immediately conclude that $P_{nm} = Q_{nm} = 0$ except for $m = n = 0$. With $\phi = 1$ at $z = 0$ and $\phi = 0$ at $z = l$, is it easy to see that $D = 1$ and $C = -1/l$. The current density $\mathbf{j} = -\sigma\nabla\phi$ thus becomes $\mathbf{j} = \frac{\sigma}{L}\mathbf{u}_z$ confirming (3.47). Other, less simple, potential boundary conditions at the contact surfaces, lead to solutions for which P_{nm} and Q_{nm} will be different from zero.

The dissipated power P_{dis} or Joule losses can be expressed as

$$P_{dis} = \int_V \sigma|\mathbf{e}(\mathbf{r})|^2 dV = \int_V \sigma(\nabla\phi \cdot \nabla\phi)dV. \tag{3.59}$$

To further transform (3.59), the following Green's theorem for the functions $f(\mathbf{r})$ and $g(\mathbf{r})$ must be applied

$$\int_V (f\nabla^2 g + \nabla f \cdot \nabla g)dV = \int_S f\frac{\partial g}{\partial n}dS, \tag{3.60}$$

with in our particular case $f = g = \phi$. Supposing that σ is constant and taking into account that ϕ satisfies Laplace's equation and boundary condition (3.50), (3.60) becomes

$$\begin{aligned}\int_V \sigma(\nabla\phi \cdot \nabla\phi)dV &= \sigma\int_{S_A} \phi\frac{\partial\phi}{\partial n}dS + \sigma\int_{S_B} \phi\frac{\partial\phi}{\partial n}dS \\ &= \sigma\phi_A\int_{S_A} \frac{\partial\phi}{\partial n}dS + \sigma\phi_B\int_{S_B} \frac{\partial\phi}{\partial n}dS. \end{aligned}\tag{3.61}$$

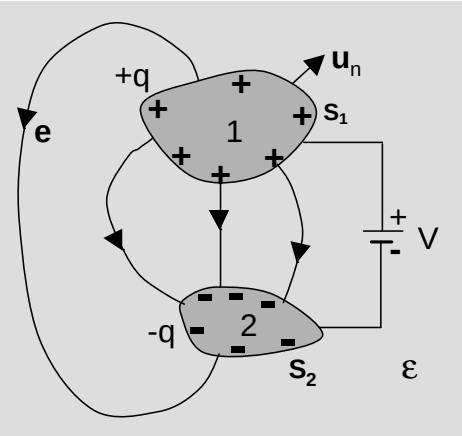


Figure 3.12: A capacitor formed by two conductors.

To obtain the above result, we used the fact that S_A and S_B are equipotential surfaces. Taking into account that the normal to S_A and the normal to S_B both point outward of the integration volume, the above expression reduces to the potential difference V between these surfaces multiplied with the flux of the current density, i.e. the total current I. With $V = RI$ we obtain the familiar Joule law from circuit theory $P_{dis} = RI^2$.

3.7 Capacitance - Capacitance matrix

In this section we restrict ourselves to perfect electric conductors. As the electric field is zero inside such a PEC, no potential difference can exist between different points on the surface of the PEC: the surface of a PEC is an equipotential surface. Now consider two arbitrary (neutral) PEC's embedded in an arbitrary but lossless dielectric medium and connected to a voltage source V (see Fig. 3.12). Connecting these conductors to a voltage source results in a positive charge accumulation q on conductor 1 and $-q$ on conductor 2. We already explained above that no volume charge can accumulate in a PEC. The charge accumulates on the surface of the PEC. The surface charge density of conductor 1 is denoted as ρ_{s1} with

$$q = \int_{S_1} \rho_{s1}(\mathbf{r}) dS = \int_{S_1} \epsilon \mathbf{u}_n \cdot \mathbf{e}(\mathbf{r}) dS = -\int_{S_1} \epsilon \frac{\partial \phi}{\partial n} dS, \tag{3.62}$$

with $\mathbf{u}_n$ the outward pointing unit normal to the surface S_1. As depicted in Fig. 3.12, the surface charges are the sources of the electric field with field lines starting at conductor 1 and ending on conductor 2. The potential ϕ now satisfies the following equations

$$\nabla^2 \phi(\mathbf{r}) = 0 \qquad \text{outside the conductors,} \tag{3.63}$$

$$\phi = V \qquad \text{on conductor surface } S_1, \tag{3.64}$$

$$\phi = 0 \qquad \text{on conductor surface } S_2. \tag{3.65}$$

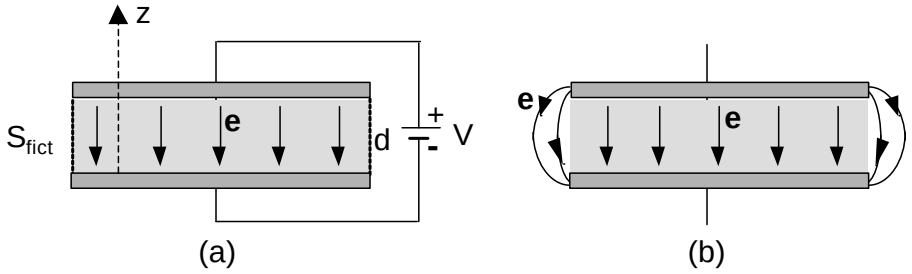


Figure 3.13: Parallel plate capacitor: idealised case (a), fringing effects (b).

The capacitance of the capacitor formed by the two conductors is defined as

$$C = \frac{q}{V} = \frac{-\int\limits_{S_1} \epsilon \frac{\partial \phi}{\partial n} dS}{-\int\limits_{1}^{2} \nabla\phi \cdot \mathbf{dl}}, \tag{3.66}$$

with the line integral taken between an arbitrary point on conductor 1 and an arbitrary point on conductor 2. The total charge q is proportional to V. Hence, the actual potential problem that needs solving is that of equations (3.63)–(3.65) with $V = 1$. The solution of this problem depends on the size, shape and position of the conductor pair and also on the properties of the dielectric. In general, (3.63)–(3.65) can only be solved numerically. This topic of great practical importance will be briefly discussed in the next section.

Let us first consider two examples that can be handled analytically. The first example is that of the idealised parallel plate capacitor of Fig. 3.13a . In this idealised case, (3.63)–(3.65) are complemented with $\frac{\partial \phi}{\partial n} = 0$ on the fictitious boundary surfaces S_{fict}. This extra condition, which does not follow from the physics of the problem, makes the solution of the potential problem much more easy as the electric field between the plates will now be uniform $\mathbf{e}(\mathbf{r}) = -(V/d)\mathbf{u}_z$. The capacitance is easily found to be

$$C_{ideal} = \frac{\epsilon S}{d}, \tag{3.67}$$

with S the surface of the plates and d the distance between them. In reality field lines will not be that well-behaved and fringing effects (see Fig. 3.13b) will reduce the capacity. When the dimensions of the plates are much larger than the distance d between them, the approximate result (3.67) is quite accurate.
As a second example consider the coaxial capacitor depicted in Fig. 3.14 . The voltage difference between the inner and the outer conductor is V. The calculation is again simplified by assuming that no fringing effects occur at the end faces $z = 0$ and $z = l$. Another way of looking at the problem, is to regard this capacitor as part of an infinitely long coaxial cable. In that case, we are interested in determining its capacitance per unit of length. In the cylindrical co-ordinate system (r, θ, z) the potential ϕ only depends on r. Hence, (3.63) reduces to

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}}{\mathrm{d}r}\phi(r)\right) = 0, \tag{3.68}$$

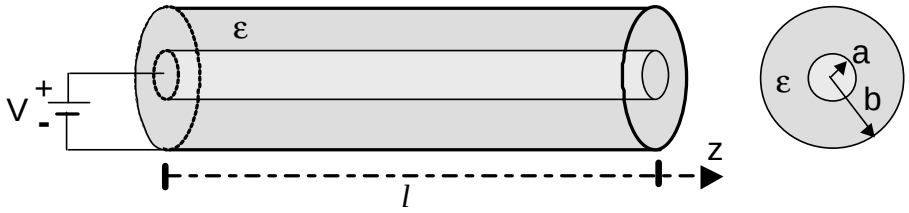


Figure 3.14: Coaxial capacitor.

the solution of which is

$$\phi(r) = P \ln r + Q, \tag{3.69}$$

with P and Q as yet unknown constants. With $\phi(r = a) = V$ and $\phi(r = b) = 0$, the final result for ϕ is

$$\phi(r) = \frac{V}{\ln \frac{a}{b}} \ln \frac{r}{b}, \tag{3.70}$$

with corresponding electric field

$$\mathbf{e}(\mathbf{r}) = -\nabla\phi(r) = -\frac{\mathrm{d}}{\mathrm{d}r}\phi(r)\mathbf{u}_r = \frac{V}{\ln \frac{b}{a}} \frac{1}{r}\mathbf{u}_r. \tag{3.71}$$

The total charge on the inner conductor then becomes

$$q = \int_0^l \int_0^{2\pi} \epsilon \frac{V}{\ln \frac{b}{a}} \frac{1}{a} a d\theta dz = 2\pi\epsilon l \frac{V}{\ln \frac{b}{a}}, \tag{3.72}$$

showing that the capacitance is

$$C_{coax} = \frac{2\pi\epsilon l}{\ln \frac{b}{a}}, \tag{3.73}$$

or $\frac{2\pi\epsilon}{\ln \frac{b}{a}}$ F/m per unit of length.

To conclude this section, the capacitance concept is extended to a configuration of $M + 1$ conductors. Take the example of the three conductors ($M = 2$) shown in Fig. 3.15 . To simplify the reasoning, one of the conductors is selected to be the zero potential reference conductor. The relationship $q = CV$ for a single conductor can now be extended to

$$\begin{aligned} q_1 &= C_{11}V_1 + C_{12}V_2, \\ q_2 &= C_{21}V_1 + C_{22}V_2. \end{aligned} \tag{3.74}$$

Equation (3.74) expresses that the total charge q_1 on conductor 1 and the total charge q_2 on conductor 2 depend linearly on the potential differences V_1 and V_2 between these conductors and the reference conductor. Remark that $q_1 + q_2 \neq 0$ as the reference conductor is also charged such that $q_1 + q_2 + q_{ref} = 0$. The matrix

$$\mathbf{C} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}, \tag{3.75}$$

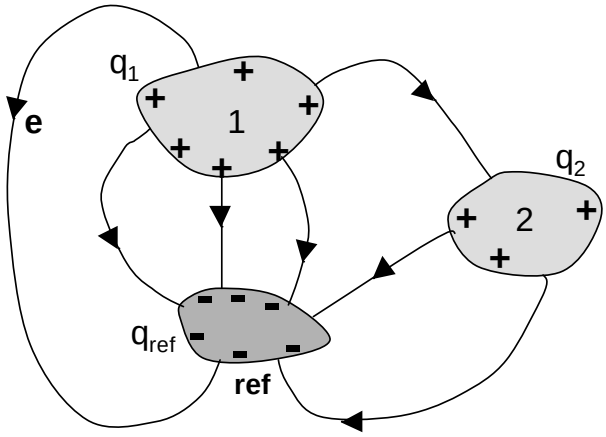


Figure 3.15: Capacitive coupling between three conductors.

is the capacitance matrix. To determine the elements of this capacitance matrix, it suffices to solve two potential problems, i.e.

$$\begin{aligned} \nabla^2\phi(\mathbf{r}) &= 0 \qquad \text{outside the conductors,} \\ \phi &= V_1 = 1 \qquad \text{on conductor 1,} \\ \phi &= V_2 = 0 \qquad \text{on conductor 2} \end{aligned} \tag{3.76}$$

and conversely,

$$\begin{aligned} \nabla^2\phi(\mathbf{r}) &= 0 \qquad \text{outside the conductors,} \\ \phi &= V_1 = 0 \qquad \text{on conductor 1,} \\ \phi &= V_2 = 1 \qquad \text{on conductor 2.} \end{aligned} \tag{3.77}$$

In both cases the reference conductor is kept on zero potential. Following (3.74), solving the first of these problems yields $q_1 = C_{11}$ and $q_2 = C_{21}$, while solving the second problem yields $q_1 = C_{12}$ and $q_2 = C_{22}$.
The above reasoning can be extended to the general $M+1$ conductor case with one of them selected to be the zero potential reference conductor and the others numbered from 1 to M. Expression (3.74) now becomes

$$\mathcal{Q} = \mathbf{C}\,\mathcal{V}, \tag{3.78}$$

with $\mathcal{V}$ and $\mathcal{Q}$, $M \times 1$ column vectors with elements V_m and q_m and with $\mathbf{C}$ the $M \times M$ capacitance matrix. It can be proved that the following general properties hold

- the capacitance matrix is symmetric, $C_{ij} = C_{ji}$;
- all the non-diagonal elements are strictly negative, $C_{ij} < 0$;
- all the diagonal elements are strictly positive, $C_{ii} > 0$;
- the matrix is strictly diagonally dominant, i.e. $C_{ii} > \sum_{j(j\neq i)} |C_{ij}|$.

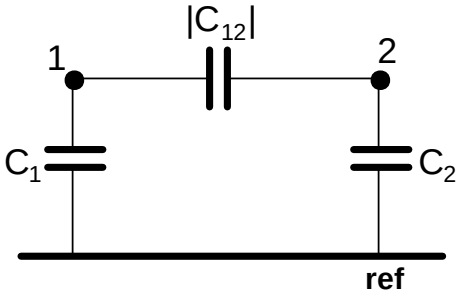


Figure 3.16: Equivalent circuit representation for the capacitive coupling between the two conductors and reference conductor example of Fig. 3.15.

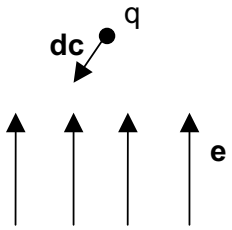


Figure 3.17: Charge q immersed in a field $\mathbf{e}$.

An equivalent capacitor circuit can be derived based on the above properties. Fig. 3.16 shows the circuit example for the two conductor plus reference conductor case of Fig. 3.15. The value of the capacitor between conductor i and conductor j is given by $|C_{ij}| = |C_{ji}|$. The value of capacitor C_i between conductor i and the reference conductor (often called coupling to "ground") is given by $C_i = C_{ii} - \sum_{j(j\neq i)} |C_{ij}|$.

3.8 Electrostatic energy

To charge the capacitor depicted in Fig. 3.12, the source has to deliver an amount of energy to transfer charges from one conductor to the other. In the absence of losses (PEC conductors and lossless dielectrics), this energy is stored in the dielectric medium as electrostatic energy. In order to calculate this energy, we must first return to the simple situation of a charge q immersed in a field $\mathbf{e}$ as depicted in Fig. 3.17 . This charge is subjected the Coulomb force $\mathbf{F} = q\mathbf{e}$. To move it over an elementary distance $\mathbf{dc}$, an external force $\mathbf{F}_{ext} = -q\mathbf{e}$, compensating the Coulomb force, is needed. The elementary amount of energy spent by this external force then is

$$du_e = \mathbf{F}_{ext} \cdot \mathbf{dc} = -q\mathbf{e} \cdot \mathbf{dc}. \tag{3.79}$$

When moving the charge between two points, say P_1 and P_2, the total energy delivered to the charge becomes

$$u_e = -q \int_{P_1}^{P_2} \mathbf{e} \cdot \mathbf{dc} = q(\phi_2 - \phi_1) = q\phi_{21}, \tag{3.80}$$

where we invoked (3.14) to obtain the final result. The total energy needed to move the charge q from P_1 to P_2 against the electric field $\mathbf{e}$ is equal to the product of that charge and the potential difference between the two points.
Let us now apply this knowledge to the capacitor problem. According to (3.80), the energy needed to move an elementary amount of charge dq from one plate to the other is

$$du_e = vdq, \tag{3.81}$$

with v representing the voltage difference between the plates when transferring dq. To completely charge the capacitor, this process must continue until the final charge q_{fin} is obtained. The total energy needed to accomplish this is found by integrating (3.81)

$$u_e = \int_0^{q_{fin}} vdq = \int_0^{q_{fin}} \frac{q}{C} dq, \tag{3.82}$$

where we used the relationship $q = Cv$. With $q_{fin} = CV$ the integration yields

$$u_e = \frac{1}{2} CV^2, \tag{3.83}$$

the well-known formula from circuit theory.
Let us now relate this result to the fields. To this end consider the electrostatic energy density $\frac{1}{2}\mathbf{e}\cdot\mathbf{d}$. The reason for giving this name to $\frac{1}{2}\mathbf{d}\cdot\mathbf{e}$ will soon become clear. We integrate this density over a large volume V encompassing the two conductors forming the capacitor.

$$u_e = \int_V \frac{1}{2}\mathbf{e}\cdot\mathbf{d}\ dV = \int_V \frac{1}{2}\epsilon\nabla\phi\cdot\nabla\phi\ dV. \tag{3.84}$$

As shown in Fig. 3.18 , the surface of V consists of three parts: the outer surface S_0 and the surfaces S_1 and S_2 of the conductors. Let us further suppose that the dielectric medium exhibits a piecewise constant permittivity (the proof can be extended to cover arbitrary place dependent values of ϵ). Using Green's theorem (3.60), (3.84) becomes

$$\begin{aligned} u_e &= \int_{S_0+S_1+S_2} \frac{1}{2}\epsilon\phi\frac{\partial\phi}{\partial n} dS - \int_V \phi\nabla^2\phi\ dV. \\ &= -\int_{S_0+S_1+S_2} \frac{1}{2}\phi\mathbf{u}_n\cdot\mathbf{d}\ dS. \end{aligned} \tag{3.85}$$

As $\nabla^2\phi = 0$, the volume integral drops out. When extending the integration volume to infinity, the electric field and the potential decrease sufficiently rapid

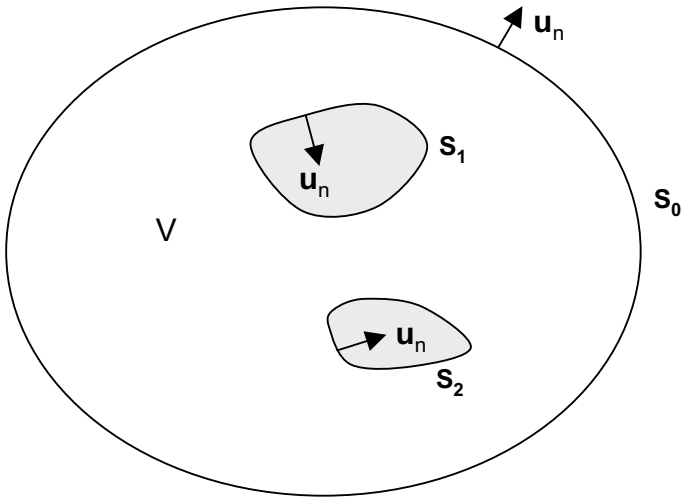


Figure 3.18: Integration volume for the electrostatic energy density.

for the contribution of S_0 to become negligible (why?). Surfaces S_1 and S_2 are equipotential surfaces and we suppose (but this can again be generalised) that the potential of S_1 is V and that of S_2 is zero. Hence, (3.85) becomes

$$\begin{aligned} u_e &= -\frac{1}{2}V \int_{S_1} \mathbf{u}_n \cdot \mathbf{d}\, dS \\ &= \frac{1}{2}Vq \\ &= \frac{1}{2}CV^2, \end{aligned} \tag{3.86}$$

where we took into account that $\mathbf{u}_n \cdot \mathbf{d} = -\rho_s$ as in Green's theorem $\mathbf{u}_n$ is the outward unit normal to S_1 pointing to the inside of the conductor (see Fig. 3.18). This is the result already obtained above for the electrostatic energy of the capacitor, making clear why $\frac{1}{2}\mathbf{e}\cdot\mathbf{d}$ is called the *electrostatic energy density* or electrostatic potential energy density. Remember that we already encountered this quantity in the dynamical case when discussing conservation of energy in relationship to Maxwell's equation (Section 2.3).
Formula (3.83) can be extended to the multiconductor case. Prove that u_e becomes

$$u_e = \frac{1}{2}\mathcal{V}^T \mathbf{C}\, \mathcal{V}, \tag{3.87}$$

with $\mathcal{V}^T$ the transposed of the voltage vector $\mathcal{V}$.

3.9 Solution of Laplace's equation

As will be clear by now, solving Laplace's equation subject to a set of boundary conditions is an important topic in Electrostatics, either to determine the resistance of a conductor or the capacitance matrix of a set of conductors. A large body of engineering and mathematics literature deals with the

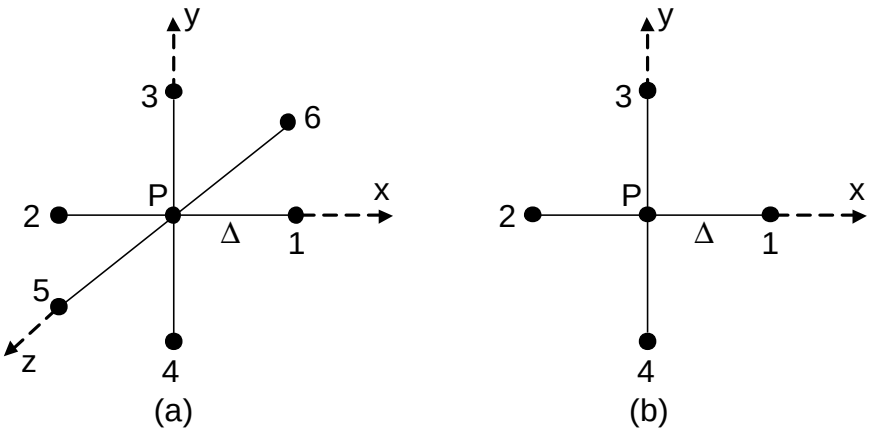


Figure 3.19: Cubic cell discretisation for the solution of Laplace's equation: three-dimensional case (a), two-dimensional case (b).

solution of Laplace's equation and many software packages to tackle the problem are available. Over the past years, the author's of this course have themselves contributed to the problem of determining the capacitance matrix for two-dimensional conductor configurations embedded in multilayered dielectrics. This section is only intended as a brief introduction to the numerical solution of Laplace's equation by means of a few examples. Two numerical techniques are introduced: the finite difference technique (FD) and the integral equation (IE) technique. Another important method is the finite element (FE) method, but a detailed discussion of FD, IE and FE techniques is the subject matter of a general course in Computational Physics or a course in Computational Electromagnetics.

3.9.1 The finite difference technique

Probably the most accessible way to tackle Laplace's equation is the finite difference technique based on the finite difference representation of the Laplace operator. In this approach, the problem space is discretised into a large number of elementary cubic cells with side Δ. Consider an arbitrary grid point P and its 6 neighbours numbered from 1 to 6 as shown in Fig. 3.19a. The potential at grid point 1 allows the following Taylor series expansion

$$\phi_1 = \phi_P + \left(\frac{\partial}{\partial x}\phi\right)_P \Delta + \left(\frac{\partial^2}{\partial x^2}\phi\right)_P \frac{\Delta^2}{2} + \cdots . \tag{3.88}$$

Similar results hold for the other neighbours of P

$$\phi_2 = \phi_P - \left(\frac{\partial}{\partial x}\phi\right)_P \Delta + \left(\frac{\partial^2}{\partial x^2}\phi\right)_P \frac{\Delta^2}{2} + \cdots , \tag{3.89}$$

$$\phi_3 = \phi_P + \left(\frac{\partial}{\partial y}\phi\right)_P \Delta + \left(\frac{\partial^2}{\partial y^2}\phi\right)_P \frac{\Delta^2}{2} + \cdots , \tag{3.90}$$

$$\phi_4 = \phi_P - \left(\frac{\partial}{\partial y}\phi\right)_P \Delta + \left(\frac{\partial^2}{\partial y^2}\phi\right)_P \frac{\Delta^2}{2} + \cdots , \tag{3.91}$$

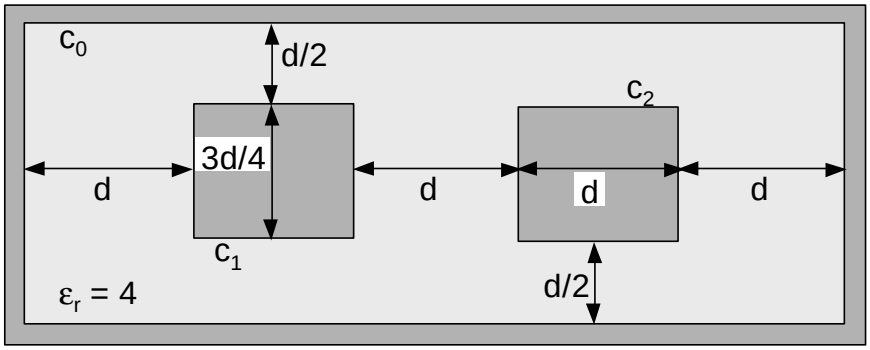


Figure 3.20: Cross section of a coaxial line with two inner signal conductors.

$$\phi_5 = \phi_P + \left(\frac{\partial}{\partial z}\phi\right)_P \Delta + \left(\frac{\partial^2}{\partial z^2}\phi\right)_P \frac{\Delta^2}{2} + \cdots, \tag{3.92}$$

$$\phi_6 = \phi_P - \left(\frac{\partial}{\partial z}\phi\right)_P \Delta + \left(\frac{\partial^2}{\partial z^2}\phi\right)_P \frac{\Delta^2}{2} + \cdots. \tag{3.93}$$

Summing up the results for the six neighbours yields a simple finite difference approximation for the Laplacian

$$\left(\nabla^2\phi\right)_P \approx \frac{1}{\Delta^2}(\phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 + \phi_6 - 6\,\phi_P). \tag{3.94}$$

For a cubic grid, the approximation error is of the order Δ^2 (why?). Laplace's equation will thus be satisfied provided

$$\phi_P = \frac{\phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5 + \phi_6}{6}, \tag{3.95}$$

i.e. the potential at each point must be the mean value of the potentials of the neighbours. In the two-dimensional case (Fig. 3.19b) a similar result is obtained

$$\phi_P = \frac{\phi_1 + \phi_2 + \phi_3 + \phi_4}{4}. \tag{3.96}$$

Of course, each point now only counts four neighbours. For a non-cubic grid, similar but less simple formulas can be derived.

An example will make clear how the finite difference approximation of the Laplacian can be used to solve a particular potential problem. Suppose we are interested in the capacitance matrix of a coaxial line with a homogeneous dielectric filling and two signal conductors with rectangular cross section inside a rectangular outer shielding conductor. The cross section of the line is shown in Fig. 3.20. The outer conductor is the zero potential reference conductor, hence all grid points on boundary c_0 have to be kept on zero potential. According to the general capacitance matrix theory of the previous section, we first have to solve (3.76). This implies that grid points on c_1 must remain at potential $\phi = 1$ while grid points on c_2 must remain at potential $\phi = 0$. To solve (3.76) the following procedure is adopted

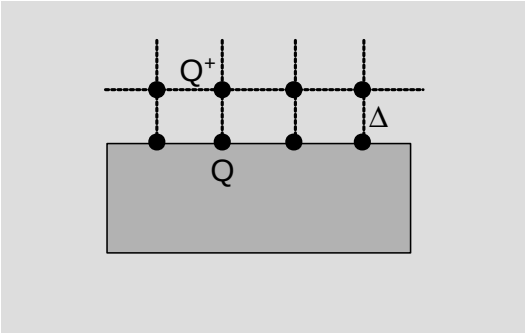


Figure 3.21: Finite difference approximation of the normal derivative.

- step 1: initialise all ϕ-values to zero except on boundary c_1 where all ϕ-values are initialised to 1;
- step 2: for each point P of the grid, except those on boundaries c_0, c_1 and c_2, calculate the value of ϕ using (3.96);
- step 3: repeat step 2 until convergence occurs;
- step 4: calculate the total charge q_1 on conductor 1 by integrating $-\epsilon\frac{\partial\phi}{\partial n}$ over its circumference c_1. To this end a simple finite difference approximation of $\frac{\partial\phi}{\partial n}$ suffices. Referring to Fig. 3.21 , the total surface charge (in C/m) on a line segment of length Δ centred on Q is $\epsilon(1-\phi_{Q^+})\Delta$.
- step 5: calculate the total charge q_2 on conductor 2.

The value of C_{11} is equal to q_1 and that of C_{21} is equal to q_2. To solve (3.77) a similar procedure is used yielding C_{22} and C_{12}.
The simple procedure described above often converges very slowly. A typical way to check convergence is to monitor the change in the potential when repeating step 2. In many cases, the absolute change in the potential values between iteration steps quickly becomes small although the final result is not yet reached and many iteration steps, each resulting in a small potential change, are still needed. Several techniques exist to enhance convergence such as successive over-relaxation. When the dielectric is piecewise homogeneous, the problem becomes much more complicated as the continuity of the normal component of $\mathbf{d}$ must now be enforced at each interface between dielectrics. For the (relative) dimensions shown in Fig. 20 and for a dielectric with $\epsilon_r = 4$, the following capacitance matrix was obtained

$$\mathbf{C} = \begin{pmatrix} 283.8 & -21.7 \\ -21.7 & 283.8 \end{pmatrix} pF/m. \tag{3.97}$$

The corresponding equivalent circuit is depicted in Fig. 3.22.

3.9.2 The integral equation technique

The finite difference approach presented below is a volume technique as it requires the discretisation of the complete problem space. Solving the parallel-plate capacitor problem of Fig. 3.13 with the FD-technique in principle requires

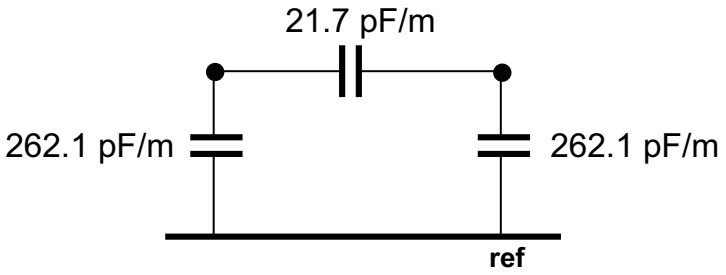


Figure 3.22: Equivalent circuit for the coaxial line of Fig. 20.

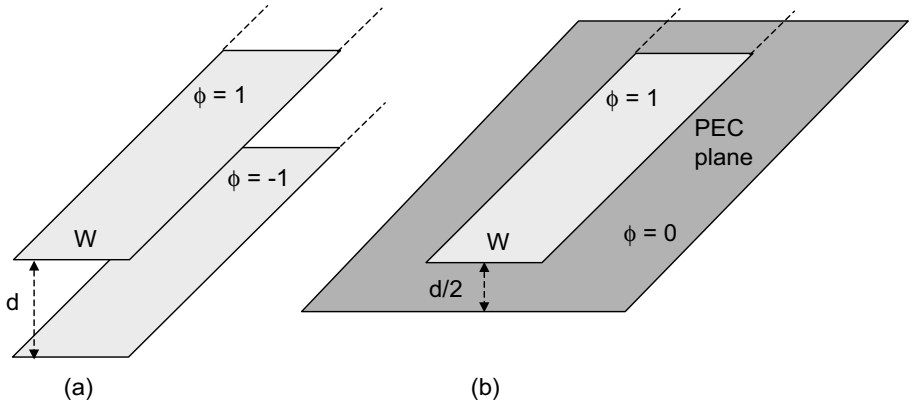


Figure 3.23: Two infinitely long parallel strips: original problem (a), equivalent problem (b).

the discretisation of the complete space. In practice, a sufficiently large discretisation volume is selected such that at its boundary surface the potential can approximately be put to zero. This implies that we are no longer analysing a two conductor problem, but a three conductor problem with the outer boundary of the discretisation volume playing the role of zero potential reference conductor. In terms of the equivalent capacitor circuit introduced in the previous section, this implies that the discretisation volume will be sufficiently large provided $C_1 = C_{11} - |C_{12}|$ and $C_2 = C_{22} - |C_{12}|$ are small with respect to $|C_{12}|$.

To avoid the volume discretisation and possible convergence issues, an integral equation can be constructed based on an expression similar to (3.20), for the potential in terms of the charge density.

Capacitance of two strips embedded in free space

As a three-dimensional problem quickly requires a large number of unknowns and hence the solution of a large linear system of equations, the integral equation technique will be illustrated by means of the two-dimensional equivalent of the parallel-plate capacitor. This two-dimensional equivalent is depicted in Fig. 3.23a . It consist of two infinitely long parallel strips of width W, separated by a distance d and placed in free space. Our purpose is to calculate the capaci-

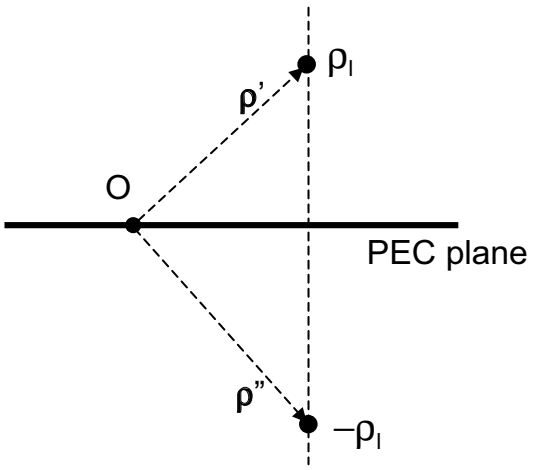


Figure 3.24: Single line charge above a PEC ground plane and corresponding image charge.

tive coupling between these strips. To simplify the analysis, suppose that the potential of the upper strip is 1 and that of the lower strip is -1. Following the image theory for charges put forward in Section 2.8, the original problem can be replaced by the equivalent one depicted in Fig. 3.23b with a single strip above a PEC plane. The capacitance per unit of length C of this plate with respect to the PEC plane is twice that of the original capacitance between the strips (why?).
Now consider a single line charge ρ_l above the PEC plane located at $\boldsymbol{\rho}'$ (see Fig. 3.24). According to image theory, the potential of this line charge is equal to that of the charge itself and its image charge. Using (3.24) we find

$$\phi(\boldsymbol{\rho}) = -\frac{\rho_l}{2\pi\epsilon} \ln \frac{|\boldsymbol{\rho} - \boldsymbol{\rho}'|}{|\boldsymbol{\rho}_{ref} - \boldsymbol{\rho}'|} + \frac{\rho_l}{2\pi\epsilon} \ln \frac{|\boldsymbol{\rho} - \boldsymbol{\rho}''|}{|\boldsymbol{\rho}_{ref} - \boldsymbol{\rho}''|}, \tag{3.98}$$

with $\boldsymbol{\rho}'' = x'\mathbf{u}_x - y'\mathbf{u}_y$ the position vector of the mirror image of $\boldsymbol{\rho}' = x'\mathbf{u}_x + y'\mathbf{u}_y$ and with the x-axis coinciding with the position of the PEC plane. As the zero reference potential coincides with the PEC plane, the distances $|\boldsymbol{\rho}_{ref} - \boldsymbol{\rho}'|$ and $|\boldsymbol{\rho}_{ref} - \boldsymbol{\rho}''|$ are identical and (3.98) can be simplified, yielding

$$\phi(\boldsymbol{\rho}) = G(\boldsymbol{\rho}, \boldsymbol{\rho}') = -\frac{\rho_l}{2\pi\epsilon} \ln \frac{|\boldsymbol{\rho} - \boldsymbol{\rho}'|}{|\boldsymbol{\rho} - \boldsymbol{\rho}''|}. \tag{3.99}$$

In the previous chapter we already introduced the notion of a Green's function. A Green's function is the solution to a particular differential equation and associated boundary conditions for a point source or line source excitation. Hence, (3.99) is the Green's function for the two-dimensional Laplace equation in a homogeneous half-space on top of a perfectly conducting ground plane.
Applying the superposition principle, (3.99) can be used to express the potential of the unknown surface charge density $\rho_l(\boldsymbol{\rho})$ on the strip c above the PEC plane as

$$\phi(\boldsymbol{\rho}) = \int_c \rho_l(\boldsymbol{\rho}') G(\boldsymbol{\rho}, \boldsymbol{\rho}') dc'. \tag{3.100}$$

This potential is automatically zero on the PEC plane and satisfies Laplace's equation. The only boundary condition still to be imposed is that the strip surface is an equipotential surface, say at potential 1. Hence,

$$\lim_{\boldsymbol{\rho}\to c}\phi(\boldsymbol{\rho}) = \lim_{\boldsymbol{\rho}\to c}\int_c \rho_l(\boldsymbol{\rho}')G(\boldsymbol{\rho},\boldsymbol{\rho})dc' = 1. \tag{3.101}$$

This equation is an *integral equation* as the unknown of the problem, the line charge density $\rho_s(\boldsymbol{\rho})$, appears under the integral sign. It has to be satisfied for any point $\boldsymbol{\rho}$ on the strip c. Remark that taking the limit under the integration sign is not automatically warranted as the integrand becomes infinite for $\boldsymbol{\rho} = \boldsymbol{\rho}'$. There are may different ways to solve the integral equation (3.101). We only discuss the most simple one. To this end, the strip is subdivided into N elementary segments of length Δ with the line charge density taking a constant value ρ_{li} on each segment i. This is of course an approximation, but for a sufficiently large number of segments N this piecewise constant charge model will converge to the correct continuous charge distribution. Expression (3.101) can be rewritten as

$$\lim_{\boldsymbol{\rho}\to c}\sum_{i=1}^{N}\rho_{li}\int_{c_i} G(\boldsymbol{\rho},\boldsymbol{\rho}')dc_i' = 1, \tag{3.102}$$

with c_i representing segment i. By the above discretisation process N unknowns were introduced. To obtain a set of equations for these unknowns, we now also impose (3.102) in N points of c. Of course, we still have the freedom to arbitrarily select these points but a logical choice is that of selecting the centres $\boldsymbol{\rho}_j$ of the segments, such that

$$\lim_{\boldsymbol{\rho}\to\boldsymbol{\rho}_j}\sum_{i=1}^{N}\rho_{li}\int_{c_i} G(\boldsymbol{\rho},\boldsymbol{\rho}')dS_i' = 1 \quad j = 1,2,...,N. \tag{3.103}$$

(3.103) is a set of N linear equations for the N unknown coefficients ρ_{li}. In matrix form this set of equation can be concisely written as

$$\mathcal{I}_{\mathcal{N}} = \mathbf{C^{-1}}\ \mathcal{Q}, \tag{3.104}$$

with $\mathcal{I}_{\mathcal{N}}$ a $N\times 1$ column matrix of elements 1, with $\mathbf{C^{-1}}$ an $N\times N$ system matrix and with $\mathcal{Q}$ a $N\times 1$ column matrix of elements ρ_{li}. The notation selected for the system matrix, $\mathbf{C^{-1}}$, is not a coincidence. Indeed, the reader will notice that this matrix plays the role of the inverse of a capacitance matrix. The elements of $\mathbf{C^{-1}}$ are given by

$$\mathbf{C^{-1}_{ji}} = \lim_{\boldsymbol{\rho}\to\boldsymbol{\rho}_j}\int_{c_i} G(\boldsymbol{\rho}_j,\boldsymbol{\rho}')dS_i'. \tag{3.105}$$

These elements have to be calculated numerically. An, admittedly simple, approximation for the non-diagonal elements ($i \neq j$) is

$$\mathbf{C^{-1}_{ji}} = \Delta G(\boldsymbol{\rho}_j,\boldsymbol{\rho}_i) = -\frac{\Delta}{2\pi\epsilon}\ln\frac{|\boldsymbol{\rho}_j-\boldsymbol{\rho}_i|}{|\boldsymbol{\rho}_j-\boldsymbol{\rho}_{i,mir}|} \quad j\neq i, \tag{3.106}$$

with $\boldsymbol{\rho}_{i,mir}$ the mirror image of $\boldsymbol{\rho}_i$. The above result is obtained by approximating the integral as the product of the value of the integrand at the centre of

c_i multiplied by the length of the segment Δ. For the diagonal elements, known as *self-patch* contributions, the integration must be handled with due care. The Green's function $G(\boldsymbol{\rho}, \boldsymbol{\rho}')$ (3.99) consists of two parts: a first part contributed by $\boldsymbol{\rho}'$ and a second part by $\boldsymbol{\rho}''$. The contribution of $\boldsymbol{\rho}''$ poses no problem and can be treated in the same way as in the case of the non-diagonal elements. Consequently

$$\mathbf{C}_{\mathbf{jj}}^{-\mathbf{1}} = -\frac{1}{2\pi\epsilon}\tau + \frac{\Delta}{2\pi\epsilon}\ln|\boldsymbol{\rho}_j - \boldsymbol{\rho}_{j,mir}|, \tag{3.107}$$

with τ given by

$$\begin{aligned} \tau &= \int_{-\Delta/2}^{\Delta/2} \ln\sqrt{x^2}dx \\ &= 2\Delta(\ln\frac{\Delta}{2} - 1). \end{aligned} \tag{3.108}$$

The table below gives a set of numerical results for the capacitance C per unit of length for different W/d ratios. These results were obtained for $N = 600$. The value in the last column is the value for the ideal parallel-strip capacitor, i.e. $\epsilon_0 W/d$.

d/W	C (pF/m)	C_{ideal} (pF/m)
10	7.49 pF/m	0.89 pF/m
5	9.20 pF/m	1.77 pF/m
1	18.47 pF/m	8.85 pF/m
0.5	28.34 pF/m	17.71 pF/m
0.1	98.06 pF/m	88.54pF/m
0.05	176.65pF/m	177.08 pF/m

The above procedure to solve an integral equation is known as the *Method of Moments (MoM)*. More in particular, the simple technique discussed above is a MoM technique with piecewise constant basis functions combined with point matching. An advantageous feature of the technique as applied above is that the resulting system matrix is symmetrical thus strongly reducing the number of matrix elements that have to be calculated and allowing the use of powerful algorithms to solve symmetrical systems of linear equations.

Signal lines embedded in a layered dielectric

A problem of practical importance is that of determining the capacitive coupling between a set of signal lines embedded in a layered dielectric medium. This situation is e.g. encountered on printed circuit boards or on chip. An example of such a problem is shown on Fig. 3.25 . The dielectric medium consists of two layers of dielectric on top of a PEC and half-infinite free space (air). Without going into details, the procedure to determine the capacitance matrix (per unit of length) of these three signal lines is

1. Determine the Green's function $G(\boldsymbol{\rho}, \boldsymbol{\rho}')$ of the problem. For this the potential due to an elementary line charge located at $\boldsymbol{\rho}'$ and embedded in the dielectric medium must be calculated. The geometry of this Green's

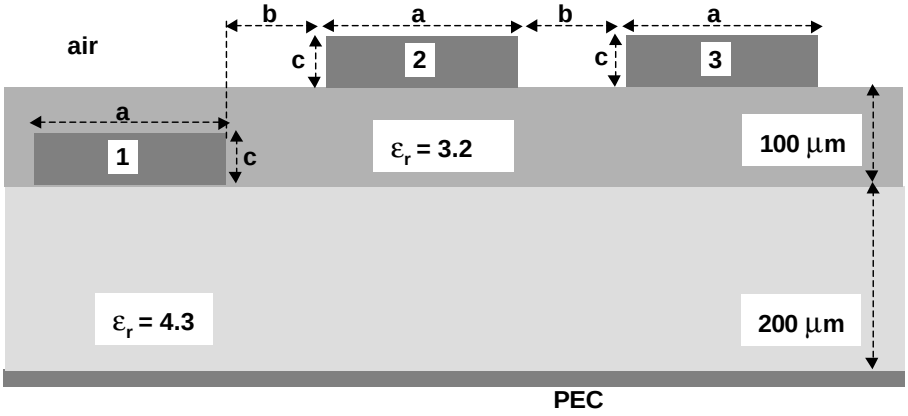


Figure 3.25: Cross section of an interconnection structure with three signal lines (a = $350\mu m$, b = $150\mu m$, c = $70\mu m$).

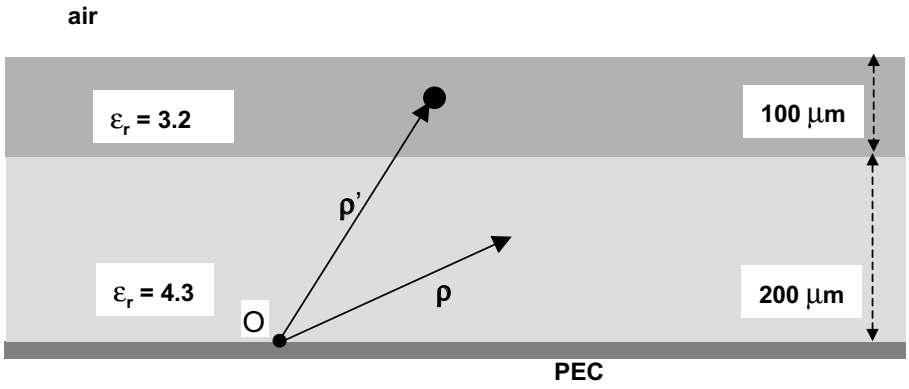


Figure 3.26: Green's function problem for the example of Fig. 3.25.

function problem is depicted in Fig. 3.26 . It turns out that this Green's function can no longer be determined analytically.

2. Use the Green's function to write down an integral equation such as (3.101) for the potential on each of the three signal conductors. Remember that in order to determine the complete capacitance matrix, three different potential problems will have to be solved with one of the signal conductors kept at potential 1 while to others are on zero potential.

3. Discretise the charges on the signal conductor circumferences, e.g. using a piecewise constant approximation.

4. Impose the integral equation in the midpoints of the discretisation intervals.

5. Solve the resulting system of linear equations.

6. Calculate the total charge (per unit of length) on each of the signal conductors. These charges are proportional to elements of the capacitance matrix.

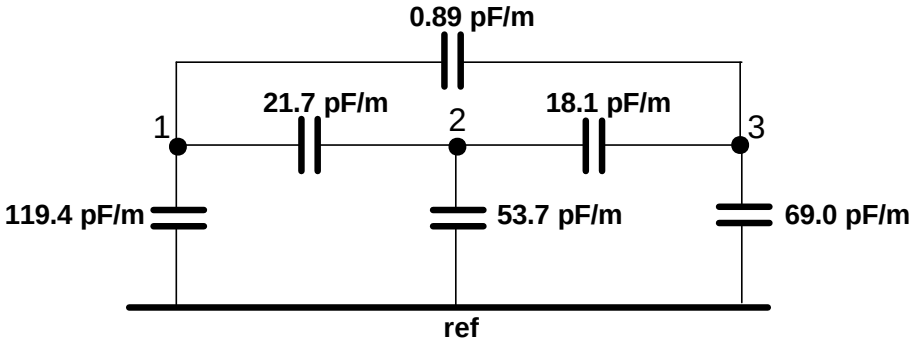


Figure 3.27: Equivalent circuit for the configuration of Fig. 3.25.

Fig. 3.27 shows the equivalent circuit of the configuration of Fig. 3.25 as calculated with the software tool *CapCad* developed by F. Olyslager of the Electromagnetics Group of the Department of Information Technology.

Chapter 4

Magnetostatics

4.1 Introduction

This chapter is devoted to Magnetostatics. For convenience we first repeat the basic equations of Magnetostatics derived in the previous chapter

$$\nabla \times \mathbf{h}(\mathbf{r}) = \mathbf{j}(\mathbf{r}), \tag{4.1}$$

$$\nabla \cdot \mathbf{b}(\mathbf{r}) = 0. \tag{4.2}$$

The chapter starts from Lorentz' force law to define the magnetic induction and to study the torque acting on a closed current loop. Next, Biot-Savart's law and the vector potential are introduced. The existence of magnetic dipoles is used to explain the difference between diamagnetic, paramagnetic and ferromagnetic materials. Finally, the notions of magnetic flux, self-inductance and mutual inductance are introduced and special attention is devoted to the calculation of the inductance matrix of a set of parallel signal lines, showing that the two-dimensional magnetostatic problem for these signal lines can be reformulated as an equivalent electrostatic problem of the type encountered in the last section of the previous chapter. Where possible, the chapter is organised such that is parallels the chapter on Electrostatics to make the reader aware of analogies and differences between electrostatic and magnetostatic fields.

4.2 Lorentz force and the magnetic induction

When considering static magnetic fields, the counterpart of the Coulomb force is the force acting on a charge q moving with velocity $\mathbf{v}$

$$\mathbf{F} = q(\mathbf{v} \times \mathbf{b}), \tag{4.3}$$

with $\mathbf{b}$ the magnetic induction. The direction of this force is perpendicular to both the velocity and the magnetic induction as depicted in Fig. 4.1 . The magnetic induction $\mathbf{b}(\mathbf{r})$ can now be *defined* as the force exerted on a unit positive charge placed at $\mathbf{r}$ and moving with velocity $\mathbf{v}$. If a charge is immersed in a magnetic field as well as in an electric field, the force acting on that charge becomes

$$\mathbf{F} = q(\mathbf{e} + \mathbf{v} \times \mathbf{b}). \tag{4.4}$$

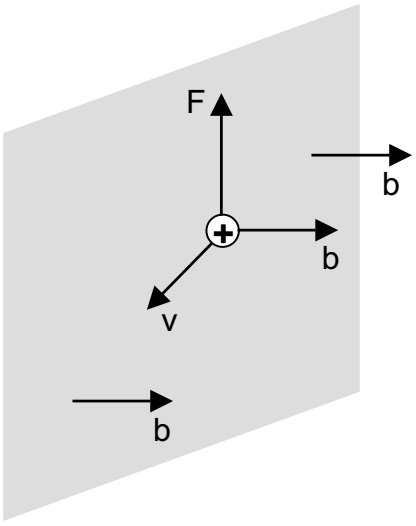


Figure 4.1: Magnetic force on a moving charge q.

This force is known as the Lorentz force and is the basic force law governing Electromagnetics. As a matter of fact, (4.4) remains valid in the dynamical case and for any velocity $\mathbf{v}$, as confirmed by accelerator experiments. The Lorentz force also plays a central role in the Special Theory of Relativity. Suppose that the velocity $\mathbf{v}$ is constant and that we switch to the co-ordinate reference frame of the moving charge. In that reference frame the charge is stationary. If we accept that in this moving reference frame Maxwell's equations still apply, as Einstein proposed in his Special Theory of Relativity (the so-called covariance principle), the charge will only feel the Coulomb force $\mathbf{F}' = q\mathbf{e}'$ with $\mathbf{F}'$ and $\mathbf{e}'$ the force and the electric field as measured in the moving reference frame. Taking into account the relativistic transformation formula for forces then shows that

$$\mathbf{e}' = \mathbf{e}_{\|} + \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}(\mathbf{e}_{\perp} + \mathbf{v} \times \mathbf{b}), \tag{4.5}$$

with $\mathbf{e}_{\|}$ the component of the electric field parallel to the velocity $\mathbf{v}$, with $\mathbf{e}_{\perp}$ the component of $\mathbf{e}$ perpendicular to the velocity and with c the velocity of light in vacuum. A similar transformation laws holds for the magnetic induction. Without going into further details, (4.5) makes clear that the distinction between an electric and a magnetic field is not as clear cut as one would be tempted to believe. The Theory of Relativity provides the appropriate unifying framework. Remember that the title of Einstein's first paper on the Theory of Relativity is "Zur Elektrodynamik Bewegten Körper", a clear indication of the central role played by Maxwell's equations in his theory.

As in last year's physics course much attention has already been paid to magnetic forces and torques, we will only briefly recapitulate the major results. In Section 3.8 it was shown that the elementary amount of energy du_e needed to move a charge immersed in an electric field over a distance $\mathbf{dc}$ is $-q\mathbf{e} \cdot \mathbf{dc}$.

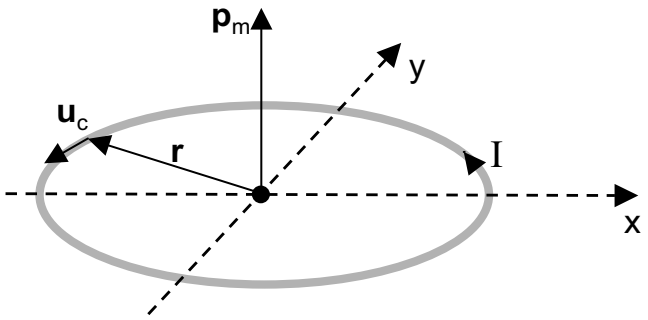


Figure 4.2: Torque acting on a current loop in a constant magnetic field.

In the sole presence of a magnetic field, this amount of energy becomes

$$du_m = -q(\mathbf{v} \times \mathbf{b}) \cdot \mathbf{dc} = -q(\mathbf{v} \times \mathbf{b}) \cdot \mathbf{v}dt = 0. \tag{4.6}$$

No work is done in moving the charge against the magnetic induction. This implies that the magnetic field cannot change the speed of the charge (as this would change its kinetic energy) but only the direction in which the charge moves.
The starting point for the calculation of forces and torques acting on a current-carrying wire is

$$\mathbf{dF} = I\mathbf{dc} \times \mathbf{b}(\mathbf{r}), \tag{4.7}$$

i.e. the elementary force acting on an elementary piece of wire **dc** carrying a current I. The reader will easily proof that the total force of a constant **b**-field on a closed current-carrying loop is zero. However, the loop will experience a magnetic torque that tends to rotate the current loop. Consider the current loop depicted in Fig. 4.2. The xy-plane coincides with the plane of the loop. Using (4.7) the torque of the force **dF** with respect to the origin is

$$\mathbf{dt} = \mathbf{r} \times (I\mathbf{dc} \times \mathbf{b}), \tag{4.8}$$

with **r** the position vector of **dc**. The total torque thus becomes

$$\mathbf{t} = I \int_c \mathbf{r} \times (\mathbf{dc} \times \mathbf{b}). \tag{4.9}$$

To further transform (4.9) note that

$$d[\mathbf{r} \times (\mathbf{r} \times \mathbf{b})] = \mathbf{dc} \times (\mathbf{r} \times \mathbf{b}) + \mathbf{r} \times (\mathbf{dc} \times \mathbf{b}). \tag{4.10}$$

As the integral over the current loop of the first term of (4.10) is zero, (4.9) can be rewritten as

$$\mathbf{t} = \frac{I}{2} \int_c [\mathbf{r} \times (\mathbf{dc} \times \mathbf{b}) - \mathbf{dc} \times (\mathbf{r} \times \mathbf{b})]. \tag{4.11}$$

We now invoke another vector identity. For three arbitrary vectors **a**, **b** and **c** we have that

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) - \mathbf{c} \times (\mathbf{b} \times \mathbf{a}) = \mathbf{b} \times (\mathbf{a} \times \mathbf{c}). \tag{4.12}$$

Using (4.12), (4.11) can finally be rewritten as

$$\mathbf{t} = \frac{I}{2} \int_c (\mathbf{r} \times \mathbf{dc}) \times \mathbf{b}. \tag{4.13}$$

Provided $\mathbf{b}$ is constant we now only need to evaluate the integral

$$\int_c (\mathbf{r} \times \mathbf{dc}). \tag{4.14}$$

We leave it to the reader to prove that the above integral yields $S/2\mathbf{u}_z$ with S the surface of the loop. Hence, the final result for the torque is

$$\mathbf{t} = SI\mathbf{u}_z \times \mathbf{b} = \mathbf{p}_m \times \mathbf{b}. \tag{4.15}$$

The vector $SI\mathbf{u}_z$ is the magnetic moment or magnetic dipole moment $\mathbf{p}_m$ already introduced in Chapter 2. Its amplitude is the product of the current I and the surface of the loop S, while its direction $\mathbf{u}_z$ is that of the unit normal to the loop surface with its orientation determined by the circulation sense of the current. Because the total force acting on the loop is zero, the torque is independent of its reference point.

4.3 The magnetic field and Biot-Savart's law

Equation (4.1) shows that the source current density or, for short, current is the source of the magnetic field $\mathbf{h}$. In the previous section only the magnetic induction $\mathbf{b}$ played a role and we need a constitutive equation to relate $\mathbf{b}$ to $\mathbf{h}$. Although we will briefly deal with the magnetic properties of materials in Section 4.5, it suffices to say that in this section and in this course we will only consider the simple linear relationship $\mathbf{b} = \mu\mathbf{h}$.
Building upon the experimental results of Hans Oersted, Jean Biot and Felix Savart were able to derive an expression for the magnetic field $\mathbf{dh}$ of an elementary steady-state current $I\mathbf{dc}$

$$\mathbf{dh} = \frac{1}{4\pi}\frac{I}{r^2}\mathbf{dc} \times \mathbf{u}_r, \tag{4.16}$$

where (see Fig. 4.3) r is the distance between the current filament and the field point P and with $\mathbf{u}_r$ the unit vector in the direction joining the location of the current filament with P. This formula is the magnetostatic equivalent of (3.6) expressing the electric field in terms of the charge q. By superposition this result is easily generalised to

$$\mathbf{h}(\mathbf{r}) = \frac{I}{4\pi} \int_c \frac{\mathbf{dc}' \times \mathbf{u}}{|\mathbf{r} - \mathbf{r}'|^2}\, dc', \tag{4.17}$$

$$\mathbf{h}(\mathbf{r}) = \frac{1}{4\pi} \int_S \frac{\mathbf{j}_s(\mathbf{r}') \times \mathbf{u}}{|\mathbf{r} - \mathbf{r}'|^2}\, dS', \tag{4.18}$$

$$\mathbf{h}(\mathbf{r}) = \frac{1}{4\pi} \int_V \frac{\mathbf{j}(\mathbf{r}') \times \mathbf{u}}{|\mathbf{r} - \mathbf{r}'|^2}\, dV'. \tag{4.19}$$

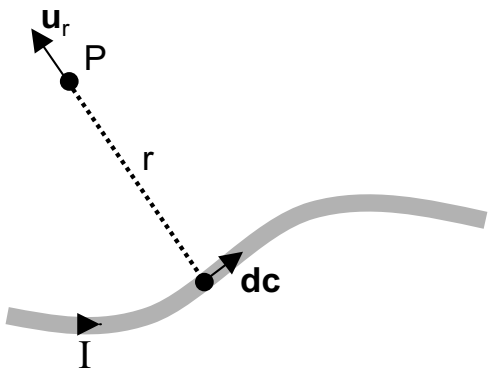


Figure 4.3: Magnetic field of an elementary current filament.

In the above expressions $\mathbf{u} = \frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}$ is the unit vector in the direction joining $\mathbf{r}'$ to $\mathbf{r}$. Eqn. (4.17) gives the magnetic field due to an arbitrary line current. The corresponding results for a surface current density and for a volume current density are given by (4.18) and (4.19).

As a first example, consider the magnetic field of an infinitely long and infinitely thin linear conductor carrying a current I, as depicted in Fig. 4.4 . Due to the symmetry of the problem, the magnetic field only depends on the distance r to the wire. Following (4.17) the magnetic field at a distance r from the wire is

$$\frac{I}{4\pi}\int_{-\infty}^{\infty}\frac{\mathbf{u}_z\times\mathbf{u}}{r^2+z^2}\,dz, \tag{4.20}$$

with

$$\mathbf{u} = \frac{-z\mathbf{u}_z + r\mathbf{u}_r}{\sqrt{r^2+z^2}}. \tag{4.21}$$

The result is

$$\mathbf{h}(r) = \frac{I}{2\pi r}\mathbf{u}_\phi, \tag{4.22}$$

with $\mathbf{u}_\phi = \mathbf{u}_z \times \mathbf{u}_r$ the unit vector in the azimuthal direction. Of course, this familiar result is more easily obtained by applying Ampère's law (2.16) repeated here for convenience,

$$\oint_c \mathbf{h}(\mathbf{r})\cdot\mathbf{dc} = \int_S \mathbf{j}(\mathbf{r})\cdot\mathbf{u}_n dS = i(\mathbf{r}). \tag{4.23}$$

To apply Ampère's law we select the surface S to be the surface of a circle with radius r centred on the wire. Equation (4.23) now becomes

$$\int_0^{2\pi} h_\phi(r) r d\phi = I, \tag{4.24}$$

immediately confirming (4.22).

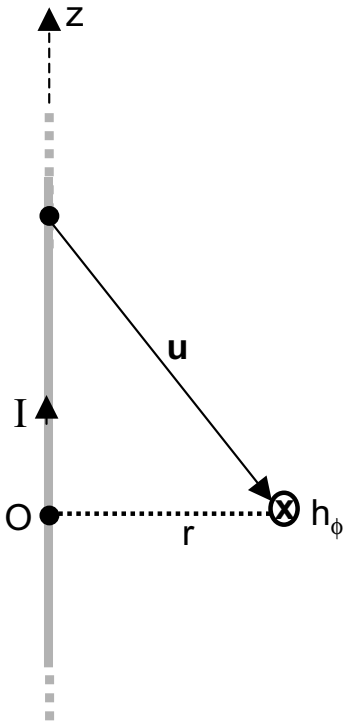


Figure 4.4: Magnetic field of a linear conductor.

Now consider the magnetic field of a current carrying infinitely long cylindrical conductor with radius a. The current is uniformly distributed over the cross-section of the cylinder. The easiest way to determine the magnetic field is by again applying Ampère's law. First a circle of radius $r < a$ is selected. We immediate find that

$$\mathbf{h}(\mathbf{r}) = h_\phi(r)\mathbf{u}_\phi = \frac{1}{2\pi r}\frac{Ir^2}{a^2}\mathbf{u}_\phi = \frac{Ir}{2\pi a^2}\mathbf{u}_\phi. \tag{4.25}$$

For a circle of radius $r > a$ we recuperate (4.22), i.e. outside the conductor the magnetic field is identical to that of a wire carrying the total current I. Outside the conductor the magnetic field decreases inversely proportional to the distance from the axis of the conductor. Inside the conductor the magnetic field increases linearly when going from the axis to the circumference.

4.4 The vector potential

In the previous chapter the electric field could be derived from a scalar potential ϕ due to the fact that the curl of a static electric field is zero. According to (4.1) this is not the case for the magnetic field. However, according to (4.2) the divergence of $\mathbf{b}$ is zero. Hence, under some very general conditions, $\mathbf{b}$ can be derived from the curl of another vector

$$\mathbf{b}(\mathbf{r}) = \nabla \times \mathbf{a}(\mathbf{r}), \tag{4.26}$$

with $\mathbf{a}(\mathbf{r})$ the vector potential already introduced in Section 2.6. Inserting (4.26) into (4.1) yields

$$\nabla \times \nabla \times \mathbf{a}(\mathbf{r}) = -\nabla^2 \mathbf{a}(\mathbf{r}) + \nabla\nabla \cdot \mathbf{a}(\mathbf{r}) = \mu \mathbf{j}(\mathbf{r}), \tag{4.27}$$

with μ a piecewise constant permeability. As in Section 2.6, it can again be argued that the vector potential is not unique. The available degree of freedom is used to enforce a gauge condition and we opt for the Lorenz gauge (2.92) with $\omega = 0$, i.e. $\nabla \cdot \mathbf{a} = 0$. Hence, in the static case, the vector potential satisfies

$$\nabla^2 \mathbf{a}(\mathbf{r}) = -\mu \mathbf{j}(\mathbf{r}), \tag{4.28}$$

which is the vector version of Poisson's equation (3.17). For infinite homogeneous space the solution of (4.28) can immediately be derived from (2.105)

$$\mathbf{a}(\mathbf{r}) = \frac{\mu}{4\pi} \int_V \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} dV'. \tag{4.29}$$

The corresponding magnetic field is found by taking the curl of (4.29)

$$\mathbf{h}(\mathbf{r}) = \nabla \times \frac{1}{4\pi} \int_V \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} dV'. \tag{4.30}$$

One must be careful when exchanging the curl operator with the integration, but for a volume density this is allowed. The result is

$$\mathbf{h}(\mathbf{r}) = \frac{1}{4\pi} \int_V \nabla \frac{1}{|\mathbf{r} - \mathbf{r}'|} \times \mathbf{j}(\mathbf{r}') dV', \tag{4.31}$$

as the curl operator only acts on the unprimed co-ordinates. Taking into account that

$$\nabla \frac{1}{|\mathbf{r} - \mathbf{r}'|} = -(\frac{1}{|\mathbf{r} - \mathbf{r}'|})^2 \mathbf{u}, \tag{4.32}$$

shows that (4.31) is identical to (4.19) derived from Biot-Savart's law.

4.5 Magnetic dipole - Magnetisation

Consider a circular current-carrying loop with radius a as depicted in Fig. 4.5. The current density in this loop can be represented as

$$\mathbf{j}(\mathbf{r}) = I\delta(r - a)\delta(z)\mathbf{u}_\phi. \tag{4.33}$$

The vector potential of this loop in a point P with cylindrical co-ordinates (r, ϕ, z) is

$$\mathbf{a}(r, \phi, z) = \frac{\mu}{4\pi} I \int_V \frac{\delta(r' - a)\delta(z')\mathbf{u}_{\phi'}}{|\mathbf{r} - \mathbf{r}'|} r' dr' d\phi' dz', \tag{4.34}$$

with

$$\mathbf{r} = r\cos\phi \mathbf{u}_x + r\sin\phi \mathbf{u}_y + z\mathbf{u}_z, \tag{4.35}$$

$$\mathbf{r}' = a\cos\phi' \mathbf{u}_x + a\sin\phi' \mathbf{u}_y. \tag{4.36}$$

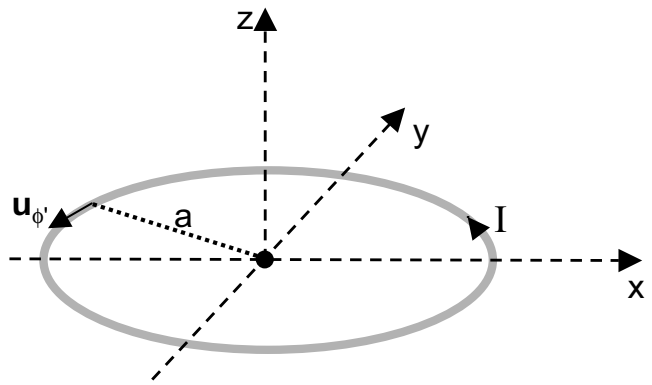


Figure 4.5: Circular current-carrying loop.

Substituting (4.35) and (4.36) into (4.34) yields

$$\mathbf{a}(r,\phi,z) = \frac{\mu}{4\pi} I \int_0^{2\pi} \frac{a(-\sin\phi'\mathbf{u}_x + \cos\phi'\mathbf{u}_y)}{\sqrt{a^2 + r^2 - 2ar\cos(\phi-\phi') + z^2}}\, d\phi'. \tag{4.37}$$

The general expression for this integral involves elliptical functions. We restrict ourselves to the case where the radius of the loop a becomes much smaller than the distance $R = \sqrt{r^2+z^2}$ from the origin of the loop to the field point, i.e. $a \ll R$. This approximation is analogues to the approximation $d \ll r$ when calculating the field of an electric dipole in Section 3.4. For $a \ll R$ the square root in the denominator of (4.37) can be approximated by

$$\frac{1}{\sqrt{a^2 + r^2 - 2ar\cos(\phi-\phi') + z^2}} \approx \frac{1}{R}\left(1 + \frac{ar}{R^2}\cos(\phi-\phi')\right) \tag{4.38}$$

and inserting this approximation into (4.37) finally leads to

$$\begin{aligned} \mathbf{a}(r,\phi,z) &= \frac{\mu}{4\pi} I \frac{\pi a^2 r}{R^3}(\mathbf{u}_z \times \mathbf{u}_r) \\ &= \frac{\mu}{4\pi R^2}(\mathbf{p}_m \times \mathbf{u}_R). \end{aligned} \tag{4.39}$$

In (4.39) $\mathbf{u}_R$ is the unit vector pointing in the direction from the origin of the loop to the field point and $\mathbf{p}_m = \pi a^2 I \mathbf{u}_z$ is the magnetic dipole moment of the circular current loop. The corresponding magnetic field is

$$\mathbf{h}(\mathbf{r}) = \frac{1}{\mu}\nabla \times \mathbf{a}(\mathbf{r}) = \frac{p_m}{4\pi R^2}(2\cos\theta\mathbf{u}_R + \sin\theta\mathbf{u}_\theta), \tag{4.40}$$

with R, θ, ϕ the spherical co-ordinates of the field point and with $\mathbf{u}_R, \mathbf{u}_\theta, \mathbf{u}_\phi$ the corresponding unit vectors. It is clear from the rotational symmetry of the problem that the final result (4.40) must be independent of ϕ. The calculations leading to (4.40) are quite tedious. The magnetic field of a magnetic dipole is very similar to that of a permanent magnet.
In Section 4.7 we will also need the magnetic field of a current loop on its axis.

We will now use the existence of magnetic dipoles to explain the magnetic properties of materials.

Magnetisation and the magnetic properties of a material result from the magnetic moments of the atoms. These moments stem from three causes:

1. atomic current loops generated by the orbital motion of the electrons around the nucleus,
2. the intrinsic magnetic moment of the spinning electrons,
3. current loops generated by the motion of the protons in the nucleus.

In most materials the contribution from the nucleus is negligible. The magnetic behaviour of a material depends on the crystalline structure of the material and leads to a three-fold classification: *diamagnetic*, *paramagnetic* and *ferromagnetic* materials. Diamagnetic materials have no permanent magnetic dipole moment, while paramagnetic and ferromagnetic materials do have a permanent magnetic dipole moment but behave very differently.
For dielectrics we introduced the electric polarisation field $\mathbf{p}_e$ which results from the polarisation of a material when embedded in an external electric field. Similarly, the magnetisation vector $\mathbf{m}$ is defined as the vector sum of the magnetic dipole moments of all atoms in a unit volume. The reader might expect that we use the notation $\mathbf{p}_m$ for this magnetisation vector, just as we used $\mathbf{p}_e$ for an individual electric dipole and for the electric polarisation field. However, we prefer the notation $\mathbf{m}$, used throughout literature. The magnetic induction corresponding to $\mathbf{m}$ is $\mu_0\mathbf{m}$ and in the presence of an external magnetic field $\mathbf{h}$ the total magnetic induction becomes

$$\mathbf{b} = \mu_0(\mathbf{h} + \mathbf{m}). \tag{4.41}$$

When a material is immersed in an incident magnetic field, the electron orbits are distorted resulting in the production of a magnetic moment. This is called diamagnetism. In some materials the atoms already have a magnetic moment in the absence of an external magnetic field. This is called paramagnetism. In both cases the relationship between the magnetisation and the external magnetic field can be expressed as

$$\mathbf{m} = \chi_m\mathbf{h}, \tag{4.42}$$

with χ_m the magnetic susceptibility. For diamagnetic and paramagnetic materials χ_m is constant for a fixed temperature and hence the relationship between $\mathbf{b}$ and $\mathbf{h}$ is linear, i.e.

$$\begin{aligned} \mathbf{b} &= \mu_0(\mathbf{h} + \mathbf{m}) = \mu_0(1 + \chi_m)\mathbf{h} \\ &= \mu\mathbf{h}, \end{aligned} \tag{4.43}$$

with $\mu = \mu_0(1 + \chi_m) = \mu_0\mu_r$ the magnetic permeability and μ_r the relative magnetic permeability. Typical diamagnetic materials are copper, silver, gold, lead or diamond with very small χ_m values of the order of -10^{-5} such that $\mu_r \approx 1$. The negative value of χ_m indicates that the induced magnetic field points in the opposite direction of the external field. Aluminium, magnesium and chromium are typical paramagnetic materials. For these materials χ_m is

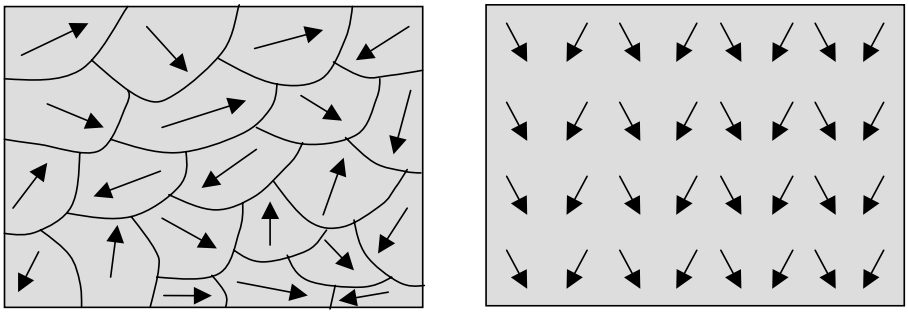


Figure 4.6: Ferromagnetic material with unmagnetised domains (a) and with magnetised domains (b).

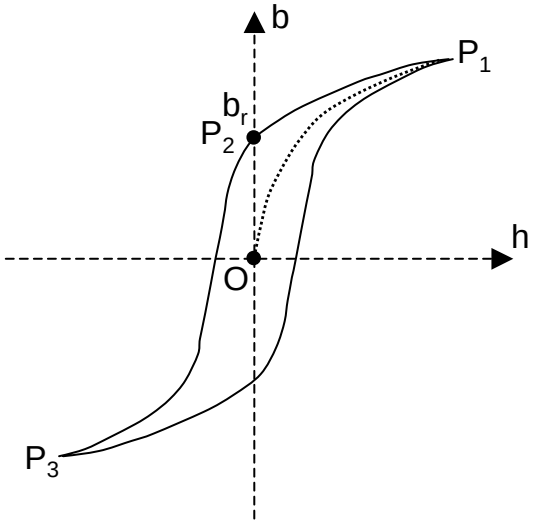


Figure 4.7: Hysteresis curve for a ferromagnetic material.

positive but still very small (also of the order of 10^{-5}).

A totally different behaviour is found in ferromagnetic materials such as iron, nickel or cobalt. These materials can be described in terms of magnetic domains of the order of $10^{-10}m^3$ within which the magnetic moments of the atoms point in the same direction. In the absence of an external magnetic field, the orientation of the domain magnetisation is random resulting in a zero net magnetisation. These domains and their walls are depicted in Fig. 4.6a . To qualitatively understand the behaviour of a ferromagnetic material, consider the unmagnetised situation which corresponds to the origin of the b, h-plane shown in Fig. 4.7 . When increasing $\mathbf{h}$, the magnetic induction $\mathbf{b}$ also increases due to the fact that more and more domains align themselves with the magnetic field (Fig. 4.6b). This phenomenon continues until saturation occurs, corresponding to point P_1 on Fig. 4.7. Further increasing $\mathbf{h}$ no longer results in an increase

in **b**. The material is now magnetised and the curve OP_1 is the magnetisation curve of the material. When decreasing the magnetic field, the resulting b, h-curve does *not* coincide with the magnetisation curve OP_1 but a different curve, P_1P_2, is traced. Even in the absence of a magnetic field a residual magnetic induction b_r remains, i.e the material now behaves as a permanent magnet due to the fact that a substantial fraction of the domains still remain aligned. Above a certain temperature, the *Curie temperature*, domain alignment disappears and the ferromagnetic material becomes paramagnetic. For iron the Curie temperature is about 1043^0K. Let's return to our hysteresis curve. When applying a magnetic field opposite to the one used when magnetising the material and for increasing amplitudes of this field, the magnetic induction decreases and the domains more and more align themselves with the magnetic field until a new saturation point P_3 is reached. Again reducing the magnetic field to zero, reversing its direction and saturating the material yields the P_3P_1 part of the curve. In this way we obtain the typical *hysteresis curve* of a ferromagnetic material. The term hysteresis means "to lag behind", i.e. the behaviour of the material does not only depend upon the present value of the magnetic field but also on the history of the magnetisation. A ferromagnetic material is clearly non-linear with large μ_r-values, e.g. 600 for nickel, $4000 - 5000$ for pure iron and even up to 10^5 for special so-called mumetals.

To conclude this discussion of ferromagnetic materials let us mention that hard ferromagnetic materials have a wide hysteresis loop with a large residual magnetisation while soft magnetic materials have a narrow hysteresis loop with a small residual magnetisation. Soft magnetic materials can much more easily be demagnetised than hard ones. To demagnetise a ferromagnetic material it must be subjected to a number of hysteresis cycles of decreasing amplitude such that the material gradually returns to its demagnetised state O.

4.6 Boundary conditions

We will be very brief here. It suffices to say that the boundary conditions in the static case are identical to those in the dynamic case (2.65) and (2.67). The normal component of the magnetic induction remains continuous at the interface between two media, i.e.

$$\mathbf{u}_n \cdot (\mathbf{b}_2 - \mathbf{b}_1) = 0, \tag{4.44}$$

with $\mathbf{u}_n$ the unit normal pointing in the direction of medium 2. The general boundary condition for the magnetic field is

$$\mathbf{u}_n \times (\mathbf{h}_2 - \mathbf{h}_1) = \mathbf{j}_s, \tag{4.45}$$

with $\mathbf{j}_s$ the surface current at the interface between medium 1 and medium 2. In the absence of such a surface current, the tangential component of the magnetic field remains continuous.
Contrary to the electrostatic case, it is not customary nor straightforward to express the above boundary conditions in terms of the vector potential **a**.

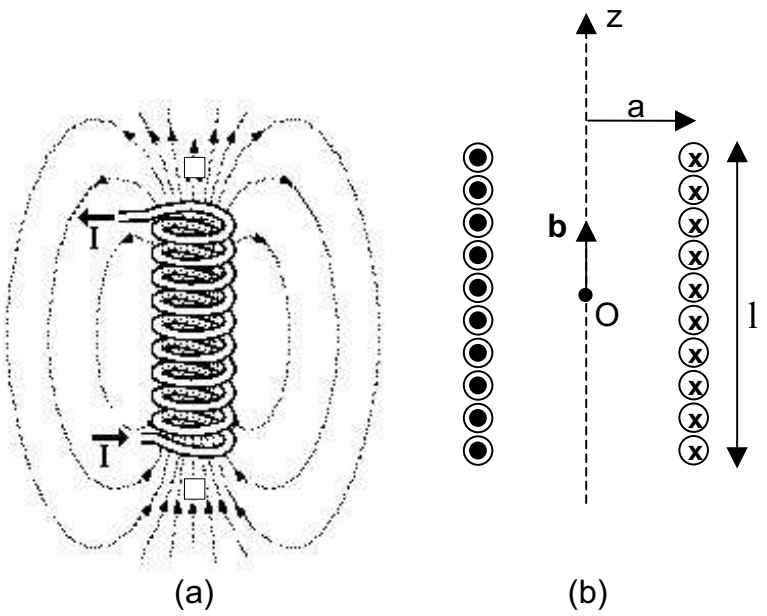


Figure 4.8: Magnetic field of a solenoid: artist impression of a solenoid (a) and geometry of the problem (b).

4.7 Inductance - Inductance matrix

In the previous chapter two circuit components, the resistor and the capacitor, were introduced from a field point of view. The total Joule losses $\int_V \mathbf{j} \cdot \mathbf{e}\, dV$ in a piece of conductor correspond to RI^2 with I the current flowing through the conductor and with R its resistance. The total amount of electrical energy stored in a capacitor formed by two arbitrary conductors is $\frac{1}{2}CV^2$ with V the voltage difference between the conductors and C their capacitance. This stored energy can also be expressed as the total electrostatic energy $\frac{1}{2}\int_V \mathbf{d}\cdot\mathbf{e}\, dV$. In this section, a third circuit element, the inductor is introduced. Consider a current-carrying loop c. The current loop creates a magnetic field and, in the absence of ferromagnetic materials, the flux ψ of the magnetic induction through the loop depends linearly on the loop current I

$$\psi = \int_S \mathbf{b} \cdot \mathbf{u}_n\, dS = LI. \tag{4.46}$$

The proportionality factor L is the *self-inductance* of the loop. Only for a few idealised configurations the self-inductance can be determined analytically. One such example is the solenoid. Fig. 4.8a shows a tightly wound solenoid with a total of N turns ($N = 10$ in the figure), while Fig. 4.8b shows the geometry of the problem. To simplify the analysis we suppose that the length l of the solenoid is much larger that its radius a. Moreover, we will concentrate on the magnetic field inside the solenoid and more particularly on the field on its axis. First, we need the field on the axis of the current loop depicted in Fig. 4.5. The

simplest way to obtain this field is to use (4.17)

$$\mathbf{h}(z) = \frac{I}{4\pi} \int_c \frac{a\mathbf{u}_\phi \times (-a\mathbf{u}_r + z\mathbf{u}_z)}{(a^2+z^2)^{\frac{3}{2}}}\, d\phi, \tag{4.47}$$

leading to

$$\mathbf{h}(z) = \frac{Ia^2}{2(a^2+z^2)^{\frac{3}{2}}}\mathbf{u}_z. \tag{4.48}$$

We can now think about the solenoid as the combination of elementary current loops each carrying a current $dI = \frac{NI}{l}dz$. Using (4.48), the magnetic field on the axis of the solenoid becomes

$$\mathbf{h}(z) = \int_{-\frac{l}{2}}^{\frac{l}{2}} \frac{NIa^2}{2l(a^2+z^2)^{\frac{3}{2}}}\, dz\ \mathbf{u}_z, \tag{4.49}$$

which is easily integrated to yield

$$\mathbf{h}(z) = \frac{NI}{2} \frac{1}{\sqrt{a^2+(\frac{l}{2})^2}}\mathbf{u}_z. \tag{4.50}$$

As the above calculation only holds for the field on the axis, (4.50) will only be a reasonable approximation for the field inside the solenoid provided $a \ll l$. Consequently, inside a tightly wound solenoid which is much longer than its radius and except near the edges of the solenoid, the magnetic field is constant and directed only the axis, i.e. $\mathbf{h}(\mathbf{r}) = \frac{NI}{l}\mathbf{u}_z$. Hence the following result is obtained for the self-inductance of the idealised solenoid or coil

$$L_{ideal} = \frac{\mu\pi a^2 N^2}{l}. \tag{4.51}$$

A particular class of problems that, just as in the electrostatic case, deserves our attention is the class of two-dimensional problems in which the currents only flow in a preferred direction, say the z-direction, but are themselves independent of z. A typical example is a pair of parallel wires separated by an axis to axis distance d and carrying opposite currents I as shown in Fig. 4.9. The radius of the wires is a. The area between the conductors is now infinite and the self-inductance can only be defined per unit of length in the z-direction as the ratio of the flux of the magnetic induction through the shaded unit length surface shown in Fig. 4.9a and the current I. This current is still flowing in a loop but this loops now extends to infinity. Applying (4.22) immediately shows that the self-inductance in H/m of the wire pair is

$$L_{wirepair} = \frac{\mu}{\pi} \ln \frac{d-a}{a}. \tag{4.52}$$

Two remarks should be made at this point. The above calculation of the self-inductance of a wire pair is correct provided the wires are perfectly conducting. If this is not the case, the definition of the surface enclosed by the wires is no longer unambiguous. We could e.g. select the surface defined by the axes of the

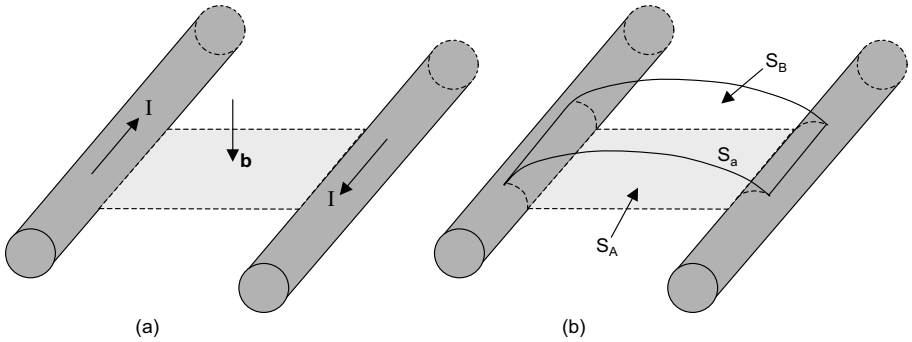


Figure 4.9: Self-inductance of a pair of parallel wires: flat flux surface (a), arbitrary flux surface (b).

Figure 4.10: Self-inductance of a coaxial cable.

wires. In that case an additional amount of flux is picked up by that part of the surface situated inside the conductors. The corresponding increase of the self-inductance is the so-called internal inductance. To keep the analysis sufficiently simple, we will restrict ourselves to perfectly conducting wires such that the internal inductance does not play a role. Even for perfect conductors one might question the uniqueness of self-inductance definition. Suppose that we do not consider the flat shaded surface of Fig. 4.9a but an arbitrary surface S_a as depicted in Fig. 4.9b. Will this surface still yield the same flux (4.52) obtained for the flat surface? To answer this question, consider the closed surface formed by the flat surface of Fig. 4.9a, the arbitrary surface S_a, part of the conductor surfaces and the two end surfaces S_A and S_B (see Fig. 4.9b). The total flux of $\mathbf{b}$ through this closed surface must be zero. As $\mathbf{b}$ is tangential to the conductor surfaces and to S_A and S_B, these surfaces do no contribute to the total flux. Hence, the flux through the flat surface of Fig. 4.9a and through the arbitrary surface of Fig. 4.9b must be equal, showing that the self-inductance is unique.

Another two-dimensional example is that of the coaxial cable depicted in Fig. 4.10. The magnetic field is the same as that of a wire: $h_\phi = \frac{I}{2\pi r}$ and the self-inductance per unit of length becomes

$$L_{coax} = \frac{\mu}{2\pi} \ln \frac{b}{a}. \tag{4.53}$$

The self-inductance concept can be extended to M current loops. Let us start with two such loops as shown in Fig. 4.11. The flux through each of the

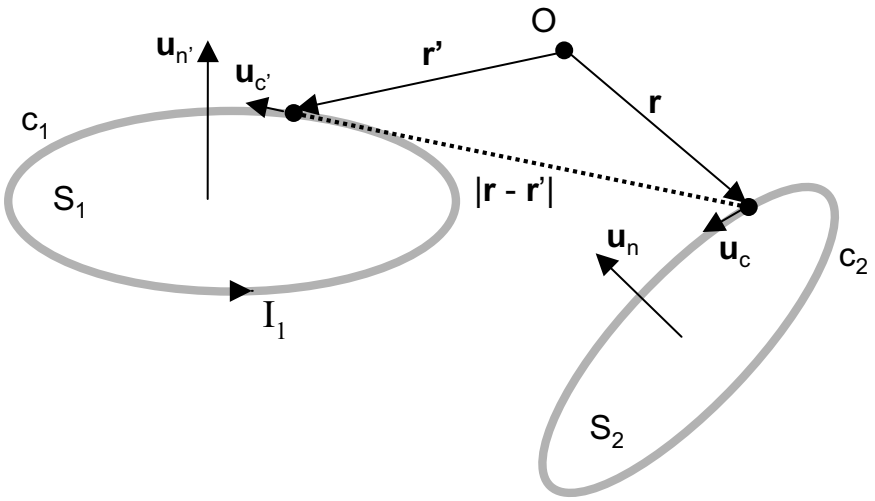


Figure 4.11: Mutual inductance between two current loops.

loops now depends linearly on both currents I_1 and I_2

$$\begin{aligned} \psi_1 &= L_{11} I_1 + L_{12} I_2, \\ \psi_2 &= L_{21} I_1 + L_{22} I_2. \end{aligned} \tag{4.54}$$

Let us determine L_{21}. With $I_2 = 0$ in (4.54) we see that L_{21} is the flux of the magnetic induction through the surface S_2 of loop 2 generated by a unit current flowing through loop 1. For simplicity we suppose that the conductors forming both loops are infinitely thin. In that case we have that

$$\psi_2 = \int_{S_2} \mathbf{b}(\mathbf{r}) \cdot \mathbf{u}_n \, dS. \tag{4.55}$$

The magnetic field can be derived from the vector potential $\mathbf{a}(\mathbf{r})$, hence (4.55) can be rewritten as

$$\begin{aligned} \psi_2 &= \int_{S_2} \nabla \times \mathbf{a}(\mathbf{r}) \cdot \mathbf{u}_n \, dS \\ &= \int_{c_2} \mathbf{a}(\mathbf{r}) \cdot \mathbf{dc}, \end{aligned} \tag{4.56}$$

where we used Stoke's theorem. For homogeneous space, the vector potential follows from (4.29), leading to the following final expression for $L_{21} = \psi_2$

$$L_{21} = \frac{\mu}{4\pi} \int_{c_1} \int_{c_2} \frac{\mathbf{dc}' \cdot \mathbf{dc}}{|\mathbf{r} - \mathbf{r}'|}, \tag{4.57}$$

with $\mathbf{r}$ and $\mathbf{r}'$ the place vectors of running integration points on c_2 respectively c_1. One immediately remarks that expression (4.57) is completely symmetric with respect to the role played by both current loops, proving that L_{21} is identical

to L_{12}. Although no proof will be given here, this conclusion holds in the presence of arbitrary linear magnetic materials. $L_{12} = L_{21}$ is called the mutual inductance of the two loops. Further remark that (4.57) solely depends upon the geometry of the loops and on their relative position in space. The matrix

$$\mathbf{L} = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}, \tag{4.58}$$

is the inductance matrix. The above reasoning can be extended to the general case of M current loops. Expression (4.54) then becomes

$$\mathcal{F} = \mathbf{L}\,\mathcal{I}, \tag{4.59}$$

with $\mathcal{I}$ and $\mathcal{F}$, $M \times 1$ column vectors with elements I_m, and ψ_m and with $\mathbf{L}$ the $M \times M$ inductance matrix. I_m is the current through loop m and ψ_m is the flux through that loop. The inductance matrix $\mathbf{L}$ is symmetric.

4.8 Magnetostatic energy

To build up the flux ψ through a current loop, energy must be spent. If we consider a perfectly conducting loop connected to a voltage source, this energy can be calculated by taking into account the relationship between the voltage over the terminals of the loop and the current flowing through the loop

$$v(t) = L\frac{di(t)}{dt}. \tag{4.60}$$

The elementary energy needed to increase the current in the loop by an amount di is

$$du_m = vdi = Ldi\frac{di}{dt} = \frac{L}{2}\frac{di^2}{dt}. \tag{4.61}$$

Integrating this between $t = 0$ when $i = 0$ and $t = \tau$ when i reaches its final value I, yields

$$u_m = \frac{1}{2}LI^2. \tag{4.62}$$

This is the magnetic energy stored in the inductor, as we know from circuit theory.

As in the electrostatic case, we want to relate this result to the fields. To this end consider the *magnetostatic energy density* $\frac{1}{2}\mathbf{b}\cdot\mathbf{h}$. We integrate this density over a large homogeneous volume V encompassing the PEC conductor forming the current loop. This volume is bounded by its outer surface S_{out} and by the surface S_{cond} of the current loop (see Fig. 4.12). For the time being we assume that the current loop has a finite cross-section. The integration gives

$$u_m = \int_V \frac{1}{2}\mathbf{h}\cdot\mathbf{b}\;dV = \int_V \frac{1}{2\mu}(\nabla\times\mathbf{a})\cdot(\nabla\times\mathbf{a})\;dV. \tag{4.63}$$

To further transform (4.63) the following Green's theorem for two arbitrary vector functions $\mathbf{f}$ and $\mathbf{g}$ is applied

$$\int_V \left((\nabla\times\mathbf{f})\cdot(\nabla\times\mathbf{g}) - \mathbf{f}\cdot\nabla\times\nabla\times\mathbf{g}\right)\;dV = \int_S (\mathbf{f}\times(\nabla\times\mathbf{g}))\cdot\mathbf{u}_n\;dS, \tag{4.64}$$

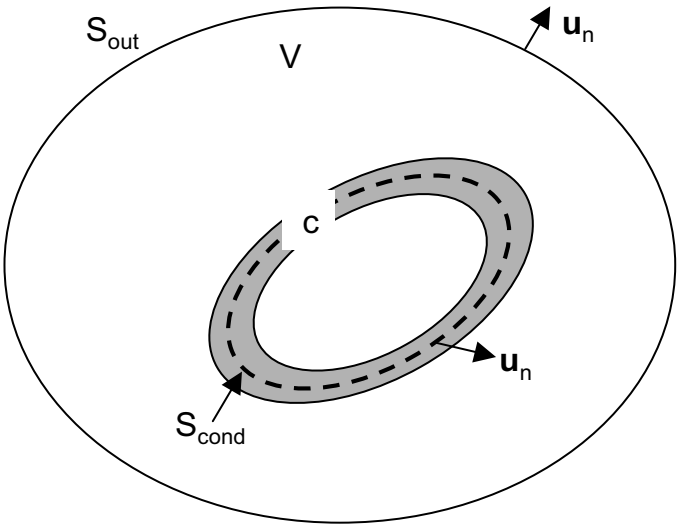


Figure 4.12: Integration volume for the magnetostatic energy.

with S the boundary surface of the volume. Applying (4.64) to (4.63) allows to rewrite u_m as

$$u_m = \frac{1}{2\mu}\int_V \mathbf{a}\cdot\nabla\times\nabla\times\mathbf{a}\;dS + \frac{1}{2\mu}\int_S (\mathbf{a}\times(\nabla\times\mathbf{a}))\cdot\mathbf{u}_n\;dS, \tag{4.65}$$

with $S = S_{out}\bigcup S_{cond}$. Outside the conductor $\nabla\times\nabla\times\mathbf{a} = 0$. Furthermore, the boundary condition for the tangential magnetic field (2.73) at the PEC surface S_{cond} states that

$$\mathbf{u}_n\times\mathbf{h} = -\frac{1}{\mu}\mathbf{u}_n\times(\nabla\times\mathbf{a}) = -\mathbf{j}_s, \tag{4.66}$$

with $\mathbf{j}_s$ the surface current and with the minus sign due to the fact that the normal $\mathbf{u}_n$ points inwards. Substituting (4.66) into (4.65) yields

$$u_m = \frac{1}{2}\int_{S_{cond}} \mathbf{a}\cdot\mathbf{j}_s\;dS \tag{4.67}$$

and where, as in the electrostatic case, the contributions of S_{out} becomes negligible when extending the integration volume to infinity. To further simplify the analysis we now suppose that the current loop becomes very thin i.e. its cross-sectional dimensions are much smaller than the diameter of the loop. This implies that the vector potential $\mathbf{a}$ will approximately remain constant over the cross-section allowing to rewrite (4.67) as

$$u_m = \frac{I}{2}\int_c \mathbf{a}\cdot\mathbf{dc}, \tag{4.68}$$

with the integration now over the current loop approximated by a thin wire loop c. To arrive at the final result Stoke's theorem can be invoked to transform the

integral over c to the corresponding integral over the surface enclosed by c

$$u_m = \frac{I}{2} \int\limits_{S_{loop}} \mathbf{u}_n \cdot \nabla \times \mathbf{a}\, dS, \tag{4.69}$$

with $\mathbf{u}_n$ now representing the normal to the loop surface. With $\nabla \times \mathbf{a} = \mathbf{b}$, (4.69) becomes

$$u_m = \frac{\psi I}{2} = \frac{1}{2} L I^2, \tag{4.70}$$

proving that $\frac{1}{2}\mathbf{b}\cdot\mathbf{h}$ represents the magnetic energy density. We also encountered this quantity in the dynamical case when discussing conservation of energy in relationship to Maxwell's equation (Section 2.3).
Formula (4.70) can be extended to the general case of M loops. Prove that u_m becomes

$$u_m = \frac{1}{2} \mathcal{I}^T \mathbf{L}\, \mathcal{I}, \tag{4.71}$$

with $\mathcal{I}^T$ the transposed of the current vector $\mathcal{I}$.

4.9 Two-dimensional signal lines

Solution of the general magnetostatic problem (4.28) is not at all straightforward and is certainly outside the scope of this introductory course. This section focusses on the important special case of the parallel signal lines already discussed in Section 3.9. In this two-dimensional configuration, all the currents are z-directed. Fig. 4.13 shows an example of such a set of perfectly conducting lines embedded in free space. We will now show that in the two-dimensional case the solution of the magnetostatic problem can be reformulated as an equivalent electrostatic problem such that the techniques developed in Section 3.9 can then be used to determine the inductance matrix.
Poisson's equation (4.28) for the vector potential in a piecewise homogeneous medium shows that for z-directed currents the vector potential will only have a z-component, i.e. $\mathbf{a}(\mathbf{r}) = \psi(\boldsymbol{\rho})\mathbf{u}_z$. As in Section 3.9, $\boldsymbol{\rho}$ is the position vector of a point in the xy-plane. The reason for selecting the notation ψ for the z-component of $\mathbf{a}$ will become clear in the sequel. The magnetic field is the curl of the vector potential, hence

$$\mathbf{h}_t(\boldsymbol{\rho}) = \frac{1}{\mu}\nabla \times (\psi(\boldsymbol{\rho})\mathbf{u}_z) = \frac{1}{\mu}\nabla_t \psi(\boldsymbol{\rho}) \times \mathbf{u}_z. \tag{4.72}$$

The magnetic field has no z-component. This is indicated by the subindex "t", standing for "transversal". The gradient also only operates on the x and y coordinates as again indicated by its subscript "t". As the currents only flow on the conductor surfaces, the curl of the magnetic field must be zero outside the conductors and hence

$$\begin{aligned} \nabla_t \times \mathbf{h}_t(\boldsymbol{\rho}) &= \frac{1}{\mu}\nabla_t \times (\nabla_t \psi \times \mathbf{u}_z) \\ &= -\frac{1}{\mu}\nabla_t^2 \psi(\boldsymbol{\rho}) = 0. \end{aligned} \tag{4.73}$$

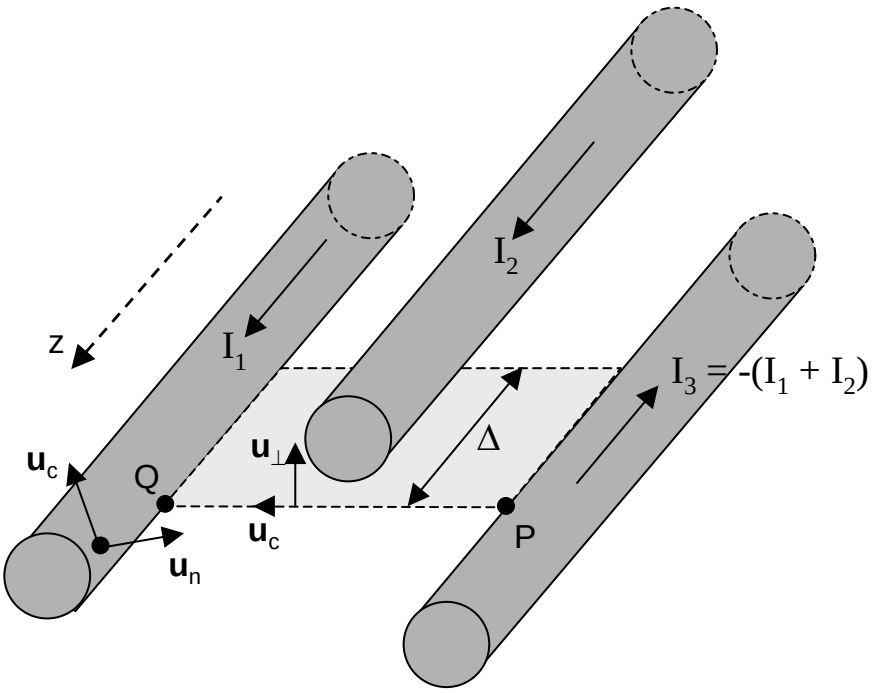


Figure 4.13: Inductance calculations for a set of parallel signal lines.

The normal component of the magnetic induction vanishes at a PEC surface. From (4.72) this implies that

$$\begin{aligned} \mathbf{u}_n \cdot (\nabla_t \psi \times \mathbf{u}_z) &= \nabla_t \psi \cdot (\mathbf{u}_z \times \mathbf{u}_n) \\ &= \nabla_t \psi \cdot \mathbf{u}_c \\ &= \frac{\partial \psi}{\partial c} = 0. \end{aligned} \tag{4.74}$$

In (4.74) $\mathbf{u}_c$ is the unit tangent vector to the circumference of the conductor (see Fig. 4.13) such that $\mathbf{u}_n \times \mathbf{u}_c = \mathbf{u}_z$. The above result shows that the tangential derivative of ψ is zero at a conductor surface. Hence ψ must take a constant value on a conductor. From the above we conclude that ψ satisfies the same differential problem as the electrostatic potential ϕ

$$\nabla_t^2 \psi = 0, \tag{4.75}$$

$$\psi = \text{constant on each conductor surface.} \tag{4.76}$$

Let us take a closer look at the physical meaning of ψ. To this end consider the flux of the magnetic induction through the rectangular surface depicted in Fig. 4.13. As the magnetic induction is independent of z, this flux becomes

$$F = \Delta \int_P^Q \mathbf{b} \cdot \mathbf{u}_\perp \, dc, \tag{4.77}$$

with the unit normal $\mathbf{u}_\perp$ perpendicular to the integration path PQ and with $\mathbf{u}_\perp = \mathbf{u}_c \times \mathbf{u}_z$. Using (4.72), (4.77) can be transformed as

$$F = \Delta \int_P^Q (\nabla_t \psi \times \mathbf{u}_z) \cdot \mathbf{u}_\perp \, dc = \Delta \left(\psi(Q) - \psi(P)\right). \tag{4.78}$$

This shows that the flux F between two conductors, per unit of length in the z-direction, can be expressed as the difference between the constant values ψ takes on these conductors. This implies that ψ truly plays the role of a potential: the magnetic potential or the magnetic flux function (hence its notation ψ). As only potential differences play a role, we will again select one of the conductors as the zero magnetic potential reference conductor.
As ψ is a potential, ψ remains continuous at the interface between two media. Furthermore, the continuity of the tangential magnetic field at the interface between two media implies that $\frac{1}{\mu}\mathbf{u}_n \times (\nabla_t \psi \times \mathbf{u}_z)$ must be continuous or

$$\psi_1 = \psi_2, \tag{4.79}$$

$$\frac{1}{\mu_1}\left(\frac{\partial \psi}{\partial n}\right)_1 = \frac{1}{\mu_2}\left(\frac{\partial \psi}{\partial n}\right)_2. \tag{4.80}$$

Comparing this with (3.36) and (3.37) finally leads to the conclusion that the solution of the inductance problem can be found by solving a capacitance problem whereby ψ plays the role of potential and with the role of ϵ taken by $1/\mu$. If we have $M+1$ conductors, we select one of them as the zero magnetic potential reference conductor. The others are numbered from 1 to M and the constant value of ψ on each of these conductors is denoted as F_m. In the electrostatic case, solution of the relevant Laplace problem leads to (3.78)

$$\mathcal{Q} = \mathbf{C}\, \mathcal{V}. \tag{4.81}$$

The role of $\mathcal{V}$ is now taken by $\mathcal{F}$, with $\mathcal{F}$ a $M \times 1$ column matrix of elements F_m. The vector $\mathcal{Q}$ is the vector of the total charges q_m (per unit of length) on each conductor. The total charge q_m on conductor m is defined as

$$q_m = \oint_{c_m} \mathbf{d} \cdot \mathbf{u}_n \, dc = -\oint_{c_m} \epsilon \nabla_t \phi \cdot \mathbf{u}_n \, dc. \tag{4.82}$$

In the magnetostatic case ϕ is replaced by ψ and ϵ by $1/\mu$ showing that the magnetostatic equivalent of q_m becomes

$$\begin{aligned} q_m &= -\oint_{c_m} \frac{1}{\mu} \nabla_t \psi \cdot \mathbf{u}_n \, dc \\ &= \oint_{c_m} \mathbf{h} \cdot \mathbf{u}_c \, dc \\ &= I_m, \end{aligned} \tag{4.83}$$

with I_m the current flowing through conductor m. From the above we conclude that solving the relevant Laplace problem for the magnetostatic case and taking into account the correct relationships between the magnetic quantities and their electrostatic counterparts, leads to

$$\mathcal{I} = \mathbf{C_m}\, \mathcal{F}, \tag{4.84}$$

with $\mathcal{I}$ and $\mathcal{F}$, $M \times 1$ column vectors with elements I_m and F_m and with $\mathbf{C_m}$ the $M \times M$ equivalent capacitance matrix of the magnetostatic problem. As the inductance matrix follows from

$$\mathcal{F} = \mathbf{L}\,\mathcal{I}, \tag{4.85}$$

we conclude that

$$\mathbf{L} = {\mathbf{C_m}}^{-1}, \tag{4.86}$$

i.e. the inductance matrix is the inverse of an equivalent capacitance matrix obtained by solving an equivalent electrostatic problem obtained by replacing the ϵ-values of the materials by corresponding $1/\mu$-values. Just as in the electrostatic case the total charge on all conductors, including the reference conductor, must be zero, in the magnetostatic case the sum of all currents must be zero. This can also be reformulated by saying that the return current flowing through the reference conductor must be equal but opposite to the sum of the currents flowing through the M signal lines.

In the two-dimensional case it can be proved that the following general properties hold

- the inductance matrix is symmetric, $L_{ij} = L_{ji}$;
- all the non-diagonal elements are strictly positive, $L_{ij} > 0$;
- all the diagonal elements are strictly positive, $L_{ii} > 0$;
- the matrix $\mathbf{L}$ is positive definite and non-singular.

It should be remarked that three-dimensional magnetic problems cannot be transformed into equivalent electrostatic problems.

For a completely homogeneous two-dimensional medium, the equivalent potential problem consists in replacing the constant value of ϵ by that of $1/\mu$. As for a homogeneous medium, the charge density is directly proportional to ϵ and hence to $1/\mu$ in the magnetostatic case, a simple relationship is obtained between $\mathbf{L}$ and $\mathbf{C}$

$$\mathbf{L} \cdot \mathbf{C} = \mathbf{C} \cdot \mathbf{L} = \epsilon\mu\mathbf{1}, \tag{4.87}$$

with $\mathbf{1}$ the $M \times M$ unit matrix. That this relationship also holds in the scalar case can easily be verified by the reader for the inductance and capacitance per unit length of the coaxial cable derived in this and in the previous chapter

$$C_{coax} = \frac{2\pi\epsilon}{\ln \frac{b}{a}}, \tag{4.88}$$

$$L_{coax} = \frac{\mu}{2\pi} \ln \frac{b}{a}, \tag{4.89}$$

the product of which is clearly $\epsilon\mu$.

To conclude this section and this chapter we return to two of the examples treated in the previous chapter: the coaxial line with two inner conductors of Fig. 3.20 and the interconnection structure with three signal lines of Fig. 3.25. In both examples the material is non-magnetic, i.e. $\mu = \mu_0$. For the coaxial line the following inductance matrix is obtained

$$\mathbf{L} = \begin{pmatrix} 157.74 & 12.06 \\ 12.06 & 157.74 \end{pmatrix} nH/m. \tag{4.90}$$

The reader can easily check that the product of $\mathbf{L}$ and the capacitance matrix from Chapter 2

$$\mathbf{C} = \begin{pmatrix} 283.8 & -21.7 \\ -21.7 & 283.8 \end{pmatrix} \; pF/m \tag{4.91}$$

satisfies (4.87) with $\epsilon_r = 4$ and $\mu_r = 1$.

The inductance matrix for the three signal line configuration of Fig. 3.25 turns out to be

$$\mathbf{L} = \begin{pmatrix} 277.73 & 87.76 & 36.77 \\ 87.76 & 328.60 & 115.77 \\ 36.77 & 115.77 & 337.98 \end{pmatrix} \; nH/m. \tag{4.92}$$

The capacitance matrix was given by

$$\mathbf{C} = \begin{pmatrix} 142.09 & -21.73 & -0.89 \\ -21.73 & 93.53 & -18.10 \\ -0.89 & -18.10 & 87.96 \end{pmatrix} \; pF/m. \tag{4.93}$$

As these signal lines are embedded in a multilayered medium, $\mathbf{C}$ and $\mathbf{L}$ no longer satisfy (4.87).

Chapter 5

Plane Waves

5.1 Introduction

A plane wave is a solution of Maxwell's equations in homogeneous, sourceless, infinite space. Furthermore, a plane wave propagating in the direction $\mathbf{u}$ only depends on the co-ordinate along that direction but remains constant in each plane perpendicular to $\mathbf{u}$. In the first part of this chapter, detailed expressions for plane waves in an isotropic, lossless medium will be derived and particular attention is devoted to polarisation. In an anisotropic medium plane wave solutions are substantially more complex compared to plane waves in an isotropic medium but this topic will not be treated in this introductory course. Next, the influence of losses is discussed and two special cases of great practical importance, plane waves in a low-loss dielectric and plane waves in a good conductor, are studied in more detail.

The second part of this chapter deals with the reflection and transmission of plane waves at the interface between two half-infinite media. First, normal incidence is considered to familiarise the reader with the physics of the problem. For oblique incidence, Snell's law is deduced and the important distinction between TE and TM polarisation is introduced together with phenomena such as total internal reflection and Brewster angle. As it is very important for the student to be completely familiar with plane waves and with their reflection and transmission at plane interfaces, a considerable amount of course time is allocated to exercises on this topic.

A plane wave is the most simple wave solution of Maxwell's equations. From this it could be assumed that plane waves have little to do with solutions of Maxwell's equations for more general and much more realistic and complex configurations. This is however not the case! It can be proved that the fields generated by arbitrary sources can be expanded in plane waves much in the same way as an arbitrary time signal can be expanded in a Fourier series, but a detailed analysis of this problem is beyond the scope of this course. In Chapter 8 it will be shown that, at a sufficiently large distance from a source or an antenna, fields locally behave as plane waves, a clear evidence of the tangible nature of plane waves.

5.2 Plane waves in a lossless dielectric

Consider a homogeneous, lossless and isotropic dielectric characterised by its permittivity $\epsilon = \epsilon_0\epsilon_r$ and its permeability $\mu = \mu_0\mu_r$. Due to the lossless nature of the dielectric ϵ_r and μ_r are real. In the Introduction is was already mentioned that plane waves only depend on a single co-ordinate. Suppose that this co-ordinate is the z-co-ordinate (the obtained results will later be generalised). In that case the curl equations (2.40) and (2.41) simplify to

$$\frac{\mathrm{d}}{\mathrm{d}z}\mathbf{u}_z \times \mathbf{e}(z) = -j\omega\mu\mathbf{h}(z), \tag{5.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}z}\mathbf{u}_z \times \mathbf{h}(z) = j\omega\epsilon\mathbf{e}(z). \tag{5.2}$$

Scalar multiplication of both equations with $\mathbf{u}_z$ immediately shows that $e_z(z) = h_z(z) = 0$ i.e. both the electric and the magnetic field are perpendicular to the z-direction. Substitution of (5.1) in (5.2) to eliminate $\mathbf{h}(z)$ yields

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\mathbf{e}(z) + k^2\mathbf{e}(z) = 0. \tag{5.3}$$

In a similar way substitution of (5.2) in (5.1) shows that

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\mathbf{h}(z) + k^2\mathbf{h}(z) = 0, \tag{5.4}$$

with $k^2 = \omega^2\epsilon\mu$. Equations (5.3) and (5.4) also follow from the general wave equations (2.114) and (2.115) for x and y independent fields. The general solution of (5.3) is given by

$$\mathbf{e}(z) = \mathbf{E}_t^+ e^{-jkz} + \mathbf{E}_t^- e^{jkz}, \tag{5.5}$$

with $\mathbf{E}_t^\pm$ constant vectors. The index t, with t standing for *transversal*, indicates that the vectors $\mathbf{E}_t^\pm$ have no z-component i.e. $\mathbf{E}_t^\pm$ are vectors in the xy-plane. Analogously, the general solution of (5.4) is

$$\mathbf{h}(z) = \mathbf{H}_t^+ e^{-jkz} + \mathbf{H}_t^- e^{jkz}. \tag{5.6}$$

As already remarked in Chapter 2 when deriving the wave equations satisfied by the electric and the magnetic field, the solutions of the one-dimensional wave equations (5.3) and (5.4) are not independent but must be linked through (5.1) or (5.2). Imposing (5.1) or (5.2) shows that

$$\mathbf{H}_t^\pm = \pm\frac{1}{Z_c}\mathbf{u}_z \times \mathbf{E}_t^\pm, \tag{5.7}$$

or

$$\mathbf{h}(z) = \frac{1}{Z_c}\mathbf{u}_z \times \mathbf{E}_t^+ e^{-jkz} - \frac{1}{Z_c}\mathbf{u}_z \times \mathbf{E}_t^- e^{jkz} \tag{5.8}$$

with $Z_c = \sqrt{\mu/\epsilon}$ in Ω the characteristic impedance of the medium. In vacuum or free space this characteristic impedance is $Z_{c0} = \sqrt{\mu_0/\epsilon_0} = 376.7303135\ \Omega \approx 377\ \Omega \approx 120\pi\ \Omega$. The magnetic field is perpendicular to the electric field and both fields are perpendicular to the propagation direction $\mathbf{u}_z$. These are the

typical characteristics of a plane wave. As the medium is lossless, both the wave number k and the characteristic impedance Z_c are real.

In the time domain (5.5) becomes

$$\mathbf{e}(z,t) = \Re[\mathbf{E}_t^+ e^{-jkz+j\omega t}] + \Re[\mathbf{E}_t^- e^{jkz+j\omega t}]. \tag{5.9}$$

The first term corresponds to a wave propagating in the positive z-direction, the second one to a wave propagating in the negative z-direction. Rewriting $\omega t \pm kz$ as $\omega(t \pm \frac{k}{\omega} z)$ shows that both waves propagate with speed $v = \omega/k$ given by $v = 1/\sqrt{\epsilon\mu}$.

Poynting's vector for the wave propagating in the positive z-direction, i.e. for the first term in (5.5) and (5.8), is found to be

$$\mathbf{p}(z) = \frac{|\mathbf{E}_t^+|^2}{2Z_c^*}\mathbf{u}_z = \frac{|\mathbf{E}_t^+|^2}{2Z_c}\mathbf{u}_z. \tag{5.10}$$

Hence, Poynting's vector has the same direction as the propagation direction of the wave. As the fields are constant in each plane perpendicular to the propagation direction, it is now warranted to interpret Poynting's vector $\mathbf{p}(z)$ locally and to say that in each point the wave carries a power density of $\mathbf{p}(z)\cdot\mathbf{u}_z$ W/m^2. For the wave propagating in the negative z-direction a similar result is obtained. Poynting's vector now becomes

$$\mathbf{p}(z) = \frac{|\mathbf{E}_t^-|^2}{2Z_c}(-\mathbf{u}_z). \tag{5.11}$$

In Chapter 1 it was shown that a complex vector corresponds to a time domain vector describing an ellipse: the vector is elliptically polarised. Let us take a closer look at the plane wave propagating in the positive z-direction. In each transversal plane the electric field describes an ellipse determined by the complex vector $\mathbf{E}_t^+$. As $\mathbf{E}_t^+$ lies in the transversal xy-plane, this plane is also the polarisation plane. The magnetic field also describes a polarisation ellipse in the same plane, but in such a way that at each moment the magnetic field remains perpendicular to the electric field. This is depicted in Fig. 5.1.

By way of example, consider the case of a left hand circularly polarised electric field. This automatically implies that the magnetic field is also left hand circularly polarised. With $\mathbf{E}_t^+ = E^+(\mathbf{u}_x + j\mathbf{u}_y)$ and for E^+ real, (5.9) leads to the following place and time dependence of the electric field

$$\mathbf{e}(z,t) = E^+\cos(\omega t - kz)\mathbf{u}_x - E^+\sin(\omega t - kz)\mathbf{u}_y \tag{5.12}$$

and similarly for the magnetic field

$$Z_c\mathbf{h}(z,t) = E^+\cos(\omega t - kz)\mathbf{u}_y + E^+\sin(\omega t - kz)\mathbf{u}_x. \tag{5.13}$$

As a function of z and at a particular moment $t = t_0$, (5.12) and (5.13) show that the endpoints of both vectors describe a helicoidal curve. This curve is periodic and the period is the wavelength $\lambda = 2\pi/k$. Fig. 5.2 shows the electric field vector $\mathbf{e}(z, t = 0)$: $\mathbf{e}(0,0) = E^+\mathbf{u}_x$, $\mathbf{e}(\lambda/4,0) = E^+\mathbf{u}_y$, $\mathbf{e}(\lambda/2,0) = -E^+\mathbf{u}_x$ and $\mathbf{e}(3\lambda/4,0) = -E^+\mathbf{u}_y$. The locus of the tip of the electric field describes a helix in space, which advances in the propagation direction (here the positive

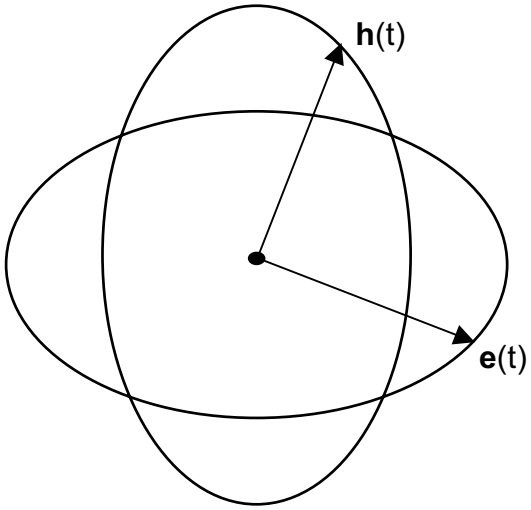


Figure 5.1: Polarisation ellipses of the electric and the magnetic field of a plane wave in a lossless medium.

z-direction) as a left handed screw. This is why clockwise circular polarisation is also called left hand circular polarisation. Right hand circular or counter clockwise circular polarisation corresponds to a plane wave with its electric field given by $\mathbf{E}_t^+ = E^+(\mathbf{u}_x - j\mathbf{u}_y)$.

Before extending our analysis to lossy media, we first generalise the obtained plane wave expressions to an arbitrary direction $\mathbf{u}$. The expressions for a plane wave propagating in the direction of $\mathbf{u}$ are given by

$$\mathbf{e}(\mathbf{r}) = \mathbf{A}e^{-jk\mathbf{u}\cdot\mathbf{r}}, \tag{5.14}$$

$$\mathbf{h}(\mathbf{r}) = \frac{1}{Z_c}\mathbf{u}\times\mathbf{A}e^{-jk\mathbf{u}\cdot\mathbf{r}} = \frac{1}{Z_c}\mathbf{u}\times\mathbf{e}(\mathbf{r}), \tag{5.15}$$

with

$$\mathbf{A}\cdot\mathbf{u} = 0. \tag{5.16}$$

Here again, electric and magnetic field are perpendicular to each other and to the propagation direction $\mathbf{u}$ and their amplitudes differ by the multiplicative factor $\frac{1}{Z_c}$. A plane wave is TEM-wave or Transversal Electric and Magnetic wave. In Chapter 5 other examples of TEM-waves will be discussed. It is easily verified that the first (second) term in (5.5) corresponds to (5.14) and the first (second) term in (5.8) to (5.15) for $\mathbf{u} = \mathbf{u}_z$ ($\mathbf{u} = -\mathbf{u}_z$) and for $\mathbf{A} = \mathbf{E}_t^+$ ($\mathbf{E}_t^-$). The reader should also verify that (5.14) and (5.15) satisfy Maxwell's equations. Furthermore, Poynting's vector becomes

$$\mathbf{p}(\mathbf{r}) = \frac{|\mathbf{A}|^2}{2Z_c}\mathbf{u}. \tag{5.17}$$

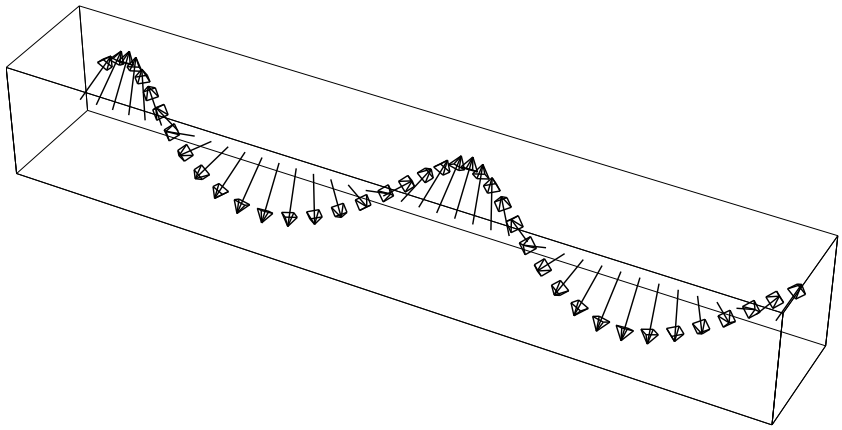


Figure 5.2: Electric field vector $\mathbf{e}(z, t = 0)$ for a left hand circular or clockwise polarised plane wave.

5.3 Plane waves in a lossy dielectric

In this section the results of the previous section are first extended to the general lossy dielectric case. The simplest way to do so is to introduce complex valued ϵ and μ. The field expressions (5.14) and (5.15)

$$\mathbf{e}(\mathbf{r}) = \mathbf{A}e^{-jk\mathbf{u}\cdot\mathbf{r}},$$
$$\mathbf{h}(\mathbf{r}) = \frac{1}{Z_c}\mathbf{u} \times \mathbf{A}e^{-jk\mathbf{u}\cdot\mathbf{r}} = \frac{1}{Z_c}\mathbf{u} \times \mathbf{e}(\mathbf{r}),$$

with

$$\mathbf{A} \cdot \mathbf{u} = 0,$$

remain unchanged, but Z_c and k are now complex valued. As in Chapter 2, the sign of the square root in $k = \omega\sqrt{\epsilon\mu}$ is chosen such that $\Re k \geq 0$ and $\Im k \leq 0$ and $Z_c = \frac{\omega\mu}{k} = \frac{k}{\omega\epsilon}$. Poynting's vector now becomes

$$\mathbf{p}(\mathbf{r}) = \frac{|\mathbf{A}|^2}{2Z_c^*}e^{-2\alpha\mathbf{u}\cdot\mathbf{r}}\mathbf{u}, \tag{5.18}$$

with α defined below. To better understand the effect of losses on the propagation of the plane wave, first write the wave number k as $k = \beta - j\alpha$. Equation (5.9) then becomes

$$\mathbf{e}(z, t) = \Re[\mathbf{E}_t^+ e^{-j\beta z + j\omega t}]e^{-\alpha z} + \Re[\mathbf{E}_t^- e^{j\beta z + j\omega t}]e^{\alpha z}. \tag{5.19}$$

The waves now propagate with speed $v = \omega/\beta$ and their amplitude decreases exponentially in their respective propagation directions. The real part of k is called the propagation constant β, while minus the imaginary part, α, is called the attenuation constant. In a lossy medium the electric and magnetic fields are no longer in phase. The phase difference is determined by the phase of Z_c. Expression (5.18) shows that when a wave propagates over a distance d in a

lossy medium that its power decreases by a factor $e^{-2\alpha d}$. Traditionally, this power loss is expressed in decibel (dB) as

$$-10\log_{10}(e^{-2\alpha d}) \approx 8.686\alpha d(dB). \tag{5.20}$$

The quantity $L = 8.686\alpha$ in dB/m is the relative power loss per metre and is a characteristic of the lossy medium. As a consequence of the presence of a complex characteristic impedance in (5.15), the polarisation ellipses of electric and magnetic field are no longer perpendicular to each other and their ellipticity differs. As a matter of fact, and this might be rather surprising at first glance, electric and magnetic field are still perpendicular to the propagation direction $\mathbf{u}$, but in general $\mathbf{e}(\mathbf{r},t)$ and $\mathbf{h}(\mathbf{r},t)$ are *no longer* perpendicular to each other. To prove this, use expression (1.20) to show that two vectors with a zero scalar product in the frequency domain $\mathbf{a}(\mathbf{r},\omega)\cdot\mathbf{b}(\mathbf{r},\omega) = 0$ does *not* automatically imply that $\mathbf{a}(\mathbf{r},t)\cdot\mathbf{b}(\mathbf{r},t) = 0$. $\mathbf{a}(\mathbf{r},t)$ and $\mathbf{b}(\mathbf{r},t)$ will be orthogonal for lossless media. Why? Also prove that for linear polarisation electric and magnetic field always remain perpendicular to each other.

Let us now restrict the discussion to a non-magnetic material with conduction losses. To make this clear, we replace ϵ in the general results obtained above by $\epsilon + \frac{\sigma}{j\omega}$ with ϵ again a real number. Hence, k and Z_c become

$$k = \omega\sqrt{\epsilon\mu}\sqrt{1+\frac{\sigma}{j\omega\epsilon}}, \tag{5.21}$$

$$Z_c = \sqrt{\frac{\mu}{\epsilon}}\frac{1}{\sqrt{1+\frac{\sigma}{j\omega\epsilon}}}. \tag{5.22}$$

As already mentioned in the Introduction, two special cases are of great practical importance.

In a *low-loss dielectric*, i.e. an isolating dielectric which, ideally, should have no losses at all, $\sigma \ll \omega\epsilon$. A Taylor series expansion of (5.21) and (5.22) shows that for $\sigma \ll \omega\epsilon$, k and Z_c can be approximated by

$$k = \beta - j\alpha = \omega\sqrt{\epsilon\mu} - j\frac{\sigma}{2}\sqrt{\frac{\mu}{\epsilon}}, \tag{5.23}$$

$$Z_c = \sqrt{\frac{\mu}{\epsilon}}. \tag{5.24}$$

Hence, for a low-loss dielectric, the propagation constant β and the characteristic impedance Z_c (approximately) have the same value as for the ideal dielectric with σ put to zero. The effect of losses manifests itself in the attenuation of the wave. The electric and magnetic field are still in phase. A typical representative of a low-loss dielectric is e.g. polyethylene ($\epsilon_r = 2.3$) with a $\frac{\sigma}{\omega\epsilon}$ value ranging from 2.10^{-4} to 3.10^{-4} between 50Hz and 1GHz.

A *good conductor* is precisely the opposite of a low-loss dielectric, i.e. $\sigma \gg \omega\epsilon$. The conduction current by far dominates the dielectric displacement. The wave number and the characteristic impedance can now be approximated by

$$k = \beta - j\alpha = \frac{(1-j)}{\delta}, \tag{5.25}$$

$$Z_c = \frac{(1+j)}{\sigma\delta}. \tag{5.26}$$

The quantity δ is a typical length known as the *skin depth*

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}}. \tag{5.27}$$

A plane wave propagating in a good conductor takes the following form

$$\mathbf{e}(\mathbf{r}) = \mathbf{A}e^{-\frac{\mathbf{u}\cdot\mathbf{r}}{\delta}}e^{-j\frac{\mathbf{u}\cdot\mathbf{r}}{\delta}},$$
$$\mathbf{h}(\mathbf{r}) = (1-j)\frac{\sigma\delta}{2}(\mathbf{u}\times\mathbf{A})e^{-\frac{\mathbf{u}\cdot\mathbf{r}}{\delta}}e^{-j\frac{\mathbf{u}\cdot\mathbf{r}}{\delta}}.$$

The amplitude of this wave decreases by a factor $e^{-\delta}$ when propagating over one skin depth. Hence, a good conductor can be used to shield electromagnetic waves (Faraday cage). As the skin depth increases for decreasing frequencies, the shielding capability of a metal plate will also decrease with frequency. The amplitude of the magnetic field is substantially higher than the amplitude of the electric field and there is a phase difference of $-\pi/4$ between them. A typical example of a good conductor is copper ($\mu = \mu_0, \epsilon = \epsilon_0, \sigma = 5.9\ 10^7$). Its skin depth is $9mm$ at 50Hz, $0.06mm$ at 1 MHz and $0.02mm$ at 1GHz. At 10GHz the skin depth of copper is $0.66\mu m$ as compared to $0.64\mu m$ for silver and $0.82\mu m$ for aluminium. Note that the definition of a low-loss dielectric or a good conductor is frequency dependent. For extremely high frequencies, even for a good conductor such a copper, dielectric displacement and conduction current will eventually become comparable.

5.4 Reflection and transmission at a plane interface

In this section the reflection and transmission of a plane wave at a plane interface between two half-infinite media is investigated (Fig. 5.3a). The coordinate system is chosen such that the plane interface coincides with the xy-plane ($z = 0$). Both media are homogeneous and isotropic with material parameters ϵ_1 and μ_1 for the medium to the left of the xy-plane and with material parameters ϵ_2 and μ_2 for the medium to the right of the xy-plane. These material parameters can take complex values such that the analysis will also be valid for conductors. To clarify the reflection and transmission physics, we first consider normal incidence. Oblique incidence is more complex due to the fact that the polarisation of the incident field plays a role. Snell's laws will be deduced and phenomena such as total internal reflection and absence of reflection (Brewster angle) will be highlighted. Finally, particular attention is devoted to the reflection at a perfect or at a good conductor, introducing the important surface impedance concept. More complex configurations such as a single material layer separating two half-spaces (Fig. 5.3b) or a stack of layers on top of a perfect conductor (Fig. 5.3c), are left to the exercises.

5.4.1 Normal incidence

Consider the situation depicted in Fig. 5.4. A linearly polarised plane wave propagating along the positive z-direction impinges on the interface between

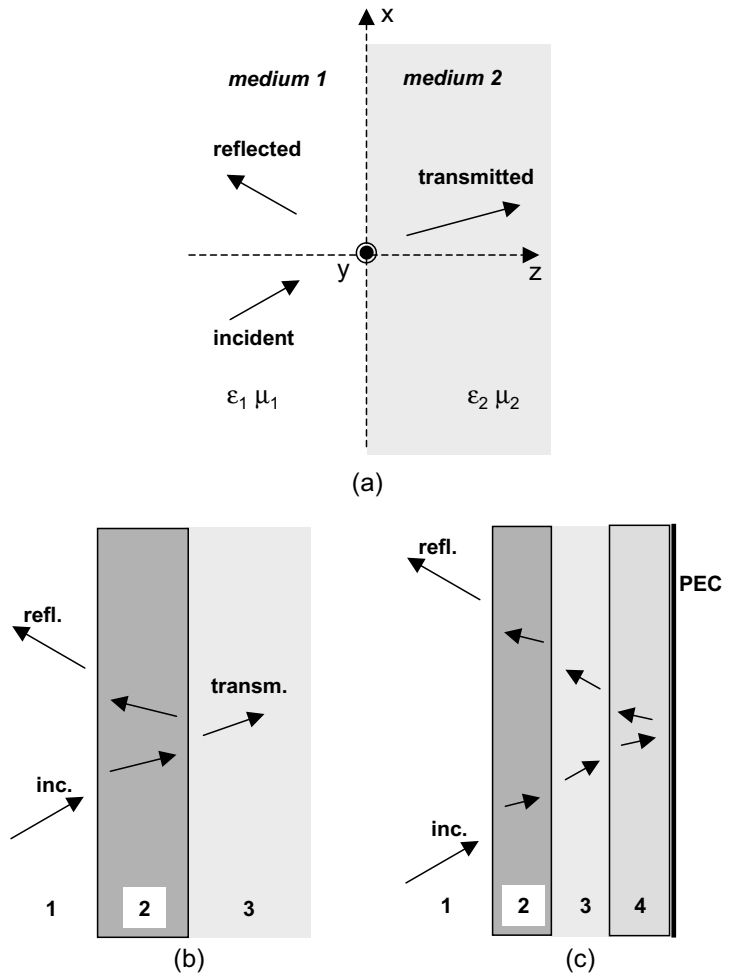


Figure 5.3: Reflection and transmission of plane waves in planar stratified medium: two half-spaces (a), a single material layer separating two half-spaces (b), a stack of planar stratified layers on top of a perfect conductor (c).

medium 1 and medium 2. This wave is called the incident wave and can be expressed as

$$e_{xi}(z) = e^{-jk_1 z}, \tag{5.28}$$

$$Z_1 h_{yi}(z) = e^{-jk_1 z}. \tag{5.29}$$

The incident wave has a unit amplitude and the electric field is polarised along the x-direction. k_1 and Z_1 are the wave number and the characteristic impedance of medium 1. As any elliptical polarisation can be written as the superposition of two linear contributions and as, at least for normal incidence, reflection and transmission is independent of polarisation, the results obtained in this section are valid for any polarisation. It hence suffices to just consider a single linear polarisation as in (5.28) and (5.29). The geometry of the problem *and* the incident field are completely independent of the transversal co-ordinates x and y.

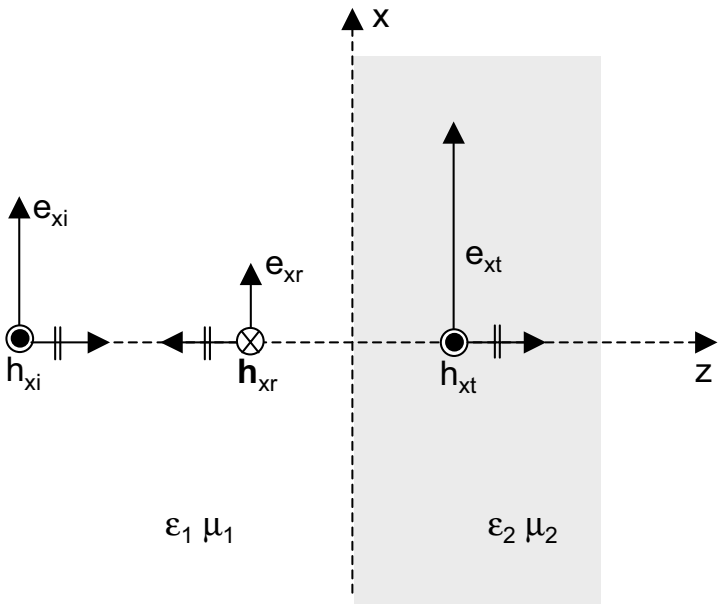


Figure 5.4: Reflection and transmission at normal incidence.

This implies that the reflected and transmitted wave will also only depend on z, i.e. these waves are plane. The reflected wave is given by

$$e_{xr}(z) = R_n e^{+jk_1 z}, \tag{5.30}$$

$$Z_1 h_{yr}(z) = -R_n e^{+jk_1 z}. \tag{5.31}$$

This wave travels in the negative z-direction and its amplitude R_n is as yet unknown. R_n is the amplitude reflection coefficient for normal incidence. Verify that the Poynting vector of the reflected wave points in the direction of $-\mathbf{u}_z$! The transmitted wave becomes

$$e_{xt}(z) = T_n e^{-jk_2 z}, \tag{5.32}$$

$$Z_2 h_{yt}(z) = T_n e^{-jk_2 z}, \tag{5.33}$$

with T_n the amplitude transmission coefficient. k_2 and Z_2 are the wave number and the characteristic impedance of medium 2. The unknown coefficients R_n and T_n can now be determined by applying the boundary conditions derived in Chapter 2, i.e. continuity of the total tangential electric and of the total tangential magnetic field (no surface current is present on the interface between the two media) at $z = 0$. In medium 1 the total field is the sum of the incident field and the reflected field. These boundary conditions yield

$$1 + R_n = T_n, \tag{5.34}$$

$$\frac{1 - R_n}{Z_1} = \frac{T_n}{Z_2} \tag{5.35}$$

and thus

$$R_n = \frac{1 - \frac{Z_1}{Z_2}}{1 + \frac{Z_1}{Z_2}}, \tag{5.36}$$

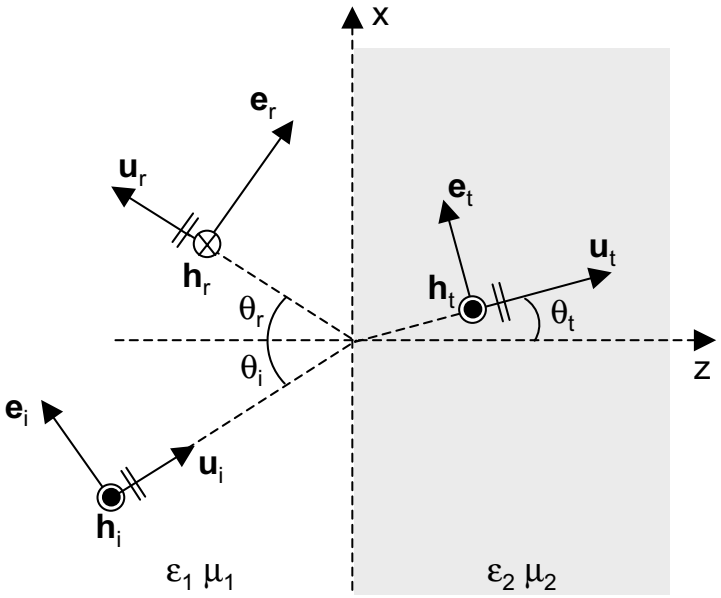


Figure 5.5: Reflection and transmission at oblique incidence.

$$T_n = \frac{2}{1 + \frac{Z_1}{Z_2}}. \tag{5.37}$$

For normal incidence, the reflection and transmission coefficient only depend on the characteristic impedances of the two media.

5.4.2 Oblique incidence and Snell's law

Fig. 5.5 depicts the general geometry pertaining to oblique incidence. The incident plane wave now impinges on medium 2 under the angle of incidence θ_i. Its propagation direction is $\mathbf{u}_i = \cos\theta_i \mathbf{u}_z + \sin\theta_i \mathbf{u}_x$. The incident fields are

$$\mathbf{e}_i = \mathbf{A} e^{-jk_1 \mathbf{u}_i \cdot \mathbf{r}} = \mathbf{A} e^{-jk_1 \cos\theta_i z} e^{-jk_1 \sin\theta_i x}, \tag{5.38}$$

$$Z_1 \mathbf{h}_i = (\mathbf{u}_i \times \mathbf{A}) e^{-jk_1 \mathbf{u}_i \cdot \mathbf{r}} = (\mathbf{u}_i \times \mathbf{A}) e^{-jk_1 \cos\theta_i z} e^{-jk_1 \sin\theta_i x}, \tag{5.39}$$

with $\mathbf{A} \cdot \mathbf{u}_i = 0$. The plane formed by $\mathbf{u}_i$ and $\mathbf{u}_z$, i.e. the normal to the interface between medium 1 and medium 2, is the plane of incidence. Remark that the geometry of the problem is translation invariant in the x and y-direction. As the incident wave is independent of y (for any given direction of incidence $\mathbf{u}_i$ the co-ordinate system can be chosen such that $\mathbf{u}_i$ has no y-component!) this will also be the case for the reflected and for the transmitted field. The only x-dependence comes from the x-dependence of the incident wave through the $e^{-jk_1 \sin\theta_i x}$ factor. Hence, this x-dependence must also be present in the reflected and in the transmitted wave. Let us first concentrate on the reflected electric field $\mathbf{e}_r$. Assume $\mathbf{e}_r$ is written as

$$\mathbf{e}_r(x, z) = \mathbf{f}(z) e^{-jk_1 \sin\theta_i x}. \tag{5.40}$$

The reflected electric field must satisfy the source free wave equation (2.114)

$$\nabla^2 \mathbf{e}_r(\mathbf{r}) + k_1^2 \mathbf{e}_r(\mathbf{r}) = 0. \tag{5.41}$$

Substituting (5.40) into (5.41) shows that

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\mathbf{f}(z) + (k_1^2 - k_1^2 \sin^2\theta_i)\mathbf{f}(z) = \frac{\mathrm{d}^2}{\mathrm{d}z^2}\mathbf{f}(z) + k_1^2 \cos^2\theta_i \mathbf{f}(z) = 0 \tag{5.42}$$

and hence $\mathbf{e}_r$ becomes

$$\mathbf{e}_r(x,z) = \mathbf{B}e^{-jk_1\mathbf{u}_r\cdot\mathbf{r}} = \mathbf{B}e^{jk_1\cos\theta_i z}e^{-jk_1\sin\theta_i x}, \tag{5.43}$$

with $\mathbf{B}$ a constant but as yet unknown vector. The reflected wave propagates in the $\mathbf{u}_r$-direction with $\mathbf{u}_r = -\cos\theta_i\mathbf{u}_z + \sin\theta_i\mathbf{u}_x$. We have excluded a second solution of (5.42), obtained by replacing $e^{jk_1\cos\theta_i z}$ in (5.43) by $e^{-jk_1\cos\theta_i z}$, because this solution propagates in the positive z-direction. Substituting (5.43) in Maxwell's equations shows that

$$Z_1\mathbf{h}_r(x,z) = (\mathbf{u}_r \times \mathbf{B})e^{-jk_1\mathbf{u}_r\cdot\mathbf{r}}, \tag{5.44}$$

$$\mathbf{B}\cdot\mathbf{u}_r = 0. \tag{5.45}$$

Expressions (5.43), (5.44) and (5.45) show that the reflected field is also a plane wave. This plane wave propagates in the direction of $\mathbf{u}_r$.

For the transmitted wave, a similar reasoning as for the reflected wave shows that the transmitted electric field can be written as

$$\mathbf{e}_t(x,z) = \mathbf{g}(z)e^{-jk_1\sin\theta_i x}, \tag{5.46}$$

with $\mathbf{g}(z)$ satisfying the wave equation

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\mathbf{g}(z) + (k_2^2 - k_1^2\sin^2\theta_i)\mathbf{g}(z) = 0. \tag{5.47}$$

Solving (5.47), the transmitted electric field becomes

$$\begin{aligned}\mathbf{e}_t(x,z) &= \mathbf{C}e^{-jk_2\mathbf{u}_t\cdot\mathbf{r}} \\ &= \mathbf{C}e^{-jk_2\cos\theta_t z}e^{-jk_1\sin\theta_i x} = \mathbf{C}e^{-jk_2\cos\theta_t z}e^{-jk_2\sin\theta_t x},\end{aligned} \tag{5.48}$$

with the angle of refraction θ_t defined by

$$k_2\sin\theta_t = k_1\sin\theta_i \tag{5.49}$$

and with $\mathbf{u}_t = \cos\theta_t\mathbf{u}_z + \sin\theta_t\mathbf{u}_x$. The square root in $\cos\theta_t = \sqrt{1-\sin^2\theta_t}$ is defined in such a way that $\Re(k_2\cos\theta_t) \geq 0$ and $\Im(k_2\cos\theta_t) \leq 0$ to make sure that $\mathbf{e}_t$ propagates in the positive z-direction and that in the presence of losses the wave decays exponentially while propagating. From Maxwell's equations we further derive that

$$Z_2\mathbf{h}_t(x,z) = (\mathbf{u}_t \times \mathbf{C})e^{-jk_1\mathbf{u}_t\cdot\mathbf{r}}, \tag{5.50}$$

$$\mathbf{C}\cdot\mathbf{u}_t = 0. \tag{5.51}$$

Expressions (5.48), (5.50) and (5.51) show that the transmitted field is a plane wave. This plane wave propagates in the direction of $\mathbf{u}_t$.

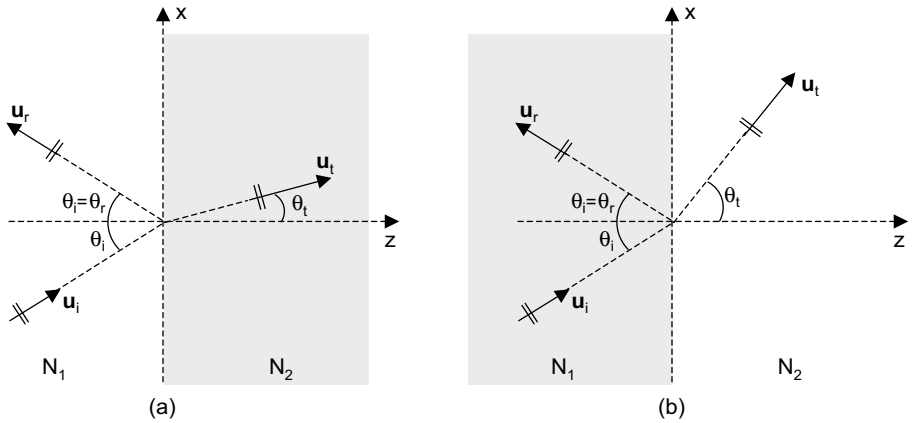


Figure 5.6: Snell's law for transmission from an optically less dense to an optically more dense medium $N_2 > N_1$ (a) and vice versa $N_2 < N_1$ (b).

The reader might object that, in the presence of losses, $\mathbf{u}_t$ is a complex vector and that it is no longer clear what propagation in the direction of $\mathbf{u}_t$ means. Therefore, to better grasp the physics of the problem, let us consider the more straightforward case of lossless materials and reconsider the one-dimensional wave equation (5.47). As k_1 and k_2 are now real, two cases must be distinguished:

1. Medium 2 is more dense than medium 1 ($k_2 > k_1$ or $N_2 > N_1$).
   If we characterise both media by their refractive index, i.e. $N_1 = \sqrt{\epsilon_{r1}\mu_{r1}}$ and $N_2 = \sqrt{\epsilon_{r2}\mu_{r2}}$, medium 2 is said to be more (optically) dense than medium 1 provided $N_2 > N_1$. Medium 1 could e.g. be air $N_1 \approx 1$ and medium 2 water $N_2 \approx 8.4$. In that case $(k_2^2 - k_1^2 \sin^2\theta_i)$ in (5.47) will always be positive and the transmitted electric field becomes

$$\begin{aligned} \mathbf{e}_t(x,z) &= \mathbf{C}e^{-jk_2\mathbf{u}_t\cdot\mathbf{r}} = \mathbf{C}e^{-jk_2\cos\theta_t z}e^{-jk_2\sin\theta_t x} \\ &= \mathbf{C}e^{-j\sqrt{k_2^2-k_1^2\sin^2\theta_i}z}e^{-jk_2\sin\theta_t x}, \end{aligned} \tag{5.52}$$

   with

$$\mathbf{u}_t = \cos\theta_t\mathbf{u}_z + \sin\theta_t\mathbf{u}_x = \mathbf{u}_z\sqrt{1-\frac{k_1^2}{k_2^2}\sin^2\theta_i} + \sin\theta_t\mathbf{u}_x. \tag{5.53}$$

   Equation (5.49) can be rewritten as

$$\sin\theta_t = \frac{N_1}{N_2}\sin\theta_i \tag{5.54}$$

   and is known as *Snell's law* (see Fig. 5.6a). This law, originally formulated in terms of rays, determines the angle under which a ray penetrates into medium 2. For transmission from an optically less dense to an optically more dense medium, θ_t will be smaller than θ_i, i.e. the ray is diffracted towards the normal. We have already seen that the reflected ray makes the same angle with the normal as the incident ray.

2. Medium 2 is less dense than medium 1 ($k_2 < k_1$ or $N_2 < N_1$).
This case is more complicated as $(k_2^2 - k_1^2 \sin^2\theta_i)$ in (5.47) will now only remain positive provided

$$\sin\theta_i < \frac{N_2}{N_1} = \sin\theta_c. \tag{5.55}$$

As long as the angle of incidence is smaller than the critical angle defined by $\sin\theta_c = N_2/N_1$, $\mathbf{e}_t$ and $\mathbf{u}_t$ are still given by (5.52) and (5.53) and Snell's law still applies. As shown in Fig. 5.6b, the transmitted ray is now diffracted away from the normal. For the transition from glass with refractive index 1.5 to air, the critical angle is 42^0; for the water to air transition the critical angle becomes much smaller, i.e. 6.8^0. If the angle of incidence approaches the critical angle $\theta_i \to \theta_c$, $\theta_t \to \frac{\pi}{2}$.

When the angle of incidence drops below the critical angle, $(k_2^2 - k_1^2 \sin^2\theta_i)$ becomes negative and the transmitted electric field is then given by

$$\begin{aligned} \mathbf{e}_t(x,z) &= \mathbf{C}e^{-jk_2\mathbf{u}_t\cdot\mathbf{r}} \\ &= \mathbf{C}e^{-\sqrt{k_1^2\sin^2\theta_i - k_2^2}z}e^{-jk_2\sin\theta_t x}, \end{aligned} \tag{5.56}$$

with

$$\mathbf{u}_t = j\mathbf{u}_z\sqrt{\frac{k_1^2}{k_2^2}\sin^2\theta_i - 1} + \sin\theta_t\mathbf{u}_x. \tag{5.57}$$

From (5.56) we see that the transmitted wave decays exponentially when propagating in medium 2. The transmitted field remains restricted to a "thin" layer in medium 2 characterised by a skin depth

$$\delta = \frac{1}{\sqrt{k_1^2\sin^2\theta_i - k_2^2}}, \tag{5.58}$$

with δ varying between ∞ for $\theta_i = \theta_c$ and $\frac{1}{\sqrt{k_1^2-k_2^2}}$ for $\theta_i = \pi/2$. The latter case is known as grazing incidence. For transmission from glass to air at $\theta_i = \pi/4$, δ is about 0.45λ, showing that the field is restricted to a thin film at the interface. In terms of rays we talk of *total internal reflection*: no rays reach medium 2 and the field is totally reflected into medium 1. This also means that the interface acts as a mirror. The reader can easily verify that the real part of Poynting's vector in medium 2 is identically zero, hence all power is reflected.

The type of plane wave (5.56) is known as an *inhomogeneous* plane wave because, contrary to the (homogeneous) plane waves introduced in the beginning of this chapter, this wave is exponentially damped while propagating, even when the medium is lossless.

Until now nothing has been said about the unknown coefficients **A**, **B** and **C**. Each of these vectorial unknowns has three degrees of freedom. However, as each of them must be perpendicular to their respective propagation directions, only four scalar unknowns remain. The boundary conditions dictate that the total tangential electric and magnetic field must be continuous at the interface

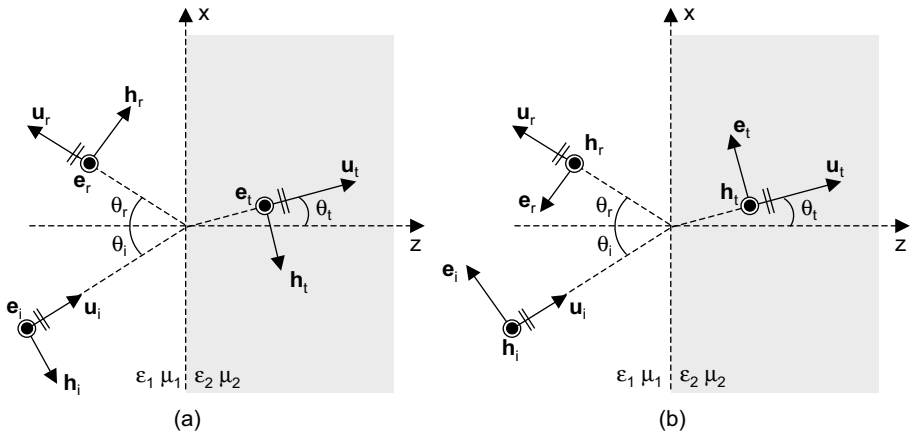


Figure 5.7: TE polarisation (a) and TM polarisation (b).

between medium 1 and medium 2, yielding the four scalar conditions to be satisfied by the remaining four unknowns. Traditionally, the problem of determining $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ is not solved in general but a distinction is made between two linear polarisations: *TE polarisation* or Transverse Electric polarisation with the electric field perpendicular (transverse) to the plane of incidence and *TM polarisation* or Transverse Magnetic polarisation with the magnetic field perpendicular (transverse) to the plane of incidence. Both situations are depicted in Fig. 5.7. An incident wave with arbitrary elliptical polarisation can be written as the superposition of a TE and a TM polarised contribution. Moreover, invoking symmetry or by going through the maths in solving for $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, it can be proved that the polarisation is not influenced by reflection and transmission.

TE polarisation

As shown in Fig. 5.7a, the electric fields are linearly polarised along the y-axis and the magnetic fields are in the xz-plane. In this case the final field expressions become

$$\begin{aligned} \mathbf{e}_i(\mathbf{r}) &= Ae^{-jk_1\mathbf{u}_i\cdot\mathbf{r}}\mathbf{u}_y, \\ Z_1\mathbf{h}_i(\mathbf{r}) &= A(\mathbf{u}_i\times\mathbf{u}_y)e^{-jk_1\mathbf{u}_i\cdot\mathbf{r}}, \end{aligned}$$

$$\begin{aligned} \mathbf{e}_r(\mathbf{r}) &= R_{TE}Ae^{-jk_1\mathbf{u}_r\cdot\mathbf{r}}\mathbf{u}_y, \\ Z_1\mathbf{h}_r(\mathbf{r}) &= R_{TE}A(\mathbf{u}_r\times\mathbf{u}_y)e^{-jk_1\mathbf{u}_r\cdot\mathbf{r}}, \end{aligned} \tag{5.59}$$

$$\begin{aligned} \mathbf{e}_t(\mathbf{r}) &= T_{TE}Ae^{-jk_2\mathbf{u}_t\cdot\mathbf{r}}\mathbf{u}_y, \\ Z_1\mathbf{h}_i(\mathbf{r}) &= T_{TE}A(\mathbf{u}_t\times\mathbf{u}_y)e^{-jk_2\mathbf{u}_t\cdot\mathbf{r}}. \end{aligned}$$

R_{TE} and T_{TE} are the scalar reflection and transmission coefficient for TE polarisation

$$R_{TE} = \frac{Z_2\cos\theta_i - Z_1\cos\theta_t}{Z_2\cos\theta_i + Z_1\cos\theta_t}, \tag{5.60}$$

$$T_{TE} = \frac{2Z_2 \cos\theta_i}{Z_2 \cos\theta_i + Z_1 \cos\theta_t}. \tag{5.61}$$

TM polarisation

In this case the magnetic fields are linearly polarised along the y-axis (see Fig. 5.7b). The relevant field expressions then become

$$\begin{aligned} Z_1\mathbf{h}_i(\mathbf{r}) &= Ae^{-jk_1\mathbf{u}_i\cdot\mathbf{r}}\mathbf{u}_y, \\ \mathbf{e}_i(\mathbf{r}) &= -A(\mathbf{u}_i \times \mathbf{u}_y)e^{-jk_1\mathbf{u}_i\cdot\mathbf{r}}, \end{aligned}$$

$$\begin{aligned} Z_1\mathbf{h}_r(\mathbf{r}) &= R_{TM}Ae^{-jk_1\mathbf{u}_r\cdot\mathbf{r}}\mathbf{u}_y, \\ \mathbf{e}_r(\mathbf{r}) &= -R_{TM}A(\mathbf{u}_r \times \mathbf{u}_y)e^{-jk_1\mathbf{u}_r\cdot\mathbf{r}}, \end{aligned} \tag{5.62}$$

$$\begin{aligned} Z_2\mathbf{h}_t(\mathbf{r}) &= T_{TM}Ae^{-jk_2\mathbf{u}_t\cdot\mathbf{r}}\mathbf{u}_y, \\ \mathbf{e}_t(\mathbf{r}) &= -T_{TM}A(\mathbf{u}_t \times \mathbf{u}_y)e^{-jk_2\mathbf{u}_t\cdot\mathbf{r}}. \end{aligned}$$

R_{TM} and T_{TM} are the scalar reflection and transmission coefficient for TM polarisation

$$R_{TM} = \frac{Z_1 \cos\theta_i - Z_2 \cos\theta_t}{Z_1 \cos\theta_i + Z_2 \cos\theta_t}, \tag{5.63}$$

$$T_{TM} = \frac{2Z_2 \cos\theta_i}{Z_1 \cos\theta_i + Z_2 \cos\theta_t}. \tag{5.64}$$

For normal incidence, $\theta_i = 0$, TE and TM polarisation coincide. The reader can easily verify that for $\theta_i = 0$, $R_{TE} = -R_{TM} = R_n$ and $T_{TE} = T_{TM} = T_n$ with R_n and T_n given by (5.36) and (5.37).

Total transmission - Brewster angle

Is it possible for a plane wave to pass from medium 1 to medium 2 without reflection at the interface? For this it suffices that R_{TE} or R_{TM} vanish. Restricting ourselves to the non-magnetic case, i.e. $\mu_1 = \mu_2 = \mu_0$, R_{TE} (5.60) and R_{TM} (5.63) can be more compactly written as

$$R_{TE} = \frac{\sin(\theta_t - \theta_i)}{\sin(\theta_t + \theta_i)}, \tag{5.65}$$

$$R_{TM} = \frac{\tan(\theta_i - \theta_t)}{\tan(\theta_t + \theta_i)}. \tag{5.66}$$

Both reflection coefficients are zero for $\theta_i = \theta_t$ but from Snell's law this implies that $\epsilon_1 = \epsilon_2$. Both media are identical and of course no reflection occurs in this trivial case. For TE polarisation no other solution is found, but for TM polarisation R_{TM} becomes zero provided the angle of incidence is equal to the *Brewster angle* defined by

$$\tan\theta_B = \sqrt{\frac{\epsilon_2}{\epsilon_1}} = \frac{N_2}{N_1}. \tag{5.67}$$

For transmission for air to glass the Brewster angle is 56.3^0.

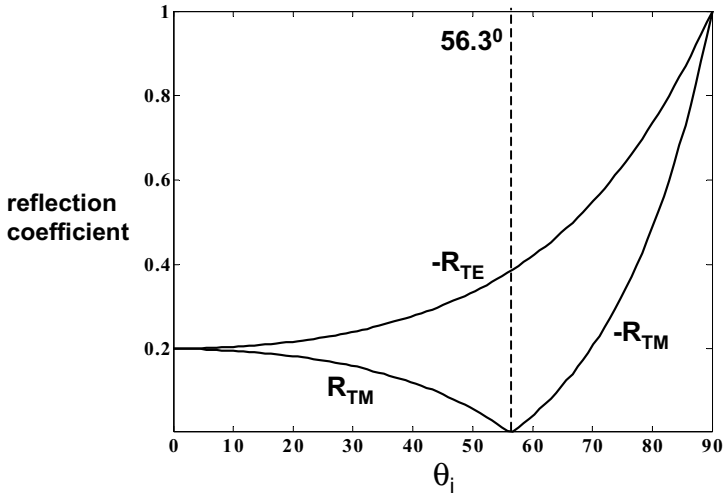


Figure 5.8: Amplitude reflection coefficients $|R_{TE}|$ and $|R_{TM}|$ as a function of angle of incidence for the air to glass transition.

Example: glass-air interface

As an example take the reflection and transmission at a glass-air interface. Fig. 5.8 shows the amplitudes of the TE and TM reflection coefficients for the air to glass transition, i.e. from a less dense medium ($N_1 = 1$) to a more dense medium ($N_2 = 1.5$). For TE polarisation minimal reflection is observed for normal incidence and this reflection increases monotonically with increasing angle of incidence, with a maximum value of 1 at grazing incidence. For TM polarisation zero reflection coefficient is obtained for the Brewster angle at 56.3^0. Fig. 5.9 shows the amplitudes of the TE and TM reflection coefficient for the glass to air transition. In this case total internal reflection occurs for angles of incidence exceeding the critical angle $\theta_c = 42^0$. The Brewster angle is now 33.7^0.

5.5 Reflection at good conductor - Surface impedance

Consider a plane wave impinging on a perfect electric conductor or PEC. This problem can be treated as a special case of the general theory presented above by assuming that $\sigma \to \infty$ and hence $Z_2 \to 0$. From (5.60), (5.61), (5.63) and (5.64) we easily deduce that

$$\begin{aligned} R_{TE,PEC} &= -1, \\ T_{TE,PEC} &= 0, \\ R_{TM,PEC} &= +1, \\ T_{TM,PEC} &= 0. \end{aligned}$$

We conclude that the wave is totally reflected by the PEC and that no waves penetrate into it. To understand why the reflection coefficient for TE polar-

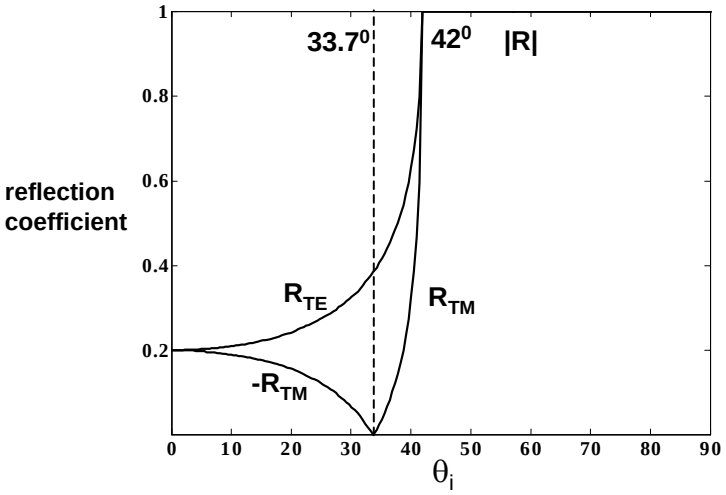


Figure 5.9: Amplitude reflection coefficients $|R_{TE}|$ and $|R_{TM}|$ as a function of angle of incidence for the glass to air transition.

isation is equal to -1 while for TM polarisation its value is $+1$, the incident and reflected fields are depicted in Fig. 5.10. For both polarisations, the total tangential electric field and the total normal magnetic field must be zero at the PEC (see Chapter 2, Section 2.4). For the TE-case this implies that $e_{yr} = -e_{yi}$, $h_{xr} = h_{xi}$ and $h_{zr} = -h_{zi}$, while for the TM-case this becomes $h_{yr} = h_{yi}$, $e_{xr} = -e_{xi}$ and $e_{zr} = e_{zi}$. The incident wave induces a surface current density $\mathbf{j}_s = \mathbf{u}_n \times (\mathbf{h}_r + \mathbf{h}_i)$. The normal $\mathbf{u}_n$ is the outward pointing normal to the PEC, i.e. $\mathbf{u}_n = -\mathbf{u}_z$. The surface current is the source of the reflected wave. For the TE-case $\mathbf{j}_s = -2h_{xi}\mathbf{u}_y$ and no surface charges $\rho_s = \mathbf{u}_n \cdot (\mathbf{d}_r + \mathbf{d}_i)$ are present as the electric field has no z-component for this polarisation. For the TM-case $\mathbf{j}_s = 2h_{yi}\mathbf{u}_x$ and $\rho_s = -2\epsilon_1 e_{zi}$. The values of the fields in the above expressions for $\mathbf{j}_s$ and ρ_s are their values on the PEC, i.e. for $z = 0$. They still depend on x through the phase factor $e^{-jk_1 \sin\theta_i x}$. Remark that the total tangential magnetic field at the PEC is twice the value of the tangential component of the incident magnetic field.

In practice no perfect conductors exist and it is important to find out how to modify the above results for good conductors. In the calculations below we will restrict ourselves to normal incidence. In Section 5.3 we already found the following approximation for the wave number k_2 and for the characteristic impedance Z_2 of a good conductor

$$k_2 = \beta - j\alpha = \frac{(1-j)}{\delta},$$
$$Z_2 = \frac{(1+j)}{\sigma\delta}, \tag{5.68}$$

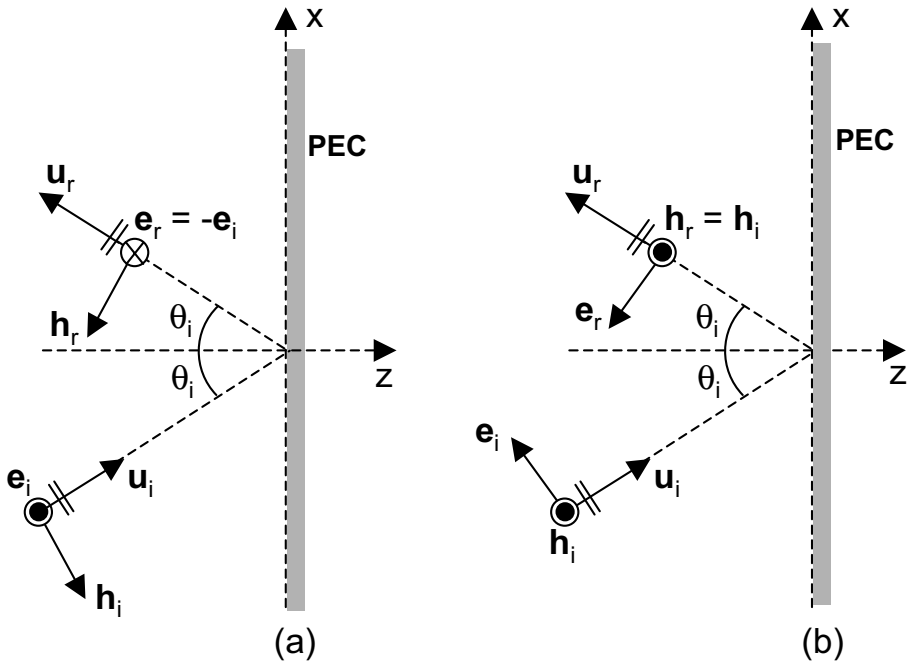


Figure 5.10: Total reflection at a perfect conductor: TE polarisation (a), TM polarisation (b).

with $\delta = \sqrt{\frac{2}{\omega\mu_2\sigma_2}}$. The reflection coefficient R_n (5.36) was

$$R_n = \frac{1 - \frac{Z_1}{Z_2}}{1 + \frac{Z_1}{Z_2}}. \tag{5.69}$$

Let us now repeat the reasoning of Section 5.4.1 but for the situation depicted in Fig. 5.11. The half-infinite conductor is replaced by an impedance surface S_{imp}. Whereas the total tangential electric field is zero at a PEC, i.e. $\mathbf{e}_{tan} = 0$, with the subscript "tan" standing for the tangential component, the boundary condition for the impedance surface S_{imp} is by definition given by

$$\mathbf{e}_{tan} = Z_s(\mathbf{u}_n \times \mathbf{h}_{tan}), \tag{5.70}$$

with Z_s the so-called *surface impedance* and with $\mathbf{e}_{tan}$ and $\mathbf{h}_{tan}$ the total tangential electric and magnetic fields at the surface. In view of the relationship between the tangential magnetic field and the surface current, (5.70) can be rewritten as

$$\mathbf{e}_{tan} = Z_s\mathbf{j}_s. \tag{5.71}$$

For a good conductor Z_s is given by (5.68)

$$Z_s = \frac{(1+j)}{\sigma\delta}. \tag{5.72}$$

According to (5.70) and (5.71), the impedance surface behaves much in the same way as a perfect conductor, in the sense that a surface current is present which

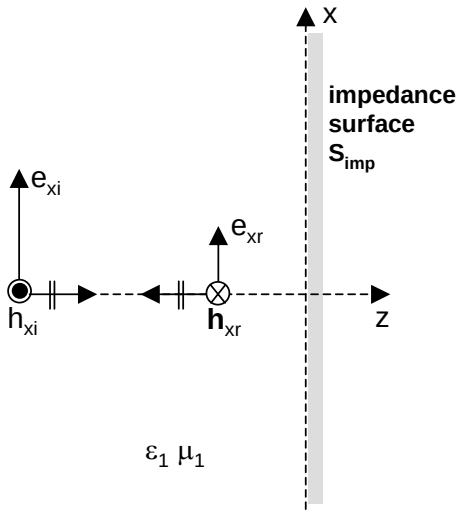


Figure 5.11: Reflection for normal incidence on a good conductor characterised by its surface impedance.

is related to the total tangential magnetic field in the same way as in the case of a perfect conductor. However, the tangential electric field is not zero but proportional to the surface current, in the same way as in a resistor the voltage is proportional to the current.

Let us now analyse the scattering by the impedance surface. Following (5.28)–(5.31) the incident and reflected fields are

$$\begin{aligned} e_{xi}(z) &= e^{-jk_1 z}, \\ Z_1 h_{yi}(z) &= e^{-jk_1 z}, \\ e_{xr}(z) &= R_n e^{+jk_1 z}, \\ Z_1 h_{yr}(z) &= -R_n e^{+jk_1 z}. \end{aligned}$$

For normal incidence, (5.70) demands that

$$e_{xi}(z=0) + e_{xr}(z=0) = \frac{Z_s}{Z_1}(h_{yi}(z=0) + h_{yr}(z=0)), \tag{5.73}$$

or

$$(1 + R_n) = \frac{Z_s}{Z_1}(1 - R_n) \tag{5.74}$$

and hence

$$R_n = \frac{1 - \frac{Z_1}{Z_s}}{1 + \frac{Z_1}{Z_s}}. \tag{5.75}$$

This is exactly the result obtained for the reflection coefficient (5.69) with $Z_2 = Z_s$. For scattering purposes, the half-infinite conducting space can be replaced by the impedance surface. For a PEC conductor Z_s becomes zero, $R_n = -1$ and according to (5.70) the tangential electric field will vanish, as required.

Although the above prove is restricted to normal incidence at a plane interface,

the conclusion can be generalised to the reflection at any arbitrary good conductor provided its curvature radius remains sufficiently large compared to the skin depth δ.

When the conductor is not perfectly conducting, the induced currents lead to Joule losses and the reflected power will be smaller than the incident power. To conclude this section we show that the surface impedance boundary condition also accounts for these losses. The current density in the conductor is $j_x(z) = \sigma_2 e_{xt}(z)$. This current density is restricted to a thin layer near the surface of the conductor, as already discussed in Section 5.3. The total current I per unit of length in the y-direction follows from the integration of j_x over z

$$I = \int_0^\infty j_x(z)dz = \int_0^\infty \sigma T_n e^{-(1+j)\frac{z}{\delta}} dz = \frac{T_n}{Z_s}. \tag{5.76}$$

The total Joule losses P_J are

$$P_J = \frac{1}{2}\Re \int_0^{+\infty} j_x e_x^* dz = \frac{1}{2}\sigma \int_0^{+\infty} |e_x|^2 dz = \frac{\sigma\delta}{4}|T_n|^2. \tag{5.77}$$

It is easy to verify that

$$P_J = \frac{1}{2}R_s|I|^2 = \frac{1}{2}\Re(Z_s)|I|^2. \tag{5.78}$$

Let us now determine the losses for the case of the impedance surface. The total current I per unit of length in the y-direction is now simply j_s and

$$P_{Jimp} = \frac{1}{2}\Re(\mathbf{j}_s \cdot \mathbf{e}_{tan}) = \frac{1}{2}\Re(Z_s)|I|^2. \tag{5.79}$$

This is precisely the value of the Joule losses P_J (5.78).

The following final conclusions can be formulated. For a *perfect conductor* the induced surface current $\mathbf{j}_s = \mathbf{u}_n \times (\mathbf{h}_i + \mathbf{h}_r)$ flows in an infinitely thin layer on the surface of the conductor ($\delta = 0$). Viewed from the outside a *good conductor* can still be described by an equivalent surface current, still obtained through $\mathbf{j}_s = \mathbf{u}_n \times (\mathbf{h}_i + \mathbf{h}_r)$. The zero tangential boundary condition $\mathbf{u}_n \times (\mathbf{e}_i + \mathbf{e}_r) = \mathbf{u}_n \times \mathbf{e} = 0$ has to be replaced by the impedance boundary condition $\mathbf{e}_{tan} = \mathbf{e}_{i,tan} + \mathbf{e}_{r,tan} = Z_s(\mathbf{u}_n \times \mathbf{h}) = Z_s \mathbf{u}_n \times (\mathbf{h}_i + \mathbf{h}_r)$ with Z_s the surface impedance $Z_s = R_s + jX_s = \frac{(1+j)}{\sigma\delta}$.
The surface impedance boundary condition is often used in numerical simulations involving good conductors.

Chapter 6

Transmission Lines

6.1 Introduction

Wireless communication uses (plane) waves together with appropriate transmit and receive antennas to transfer information. Plane waves were studied in the previous chapter and antennas will be treated in Chapter 8. *Wired* communication uses *transmission lines* to transfer information between a transmitter and a receiver. Typical examples are telephone wires, coaxial cables, optical fibres, UTP (universal twisted pair) cables used for computer networks, microstrip and striplines for Printed Circuit Board (PCB) interconnect, etc. Fig. 6.1 shows a few transmission line examples. A microstrip line consists of a metal ground plane, a dielectric substrate material and a signal line on top of the substrate. A stripline is a similar structure, but in this case the dielectric substrate is sandwiched between two metal plates and the signal line is located in between the plates. Usually, microstrip and stripline interconnections are combined to form a multilayered PCB. Other, less familiar examples are the hollow rectangular waveguide, e.g. used to transmit high power signals in radar applications, or the coplanar waveguide suited for microwave or millimetre wave applications. The transmission line concept is a very broad concept: nerve fibres in our body e.g. also act as transmission lines.

In electrical engineering, transmission lines are not only used for the transmission of information but also for energy transport, the so-called power lines. Their purpose is quite different. In energy transport losses must be minimised and the distribution of power must be guaranteed. Power lines operate at low frequencies (e.g. 50 Hz), high voltages and currents. For signal transmission, signal integrity is of utmost importance. This implies that distortion of the signal, losses, reflection and crosstalk between signal conductors must be minimised. For signal transmission, bandwidths up to several tens of GHz are feasible and in most cases the transmitted power remains low (of a few tens of Watts at maximum).

For electrical transmission lines, the bandwidth is inversely proportional to the length. Not only will losses and the frequency dependent signal velocity more and more deteriorate the signal when further propagating down the line, but the magnitude of these effects also increases with frequency. Depending on

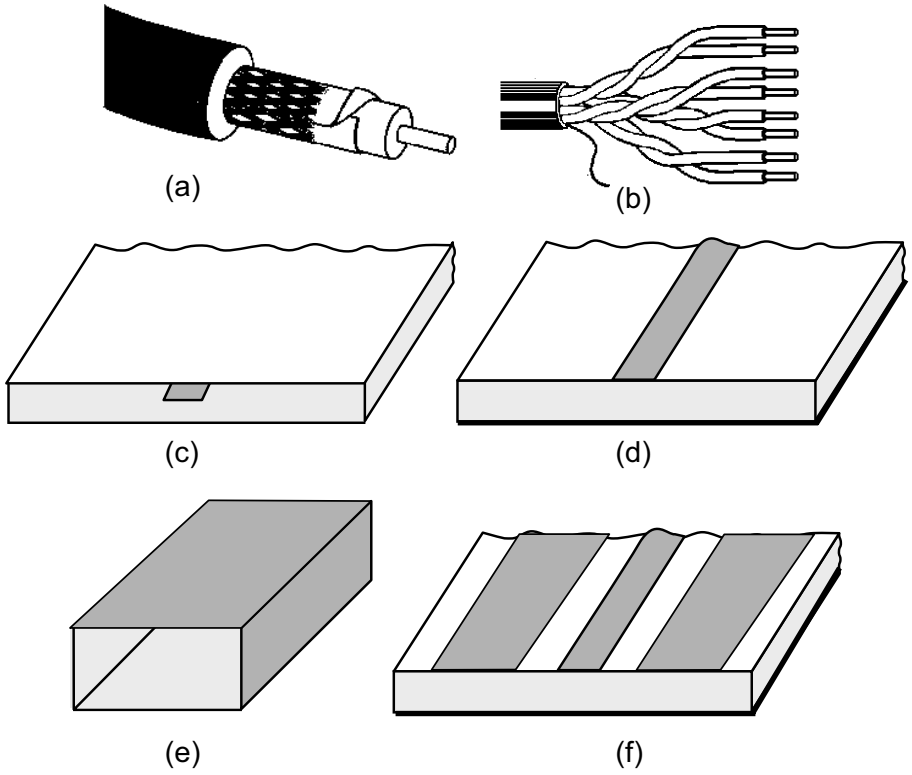


Figure 6.1: A few examples of transmission lines: coaxial cable (a), twisted pair (b), stripline (c), microstrip line (d), rectangular waveguide (e), coplanar waveguide (f).

the application, the length of the transmission line can vary between hundreds of kilometres, e.g. in the case of power lines or long distance telephone lines (when not replaced by fibres) and less than a millimetre on chip. It is not the length of the line in absolute terms that plays a role, but its length expressed as a fraction of the wavelength. As will become clear in the sequel, transmission line effects become visible when their length exceeds a tenth of the wavelength. Hence, for clock speeds in computers as high as 2 GHz, and taking into account that harmonics up to 20 GHz must be considered, even "short" interconnections of a few millimetre start behaving as transmission lines.

Optical transmission becomes increasingly important, not only using the already very widespread optical fibres, but many types of integrated optical waveguides. Optical interconnections admit larger bandwidths, reduced losses and dispersion effects. In this course we will restrict ourselves to electrical interconnections. Optical waveguides and optical communication are treated in other courses.

The most general meaning of a transmission line is a wired connection between two arbitrary points. However, in transmission line theory, a more restricted definition is used: a transmission line is defined by its cross-section and this cross-section is considered to be invariant in the signal propagation di-

rection. This signal propagation direction is perpendicular to the cross-section and is also known as the longitudinal direction. The cross-sectional plane is indicated as the transversal plane. Sometimes transmission line theory also deals with cross-sections that vary in the longitudinal direction. These are the so-called non-uniform lines as opposed to uniform lines with a constant cross-section.

To completely understand the behaviour of transmission lines and to characterise their signal propagation properties, Maxwell's equations must be solved. This is in general not an easy task. Chapter 7 is entirely devoted to the field analysis of transmission lines and waveguides. In this chapter, transmission lines will first be studied from a circuit point of view. Circuit theory uses *lumped elements* (resistors, capacitors, inductors, transformers, ...), active circuits (transistors, amplifiers, ...), voltage and current sources together with Kirchoff's voltage and current laws to describe the behaviour of a circuit. When extending circuit analysis to transmission lines, it would be advantageous to avoid solving Maxwell's equations in their full generality - using fields - while still being able to describe signal propagation or wave effects in terms of voltages and currents. It turns out that this is indeed possible. In circuit theory a transmission line is a circuit with *distributed parameters* described by a set of partial differential equations, the *telegrapher's equations*, for the voltage and the current. In these equations voltage and current still depend on time (frequency) but also on the co-ordinate in the longitudinal direction.

In this chapter the telegrapher's equations for a transmission line will be derived starting from a lumped element representation of the line. Next, the general solution of these equations is studied. A large part of the chapter is devoted to the discussion of wave phenomena on transmission lines in sinusoidal regime introducing concepts such as characteristic impedance, reflection coefficient, mismatch, input impedance and standing waves. Particular attention is devoted to the Smith chart which is used to represent the behaviour of an input impedance at high frequencies. In the second part of the chapter we briefly consider transients on and bit sequences propagating along transmission lines. In this chapter only single transmission lines will be considered. Coupled transmission lines allow to elegantly describe crosstalk between neighbouring signal conductors, but their analysis is beyond the scope of this course.

6.2 The telegrapher's equations

The classical way to introduce a transmission line in circuit theory is to say that the capacitance, inductance, resistance and conductance of the line are distributed over its length. Fig. 6.2 shows the circuit representation of an elementary section dz of a transmission line. In the sequel, the longitudinal direction coincides with the z-direction. The transmission line itself is obtained by the concatenation of elementary sections as shown in Fig. 6.3. C in F/m is the capacitance per unit of length of the transmission line. Similarly L in H/m, R in Ω/m and G in Ω^{-1}/m resp. are the inductance, the resistance and the conductance per unit of length. The resistance per unit of length R expresses the conductor losses while the conductance per unit of length G expresses the

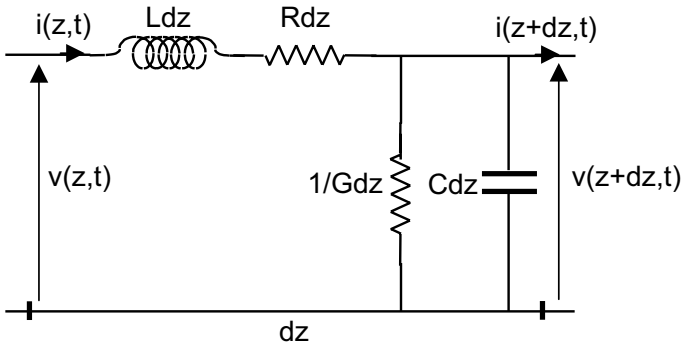


Figure 6.2: Circuit representation of an elementary section dz of a transmission line.

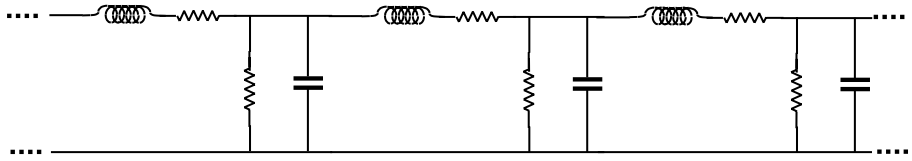


Figure 6.3: A transmission line as the concatenation of elementary sections.

dielectric or substrate losses. For a uniform line these quantities are independent of z. Finally, Fig. 6.4 shows the representation of the transmission line as a circuit element: two wires of length d supporting a place and time or frequency dependent voltage and current.

The lumped element representation of a transmission line as depicted in Fig. 6.3 is based on the availability of C, L, R and G. We have introduced C and R in the electrostatic case and L in the magnetostatic case. G could easily be found by replacing the real valued dielectric permittivity ϵ in capacitance calculations by $\epsilon + \frac{\sigma}{j\omega}$ with σ the small conductivity of the dielectric. However, the question arises if these quantities still make sense at arbitrary frequencies and if they apply for the complete range of transmission lines considered in the Introduction. The answer to this question requires a thorough field analysis of waveguides and transmission lines and an analysis of the way in which the field behaviour can be translated into a voltage-current description. The next chapter introduces the reader to this topic. One example we already encountered in the previous chapters was the coaxial cable. From (4.88) and (4.89) we know that

$$C_{coax} = \frac{2\pi\epsilon}{\ln\frac{b}{a}},$$
$$L_{coax} = \frac{\mu}{2\pi}\ln\frac{b}{a},$$

while in Chapter 7 it will be proved that

$$R_{coax} = \frac{R_s}{2\pi\left(\frac{1}{a}+\frac{1}{b}\right)},$$
$$G_{coax} = \frac{2\pi\sigma}{\ln\frac{b}{a}}. \tag{6.1}$$

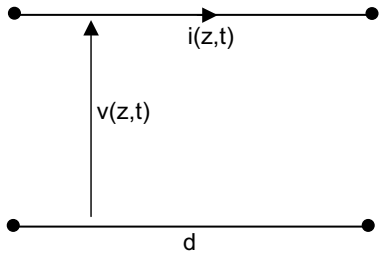


Figure 6.4: Circuit representation of a transmission line.

The resistance R_{coax} per unit of length is the high frequency value with $R_s = \frac{1}{\sigma\delta}$ the real part of the surface impedance Z_s (5.72). The conductivity σ in (6.1) is the conductivity of the dielectric filling material of the coaxial cable.

For the moment, we will take the lumped element representation of Fig. 6.3 for granted and use it as the starting point of our analysis of signal propagation along transmission lines. In Fig. 6.2, the potential difference between z and $z + dz$ can be expressed as

$$v(z + dz, t) - v(z, t) = -Rdz\, i(z, t) - Ldz\, \frac{\partial}{\partial t} i(z, t), \tag{6.2}$$

while the current difference is given by

$$i(z + dz, t) - i(z, t) = -Gdz\, v(z + dz, t) - Cdz\, \frac{\partial}{\partial t} v(z + dz, t). \tag{6.3}$$

Taking the limit for $dz \to 0$ and neglecting higher-order terms, yields

$$\frac{\partial v(z, t)}{\partial z} = -Ri(z, t) - L\frac{\partial i(z, t)}{\partial t}, \tag{6.4}$$

$$\frac{\partial i(z, t)}{\partial z} = -Gv(z, t) - C\frac{\partial v(z, t)}{\partial t}. \tag{6.5}$$

These equations are the *telegrapher's equations* for a single transmission line. Remark that voltage and current are still continuous functions of z and t. We have already encountered similar differential equations in the previous chapter. For real-valued ϵ and μ, and for a linearly polarised plane wave with the electric field along the x-axis, the time domain counterparts of equations (5.1) and (5.2) are

$$\frac{\partial e_x(z, t)}{\partial z} = -\mu\frac{\partial h_y(z, t)}{\partial t}, \tag{6.6}$$

$$\frac{\partial h_y(z, t)}{\partial z} = -\epsilon\frac{\partial e_x(z, t)}{\partial t}, \tag{6.7}$$

with μ and ϵ playing the role of L and C in (6.4) and (6.5).

In the last section of this chapter, the time domain solution of (6.4) and (6.5) will be considered, but we first turn to the sinusoidal regime. The telegrapher's equations become

$$\frac{\mathrm{d}V(z)}{\mathrm{d}z} = -(R + j\omega L)I(z) = -Z(\omega)I(z), \tag{6.8}$$

$$\frac{\mathrm{d}I(z)}{\mathrm{d}z} = -(G + j\omega C)V(z) = -Y(\omega)V(z). \tag{6.9}$$

As in this chapter both frequency and time domain phenomena will be considered, we have introduced capital V and capital I to denote the phasors of v and i. $Z(\omega)$ is the circuit impedance per unit of length and $Y(\omega)$ is the circuit admittance per unit of length. Although in going from (6.4) and (6.5) to (6.8) and (6.9) we assumed R, L, G and C to be constant, one could also consider the frequency domain telegrapher's equations (6.8) and (6.9) as the starting point of the analysis and still allow R, L, G and C to be frequency dependent. From the field analysis in the next chapter it will indeed become clear that this approach is to be preferred. In the previous chapter e.g., we have already shown that the skin-effect losses in conductors can be described by the surface impedance Z_s (5.72). This surface impedance is inversely proportional to the frequency. Remark that transforming (6.8) and (6.9) back to the frequency domain for a frequency dependent circuit impedance and circuit admittance results in convolution integrals in the time domain.

Taking the derivative of (6.8) with respect to z and using (6.9) immediately leads to the one-dimensional wave equation satisfied by $V(z)$

$$\frac{\mathrm{d}^2V(z)}{\mathrm{d}z^2} - (R + j\omega L)(G + j\omega C)V(z) = 0, \tag{6.10}$$

the solution of which is

$$V(z) = Ae^{-jkz} + Be^{jkz}, \tag{6.11}$$

with the wavenumber k defined as

$$k = \beta - j\alpha = \sqrt{-(R + j\omega L)(G + j\omega C)}. \tag{6.12}$$

Just as in the case of plane waves in a lossy dielectric, the square root in (6.12) is chosen such that its real part, the propagation constant β, is positive and that its imaginary part is negative such that the attenuation constant α is positive. The voltage on a transmission line is the superposition of a voltage wave propagating in the positive z-direction (the wave with complex amplitude A) and a voltage wave propagating in the negative z-direction (the wave with complex amplitude B). It is easy to show that the current $I(z)$ also satisfies the same wave equation as the voltage

$$\frac{\mathrm{d}^2I(z)}{\mathrm{d}z^2} - (R + j\omega L)(G + j\omega C)I(z) = 0. \tag{6.13}$$

However, voltage and current are related through (6.8) and (6.9). Substituting (6.11) in (6.8) shows that

$$I(z) = Y_c(Ae^{-jkz} - Be^{jkz}) = \frac{1}{Z_c}(Ae^{-jkz} - Be^{jkz}), \tag{6.14}$$

with Z_c the *characteristic impedance* of the line and Y_c the *characteristic admittance*

$$Z_c = \frac{1}{Y_c} = \sqrt{\frac{R + j\omega L}{G + j\omega C}}. \tag{6.15}$$

The current on the transmission line is also the superposition of a wave propagating in the positive z-direction and a wave propagating in the negative z-direction. For the wave propagating in the positive z-direction the ratio of the

voltage to the current remains constant along the line and is equal to the characteristic impedance Z_c. This is also the case for the wave propagating in the negative z-direction except for a minus sign. This is a consequence of taking the positive z-direction as the reference direction for the current. It is important not to confuse the characteristic impedance or admittance with the circuit impedance and circuit admittance of the line, as defined through (6.8) and (6.9)! For a lossy transmission line, the wave number and the characteristic impedance are complex resulting in waves that are attenuated while propagating. In the sequel we will restrict the analysis to lossless lines with $R = G = 0$ and with real-valued L and C. In that case the wave number and the characteristic impedance are real

$$k = \omega\sqrt{LC}, \tag{6.16}$$

$$Z_c = R_c = \sqrt{\frac{L}{C}}. \tag{6.17}$$

The real characteristic impedance is denoted by R_c. On a lossless line voltage and current waves are in phase. When L and C are constant, the frequency independent signal propagation speed c is

$$c = \frac{1}{\sqrt{LC}}. \tag{6.18}$$

When the signal propagation speed does not depend on frequency, the signal propagation is said to be *non-dispersive*. When L and C depend on frequency, c (6.18) no longer represents the signal propagation velocity but only the phase velocity. The actual signal propagation velocity is given by the group velocity c_g defined as

$$c_g = \frac{d\omega}{dk}. \tag{6.19}$$

The signal propagation is now dispersive: different frequencies propagate at different velocities. In this chapter we will not go into detail at this point. The difference between phase and group velocity will be further discussed in Chapter 7.

6.3 Voltage reflection coefficient

Fig. 6.5 shows a transmission line of length d terminated by a complex load impedance Z_L. Generally speaking, this load impedance is the equivalent input impedance of the circuit the transmission line is connected to. Suppose e.g. that the transmission line is a coaxial cable connected to a transmitter antenna. We will prove in Chapter 8 that the antenna can be replaced by an equivalent impedance, the radiation impedance. The transmission line could also be a stripline on a board terminated by via-hole connected to a load resistor. In that case, the equivalent impedance of the via-hole connected to the load resistor forms the load impedance of the stripline. By convention, the origin of the z-axis is chosen such that it coincides with the place of the load, also called the *load impedance plane*. At the beginning of the line ($z = -d$), also referred to as the *input plane*, the transmission line is connected to a driving circuit represented by its Thevenin equivalent with voltage source V_g and generator

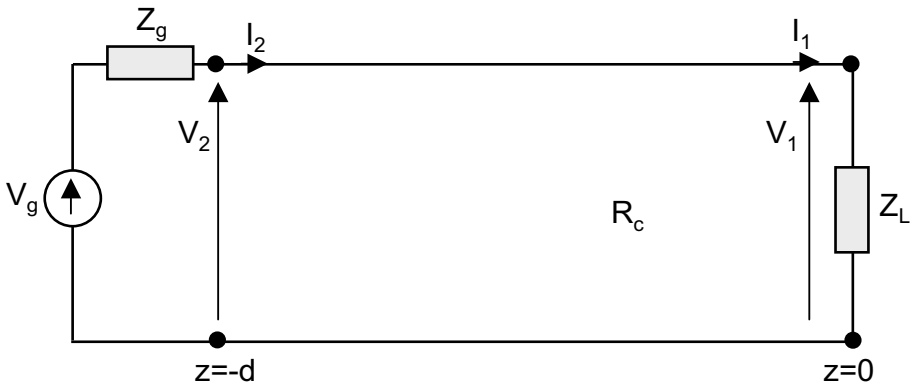


Figure 6.5: Transmission line with Thevenin generator and load impedance.

impedance Z_g.

The voltage and the current along the line are given by (6.11) and (6.14)

$$V(z) = Ae^{-jkz} + Be^{jkz}, \tag{6.20}$$

$$I(z) = \frac{1}{R_c}(Ae^{-jkz} - Be^{jkz}), \tag{6.21}$$

with

$$A = ae^{j\alpha}, \tag{6.22}$$

$$B = be^{j\beta} \tag{6.23}$$

(with β in (6.23) not to be confused with the propagation constant β which in the lossless case is equal to k).
The voltage V_1 and the current I_1 at $z = 0$ are not independent as

$$Z_L = \frac{V_1}{I_1} = \frac{V(z=0)}{I(z=0)} = R_c\frac{A+B}{A-B}. \tag{6.24}$$

From (6.24) the ratio of A to B can be derived as

$$\frac{B}{A} = \frac{b}{a}e^{j(\beta-\alpha)} = K_L = \frac{Z_L - R_c}{Z_L + R_c}. \tag{6.25}$$

K_L is the voltage reflection coefficient or for short, the *reflection coefficient* of the load. It is the ratio of the amplitude of the voltage wave reflected by the load to the amplitude of the voltage wave propagating towards the load. The characteristic impedance R_c defines a reference level for the impedance scale. For that reason the normalised impedance $Z'_L = Z_L/R_c$ is introduced such that

$$K_L = \frac{Z'_L - 1}{Z'_L + 1}. \tag{6.26}$$

Expression (6.26) shows that there is a one to one correspondence between the normalised load impedance and the reflection coefficient. The Smith chart, discussed in Section 6.8 exploits this relationship.

When the load impedance is equal to the characteristic impedance, the reflection coefficient is zero. In that case the line is *matched*. The energy of the wave propagating towards the load is completely dissipated in the load. When transmitting high bit rate signals, it is advantageous to use a matched transmission line to avoid unwanted reflections at the receiver. Table 6.1 shows how the value of the reflection coefficient varies when the load ranges from an open circuit ($Z_L = \infty$) to a short circuit ($Z_L = 0$). If the line is left open, the reflection coefficient $K_L = 1$, i.e. the incident voltage and the incident current are completely reflected ($B = A$) such that the voltage over the open end equals twice the incident voltage ($A + B = 2A$) and the current becomes zero ($A - B = 0$). If the line is short circuited, the reflection coefficient $K_L = -1$, i.e. the incident voltage and the incident current are still completely reflected ($B = -A$) but with opposite sign such that the voltage over the short circuit ($A + B = 0$) becomes zero and the current doubles ($A - B = 2A$). While the load impedance varies in amplitude between 0 and ∞, the amplitude of the reflection coefficient remains bounded, i.e. $|K_L| \leq 1$.

$Z_L(\Omega)$	K_L	VSWR
∞	1	∞
11 R_c	5/6	11
5 R_c	2/3	5
2 R_c	1/3	2
R_c	0	1
1/2 R_c	-1/3	2
1/5 R_c	-2/3	5
1/11 R_c	-5/6	11
0	-1	∞

Table 6.1: Load impedance, reflection coefficient and VSWR.

The load impedance is not necessarily passive. For a passive load $|K_L| \leq 1$, for an active load $|K_L| > 1$.

6.4 Input impedance

At the driver side of the transmission line it is important to know how to model the loaded transmission line as an equivalent impedance. The equivalent circuit is depicted in Fig. 6.6. The input impedance Z_i is given by

$$Z_i = \frac{V_2}{I_2} = \frac{V(-d)}{I(-d)} = R_c \frac{A\ e^{jkd} + B\ e^{-jkd}}{A\ e^{jkd} - B\ e^{-jkd}}. \tag{6.27}$$

Dividing the denominator and the numerator by A and using (6.25) yields the following expression for the normalised input impedance

$$Z_i' = \frac{Z_i}{R_c} = \frac{Z_L' + j\tan(kd)}{1 + jZ_L'\tan(kd)}. \tag{6.28}$$

This normalised input impedance depends on the normalised load impedance and on the normalised length of the transmission $kd = 2\pi d/\lambda$. The transmission line acts as an *impedance transformer* between the load and the driver. As

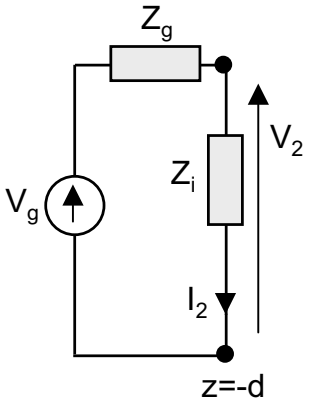


Figure 6.6: Equivalent circuit of the transmission showing the input impedance Z_i.

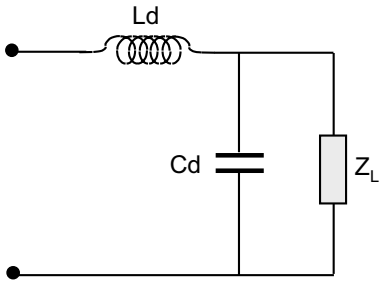


Figure 6.7: Equivalent circuit of a short transmission line.

(6.28) only depends on the tangent of kd, replacing d by $d + n(\lambda/2)$, with n an integer number, does not change the resulting input impedance. Hence, when the length of the line is a multiple of $\lambda/2$, the input impedance remains identical to the load impedance. The input impedance of a matched line($Z'_L = 1$) is equal to the characteristic impedance of that line.

For short transmission lines ($kd \ll 1$), $\tan(kd)$ can be replaced by kd and (6.28) reduces to

$$Z_i = \frac{Z_L + j(R_c kd)}{1 + jZ_L(kd/R_c)}. \tag{6.29}$$

Substitution of k and R_c by their values (6.16) and (6.17) finally yields

$$Z_i = \frac{Z_L + j\omega Ld}{1 + jZ_L\omega Cd}. \tag{6.30}$$

Neglecting terms in $(kd)^2$, this is the input impedance of the circuit shown in Fig. 6.7. The total distributed capacitance and distributed inductance of the line are collected into the lumped capacitance and inductance of the equivalent circuit of Fig. 6.7. The difference between $\tan(kd)$ and kd is about 0.1 for $d = \lambda/10$ and 0.01 for $d = \lambda/20$. Consequently, transmission line effects manifesting themselves by the difference between (6.28) and (6.30), certainly

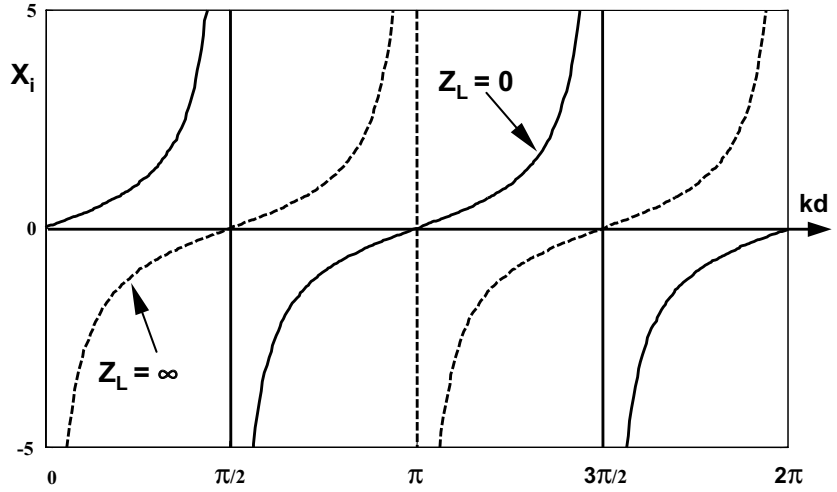


Figure 6.8: Input impedance of an open or a short circuited line.

become relevant when the length of the line exceeds $\lambda/10$.

When the transmission line is left open ($Z_L = \infty$) its input impedance becomes

$$Z_i' = \frac{1}{j \tan kd}, \tag{6.31}$$

while for a short circuited line

$$Z_i' = j \tan kd. \tag{6.32}$$

In both cases the input impedance is purely reactive $Z_i = jX_i$ (see Fig. 6.8). This implies that an open line or a short circuited line can be used to synthesise a capacitance or an inductance. In microwave and millimetre wave applications this is very useful as lumped capacitances or inductances are not easily fabricated at high frequencies.
Further remark that for a reactive load $Z_L = jX_L$, the input impedance always remains reactive, while a real load $Z_L = R_L$ does not automatically guarantee a real input impedance.

Knowledge of the input impedance allows to determine the voltage and the current at the driver ($z = -d$) as (see Fig. 6.6)

$$V(-d) = V_g \frac{Z_i}{Z_i + Z_g} = Ae^{jkd} + Be^{-jkd}, \tag{6.33}$$

$$I(-d) = \frac{V_g}{Z_i + Z_g} = \frac{1}{R_c}(Ae^{jkd} - Be^{-jkd}). \tag{6.34}$$

Hence, A and B are found to be

$$A = \frac{V_g}{2(Z_i + Z_g)}(Z_i + R_c)e^{-jkd}, \tag{6.35}$$

$$B = \frac{V_g}{2(Z_i + Z_g)}(Z_i - R_c)e^{jkd}. \tag{6.36}$$

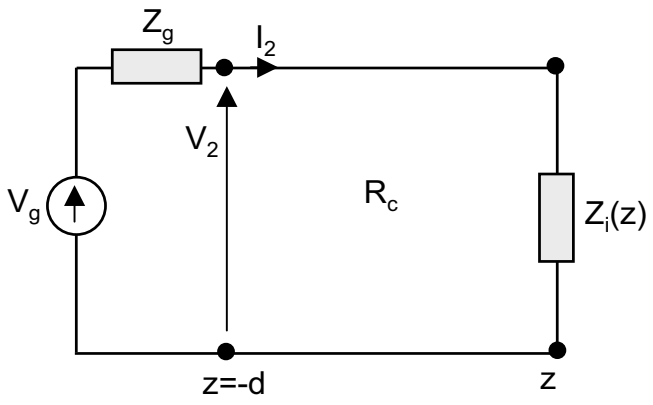


Figure 6.9: Generalised input impedance.

6.5 Generalised reflection coefficient

In the previous section the reflection coefficient K_L of the load was introduced. According to (6.25), K_L is the ratio of the amplitude of the voltage wave propagating away from (reflected by) the load to the amplitude of the voltage wave propagating towards the load. The generalised voltage reflection coefficient $K(z)$ can now be defined as the ratio of the amplitude of the wave propagating in the negative z-direction to the amplitude of the wave propagating in the positive z-direction at each point of the line

$$K(z) = \frac{Be^{jkz}}{Ae^{-jkz}} = \frac{B}{A}e^{2jkz}. \tag{6.37}$$

Taking into account (6.25), this becomes

$$K(z) = K_L e^{2jkz} = K_L e^{-2jk|z|}. \tag{6.38}$$

As the origin of the z-axis coincides with the load plane, z in (6.38) is negative. Remark that for a lossless line the amplitude of the reflection coefficient $K(z)$ remains constant, i.e. $|K(z)| = |K_L|$. Furthermore, the reflection coefficient is a periodic function of the wavelength with period $\lambda/2$. For a lossy line the amplitude of the reflection coefficient decreases exponentially with a factor $e^{-2\alpha\tau}$ when moving in the direction from the load to the driver over a distance τ.

In Fig. 6.6 the transmission line is replaced by its equivalent input impedance. In a similar way, we can only replace part of the line by its equivalent input impedance $Z_i(z)$ as shown in Fig. 6.9. This generalises the concept of input impedance. Proof that the following relationships exist between the generalised reflection coefficient and the generalised input impedance

$$K(z) = \frac{Z_i'(z) - 1}{Z_i'(z) + 1}, \tag{6.39}$$

$$Z_i'(z) = \frac{1 + K(z)}{1 - K(z)}. \tag{6.40}$$

6.6 Power flow

The time-average power transported in the positive z-direction is given by

$$P = \frac{1}{2} Re(I^* V). \tag{6.41}$$

For a *lossless* line, substitution of (6.20) and (6.21) in (6.41) shows that

$$P = \frac{1}{2}\frac{|A|^2}{R_c} - \frac{1}{2}\frac{|B|^2}{R_c}. \tag{6.42}$$

The first term is the power transported by the wave propagating towards the load. This is the incident power P_i. The second term represents the power transported by the wave propagating towards the generator. This is the reflected power P_r. Remark that (6.42) features no terms depending on both A and B. This is an example of *power orthogonality*. This means that the power of the wave resulting from the superposition of the incident wave and the reflected wave is equal to the algebraic sum of the powers of the individual waves. Such power orthogonality is typical for a lossless transmission line (or for that matter for plane waves propagating in a lossless medium). Prove that in the presence of losses, i.e. when k and Z_c are complex, additional terms, depending on both A and B, will occur in (6.42), destroying power orthogonality.
Expression (6.42) can be rewritten as

$$P = P_i - P_r = P_i(1 - |K_L|^2), \tag{6.43}$$

with

$$P_i = \frac{1}{2}\frac{|A|^2}{R_c}, \tag{6.44}$$

$$P_r = |K_L|^2 P_i. \tag{6.45}$$

If the line is matched, no power is reflected ($K_L = 0$). For an open line and a short circuited line, $|K_L| = 1$ and the power is totally reflected ($P_r = P_i$).

6.7 Standing waves and VSWR

The time-domain counterparts of (6.20) and (6.21) are

$$v(z,t) = a\cos(\omega t - kz + \alpha) + b\cos(\omega t + kz + \beta), \tag{6.46}$$

$$i(z,t) = \frac{1}{R_c}[a\cos(\omega t - kz + \alpha) - b\cos(\omega t + kz + \beta)]. \tag{6.47}$$

The incident and the reflected wave interfere along the transmission line and give rise to a *standing wave*. At the co-ordinates satisfying

$$-kz_n + \alpha = +kz_n + \beta + n2\pi, \qquad n = ..., -1, 0, 1, ... \tag{6.48}$$

this interference is constructive for the voltages waves and destructive for the current waves. At these points the amplitude of the sinusoidally varying voltage

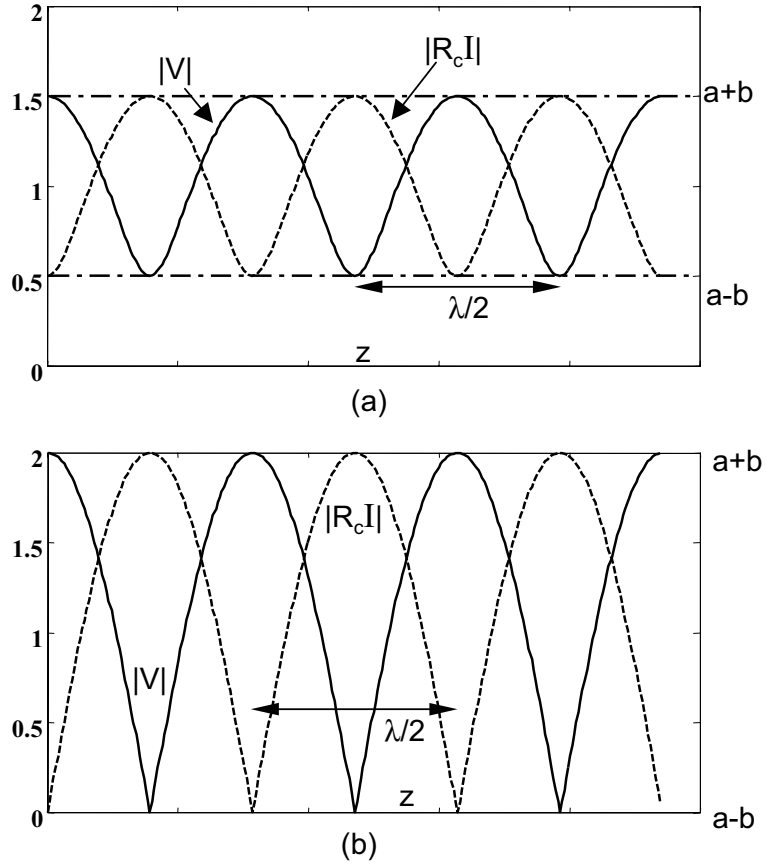


Figure 6.10: Standing wave pattern along a transmission line: for partial reflection ($a \neq b$) (a) and for total reflection ($a = b$) (b).

reaches a maximum equal to $(a+b)$ while the amplitude of the current is minimal $((a-b)/R_c)$. At the co-ordinates satisfying

$$-kz_n + \alpha = +kz_n + \beta + \pi + n2\pi, \qquad n = ..., -1, 0, 1, ... \tag{6.49}$$

the interference is destructive for the voltages waves and constructive for the current waves. At these points the amplitude of the voltage reaches a minimum equal to $(a-b)$ while the amplitude of the current is maximal $((a+b)/R_c)$. The distance between two maxima or two minima of the voltage (or the current) is equal to half a wavelength. A maximum and a minimum are separated by a quarter of a wavelength. The variation of the absolute value of the total voltage $|V(z)|$ and the absolute value of the total current $|I(z)|$ constitutes the standing wave pattern. Fig. 6.10 shows this standing wave pattern for partial reflection ($a \neq b$) and for total reflection ($a = b$). An important quantity is the *VSWR* or the *V*oltage *S*tanding *W*ave *R*atio S defined as

$$S = \frac{a+b}{a-b} = \frac{1+|K|}{1-|K|} = \frac{1+|K_L|}{1-|K_L|}. \tag{6.50}$$

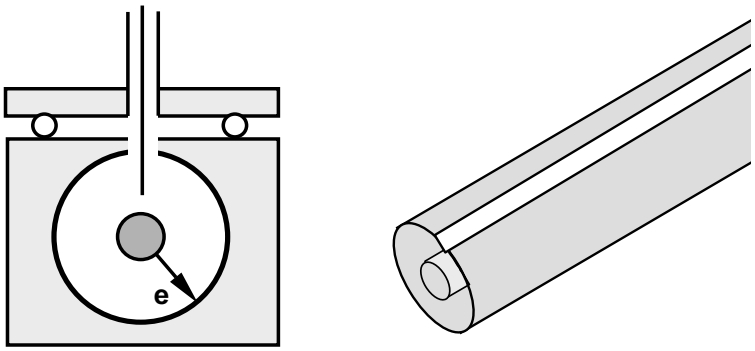


Figure 6.11: Standing-wave detector for a coaxial line.

The third column of Table 6.1 gives the VSWR-values corresponding to various values of the load impedance. Fig. 6.11 shows a standing-wave detector for a coaxial line. A small wire antenna penetrating through a z-directed slot in the outer conductor of a coaxial cable measures a signal proportional to the radial electric field in the coaxial cable. This is a direct measure for the voltage difference between the inner and the outer conductor. The probe is moved along the z-axis and the ratio between the maximum and the minimum signal detected by the probe is the VSWR (6.50). This immediately leads to absolute value of the reflection coefficient

$$|K| = |K_L| = \frac{S-1}{S+1}. \tag{6.51}$$

To completely determine K_L and hence the load impedance Z_L, we also need the phase of K_L. According to (6.25), the phase of K_L is equal to $\beta - \alpha$. From (6.48), this phase follows from the knowledge of the position of the n-th voltage maximum and the knowledge of the position of the load plane as

$$2kz_n = \alpha - \beta - n2\pi. \qquad n = ..., -1, 0, 1, ... \tag{6.52}$$

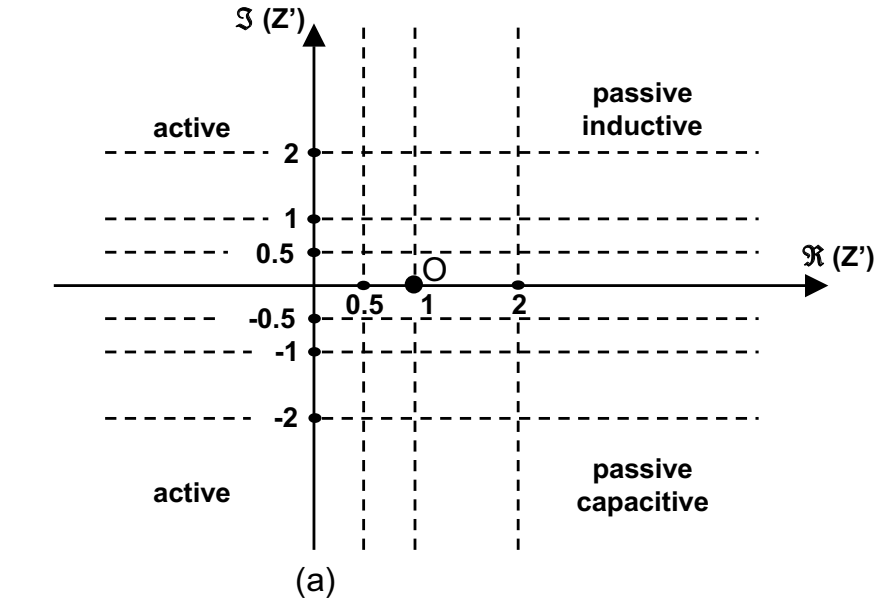
Similarly, from (6.49), knowledge of the position of the n-th voltage minimum and knowledge of the position of the load plane gives

$$2kz_n = \alpha - \beta - \pi - n2\pi. \qquad n = ..., -1, 0, 1, ... \tag{6.53}$$

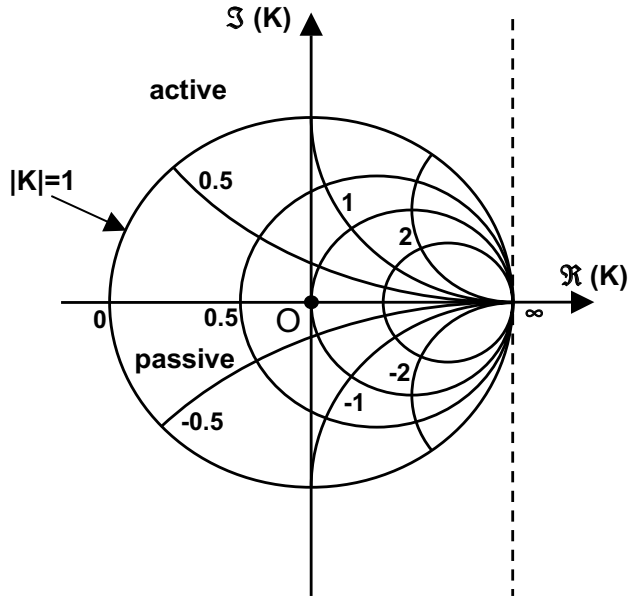
6.8 The Smith chart

6.8.1 Generalities

We are now ready to introduce the most important instrument to represent impedances and wave phenomena on transmission lines. This instrument is the Smith chart. Fig. 6.12a shows the complex impedance plane Z or more precisely the normalised impedance plane $Z' = Z/R_c$. Instead of representing an impedance in this plane, the complex reflection coefficient plane K is used with $K = (Z'-1)/(Z'+1)$. This reflection coefficient plane is depicted in Fig. 6.12b. The transformation between the Z'-plane and the K-plane is a bilinear transformation. We will not study the general properties of such a transformation here. The right half-plane (passive impedances) is transformed into the inside



(a)



(b)

Figure 6.12: Impedance plane (a) and reflection coefficient plane (b).

of the unit circle; the left half-plane (active impedances) is mapped outside this circle. Furthermore, the bilinear transformation maps lines of constant real or constant imaginary part of the impedance into circles in the K-plane. Circles corresponding to $\Re(Z') = cnt.$ have their origin on the real axis of the K-plane, while circles corresponding to $\Im(Z') = cnt.$ have their origin on the vertical line going through the point $(0, 1)$, i.e. a line parallel to the imaginary axis of the K-plane. The origin O of the K-plane corresponds to the point $Z' = (1, 0)$ of the normalised impedance plane. This is the point corresponding to a matched load ($Z = R_c$, $K = 0$). Fig. 6.12b shows three circles of constant real part for the impedance: $\Re(Z') = 0.5$, $\Re(Z') = 1$ and $\Re(Z') = 2$. For $R_c = 50\Omega$ and $Z = R + jX$ this corresponds to $R = 25\Omega$, 50Ω and 100Ω. Fig. 6.12b also shows six circles of constant imaginary part: $\Im(Z') = \pm 0.5$, $\Im(Z') = \pm 1$, $\Im(Z') = \pm 2$. For $R_c = 50\Omega$ this corresponds to $X = \pm 25\Omega$, $X = \pm 50\Omega$ and $X = \pm 100\Omega$. The real axis of the K-plane ($\Im(K) = 0$) corresponds to real impedances ($X = 0$). An open circuit ($Z = \infty$) corresponds to the point $(1, 0)$, while a short circuit ($Z = 0$) corresponds to the point $(-1, 0)$. Purely reactive impedances ($R = 0$) are mapped on the outer circle with radius 1, i.e. $|K| = 1$.

Although this might at first be rather confusing, the reflection plane representation is simultaneously used to represent normalised impedances Z' *as well as* normalised admittances $Y' = 1/Z'$ with $Y = Y'/R_c$. The above reasoning can be repeated starting from the admittance plane in place of the impedance plane. When using the reflection coefficient plane to represent impedances, the upper half of the K-plane corresponds to inductive loads and the lower half-plane to capacitive loads. However, when using the reflection coefficient plane to represent normalised admittances, this role is inverted: the upper part now stands for capacitive loads and the lower part for inductive loads. For impedances, the point $(1, 0)$ corresponds to an open circuit, as already remarked above. When working with admittances this point corresponds to $Y' = \infty$, i.e. a short circuit! In the same way the point $(-1, 0)$ either represents a short circuit ($Z' = 0$) or an open circuit $Y' = 0$.

The reflection coefficient plane equipped with circles of constant real and imaginary part for the impedance (or for the admittance) is the *Smith chart* shown in Fig. 6.13. The resolution of this chart for $R' > 2$ and/or $|X'| > 2$ is very limited. This shows that the reflection coefficient only changes marginally when $|Z'|$ becomes sufficiently larger than R_c. The Smith chart maps the impedance plane in such a way that prominence is given to impedances (admittances) the absolute value of which does not differ too much from the characteristic impedance R_c (or $1/R_c$ for admittances). As matching is very important at high frequencies, this is a desirable feature of the Smith chart. An arbitrary impedance or admittance can easily be located on the Smith chart using the circles of constant real or imaginary part. Point A on Fig. 6.13 e.g. corresponds to $Z' = 0.8 + j1.2$. The vector from the origin O to A is the complex reflection coefficient K with amplitude OA and with the phase ϕ given by the angle with the real axis.

As stated before, a microwave engineer prefers the Smith chart over the traditional impedance plane. Fig. 6.14 shows the variation of the impedance of a 50Ω chip resistor as a function of frequency, resulting from a measurement

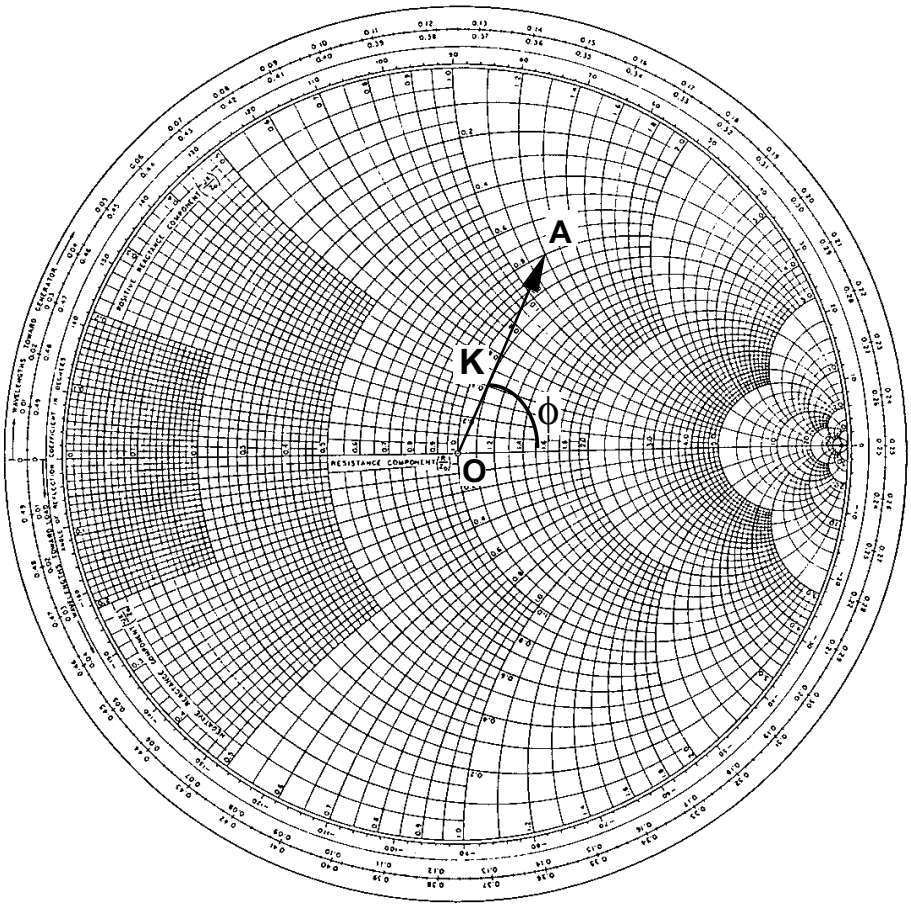


Figure 6.13: The Smith chart.

between 50 MHz and 5 GHz. The locus of the impedance starts close to the origin of the Smith chart. However, when increasing the frequency the impedance value clearly starts to deviate from its nominal 50Ω value. The impedance first becomes capacitive (lower half-plane) and its real part decreases. This is due to parasitic effects. The net capacitive effect then starts to decrease, due to the fact that the inductive parasitics gain in importance with increasing frequency. In the GHz range, the impedance locus again crosses the real axis. At the crossing point the impedance is real ($|K| = 0.375$, $Z = 23\Omega$). For even higher frequencies the inductive effect dominates (locus in the upper half-plane).

6.8.2 Elementary Smith chart applications

The impedance transformation (6.28) can easily be represented on the Smith chart. Point A on Fig. 6.15 represents the load impedance Z_L, with K_L the corresponding reflection coefficient. According to (6.38), the reflection coefficient at the input plane is

$$K(z = -d) = K_L e^{-2jkd}. \tag{6.54}$$

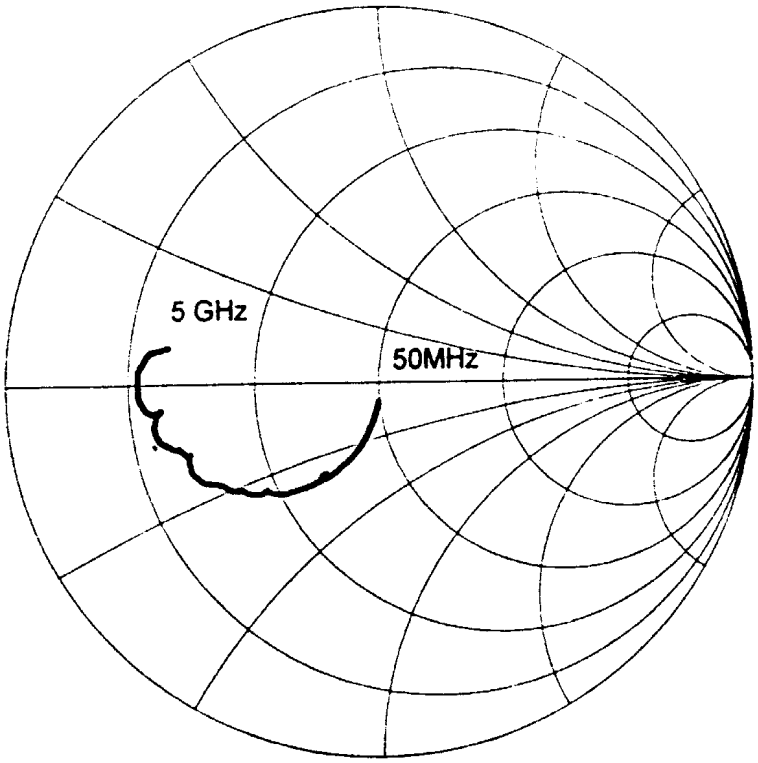


Figure 6.14: Impedance of a 50Ω chip resistor.

For a lossless line, the amplitude of the reflection coefficient remains constant. Hence, while moving from the load plane $z = 0$ towards the input plane $z = -d$, the generalised input impedance (6.40) will describe a circle with radius $|K_L|$ on the Smith chart. This circle is depicted in Fig. 6.15. According to (6.54), when moving from the load plane to the input plane, the phase decreases. This corresponds to counter clockwise movement along the circle. Moving from the input plane towards the load plane corresponds to clockwise movement along the circle. The directions "towards load" and "towards generator" are indicated on the outside of the Smith chart. Movement over a distance d corresponds to a change in phase of $2kd$, i.e. a half wave length corresponds to 360^0 or a full circle, confirming that all phenomena on a transmission line are periodic with period $\lambda/2$. At the outside of the Smith chart both a scale in degrees and a scale in wavelengths are added. The point B on Fig. 6.15 corresponds to the input impedance.

Substituting $d = \lambda/4$ in (6.28) yields $Z_i' = 1/Z_L'$, i.e. the *normalised* admittance Y' corresponding to a given *normalised* impedance Z' is obtained by moving over a quarter of a wavelength along a transmission line. On the Smith chart of Fig. 6.15 this implies that the point B', diametrically opposed to B, represents the normalised input admittance. Similarly, A' represents the normalised load admittance.

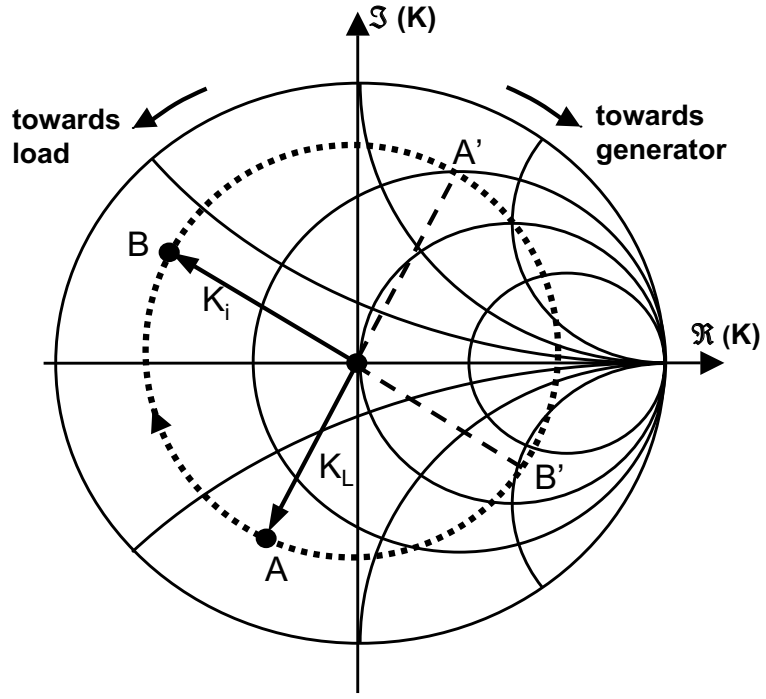


Figure 6.15: Determination of the input impedance and the input admittance.

From (6.20), (6.21) and (6.38) we easily derive that

$$|V(z)| = a|1 + K(z)|, \tag{6.55}$$

$$R_c|I(z)| = a|1 - K(z)|, \tag{6.56}$$

with $K(z)$ the generalised reflection coefficient. Fig. 6.16 shows an example of a reflection coefficient $K(z)$. The end point A of the vector $\overrightarrow{OA}$ representing K on the Smith chart moves on the dashed circle for different z-values. Similarly, the vector $\overrightarrow{OA'}$ represents $-K$ and A' moves on the same dashed circle as A. From (6.55) and (6.56) we see that the amplitude of the vector $\overrightarrow{BA}$ represents $|V(z)|/a$ and that the amplitude of the vector $\overrightarrow{BA'}$ represents $|R_cI(z)|/a$. It is now graphically very clear, that when A and A' simultaneously move on the dashed circle, that $|V|$ and $|R_cI|$ change periodically (with period $\lambda/2$) and that their maximum and minimum values are determined by the intersection points M_{max} and M_{min} of the circle and the real axis. When the voltage reaches a maximum, the current will be minimum and vice versa.

6.9 Matching

A transmission line is matched when the input impedance Z_i is equal to the characteristic impedance of the line. The importance of matching follows from the fact that according to (6.43) part of the incident power is reflected in the absence of matching. When a transmission line is feeding an antenna e.g., the overall system efficiency is drastically reduced when the power generated by

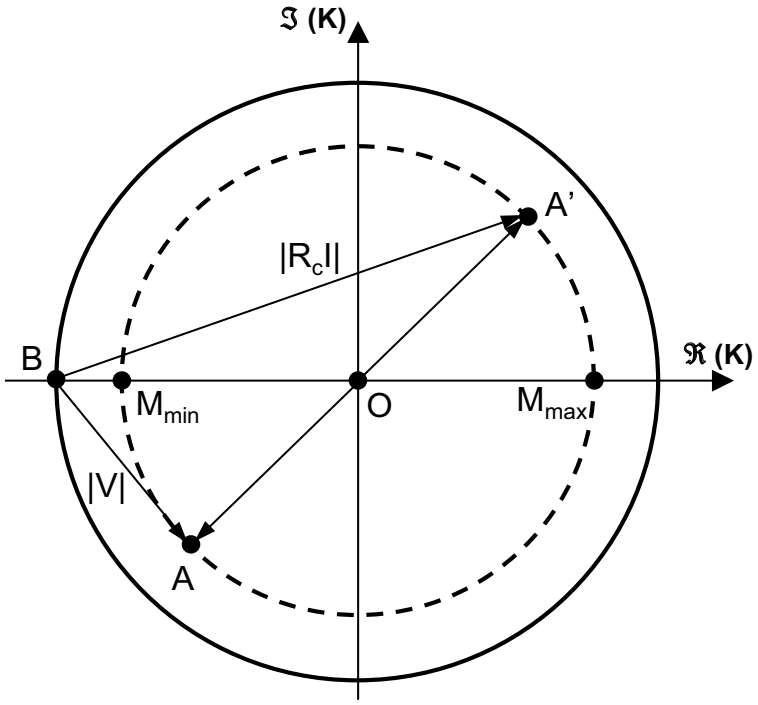


Figure 6.16: Interpretation of the standing wave pattern on the Smith chart.

the source does not completely reach the antenna but gets lost by a mismatch between antenna and generator. Signal reflection can also have very detrimental effects on high bit rate data transfer. Mismatch also entails a standing wave pattern resulting in higher maximum voltage values as compared to the matched situation. For power lines this could lead to dielectric breakdown. All this leads to the important topic of impedance matching and matching circuits.

An *impedance matching* or *tuning* circuit is often part of a much larger design (e.g. a feeding network for a microstrip patch antenna). The basic idea is illustrated in Fig. 6.17. An impedance matching circuit is placed between the feed line and the load impedance with $R_c \neq Z_L$. The matching circuit causes additional reflections to occur between the matching circuit and the load. However, if carefully designed, the original reflection towards the generator is now canceled out and the input impedance becomes R_c as shown in the equivalent circuit at the bottom of Fig. 6.17, i.e. *matching* is obtained.

This course does not leave enough room for the further study of matching concepts. The most important techniques used for (lossless) matching are

- lumped element matching;
- single stub matching;
- double stub matching;
- the quarter-wavelength transformer;
- multisection matching;
- tapered lines.

Lumped element matching, single and double stub matching and the use of a

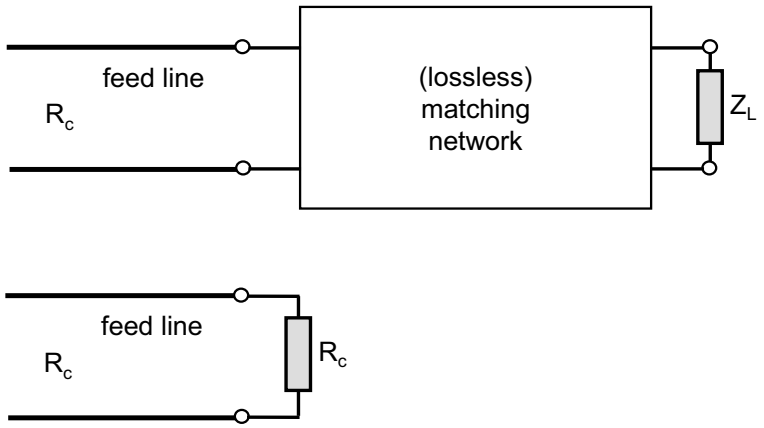


Figure 6.17: Impedance matching circuit.

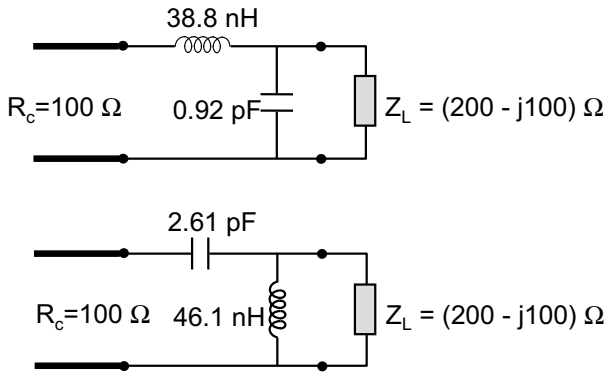


Figure 6.18: Example of lumped element matching.

quarter-wavelength transformer are single frequency or narrow band solutions. Multisection matching and tapered lines can be advantageously used to obtain matching over a larger frequency band.

Fig. 6.18 shows an example of lumped element matching. Discrete capacitances and inductances are used in a simple matching network. In the example, two solutions are shown to match a load $Z_L = (200 - j100)\Omega$ to a feed line with a characteristic impedance of 100Ω at a frequency of 500 MHz. Lumped elements can only be used at low frequencies.

In single and double stub matching, open or short circuited lines are put in parallel to the feed line to obtain matching. Remember that (see (6.31) or (6.32)) open or short circuited lines play the role of a reactance. Just to give the reader an idea of what we are talking about, Fig. 6.19a shows an example of single stub matching at 2 GHz (with the length of the stub and its distance from the load as design parameters) while Fig. 6.19b shows an example of double stub matching, also at 2 GHz (with the length of both stubs as design parameters, while the distance between the stubs takes a fixed value).

A quarter-wavelength transformer is a transmission line of length $\lambda/4$ placed between the load and the feed line as shown in Fig. 6.20. The quarter-wavelength

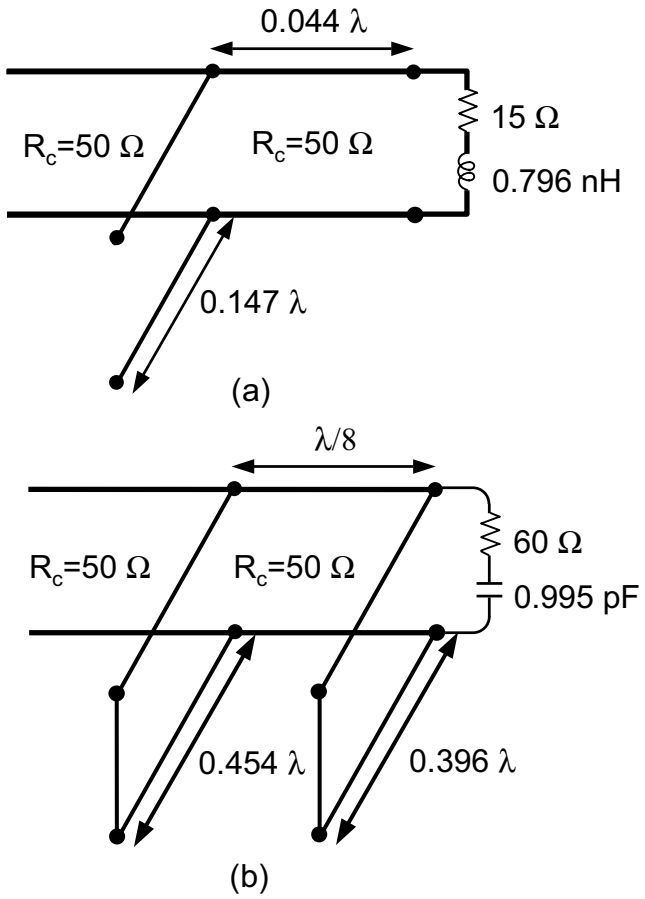


Figure 6.19: Single and double stub matching examples.

matching is restricted to real loads $Z_L = R_L$. It is very easy to prove that matching is obtained when the characteristic impedance of the quarter-wavelength line is chosen such that

$$R_{quarter} = \sqrt{R_L R_c}, \tag{6.57}$$

i.e. its characteristic impedance is the geometrical mean of the load resistance and the characteristic impedance of the feed line.

In all of the above cases exact matching is obtained at a single frequency. Only in a narrow frequency band around this frequency matching remains acceptable. Broadband solutions require more degrees of freedom. Fig. 6.21a shows the principle of multisection matching. Here, several transmission line sections with different impedances and/or lengths are used for matching. Another solution is shown in Fig. 6.21b. A tapered line, i.e. a line with geometrical properties that change as a function of z and hence with a continuously changing impedance, is used for matching.

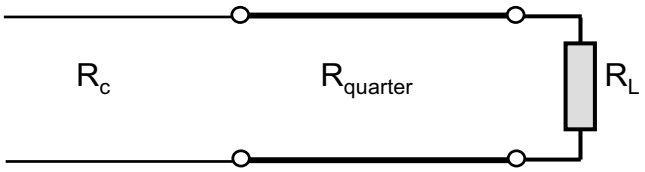


Figure 6.20: Quarter-wavelength transformer.

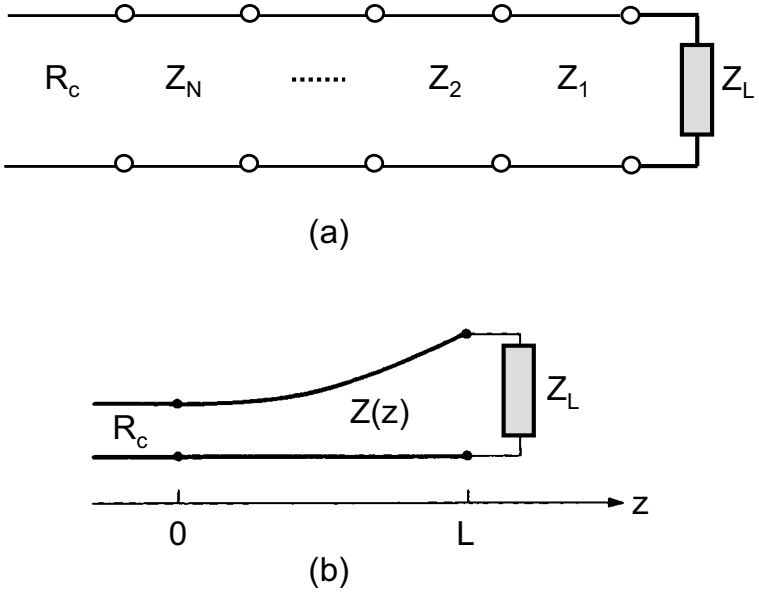


Figure 6.21: Principle of multisection matching (a) or matching using a tapered line (b).

6.10 Transients on a transmission line

The above analysis of transmission lines in the frequency domain is particularly suited for narrow band applications. When considering digital systems, transients play a central role. In this section, the reader will gain some first insights into the complex problem of transients on transmission lines. The effect of losses and dispersion (frequency dependent signal velocity) is not taken into account. Moreover, the analysis will be further restricted to real load impedances. In practice, transmission lines in digital systems connect non-linear drivers and receivers. This substantially increases the difficulty of the analysis.

The starting point of the time domain analysis is the time domain counterpart of (6.10) with $R = G = 0$, i.e.

$$\frac{\partial^2}{\partial z^2} v(z,t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} v(z,t) = 0, \tag{6.58}$$

with the frequency independent velocity c given by

$$c = \frac{1}{\sqrt{LC}}. \tag{6.59}$$

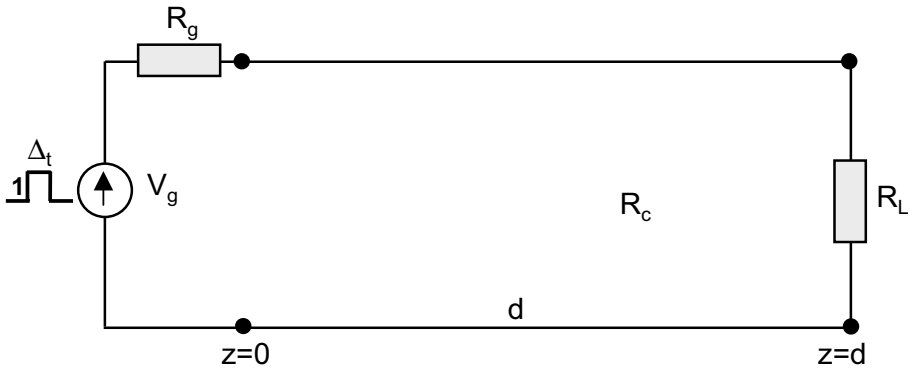


Figure 6.22: Pulses on a transmission line.

The general solution of (6.58) is

$$v(z,t) = v^+(z-ct) + v^-(z+ct), \tag{6.60}$$

with v^+ an arbitrary function of argument $z-ct$ representing a wave propagating in the positive z-direction and with v^- also an arbitrary function, now of argument $z+ct$, representing a wave propagating in the negative z-direction. By substituting (6.60) into (6.58), the reader will indeed verify that the proposed solution (6.60) satisfies the one-dimensional wave equation (6.58). According to (6.4) and (6.5) the corresponding solution for the current is

$$i(z,t) = \frac{1}{R_c}\left(v^+(z-ct) - v^-(z+ct)\right), \tag{6.61}$$

with $R_c = \sqrt{L/C}$, the already familiar characteristic impedance of the line.

Fig. 6.22 shows a source V_g with internal impedance R_g connected to a transmission line of length d terminated in a resistive load R_L. At $t = 0$, this source generates a unit amplitude pulse of duration Δ_t. Rise and fall time of the pulse are infinite. A bit stream is obtained by the superposition of such pulses. We now investigate what happens to this pulse when transmitted to the receiver, i.e. to the load impedance R_L. First we have to determine what happens at the generator at $t = 0$. According to (6.60) and (6.61) a voltage-current wave $(v^+, v^+/R_c)$ traveling in the positive z-direction will be generated. The equivalent circuit for this situation is depicted in Fig. 6.23. The time domain behaviour is completely different from what happens in sinusoidal regime. In sinusoidal regime, all transients have died out and the input impedance takes the value (6.28) derived in Section 6.4. However, when considering transients, the input impedance of the line is its characteristic impedance as in the time domain the line only supports an instantaneous wave for which the ratio of the voltage to the current is the characteristic impedance. Applying Ohm's law to the circuit of Fig. 6.23 we conclude that a voltage pulse with amplitude

$$\kappa = \frac{R_c}{R_c + R_g} \tag{6.62}$$

and duration Δ_t will start traveling along the line in the direction of the load. After a time $\tau = d/c$, with c the signal speed (6.59), this pulse reaches the load.

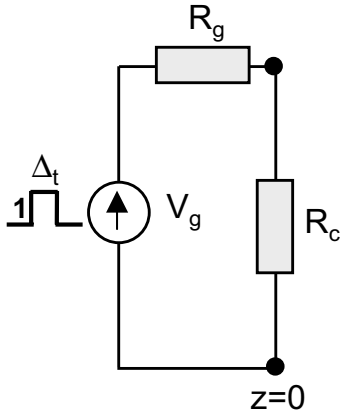


Figure 6.23: Equivalent circuit for the configuration of Fig. 6.22 at $t = 0$.

τ is known as the *time delay* of the line. As the ratio of the voltage to the current of this pulse is R_c and with the exception of a matched load ($R_L = R_c$), Ohm's law will not be satisfied at the load. Consequently a reflected voltage-current wave $(v^-, -v^-/R_c)$ will be generated. Ohm's law at the load will then be satisfied provided

$$\frac{v^+ + v^-}{\frac{1}{Rc}(v^+ - v^-)} = \frac{v_{total}}{i_{total}} = R_L. \tag{6.63}$$

Solving for v^- gives

$$v^- = \frac{R_L - R_c}{R_L + R_c} v^+ = K_L v^+, \tag{6.64}$$

i.e. the pulse reflected at the load equals K_L times the incident pulse, with K_L the familiar reflection coefficient of the load. The amplitude of the reflected pulse is κK_L. It will again take τ seconds for this pulse to travel back to the generator. As after generating the pulse the voltage of the generator drops to zero, the generator side of the transmission line reduces to a load impedance R_g. The reflection mechanism explained above for the load can now be applied to the generator impedance. This implies that the pulse incident on the generator will itself be reflected back to the load with reflection coefficient

$$K_g = \frac{R_g - R_c}{R_g + R_c}. \tag{6.65}$$

The complete reflection process now repeats itself until the amplitude of the pulse becomes negligible. Fig. 6.24 shows the history of the pulses at three different places: at the beginning of the line, in the middle of the line and at the end of the line. The amplitudes of the pulses in the figure correspond to $K_L = 0.6$ and $K_g = 0.3$. The displayed amplitude values are the actual voltages values divided by κ. The amplitude values at input plane are

$$\begin{aligned} v(z=0,t)/\kappa &= 1 && 0 < t < 2\tau, \\ &= K_L(1+K_g) && 2\tau \le t < 4\tau, \\ &= K_L(1+K_g)K_gK_L && 4\tau \le t < 6\tau, \end{aligned}$$

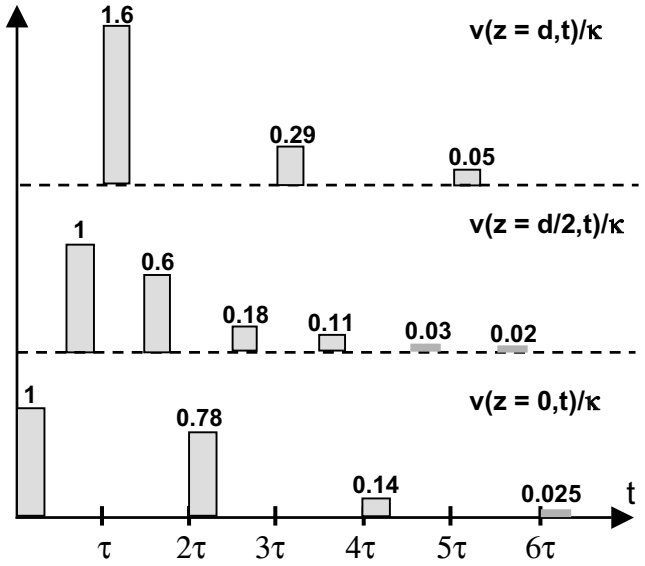


Figure 6.24: History of the pulses propagating over the transmission line.

$$\begin{aligned} &= K_L(1+K_g)K_g^2K_L^2 && 6\tau \le t < 8\tau, \\ &= \ldots && (6.66) \end{aligned}$$

To obtain these values, remark that the total voltage at $z = 0$ is always the sum of an incident and a reflected pulse (except of course for the very first pulse). This is expressed by the presence of the multiplicative factor $(1+K_g)$. Similarly, the amplitude values at the load plane are

$$\begin{aligned} v(z=d,t)/\kappa &= 0 && t < \tau, \\ &= (1+K_L) && \tau \le t < 3\tau, \\ &= (1+K_L)K_gK_L && 3\tau \le t < 5\tau, \\ &= (1+K_L)K_g^2K_L^2 && 5\tau \le t < 7\tau, \\ &= \ldots && (6.67) \end{aligned}$$

Here also, the voltage is always the sum of an incident and a reflected pulse. The multiplicative factor is now $(1 + K_L)$. The individual pulses can e.g. be detected in the middle of the line. Here the voltage values are

$$\begin{aligned} v(z=d/2,t)/\kappa &= 0 && t < \tau/2, \\ &= 1 && \tau/2 \le t < 3\tau/2, \\ &= K_L && 3\tau/2 \le t < 5\tau/2, \\ &= K_LK_g && 5\tau/2 \le t < 7\tau/2, \\ &= K_L^2K_g && 7\tau/2 \le t < 9\tau/2, \\ &= K_L^2K_g^2 && 9\tau/2 \le t < 11\tau/2, \\ &= \ldots && (6.68) \end{aligned}$$

A bit stream is a concatenation of pulses. It is clear that the echoes of a

single pulse when using a non-matched line can adversely interfere with subsequent pulses of the bit stream. When the time delay of the line remains small with respect to the duration Δ_t of a single bit or, equivalently, with respect to the inverse of the bit rate, echoes have died out before they can interfere with a next bit. This will be the case for sufficiently low bit rates and/or for sufficiently short lines. However, for increasing bit rates and increasing line lengths, the effect of signal reflection becomes important and matching will be necessary.

Suppose now that the source in Fig. 6.22 does not generate a pulse but a step signal with unit amplitude. Repeat the above reasoning to show that the voltage $v_L(t)$ over the load R_L takes the following values as a function of time

$$\begin{aligned} v_L(t) &= 0 && 0 < t < \tau, \\ &= \kappa(1+K_L) && \tau \leq t < 3\tau, \\ &= \kappa(1+K_L)(1+K_gK_L) && 3\tau \leq t < 5\tau, \\ &= \kappa(1+K_L)(1+K_gK_L+K_g^2K_L^2) && 5\tau \leq t < 7\tau, \\ &= \kappa(1+K_L)(1+K_gK_L+K_g^2K_L^2+K_g^3K_L^3+...) && t > 7\tau. \end{aligned} \quad (6.69)$$

To obtain the final voltage for $t \to \infty$, the geometrical series in the expression for $v_L(t)$ can be calculated explicitly, yielding

$$\begin{aligned} \lim_{t\to\infty} v_L(t) &= \kappa(1+K_L)\frac{1}{1-K_LK_g} \\ &= \frac{R_L}{R_L+R_g}. \end{aligned} \quad (6.70)$$

For a unit step signal, the limiting result for $t \to \infty$ must be identical to the result obtained for a DC voltage of 1V. At DC, transmission line effects do not come into play. Hence, the load voltage is indeed given by (6.70).

Good quality copy of the Smith chart

Chapter 7

Multiconductor Lines and Waveguides

7.1 Introduction

As announced in the Introduction of Chapter 6, the present chapter is entirely devoted to the field analysis of multiconductor lines and waveguides. From Chapter 6 we already know that multiconductor transmission lines and waveguides result from idealising actual interconnections by ignoring geometrical variations and/or variations in the material properties in one direction, the signal propagation direction. In the first part of this chapter, it is proved how this assumption leads to general eigenmode equations showing that only specific field patterns or modes can be supported by the waveguide. Each of these modes has its own propagation and attenuation constant.

The general eigenmode equations are first solved for multiconductor lines in the low frequency limit. This leads to the so-called quasi-TEM modes. The connection between the quasi-TEM field analysis and the capacitance and inductance matrix calculations of Chapters 3 and 4 is outlined and it is shown how the quasi-TEM analysis naturally leads to the telegrapher's equations for the voltages and currents propagating along the multiconductor lines. This provides the link with the transmission line analysis of the previous chapter.

Next, general expressions for the TEM, TE and TM modes in a homogeneously filled waveguide are derived and the mode orthogonality concept is introduced. Mode solutions that are neither TEM, TE or TM are called hybrid modes. Their analysis falls outside the scope of this course.

The last part of this chapter is devoted to the detailed analysis of a number of waveguide examples: the parallel-plate waveguide, the rectangular waveguide, the coaxial cable, the microstrip and the stripline. Optical waveguides such as the slab waveguide and the optical fibre are not treated here. They will be discussed in the "Optoelectronics" course.

The larger part of this chapter deals with lossless waveguides. In practice, dielectric losses and conductor losses will attenuate the propagating signals. In the section devoted to the parallel-plate waveguide, a detailed discussion is provided on how to determine these losses starting from the solution for the lossless problem. This is the typical perturbational approach used to assess losses in

multiconductor lines and waveguides. This approach is then also applied to the rectangular waveguide and to the coaxial cable.

Particular attention is also devoted to the important notion of mode dispersion, i.e. the fact that the velocity of the signal propagation is frequency dependent, except for the TEM and quasi-TEM modes. This leads to the distinction between the phase velocity and the group velocity. Here again advantage is taken of the analytical results for the parallel-plate waveguide to introduce the notion of dispersion relation and to show how the group and phase velocity derive from this dispersion relation.

For a detailed and very general analysis of waveguides and transmission lines, the reader is referred to two books "Electromagnetic and Circuit Modelling of Multiconductor Transmission Lines" by N. Faché, F. Olyslager and D. De Zutter (1993) and "Electromagnetic Waveguides and Transmission Lines" by F. Olyslager (1999), both published in the Oxford Engineering Science Series and to the many journal papers by the authors of this course published in the past ten years.

7.2 General eigenmode equations

In this chapter, as in the previous one, we select the z-axis to be the signal propagation direction. This signal propagation direction is also called longitudinal direction. The cross-sectional plane is known as the transversal plane. All material properties are independent of z but still depend upon the transversal co-ordinate $\boldsymbol{\rho} = x\mathbf{u}_x + y\mathbf{u}_y$. We restrict the analysis to isotropic media. All field components are written as the sum of a longitudinal part and a transversal part

$$\mathbf{e}(\mathbf{r}) = e_z(\mathbf{r})\mathbf{u}_z + \mathbf{e}_t(\mathbf{r}), \tag{7.1}$$

$$\mathbf{h}(\mathbf{r}) = h_z(\mathbf{r})\mathbf{u}_z + \mathbf{h}_t(\mathbf{r}). \tag{7.2}$$

Just as in the case of plane waves (Chapter 5), we search for solutions of the sourceless Maxwell's equations. These solutions are the *eigenmodes* of the waveguide. Similarly, plane waves are the eigenmodes of homogeneous infinite space. Substituting (7.1) and (7.2) in Maxwell's equations yields

$$\nabla_t \times e_z(\mathbf{r})\mathbf{u}_z + \mathbf{u}_z \times \frac{\partial}{\partial z}\mathbf{e}_t(\mathbf{r}) = -j\omega\mu(\boldsymbol{\rho})\mathbf{h}_t(\mathbf{r}), \tag{7.3}$$

$$\nabla_t \times \mathbf{e}_t(\mathbf{r}) = -j\omega\mu(\boldsymbol{\rho})h_z(\mathbf{r})\mathbf{u}_z, \tag{7.4}$$

$$\nabla_t \times h_z(\mathbf{r})\mathbf{u}_z + \mathbf{u}_z \times \frac{\partial}{\partial z}\mathbf{h}_t(\mathbf{r}) = j\omega\epsilon(\boldsymbol{\rho})\mathbf{e}_t(\mathbf{r}), \tag{7.5}$$

$$\nabla_t \times \mathbf{h}_t(\mathbf{r}) = j\omega\epsilon(\boldsymbol{\rho})e_z(\mathbf{r})\mathbf{u}_z. \tag{7.6}$$

Conduction currents are included by replacing ϵ by $\epsilon + \frac{\sigma}{j\omega}$. To tackle the above set of coupled partial differential equations with respect to z, remark that it must be possible to write each field component $\mathbf{e}_t(\mathbf{r})$, $e_z(\mathbf{r})$, $\mathbf{h}_t(\mathbf{r})$ and $h_z(\mathbf{r})$ as the product of a factor only depending on z and a factor depending on $\boldsymbol{\rho}$. The z-dependence must be chosen such that in each of the equations, the z-dependence of all terms in that equation is identical. If this turns out to be possible, separate equations will be obtained for the z-dependence and for the transversal $\boldsymbol{\rho}$-dependence. This approach is known as the separation of variables

procedure and is most often used when analytically solving differential equations in special co-ordinate systems. Suppose that the z-dependence of e_z in (7.3) is given by $f(z)$ and that of $\mathbf{e}_t$ in the same equation by $g(z)$. The separation of variables procedure applied to (7.3)–(7.6) implies that

$$e_z \sim f(z), \qquad \mathbf{e}_t \sim g(z), \tag{7.7}$$
$$h_z \sim g(z), \qquad \mathbf{h}_t \sim f(z), \tag{7.8}$$

and

$$\frac{d}{dz} f(z) \sim g(z), \tag{7.9}$$
$$\frac{d}{dz} g(z) \sim f(z). \tag{7.10}$$

The symbol $\sim$ stands for "proportional to". Taking the derivative of (7.9) with respect to z and using (7.10) shows that

$$\frac{d^2}{dz^2} f(z) \sim f(z). \tag{7.11}$$

A similar equation holds for $g(z)$. The general solution to (7.11) is

$$f(z) = Ae^{-j\gamma z} + Be^{j\gamma z}. \tag{7.12}$$

This z-dependence is identical to that of plane waves propagating along the z-direction. Eqn. (7.12) shows that the eigenmode solutions of (7.3)–(7.6) take the form

$$\begin{aligned} \mathbf{e}(\mathbf{r}) &= \mathbf{E}(\boldsymbol{\rho})e^{-j\gamma z} = \mathbf{E}(\boldsymbol{\rho})e^{-j\beta z}e^{-\alpha z}, \\ \mathbf{h}(\mathbf{r}) &= \mathbf{H}(\boldsymbol{\rho})e^{-j\gamma z} = \mathbf{H}(\boldsymbol{\rho})e^{-j\beta z}e^{-\alpha z}. \end{aligned} \tag{7.13}$$

This is the typical form of an eigenmode of the waveguide or multiconductor line. The factor $\gamma = \beta - j\alpha$ in the exponential is the propagation factor of the eigenmode, $\mathbf{E}(\boldsymbol{\rho})$ and $\mathbf{H}(\boldsymbol{\rho})$ are the modal field distributions. The real part of γ, β, is the propagation constant, while minus the imaginary part of γ, α, is the attenuation constant. Provided β is positive, (7.13) represents an eigenmode propagating in the positive z-direction. The result obtained for $f(z)$ in (7.12) is a combination of an eigenmode propagating in the positive z-direction and another one propagating in the negative z-direction. The time domain counterpart of (7.13) is

$$\mathbf{e}(\mathbf{r}, t) = \Re[\mathbf{E}(\boldsymbol{\rho})e^{-j\gamma z + j\omega t}] = \Re[\mathbf{E}(\boldsymbol{\rho})e^{-j\beta z + j\omega t}]e^{-\alpha z}, \tag{7.14}$$

showing that the eigenmode propagates at speed

$$v_p = \frac{\omega}{\beta}. \tag{7.15}$$

In general v_p will be frequency dependent. In that case the eigenmode is said to be *dispersive*: different frequencies propagate at different speeds. This results in the broadening of pulses and ultimately (together with the attenuation) limits the distance over which bits can be propagated before they are smoothened out

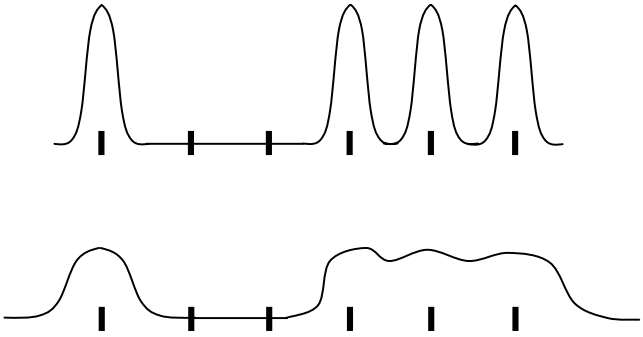


Figure 7.1: Effect of eigenmode dispersion on a propagating bit sequence.

and become indistinguishable from neighbouring bits as illustrated in Fig. 7.1. The speed v_p is the *phase velocity*. This is not the speed of signal propagation nor of the propagation of energy. The correct signal and energy speed is the *group velocity*

$$v_g = \frac{1}{\frac{d\beta}{d\omega}} = \frac{d\omega}{d\beta}. \tag{7.16}$$

We will come back to this difference between phase and speed velocity when discussing the parallel-plate waveguide. For non-dispersive eigenmodes $v_g = v_p$. The reader easily checks from (7.15) and (7.16) that $v_g = v_p$ automatically leads to $\beta = \omega/V$ with $v_g = v_p = V$ and where V takes a constant frequency independent value. Plane waves in homogeneous space are also non-dispersive and V is the speed of light, $V = \frac{1}{\sqrt{\epsilon\mu}}$.
Substitution of (7.13) in (7.3)–(7.6) yields

$$\nabla_t \times E_z(\boldsymbol{\rho})\mathbf{u}_z - j\gamma\mathbf{u}_z \times \mathbf{E}_t(\boldsymbol{\rho}) = -j\omega\mu(\boldsymbol{\rho})\mathbf{H}_t(\boldsymbol{\rho}), \tag{7.17}$$

$$\nabla_t \times \mathbf{E}_t(\boldsymbol{\rho}) = -j\omega\mu(\boldsymbol{\rho})H_z(\boldsymbol{\rho})\mathbf{u}_z, \tag{7.18}$$

$$\nabla_t \times H_z(\boldsymbol{\rho})\mathbf{u}_z - j\gamma\mathbf{u}_z \times \mathbf{H}_t(\boldsymbol{\rho}) = j\omega\epsilon(\boldsymbol{\rho})\mathbf{E}_t(\boldsymbol{\rho}), \tag{7.19}$$

$$\nabla_t \times \mathbf{H}_t(\boldsymbol{\rho}) = j\omega\epsilon(\boldsymbol{\rho})E_z(\boldsymbol{\rho})\mathbf{u}_z. \tag{7.20}$$

One easily verifies that when γ, $\mathbf{E}_t(\boldsymbol{\rho})$, $\mathbf{H}_t(\boldsymbol{\rho})$, $E_z(\boldsymbol{\rho})$ and $H_z(\boldsymbol{\rho})$ is a solution to (7.17)–(7.20), $-\gamma$, $\mathbf{E}_t(\boldsymbol{\rho})$, $-\mathbf{H}_t(\boldsymbol{\rho})$, $-E_z(\boldsymbol{\rho})$ and $H_z(\boldsymbol{\rho})$ is also a solution. With each mode propagating in the positive z-direction corresponds a mode propagating in the negative z-direction with a propagation factor with opposite sign. This property is known as "bidirectionallity". It does not automatically hold when replacing the isotropic material with anisotropic one.

Analytical solutions to the above equations only exist for a few simple geometries. In this chapter, three such examples, the parallel-plate waveguide, the coaxial cable and the rectangular waveguide will be treated in detail.

7.3 Multiconductor transmission lines

7.3.1 Geometry of the problem

The general geometry considered in this section is depicted in Fig. 7.2. Fig.

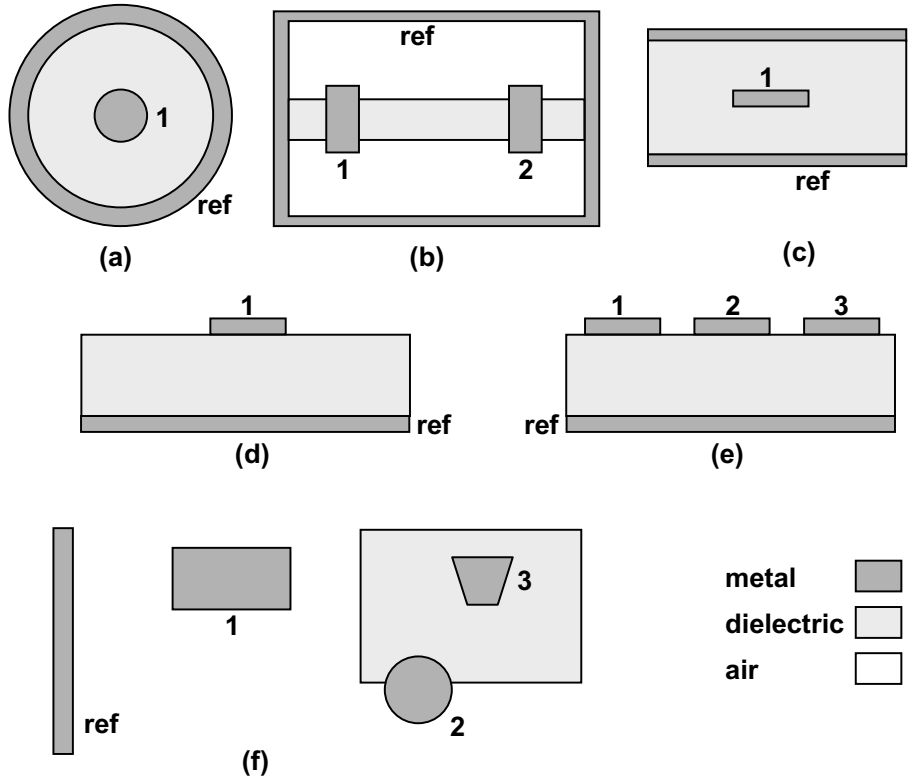


Figure 7.2: Cross-sections of general multiconductor transmission lines: traditional coaxial line (a), coaxial line with square cross-section and two signal conductors (b), stripline (c), microstrip line (d), three coupled microstrip lines (e), arbitrary cross-section (f).

7.2 shows the cross-sections of a few examples of multiconductor transmission lines. To keep the analysis tractable, we suppose that the cross-section consists of piecewise homogeneous, isotropic and lossless dielectrics and of perfect conductors. When treating the parallel-plate waveguide, later in this chapter, it will be shown how skin-effect losses in the conductors and how the imperfect nature of the dielectrics can be accounted for. The considered configurations can be closed (the coaxial cable in Fig. 7.2a, the coaxial two signal line example of Fig. 7.2b and the stripline of Fig. 7.2c), half-open (the single microstrip of Fig. 7.2d and the coupled microstrips of Fig. 7.2e) or completely open (the example of Fig. 7.2f). A common and for the further analysis important feature of all configurations, is that at least two separate conductors can be identified such that at DC the signal current has a return path. One of the conductors plays the role of reference conductor, as also indicated on the figure.

7.3.2 Quasi-TEM equations

The starting point of the analysis is the eigenmode equations (7.17)–(7.20). To underline the absence of losses, γ is replaced by β and the curl operators

acting on $E_z\mathbf{u}_z$ and $H_z\mathbf{u}_z$ are expressed in terms of the gradient operator

$$\nabla_t E_z(\boldsymbol{\rho}) + j\beta\mathbf{E}_t(\boldsymbol{\rho}) = -j\omega\mu(\mathbf{u}_z \times \mathbf{H}_t(\boldsymbol{\rho})), \tag{7.21}$$

$$\nabla_t \times \mathbf{E}_t(\boldsymbol{\rho}) = -j\omega\mu H_z(\boldsymbol{\rho})\mathbf{u}_z, \tag{7.22}$$

$$\nabla_t H_z(\boldsymbol{\rho}) + j\beta\mathbf{H}_t(\boldsymbol{\rho}) = j\omega\epsilon(\mathbf{u}_z \times \mathbf{E}_t(\boldsymbol{\rho})), \tag{7.23}$$

$$\nabla_t \times \mathbf{H}_t(\boldsymbol{\rho}) = j\omega\epsilon(\boldsymbol{\rho})E_z\mathbf{u}_z. \tag{7.24}$$

These equations are complemented by the corresponding divergence equations

$$\nabla_t \cdot (\epsilon\mathbf{E}_t(\boldsymbol{\rho})) - j\beta\epsilon E_z(\boldsymbol{\rho}) = 0, \tag{7.25}$$

$$\nabla_t \cdot (\mu\mathbf{H}_t(\boldsymbol{\rho})) - j\beta\mu H_z(\boldsymbol{\rho}) = 0. \tag{7.26}$$

To simplify the notation, the $\boldsymbol{\rho}$-dependence of ϵ and μ is left out.

As stated in the Introduction, the full-wave analysis of (7.21)–(7.24) is in general very complex and will not be pursued here. A good understanding of multiconductor lines can already be obtained by a *quasi-TEM* or *low-frequency* analysis. The meaning of quasi-TEM will become clear later on. The low-frequency analysis is based on the expansion of each field component and of the propagation constant β as a power series in the circular frequency ω

$$\beta(\omega) = \beta_0 + \beta_1\omega + \beta_2\omega^2 + \dots, \tag{7.27}$$

$$\mathbf{E}_t(\omega) = \mathbf{E}_{t0} + \mathbf{E}_{t1}\omega + \mathbf{E}_{t2}\omega^2 + \dots, \tag{7.28}$$

$$E_z(\omega) = E_{z0} + E_{z1}\omega + E_{z2}\omega^2 + \dots, \tag{7.29}$$

$$\mathbf{H}_t(\omega) = \mathbf{H}_{t0} + \mathbf{H}_{t1}\omega + \mathbf{H}_{t2}\omega^2 + \dots, \tag{7.30}$$

$$H_z(\omega) = H_{z0} + H_{z1}\omega + H_{z2}\omega^2 + \dots. \tag{7.31}$$

Since the frequency domain is the Fourier transformation with respect to the time domain and since in the time domain all quantities are real, $\beta(\omega) = -\beta^*(-\omega)$. Hence, $\beta(\omega)$ is an odd function in ω, implying that $\beta_0, \beta_2, \dots$ are zero and that the first term in (7.27) is proportional to ω. Without going into further details here, a similar reasoning for the field components shows that their dominant term is of zeroth-order i.e. $\mathbf{E}_{t0}, \mathbf{H}_{t0}, E_{z0}$ and H_{z0} are all different from zero. A more detailed explanation, in particular in the presence of losses, can be found in the second book cited in the Introduction. We will now show that the knowledge of the zeroth-order contributions for the fields together with β_1 suffices to predict the behaviour of multiconductor transmission lines at sufficiently low frequencies. The meaning of "sufficiently low" will become clear when discussing the validity range of the present analysis.

The next step is to insert the above power series expansions into the field equations (7.21)–(7.26) and to identify terms of the same order in ω at both sides of each equation. We turn our attention to the dominant zeroth-order contributions. The zeroth-order electric field satisfies

$$\nabla_t E_{z0}(\boldsymbol{\rho}) = 0, \tag{7.32}$$

$$\nabla_t \times \mathbf{E}_{t0}(\boldsymbol{\rho}) = 0, \tag{7.33}$$

$$\nabla_t \cdot (\epsilon\mathbf{E}_{t0}(\boldsymbol{\rho}) = 0. \tag{7.34}$$

For the zeroth-order magnetic field similar equations are found

$$\nabla_t H_{z0}(\boldsymbol{\rho}) = 0, \tag{7.35}$$

$$\nabla_t \times \mathbf{H}_{t0}(\boldsymbol{\rho}) = 0, \tag{7.36}$$

$$\nabla_t \cdot (\mu\mathbf{H}_{t0}(\boldsymbol{\rho})) = 0. \tag{7.37}$$

From (7.32) and (7.35) follows that E_{z0} and H_{z0} must be constant throughout space. As this is impossible for energy reasons, it follows that $E_{z0} = 0$ and $H_{z0} = 0$. This explains why the low-frequency analysis is also called the quasi-TEM analysis. The zeroth-order electric and magnetic field have no longitudinal field component. The zeroth-order fields are completely transversal, hence, *T*ransversal *E*lectro*M*agnetic or *TEM*. This TEM character is only valid for sufficiently low frequencies and the modal solution is a *quasi-TEM* solution. As a matter of fact, for lossless media, it can be shown that the actual longitudinal field components E_z and H_z are of order $O(\omega^2)$.
The zeroth-order electric field problem is the static potential problem analysed in Chapter 3. The electric field can be derived from an electric potential

$$\mathbf{E}_{t0}(\boldsymbol{\rho}) = -\nabla_t \phi(\boldsymbol{\rho}), \tag{7.38}$$

which combined with (7.34) shows that

$$\nabla_t^2 \phi(\boldsymbol{\rho}) = 0. \tag{7.39}$$

The electric potential must be constant on each of the conductors. One of the conductors is selected to be the zero electric potential reference conductor (remember that each configuration counts at least two separate conductors). The constant electric potential on the other conductors is V_m with $m = 1, 2, ...M$, with $M + 1$ the total number of conductors and with conductor $M + 1$ taking the role of reference conductor. The total charge on each conductor is denoted as q_m. Solution of the appropriate Laplace problem (see Sections 3.7 and 3.9.2) leads to

$$\mathcal{Q} = \mathbf{C}\, \mathcal{V}, \tag{7.40}$$

with $\mathcal{V}$ and $\mathcal{Q}$, $M \times 1$ column vectors with elements V_m and q_m and with $\mathbf{C}$ the $M \times M$ capacitance matrix.

The zeroth-order magnetic field problem is the static problem analysed in Chapter 4. For two-dimensional configurations it was proved in Section 4.9 that the magnetic field can be derived from a magnetic potential

$$\mathbf{h}_t(\boldsymbol{\rho}) = \frac{1}{\mu} \nabla_t \psi(\boldsymbol{\rho}) \times \mathbf{u}_z, \tag{7.41}$$

which combined with (7.36) shows that

$$\nabla_t^2 \psi(\boldsymbol{\rho}) = 0. \tag{7.42}$$

The magnetic potential or flux function must also be constant on each of the conductors. The same conductor as in the electrostatic case is selected to be the zero magnetic potential reference conductor. The constant magnetic potential on the other conductors is F_m with $m = 1, 2, ...M$. The total current flowing on each conductor is denoted as I_m. Solution of the appropriate equivalent electrostatic problem (see Section 4.9) leads to

$$\mathcal{F} = \mathbf{L}\, \mathcal{I}, \tag{7.43}$$

with $\mathcal{F}$ and $\mathcal{I}$, $M \times 1$ column vectors with elements F_m and I_m and with $\mathbf{L}$ the $M \times M$ inductance matrix. This inductance matrix is the inverse of the capacitance matrix $\mathbf{C_m}$ of the equivalent electrostatic problem.

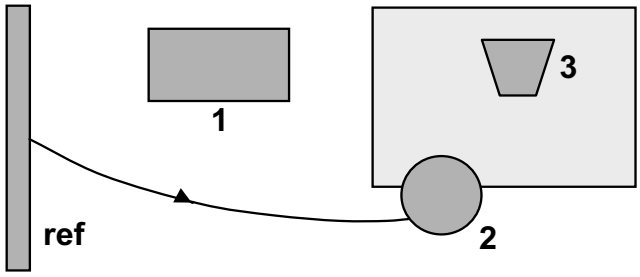


Figure 7.3: Integration path to obtain the voltage-flux relationship.

7.3.3 Quasi-TEM modes

Up to now we have considered the zeroth-order electrostatic and magnetostatic problem. Strictly speaking they only apply for $\omega = 0$ and are independent of each other. Moreover, in the series expansion process and by only retaining the zeroth-order fields, the propagation constant β is no longer present in the zeroth-order equations. It would thus seem that not too much progress has been made in solving the general equations (7.21)–(7.24). This is however not the case! To grasp this, let us first turn back to (7.21) and integrate both sides of this equation along an arbitrary line in the transversal plane between the reference conductor and conductor m. Fig. 7.3 shows an example of such an integration path. The resulting integral is

$$\int_{ref}^{m} \nabla_t E_z \cdot \mathbf{dl} + j\beta \int_{ref}^{m} \mathbf{E}_t \cdot \mathbf{dl} = -j\omega\mu \int_{ref}^{m} (\mathbf{u}_z \times \mathbf{H}_t) \cdot \mathbf{dl}. \tag{7.44}$$

The first integral, $E_z(m) - E_z(ref)$ is always zero as E_z is tangential to the conductor surfaces. For the second and third integral, the transversal fields are replaced by their zeroth-order approximation and from (7.38) and (7.41) it is easily seen that (7.44) yields

$$j\beta V_m = j\omega F_m. \tag{7.45}$$

This is a relationship between the electric and the magnetic potential. Its meaning can best be understood by considering its time domain differential equation counterpart. Remark that the factor $-j\beta$ stems from the derivation with respect to z of the $e^{-j\beta z}$-dependence, while the factor $j\omega$ stems from the derivation with respect to time of the $e^{j\omega t}$-dependence. This implies that (7.45) can be rewritten as

$$\frac{\partial}{\partial z} v_m(z,t) = -\frac{\partial}{\partial t} f_m(z,t), \tag{7.46}$$

where capital symbols were replaced by their non-capital counterparts to emphasise that (7.46) applies to the time domain. Although the conductor potentials are constant in a single cross-section, they still vary in the longitudinal direction. Following Maxwell, these variations are linked to magnetic flux variations as a function of time.
Secondly, we turn our attention to (7.23). This equation is integrated over the boundary contour of each conductor m as illustrated in Fig. 7.4

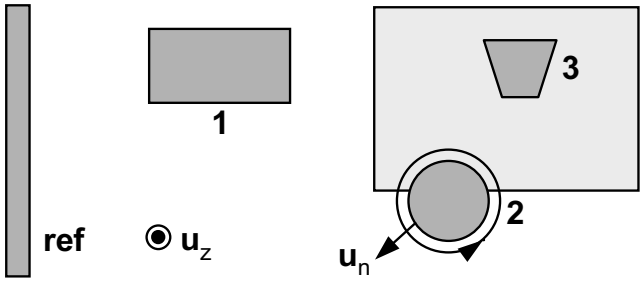


Figure 7.4: Integration path to obtain the current-charge relationship.

$$\oint_m \nabla_t H_z \cdot \mathbf{dl} + j\beta \oint_m \mathbf{H}_t \cdot \mathbf{dl} = j\omega\epsilon \oint_m (\mathbf{u}_z \times \mathbf{E}_t) \cdot \mathbf{dl}. \tag{7.47}$$

Obviously, the first integral is zero. From Ampère's law with $E_{z0} = 0$ and again only using the zeroth-order fields, the second integral is seen to be the total current I_m flowing on conductor m, while the third integral is recognised to be q_m/ϵ with q_m the total charge on conductor m per unit of length in the z-direction. The final result is

$$j\beta I_m = j\omega q_m, \qquad \frac{\partial}{\partial z} i_m(z,t) = -\frac{\partial}{\partial t} q_m(z,t). \tag{7.48}$$

This expresses that changes of the total current along the longitudinal direction are linked to the changes in charge accumulation as a function of time.

We are now ready to take the final step in the quasi-TEM analysis that must allow us to determine the still unknown value(s) of the propagation constant(s) β. Substituting (7.40) and (7.43) into (7.45) (or (7.46)) and (7.48) finally shows that

$$\frac{\partial}{\partial z}\mathcal{V}(z,t) = -\mathbf{L}\frac{\partial}{\partial t}\mathcal{I}(z,t), \tag{7.49}$$

$$\frac{\partial}{\partial z}\mathcal{I}(z,t) = -\mathbf{C}\frac{\partial}{\partial t}\mathcal{V}(z,t), \tag{7.50}$$

or in the frequency domain

$$\frac{\partial}{\partial z}\mathcal{V}(z,\omega) = -j\omega\mathbf{L}\mathcal{I}(z,\omega), \tag{7.51}$$

$$\frac{\partial}{\partial z}\mathcal{I}(z,\omega) = -j\omega\mathbf{C}\mathcal{V}(z,\omega). \tag{7.52}$$

These equations are completely similar to the telegrapher's equations (6.8) and (6.9) derived for a transmission line in Section 6.2 of the previous chapter. As a matter of fact, these equations are the telegrapher's equations for a set of coupled lines. The lumped element representation for an elementary section of a single transmission line, as depicted in Fig. 6.2, can be extended to the multiconductor case. Suppose e.g. that $M = 2$, i.e. a configuration with two signal conductors and a reference conductor (as in the example of Fig. 7.2b). Fig. 7.5 shows the lumped element representation for this case.

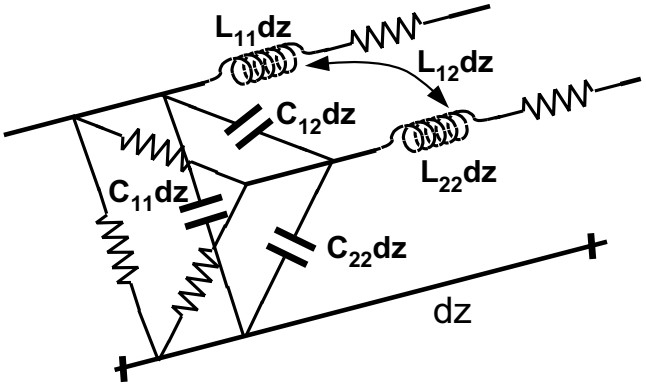


Figure 7.5: Lumped element representation of two coupled lines.

The first-order propagation constants β_1, with $\beta = j\omega\beta_1$, are the eigenvalues of

$$(\mathbf{LC})\ \mathcal{V} = {\beta_1}^2\mathcal{V}, \tag{7.53}$$

$$(\mathbf{CL})\ \mathcal{I} = {\beta_1}^2\mathcal{I}. \tag{7.54}$$

One can proof that the eigenvalues of **LC** and **CL** are identical but this is not the case for their eigenvectors. From linear algebra and from the fact that $(\mathbf{LC})^T = (\mathbf{C})^T(\mathbf{L})^T = (\mathbf{CL})$, we know that the eigenvectors of (7.53) are orthogonal to the eigenvectors of (7.54).

Equations (7.53) and (7.54) and the above mathematical analysis allow the following physical interpretation. In the quasi-TEM approximation, a multiconductor transmission line configuration consisting of $M + 1$ conductors, has M different eigenmodes. The propagations constants $\beta_m = j\omega\beta_{1m}$, with $m = 1, 2, 3, ..., M$ of these modes follow from the eigenvalues ${\beta_1}^2$ of the product of the capacitance matrix and the inductance matrix. According to (7.15) and (7.16) $v_p = v_g = \frac{1}{\beta_{1m}}$. The quasi-TEM modes are non-dispersive. This is of course again only true for the frequency range in which the quasi-TEM approximation remains valid. Each mode m has its own characteristic field distribution. In the quasi-TEM limit and for perfect conductors, the longitudinal field components e_z and h_z are negligible. The characteristic electric field distribution in the cross-section $\mathbf{E}_{tm}(\boldsymbol{\rho})$ follows from the solution of an electrostatic potential problem in the cross-section. For eigenmode m this electrostatic problem is the problem for which the potentials of the conductors are the elements of eigenvector $\mathcal{V}_m$ of (7.53). The characteristic magnetic field distribution in the cross-section $\mathbf{H}_{tm}(\boldsymbol{\rho})$ follows from the solution of a magnetostatic potential problem in the cross-section. For eigenmode m this magnetostatic problem is the problem for which the currents on the respective conductors are the elements of eigenvector $\mathcal{I}_m$ of (7.54). Finally, remark that voltages, currents and fields still depend on z. According to (7.13) the field patterns are identical in each cross-section except for the propagation factor $e^{-j\omega\beta_1 z}$ (when only considering modes propagating in the positive z-direction).

More generally speaking, for time domain signal analysis, the above eigenmode analysis remains valid, but the longitudinal z-dependence and time dependence now satisfy (7.49) and (7.50). Whereas the above analysis started

from Maxwell's equations, from a circuit point of view, we have succeeded in capturing the cross-sectional properties of the system in the capacitance and inductance matrix. Once they have been determined, the interconnection system can be described in terms of voltages and currents which are not only a continuous function of time but also of the longitudinal co-ordinate z. Voltages and currents are related through the telegrapher's equations. These telegrapher's equations can be handled in the same way as presented in the previous chapter for the single signal line case ($M = 1$), although concepts such as characteristic impedance and matching become more complicated. Moreover, solving the telegrapher's equations also allows to calculate crosstalk between signal lines.

It is of course possible to excite the multiconductor lines in such a way that the voltage distribution does not coincide with one of the eigenmodes. The orthogonality of the voltage and current eigenvectors can be used to expand such an arbitrary voltage-current distribution as a sum of eigenmodes.

When the conductors are embedded in a homogeneous medium, we have already proved in Section 4.9 that the inductance and the capacitance matrix are no longer independent but satisfy $\mathbf{LC} = \mathbf{CL} = \epsilon\mu\mathbf{I}$. Consequently, in that particular case all the eigenmodes have the same propagation velocity $v = \frac{1}{\sqrt{\epsilon\mu}} = \frac{c}{\sqrt{\epsilon_r\mu_r}}$, i.e. the velocity of light in the homogeneous medium (with c the velocity of light in free space). In the non-homogeneous case, and for non-magnetic media (which is most often the case in practice), the velocity v of each mode is also expressed as $v = \frac{c}{\sqrt{\epsilon_{r,eff}}}$ where $\epsilon_{r,eff}$ is the so-called *effective relative permittivity* of the background medium. It is some kind of "average" dielectric constant as experienced by the fields of a particular mode and this average will in general be different for each of the modes.

We conclude this section on the quasi-TEM analysis of PEC multiconductor lines with a remark on its validity range. For a more detailed analysis, to be found in second book cited in the Introduction, the quasi-TEM analysis remains valid provided

$$4\pi^2(\frac{\Delta}{|\lambda|})^2\frac{|\phi - \sqrt{\epsilon\mu}\psi|}{|\psi|} \ll 1, \tag{7.55}$$

with λ the wavelength in the material and with Δ a typical dimension of the cross-section of the multiconductor structure. The factor involving the potentials gives a measure for their relative difference at each point. Provided the contrast between the different media is not too large, this factor will remain bounded. Hence, the rule of thumb to be derived from (7.55) is that the wavelength must remain large enough with respect to a typical cross-sectional dimension. This essentially means that in the cross-section the whole field problem remains static while propagation effects only manifest themselves in the longitudinal direction. Once the cross-sectional dimensions become comparable to the wavelength (typically $\frac{\lambda}{10}$) the quasi-static analysis will no longer be valid and dynamical effects in the cross-section come into play.

7.4 TEM modes

As explained in the previous section, the designation quasi-TEM derives from the fact that the longitudinal field components are negligible at low frequencies. One can now ask the question whether the modal field equations (7.17)–(7.20) allow solutions for which E_z and H_z are *exactly* zero and this independent of frequency. For $E_z = H_z = 0$, (7.17)–(7.20) become

$$-j\gamma \mathbf{u}_z \times \mathbf{E}_t(\boldsymbol{\rho}) = -j\omega\mu(\boldsymbol{\rho})\mathbf{H}_t(\boldsymbol{\rho}), \tag{7.56}$$

$$\nabla_t \times \mathbf{E}_t(\boldsymbol{\rho}) = 0, \tag{7.57}$$

$$-j\gamma \mathbf{u}_z \times \mathbf{H}_t(\boldsymbol{\rho}) = j\omega\epsilon(\boldsymbol{\rho})\mathbf{E}_t(\boldsymbol{\rho}), \tag{7.58}$$

$$\nabla_t \times \mathbf{H}_t(\boldsymbol{\rho}) = 0. \tag{7.59}$$

Remark that (7.57) and (7.59), together with the corresponding divergence equations, are identical to the equations (7.33), (7.34), (7.36) and (7.37) satisfied by $\mathbf{E}_{t0}(\boldsymbol{\rho})$ and $\mathbf{H}_{t0}(\boldsymbol{\rho})$ in the quasi-TEM analysis. Vector multiplying (7.56) with $\mathbf{u}_z$ and comparison with (7.59) shows that both equations can only be simultaneously satisfied provided

$$\frac{j\gamma}{j\omega\epsilon(\boldsymbol{\rho})} = \frac{-j\omega\mu(\boldsymbol{\rho})}{-j\gamma}. \tag{7.60}$$

This is possible for materials for which $\epsilon(\boldsymbol{\rho})\mu(\boldsymbol{\rho})$ assumes a constant value over the complete cross-section. In practice this will only be the case for a homogeneous cross-section with constant ϵ and μ. Hence, we conclude that *TEM modes*, i.e. *T*ransversal *E*lectro*M*agnetic modes with zero longitudinal field components only exist in homogeneous media. From (7.60), these modes all have the same propagation factor $\gamma = \omega\sqrt{\epsilon\mu}$. Substituting this value of γ in (7.56) or (7.58) shows that

$$\mathbf{H}_t(\boldsymbol{\rho}) = \sqrt{\frac{\epsilon}{\mu}}\left(\mathbf{u}_z \times \mathbf{E}_t(\boldsymbol{\rho})\right) = \frac{1}{Z_c}\left(\mathbf{u}_z \times \mathbf{E}_t(\boldsymbol{\rho})\right), \tag{7.61}$$

with $Z_c = \sqrt{\frac{\mu}{\epsilon}}$ the characteristic impedance of the TEM modes. Remark that the field expressions for the TEM modes are completely similar to those for a plane wave: in each point of the cross-section electric and magnetic field are perpendicular to each other and to their propagation direction. The propagation speed is the speed of light in the medium and the ratio of the field amplitudes is equal to the characteristic impedance of the medium. Typical examples of structures supporting a TEM mode are the coaxial cable of Fig. 7.3a and the stripline of Fig. 7.3c (provided the dielectric is homogeneous).

The above analysis is only valid provided the cross-section counts at least two separate conductors, one of which is the reference conductor ($M \geq 1$). If only a single conductor is present, a TEM mode cannot be supported as, in particular at low frequencies, the current has no return path. In Section 7.8 the rectangular waveguide will be discussed, clearly showing that no TEM mode exists in that case. A homogeneous structure with $M + 1$ conductors supports M TEM modes. The coaxial line with two inner signal conductors of Fig. 3.20 is an example of a structure supporting two TEM modes ($M = 2$).

It must be emphasised that the above analysis is *not* restricted to low frequencies as was the case for the quasi-TEM analysis of multiconductor lines in Section 7.3. For increasing frequencies, the quasi-TEM solutions of the previous

section lose their TEM character as the longitudinal field components E_z and H_z are no longer negligible. This is not the case for pure TEM modes. For homogeneous media the analysis of Section 7.3 confirms the fact that TEM modes all have the same velocity and that the ratio of the field amplitudes is equal to the characteristic impedance of the medium. Indeed, we have proved in Section 4.9 that the inductance and the capacitance matrix for a homogeneous medium satisfy $\mathbf{L} \cdot \mathbf{C} = \mathbf{C} \cdot \mathbf{L} = \epsilon\mu\mathbf{1}$. In that case (7.53) and (7.54) show that $\epsilon\mu = \beta_1^2$, i.e. all modes have the same velocity $v = 1/\sqrt{\epsilon\mu}$. Substituting $j\beta = j\omega\sqrt{\epsilon\mu}$ into (7.45) shows that on the conductors $\sqrt{\epsilon\mu}V_m = F_m$. For a completely homogeneous medium this implies that $\sqrt{\epsilon\mu}\phi(\boldsymbol{\rho}) = \psi(\boldsymbol{\rho})$. Taking the gradient on both sides of this equality and taking into account (7.38) and (7.41), confirms (7.61). Finally remark that for $\sqrt{\epsilon\mu}\phi(\boldsymbol{\rho}) = \psi(\boldsymbol{\rho})$ the left-hand member of (7.55) is exactly zero.

7.5 TE and TM modes

In the previous section we concluded that TEM modes only exist for homogeneous media. In this section we again restrict ourselves to homogeneous media. However, the number of perfect conductors is now arbitrary. Hence the analysis applies to the coaxial cable (Fig. 7.2a) or the stripline (Fig. 7.2c) ($M = 1$), to the example of Fig. 3.20 ($M = 2$), but also to the rectangular waveguide ($M = 0$) (Fig. 6.1e) and even to free space. The starting point of the analysis is (7.21)–(7.24) with a constant value for ϵ and μ throughout space. As we do not need to restrict ourselves to real ϵ and μ values here, $j\beta$ is replaced by $j\gamma$, yielding

$$\nabla_t E_z(\boldsymbol{\rho}) + j\gamma \mathbf{E}_t(\boldsymbol{\rho}) = -j\omega\mu(\mathbf{u}_z \times \mathbf{H}_t(\boldsymbol{\rho})), \tag{7.62}$$

$$\nabla_t \times \mathbf{E}_t(\boldsymbol{\rho}) = -j\omega\mu H_z(\boldsymbol{\rho})\mathbf{u}_z, \tag{7.63}$$

$$\nabla_t H_z(\boldsymbol{\rho}) + j\gamma \mathbf{H}_t(\boldsymbol{\rho}) = j\omega\epsilon(\mathbf{u}_z \times \mathbf{E}_t(\boldsymbol{\rho})), \tag{7.64}$$

$$\nabla_t \times \mathbf{H}_t(\boldsymbol{\rho}) = j\omega\epsilon(\boldsymbol{\rho})E_z\mathbf{u}_z. \tag{7.65}$$

Eliminating $\mathbf{H}_t$ between (7.62) and (7.64) gives

$$\mathbf{E}_t = \frac{1}{k^2 - \gamma^2}(-j\gamma\nabla_t E_z + j\omega\mu\mathbf{u}_z \times \nabla_t H_z), \tag{7.66}$$

while eliminating $\mathbf{E}_t$ between the same equations yields

$$\mathbf{H}_t = \frac{1}{k^2 - \gamma^2}(-j\omega\epsilon\mathbf{u}_z \times \nabla_t E_z - j\gamma\nabla_t H_z), \tag{7.67}$$

with $k^2 = \omega^2\epsilon\mu$. This shows that the transversal field components can be expressed in terms of the longitudinal ones. Moreover, we know that each cartesian field component satisfies the sourceless wave equation or Helmholtz equation. Taking into account (7.13), the wave equations for E_z and H_z become

$$\nabla_t^2 E_z + (k^2 - \gamma^2)E_z = 0, \tag{7.68}$$

$$\nabla_t^2 H_z + (k^2 - \gamma^2)H_z = 0. \tag{7.69}$$

From (7.66)–(7.69) it is clear that the eigenmodes for a *homogeneous* medium can be subdivided into two independent classes: the *TE modes* or *T*ransversal

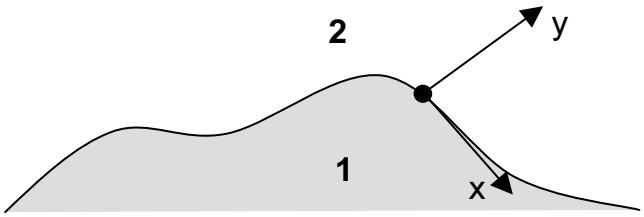


Figure 7.6: Boundary conditions between two media.

*E*lectric modes with $E_z = 0$, i.e. with no longitudinal electric field component and the *TM modes* or *T*ransversal *M*agnetic modes with no longitudinal magnetic field component, i.e. $H_z = 0$.

The reader might argue that the above analysis remains valid for a piecewise homogeneous medium. However, this is not the case as the boundary conditions at the interface between two media entail a coupling between E_z and H_z. To understand this, consider the interface between two media shown in Fig. 7.6 with the co-ordinate axes chosen such that the x-axis is tangential to the boundary and with the y-axes normal to it. From (7.66) it then follows that

$$E_x = \frac{1}{k^2 - \gamma^2}(-j\gamma \frac{\partial}{\partial x} E_z - j\omega\mu \frac{\partial}{\partial y} H_z). \tag{7.70}$$

Suppose now that a TM mode would exist implying that $H_z = 0$. In that case (7.70) reduces to

$$E_x = \frac{-j\gamma}{k^2 - \gamma^2} \frac{\partial}{\partial x} E_z. \tag{7.71}$$

E_z is itself a tangential field component and remains continuous and this will also be the case for the tangential derivative $\frac{\partial}{\partial x} E_z$. Hence, E_x can only be continuous if k^2 remains constant, i.e. the medium must be homogenous. A similar reasoning can of course be used to show that no TE modes can exist in a piecewise homogeneous medium.

From (7.62)-(7.69) the final equations for the TE modes are

$$\mathbf{E}_t = \frac{j\omega\mu}{k^2 - \gamma^2}(\mathbf{u}_z \times \nabla_t H_z), \tag{7.72}$$

$$\mathbf{H}_t = \frac{\gamma}{\omega\mu}(\mathbf{u}_z \times \mathbf{E}_t), \tag{7.73}$$

$$E_z = 0. \tag{7.74}$$

The transversal fields can be derived from the scalar function H_z which plays the role of some kind of potential. H_z is found by solving

$$\nabla_t^2 H_z + (k^2 - \gamma^2) H_z = 0, \tag{7.75}$$

$$\frac{\partial H_z}{\partial n} = 0 \qquad \text{on the conductor surfaces.} \tag{7.76}$$

Boundary condition (7.76) directly follows from (7.72) and from the fact that the tangential electric field $\mathbf{u}_n \times \mathbf{E}_t$ at the PEC conductors must be zero. The Helmholtz equation (7.75) together with the zero normal derivative boundary

condition (7.76) is known as the *Neumann* problem. This problem will only have non-zero solutions for specific values of γ. These values are the eigenvalues of the problem and the propagation factors of the eigenmodes. Eqn. (7.73) shows that, just as for a plane wave, transversal electric and magnetic field are still perpendicular to each other. The ratio of their amplitudes is given by the TE mode impedance Z_{TE}

$$Z_{TE} = \frac{\omega\mu}{\gamma}. \tag{7.77}$$

In general, $Z_{TE} \neq \sqrt{\mu/\epsilon}$.

From (7.62)-(7.69) the final equations for the TM modes are

$$\mathbf{E}_t = -\frac{\gamma}{\omega\epsilon}(\mathbf{u}_z \times \mathbf{H}_t), \tag{7.78}$$

$$\mathbf{H}_t = -\frac{j\omega\epsilon}{k^2-\gamma^2}(\mathbf{u}_z \times \nabla_t E_z), \tag{7.79}$$

$$H_z = 0. \tag{7.80}$$

The transversal fields can now be derived from the scalar function E_z. E_z is found by solving

$$\nabla_t^2 E_z + (k^2-\gamma^2)E_z = 0, \tag{7.81}$$

$$E_z = 0 \qquad \text{on the conductor surfaces.} \tag{7.82}$$

The Helmholtz equation (7.81) together with the fact that the function must be zero on the boundaries is known as the *Dirichlet* problem. Here again non-zero solutions will only exist for specific values of γ which, in general, will be different from the eigenvalues for the Neumann problem. Transversal electric and magnetic field are perpendicular to each other. The ratio of their amplitudes is now given by the TM mode impedance Z_{TM}

$$Z_{TM} = \frac{\gamma}{\omega\epsilon}. \tag{7.83}$$

The reader easily verifies that in the TM case $\mathbf{u}_n \times \mathbf{E}_t = 0$.

The propagation factors γ of the TE and the TM modes are in general not proportional to the frequency and these modes are dispersive. In non-homogeneous media no decoupling between TE and TM modes is possible. Mode solutions have both non-zero E_z and H_z components. These modes are called *hybrid* modes.

7.6 Mode orthogonality

In Section 7.3.3 we remarked that the eigenmodes of (7.53) and (7.54) can be used as a basis for expanding arbitrary voltage-current waves propagating along a multiconductor line. In the general waveguide case, the eigenmodes for the fields are given by (7.13). It can be proved that for two such eigenmode solutions, $(\mathbf{E}_1, \mathbf{H}_1)$ with propagation factor γ_1 and $(\mathbf{E}_2, \mathbf{H}_2)$ with propagation factor γ_2, the following orthogonality property holds

$$\int_S (\mathbf{E}_1 \times \mathbf{H}_2).\mathbf{u}_z dS = \int_S (\mathbf{E}_{1t} \times \mathbf{H}_{2t}).\mathbf{u}_z dS = 0, \tag{7.84}$$

with S the cross-section of the waveguide. For on open or an half-open waveguide this cross-section extends to infinity. The proof is based on Lorentz reciprocity theorem and will not be given here. Based on this very general orthogonality property, an arbitrary field propagating in a waveguide can be expanded in the eigenmodes of that waveguide.

To convince the reader of the plausibility of (7.84) we now proof the orthogonality (7.84) for the TM modes of the previous section. From (7.78) and (7.79) we derive that

$$\mathbf{E}_t = -\frac{j\gamma}{k^2-\gamma^2}\nabla_t E_z, \tag{7.85}$$

$$\mathbf{u}_z \times \mathbf{H}_t = \frac{j\omega\epsilon}{k^2-\gamma^2}\nabla_t E_z. \tag{7.86}$$

Substituting (7.85) and (7.86) into the left-hand member of (7.84) yields

$$\begin{aligned}\int_S (\mathbf{E}_{1t} \times \mathbf{H}_{2t}).\mathbf{u}_z dS &= -\int_S \mathbf{E}_{1t}.(\mathbf{u}_z \times \mathbf{H}_{2t}) dS \\ &= -\frac{\gamma_1\omega\epsilon}{(k^2-\gamma_1^2)(k^2-\gamma_2^2)}\int_S \nabla_t E_{z1}.\nabla_t E_{z2} dS.\end{aligned} \tag{7.87}$$

We turn our interest to the integral in (7.87). Using (3.60) this integral can be rewritten as

$$-\int_S E_{z1}\nabla_t^2 E_{z2} dS + \int_c E_{z1}\frac{\partial}{\partial n}E_{z2} dc. \tag{7.88}$$

The integral over the boundaries c of the conductors vanishes as E_{z1} is zero on these conductors. Invoking (7.68), (7.88) finally gives

$$(k^2-\gamma_2^2)\oint_S E_{z1}E_{z2} dS. \tag{7.89}$$

It is clear that in going from (7.87) to (7.88) the role of E_{z1} and E_{z2} could have been inverted, yielding

$$(k^2-\gamma_1^2)\oint_S E_{z1}E_{z2} dS. \tag{7.90}$$

Provided $\gamma_1 \neq \gamma_2$, results (7.89) and (7.90) imply that the surface integral must be zero, proving the orthogonality of the modes.
Remark that the mode orthogonality of the voltage and current eigenvectors of Section 7.3.3 is a special case of the general mode orthogonality (7.84). The reader can proof himself that in the TEM and in the quasi-TEM case the integration in (7.84) indeed reduces to the scalar product of a voltage and a current eigenvector.

7.7 The parallel-plate waveguide

This waveguide is the most simple waveguide one can imagine. As shown in Fig. 7.7a, this waveguide consists of an infinite sheet of material of thickness

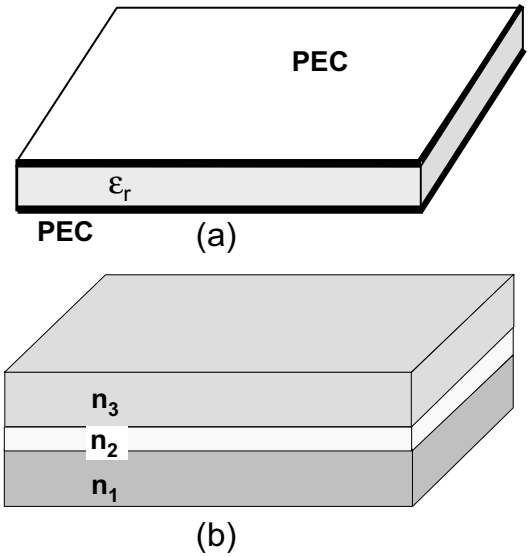


Figure 7.7: Parallel-plate waveguide (a), optical slab waveguide (b).

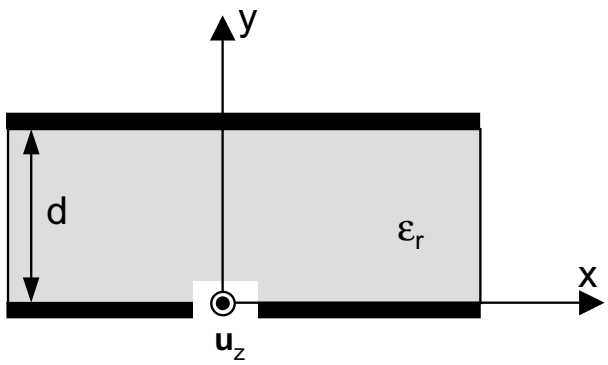


Figure 7.8: Geometry of the parallel-plate waveguide problem.

d between two infinite PEC plates. The metal plates and the dielectric of a stripline in a Printed Circuit Board act as a parallel-plate waveguide and play a central role in ground bounce phenomena. When a circuit switches, transient currents induce a wave propagating between the metal plates. This wave can be picked up by another circuit, inducing unwanted noise. The analysis of the parallel-plate waveguide also serves as an introduction to the analysis of an important optical waveguide structure, i.e. the slab waveguide depicted in Fig. 7.7b. The slab waveguide consists of three non-magnetic dielectric layers with different refractive index $n = \sqrt{\epsilon_r}$ and with $n_2 > n_1 > n_3$ such that light can be confined to the central layer.

To analyse the parallel-plate waveguide (see Fig. 7.8), we restrict ourselves to x-independent eigenmodes and to a homogeneous filling material.

7.7.1 TEM mode

As the parallel-plate waveguide counts two independent conductors, a single TEM mode will exist. The potential $\phi(y)$ satisfies

$$\frac{\mathrm{d}^2}{\mathrm{d}y^2}\phi(y) = 0, \tag{7.91}$$

with $\phi(y=0)=0$ and $\phi(y=d)=V$. V is the potential difference between the plates. The solution is $\phi(y) = Vy/d$. From $\mathbf{E}_t = -\nabla_t\phi$ and (7.61) the TEM fields are

$$\mathbf{E}_t(y) = -\frac{\mathrm{d}}{\mathrm{d}y}\phi(y)\mathbf{u}_y = -\frac{V}{d}\mathbf{u}_y, \tag{7.92}$$

$$\mathbf{H}_t(y) = \frac{V}{dZ_c}\mathbf{u}_x, \tag{7.93}$$

with $Z_c = \sqrt{\frac{\mu}{\epsilon}}$. The current flows in the z-direction. Its value per unit of length in the x-direction is given by $V/(dZ_c)$ on the top plate and $-V/(dZ_c)$ on the bottom plate. The reader will easily show that the capacitance and inductance per unit of length in the z-direction and per unit of length in the x-direction are given by

$$C_{pp} = \frac{\epsilon}{d}, \tag{7.94}$$

$$L_{pp} = d\mu \tag{7.95}$$

and hence $L_{pp}C_{pp} = \epsilon\mu$. The propagation factor γ of the TEM mode is $k = \omega\sqrt{\epsilon\mu}$. This mode is non-dispersive.

7.7.2 Losses

This is the appropriate point to introduce the effect of losses. Although the actual results obtained in this subsection hold for the TEM mode of the parallel-plate waveguide, the way to obtain these losses is readily generalised to other modes and other waveguide types.

Dielectric losses are obtained by replacing ϵ by $\epsilon + \frac{\sigma}{j\omega}$ and C by $C + \frac{G}{j\omega}$ in (7.94). From (6.9) it is indeed obvious that G can be considered as the imaginary part of C. This approach to obtain G will always be valid for homogeneous media as in that case C must be proportional to ϵ. For homogeneous media this boils down to replacing ϵ in C by σ to immediately obtain G. For the parallel-plate waveguide the result is

$$G_{pp} = \frac{\sigma}{d}. \tag{7.96}$$

It is much less straightforward to introduce *conductor losses* as these losses are frequency dependent. In this chapter, only perfect conductors were considered such that currents are flowing on the conductor surfaces. For high frequency applications it is customary to only consider *skin-effect losses*. For digital signals this does not suffice as their frequency content goes from DC to the GHz range. This more complex problem will be treated in advanced courses.

In Section 5.5 the surface impedance concept was introduced to model skin-effect losses for plane wave scattering. To calculate skin-effect conductor losses in waveguides, the following approach is used

1. The field problem is first solved in the absence of conductor losses, i.e. for PEC conductors. For simplicity, assume that all subsequent calculations are performed at $z = 0$.

2. The surface currents on the conductor surfaces are derived through $\mathbf{J}_s(\boldsymbol{\rho}) = \mathbf{u}_n \times \mathbf{H}(\boldsymbol{\rho})$.

3. The Joule losses $P_{losses}\Delta z$ in the conductors for an elementary length Δz are found by generalising (5.77) to

$$P_{losses} = \frac{1}{2}\Re \int_c \mathbf{J}_s \cdot \mathbf{E}^* dc = \frac{1}{2} R_s \int_c |\mathbf{J}_s|^2 dc, \tag{7.97}$$

   with $R_s = \frac{1}{\sigma\delta}$ the real part of the surface impedance $Z_s = R_s + jX_s$ and with δ the skin-depth $\delta = \sqrt{\frac{2}{\omega\mu\sigma}}$.

4. The total power propagated by the considered mode is calculated by integrating Poynting's vector over the waveguide cross-section

$$P_{prop} = \frac{1}{2}\Re \int_S (\mathbf{E}(z) \times \mathbf{H}^*(z)) \cdot \mathbf{u}_z dS. \tag{7.98}$$

5. The decrease of the propagated power as a function of z is related to the dissipated power through

$$\frac{\mathrm{d}}{\mathrm{d}z} P_{prop} = -P_{losses}. \tag{7.99}$$

6. The Joule losses are themselves proportional to the power propagated by the considered mode. This proportionality between the dissipated power (7.97) and the propagated power (7.98) can be expressed as

$$P_{losses} = 2\alpha P_{prop}. \tag{7.100}$$

   The meaning of the proportionality factor 2α will soon become clear.

7. Substituting (7.100) into (7.99) shows that

$$P_{prop}(z) = P_{prop}(z = 0)e^{-2\alpha z}, \tag{7.101}$$

   i.e. the propagated power decreases exponentially as a function of z. The fields themselves have a $e^{-\alpha z}$ dependence.

The conclusion is that due to the losses the propagation factor now also has an imaginary part i.e. $\gamma = \beta - j\alpha$ with the attenuation constant α given by

$$\alpha = \frac{P_{losses}}{2P_{prop}}. \tag{7.102}$$

For the TEM mode of the parallel-plate waveguide it is easily found that

$$\alpha_{pp} = \frac{R_s}{d^2 Z_c}. \tag{7.103}$$

To complete the discussion on conductor losses, let us go back to the expression for the propagation factor of voltage and current waves on a transmission line (6.12)

$$\gamma = \beta - j\alpha = \sqrt{-(R + j\omega L)(G + j\omega C)}, \tag{7.104}$$

where we have replaced k by γ. For small losses, i.e. the case considered in this subsection, we have that $R \ll j\omega L$ and $G \ll j\omega C$. This leads to the following approximation for γ

$$\gamma = \beta - j\alpha \approx \omega\sqrt{LC} - j\frac{G}{2}\sqrt{\frac{L}{C}} - j\frac{R}{2}/\sqrt{\frac{L}{C}}. \tag{7.105}$$

This expression yields a relationship between the attenuation factor

$$\alpha \approx \frac{G}{2}\sqrt{\frac{L}{C}} + \frac{R}{2}/\sqrt{\frac{L}{C}} \tag{7.106}$$

of the transmission line representation of the considered mode and the values of R, G, L and C. For the TEM mode of the parallel-plate waveguide (7.106) implies that

$$\begin{aligned} \alpha_{pp} &= \alpha_{cond} + \alpha_{diel} \\ &= \frac{R_s}{d^2 Z_c} + \frac{\sigma Z_c}{2} \end{aligned} \tag{7.107}$$

with $Z_c = \sqrt{\mu/\epsilon}$ and that

$$R_{pp} = \frac{2R_s}{d}. \tag{7.108}$$

7.7.3 TE and TM modes

In the absence of losses, the TE eigenmodes (7.72)–(7.76) satisfy

$$\frac{\mathrm{d}^2}{\mathrm{d}y^2}H_z + (k^2 - \gamma^2)H_z = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}y}H_z(y=0) = \frac{\mathrm{d}}{\mathrm{d}y}H_z(y=d) = 0. \tag{7.109}$$

The general solution of this one-dimensional Helmholtz equation is

$$H_z(y) = A\sin\sqrt{k^2 - \gamma^2}y + B\cos\sqrt{k^2 - \gamma^2}y. \tag{7.110}$$

Consequently,

$$\frac{\mathrm{d}}{\mathrm{d}y}H_z(y) = \sqrt{k^2 - \gamma^2}(A\cos\sqrt{k^2 - \gamma^2}y - B\sin\sqrt{k^2 - \gamma^2}y). \tag{7.111}$$

The boundary conditions imply that $A = 0$ and that $\sqrt{k^2 - \gamma^2}d = n\pi$ with $n > 0$. The solution for $n = 0$ leads to a constant value of H_z which is physically impossible. The propagation factors of the TE modes are given by

$$\gamma_n = \sqrt{k^2 - \frac{n^2\pi^2}{d^2}} \tag{7.112}$$

and the modal fields are

$$H_{zn}(y) = B_n \cos \frac{n\pi y}{d}, \tag{7.113}$$

$$\mathbf{E}_{tn}(y) = \frac{j\omega\mu d}{n\pi} B_n \sin \frac{n\pi y}{d} \mathbf{u}_x, \tag{7.114}$$

$$\mathbf{H}_{tn}(y) = \frac{jd\gamma_n}{n\pi} B_n \sin \frac{n\pi y}{d} \mathbf{u}_y. \tag{7.115}$$

For small enough values of n, $n\pi/d < k$ and γ_n remains positive. The corresponding TE mode is said to be *propagating* with propagation factor

$$\gamma_n = \beta_n = \sqrt{k^2 - \frac{n^2\pi^2}{d^2}}. \tag{7.116}$$

For $n\pi/d > k$, γ_n becomes imaginary and the eigenmode is said to be *evanescent*. This mode does not really propagate and is characterised by its attenuation constant

$$\alpha_n = \sqrt{\frac{n^2\pi^2}{d^2} - k^2}, \tag{7.117}$$

with $\gamma_n = -j\alpha_n$. As the frequency increases, more and more modes will start propagating. The frequency for which a particular eigenmode starts propagating is the *cut-off frequency* f_{cn} of that mode. At this frequency $n\pi/d = k$, hence

$$f_{cn} = \frac{n}{2d\sqrt{\epsilon\mu}}. \tag{7.118}$$

The cut-off frequency of the TEM mode is zero.

The TM eigenmodes (7.78)–(7.82) satisfy

$$\frac{\mathrm{d}^2}{\mathrm{d}y^2} E_z + (k^2 - \gamma^2) E_z = 0, \qquad E_z(y=0) = E_z(y=d) = 0. \tag{7.119}$$

The general solution again takes the form

$$E_z(y) = A \sin \sqrt{k^2 - \gamma^2}\, y + B \cos \sqrt{k^2 - \gamma^2}\, y, \tag{7.120}$$

but the boundary conditions now imply that $B = 0$ and that $\sin \sqrt{k^2 - \gamma^2}\, d = 0$. The modal propagation factors are also given by (7.112) and the modal fields become

$$E_{zn}(y) = A_n \sin \frac{n\pi y}{d}, \tag{7.121}$$

$$\mathbf{H}_{tn}(y) = \frac{j\omega\epsilon d}{n\pi} A_n \cos \frac{n\pi y}{d} \mathbf{u}_x, \tag{7.122}$$

$$\mathbf{E}_{tn}(y) = -\frac{jd\gamma_n}{n\pi} A_n \cos \frac{n\pi y}{d} \mathbf{u}_y. \tag{7.123}$$

The cut-off frequencies of the modes are identical to those of the TM modes (7.118).

It can easily be seen that the eigenmodes are orthogonal. Moreover, all the modes taken together constitute a complete Fourier series expansion for the cross-section (i.e. the segment $0 \leq y \leq d$) of the waveguide, proving that any field in the waveguide can be expanded as a set of modes.

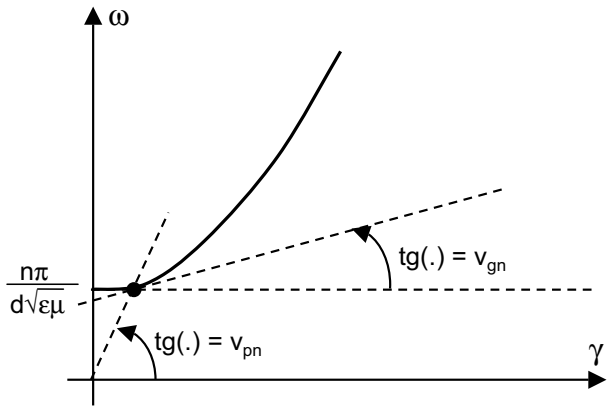


Figure 7.9: Dispersion relation.

7.7.4 Phase and group velocity

As already mentioned at the end of Section 6.2 and in Section 7.2, one must be careful to distinguish between phase and group velocity when the propagation factor γ is not proportional to the frequency. This is clearly the case for the TE and TM modes of the parallel-plate waveguide. Fig. 7.9 shows the *dispersion relation* for a TE or a TM mode. Generally speaking, the dispersion relation gives the circular frequency ω as a function of the wave number or propagation factor γ. The phase velocity of a propagating eigenmode is given by

$$v_{pn} = \frac{\omega}{\gamma_n} = \frac{\omega}{\sqrt{k^2 - \frac{n^2\pi^2}{d^2}}} = \frac{v}{\sqrt{1 - (\frac{n\pi}{kd})^2}}, \tag{7.124}$$

with $v = 1/\sqrt{\epsilon\mu}$ the velocity of plane waves in an infinite homogeneous medium. In Fig. 7.9 the phase velocity at a particular frequency ω is the tangent of the angle of the line connecting the origin to the considered frequency point on the dispersion curve. When the filling material of the parallel-plate waveguide is free space, i.e. $v = c$ with c the velocity of light in vacuum, v_{pn} exceeds the velocity of light. According to the Special Theory of Relativity this is impossible. This nonphysical result is due to the fact that we consider the sinusoidal regime with a sinusoidal wave extending from $z = -\infty$ to $z = \infty$. This wave cannot be used to transmit information nor energy. A correct picture of the signal propagation is e.g. obtained by considering a disturbance of the sinusoidal signal in the form of frequency modulation. Consider a central frequency ω_0. The actual frequency $\omega(t)$ varies very slowly around this central frequency. By way of example consider H_{zn}. The time and z-dependence of this field component is given by

$$\Re[e^{j(\omega(t)t - \gamma_n(\omega)z)}]. \tag{7.125}$$

For slow variations of ω around ω_0 we have that

$$\omega(t) = \omega_0 + \Delta\omega(t), \tag{7.126}$$

$$\gamma_n = \gamma_{n0} + [\frac{d\gamma_n}{d\omega}]_{\omega=\omega_0}\Delta\omega(t). \tag{7.127}$$

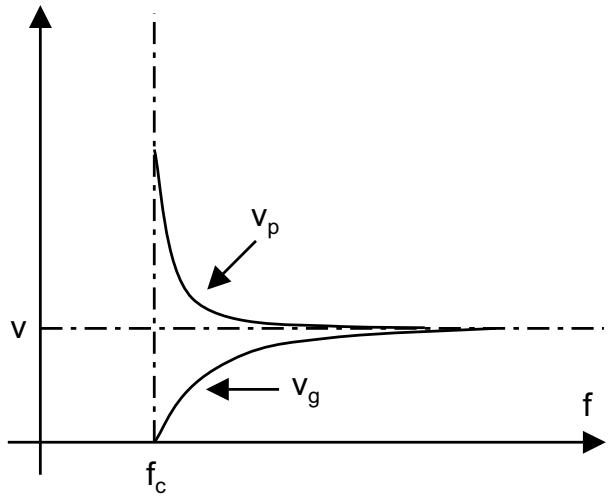


Figure 7.10: Group velocity versus phase velocity.

Substituting this into (7.125) yields

$$\Re[e^{j(\omega_0 t-\gamma_{n0} z)} e^{-j\Delta\omega(z/v_g-t)}]. \tag{7.128}$$

The first phase factor stands for the rapid phase changes of the carrier wave. These phase changes propagate with the phase velocity (7.124). Superimposed on these rapid phase changes are the slower phase changes that carry the information. The actual signal propagation velocity or *group velocity* is given by the velocity associated with the slow phase changes. From (7.127) and (7.128) this group velocity is given by

$$v_{gn} = v\sqrt{1-(\frac{n\pi}{kd})^2}. \tag{7.129}$$

This group velocity does not exceed the velocity of light. In Fig. 7.9 the group velocity at a particular frequency ω is the slope of the dispersion curve at that frequency. Fig. 7.10 shows the variation of the group velocity and the phase velocity as a function of frequency. Below cut-off the mode does not propagate. At cut-off, the group velocity is zero while the phase velocity is infinite. When the frequency increases, the group velocity increases and the phase velocity decreases. For very high frequencies both velocities coincide and take the value $v_g = v_p = v = 1/\sqrt{\epsilon\mu}$. This can be understood as follows. For very high frequencies the distance between the plates of the waveguide is very large with respect to wavelength. The modes no longer "feel" the presence of these plates and their velocity is that of a plane wave in an infinite medium. Remember that the TEM mode is non-dispersive and propagates with velocity $v_g = v_p = v = 1/\sqrt{\epsilon\mu}$.

7.8 The rectangular waveguide

Rectangular waveguides are used in microwave power applications and in microwave and millimetre wave antenna applications, in particular in antennas

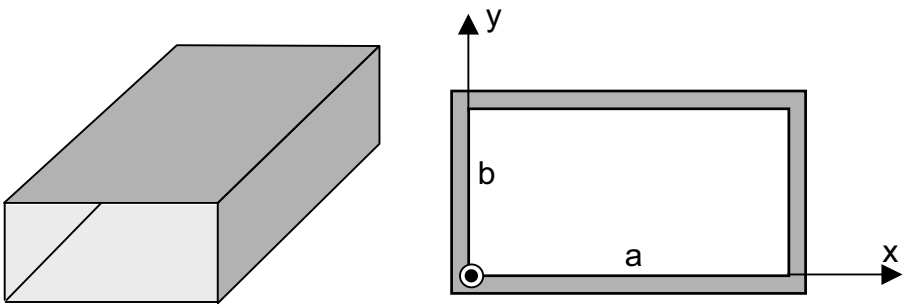


Figure 7.11: Rectangular waveguide.

and antenna arrays on satellites. Fig. 7.11 shows such a rectangular waveguide. The waveguide dimensions a and b are standardised by the Electronic Industry Association (EIA). The table below gives an example of standardised air-filled waveguides at 2.45GHz, typically used in microwave ovens.

type	width a (cm)	height b (cm)	cut-off frequency (GHz)	frequency band (GHz)
EIA-WR-340	8.636	4.318	1.737	2.17 - 3.30
EIA-WR-184	7.214	3.404	2.079	2.60 - 3.95

The TE modes follow from the solution of (7.75) subject to the boundary conditions (7.76). One easily verifies that the TE modes are given by

$$H_{z,nm}(x,y) = B_{nm} \cos\frac{n\pi x}{a} \cos\frac{m\pi y}{b}, \tag{7.130}$$

$$\begin{aligned}\mathbf{E}_{t,nm}(x,y) = \frac{j\omega\mu B_{nm}}{\frac{n^2\pi^2}{a^2}+\frac{m^2\pi^2}{b^2}}[&-\frac{n\pi}{a}\sin\frac{n\pi x}{a}\cos\frac{m\pi y}{b}\mathbf{u}_y \\ &+\frac{m\pi}{b}\cos\frac{n\pi x}{a}\sin\frac{m\pi y}{b}\mathbf{u}_x],\end{aligned} \tag{7.131}$$

$$\begin{aligned}\mathbf{H}_{t,nm}(x,y) = \frac{j\gamma_{nm} B_{nm}}{\frac{n^2\pi^2}{a^2}+\frac{m^2\pi^2}{b^2}}[&\frac{n\pi}{a}\sin\frac{n\pi x}{a}\cos\frac{m\pi y}{b}\mathbf{u}_x \\ &+\frac{m\pi}{b}\cos\frac{n\pi x}{a}\sin\frac{m\pi y}{b}\mathbf{u}_y].\end{aligned} \tag{7.132}$$

The eigenvalues of the Neumann problem and hence the propagation factors γ_{nm} are given by

$$\gamma_{nm} = \sqrt{k^2 - \frac{n^2\pi^2}{a^2} - \frac{m^2\pi^2}{b^2}}, \tag{7.133}$$

with $n \geq 0$ and $m \geq 0$ and with n and m not simultaneously equal to zero. For $n = m = 0$ the nonphysical result with a constant H_z value is obtained.
For a lossless waveguide the cut-off frequency $f_{c,nm}$ of mode (n, m) is

$$f_{c,nm} = \frac{1}{2\sqrt{\epsilon\mu}}\sqrt{\frac{n^2}{a^2}+\frac{m^2}{b^2}}. \tag{7.134}$$

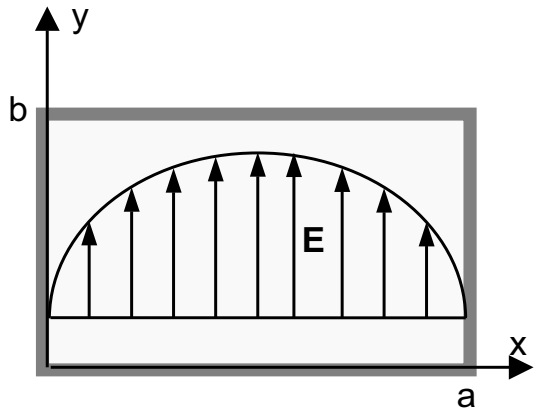


Figure 7.12: Electric field for the TE_{10} mode.

For $a > b$, the lowest cut-off frequency is the cut-off frequency of the TE_{10} eigenmode given by

$$f_{c,10} = \frac{1}{2a\sqrt{\epsilon\mu}}. \tag{7.135}$$

Below this frequency no TE modes (and no TM modes as we will soon see) propagate in the rectangular waveguide!

In practice, waveguides with $a = 2b$ are used. This implies that in the frequency range

$$\frac{1}{2a\sqrt{\epsilon\mu}} < f < \frac{1}{a\sqrt{\epsilon\mu}} \tag{7.136}$$

only a single mode, the TE_{10} mode, propagates. Above $f = 1/(a\sqrt{\epsilon\mu})$ the TE_{20} and TE_{01} are above cut-off and start propagating.

In practice the rectangular waveguide is only used in monomode regime, i.e. in the frequency range (7.136) such that only the TE_{10} mode propagates. Different frequency ranges are covered by changing the dimensions of the waveguide. The fields of the TE_{10} mode are given by

$$E_{x,10} = E_{z,10} = H_{y,10} = 0, \tag{7.137}$$

$$E_{y,10} = -\frac{j\omega\mu}{\pi/a} B_{10} \sin\frac{\pi x}{a}, \tag{7.138}$$

$$H_{x,10} = \frac{j\sqrt{k^2 - (\frac{\pi}{a})^2}}{(\pi/a)} B_{10} \sin\frac{\pi x}{a}, \tag{7.139}$$

$$H_{z,10} = B_{10} \cos\frac{\pi x}{a}. \tag{7.140}$$

To obtain the complete field expressions, the above results must be multiplied with a factor $e^{-j\sqrt{k^2-(\frac{\pi}{a})^2}z}$ expressing the z-dependence. Fig. 7.12 shows the electric field for the TE_{10} mode. Applying the perturbation method of Section 7.7.2, the attenuation factor of the TE_{10} mode due to conductor losses is found to be

$$\alpha_{cond,10} = \frac{R_s}{Z_c}\frac{1}{b}\frac{1 + \frac{2b}{a}(\frac{f_{c,10}}{f})^2}{\sqrt{1 - (\frac{f_{c,10}}{f})^2}}, \tag{7.141}$$

with $f_{c,10}$ the cut-off frequency (7.134) of the TE_{10} mode and with f the frequency. Close to the cut-off frequency the attenuation becomes very large. In practice one has to avoid using a waveguide too close to its cut-off frequency. The attenuation factor due to dielectric losses is given by

$$\alpha_{diel,10} = \frac{\sigma\omega\mu}{2\sqrt{k^2 - (\frac{\pi}{a})^2}} \tag{7.142}$$

It is left to the reader as an exercise to prove (7.142) by applying (7.102) to dielectric losses.

To complete the analysis, the TM modes, obtained from the solution of (7.81) subject to the boundary conditions (7.82), are given

$$E_{z,nm}(x,y) = A_{nm} \sin\frac{n\pi x}{a} \sin\frac{m\pi y}{b}, \tag{7.143}$$

$$\begin{aligned}\mathbf{H}_{t,nm}(x,y) = \frac{j\omega\epsilon A_{nm}}{\frac{n^2\pi^2}{a^2} + \frac{m^2\pi^2}{b^2}}[&-\frac{n\pi}{a}\cos\frac{n\pi x}{a}\sin\frac{m\pi y}{b}\mathbf{u}_y \\ &+\frac{m\pi}{b}\sin\frac{n\pi x}{a}\cos\frac{m\pi y}{b}\mathbf{u}_x],\end{aligned} \tag{7.144}$$

$$\begin{aligned}\mathbf{E}_{t,nm}(x,y) = \frac{j\gamma_{nm} A_{nm}}{\frac{n^2\pi^2}{a^2} + \frac{m^2\pi^2}{b^2}}[&-\frac{n\pi}{a}\cos\frac{n\pi x}{a}\sin\frac{m\pi y}{b}\mathbf{u}_x \\ &-\frac{m\pi}{b}\sin\frac{n\pi x}{a}\cos\frac{m\pi y}{b}\mathbf{u}_y].\end{aligned} \tag{7.145}$$

The propagation factors are identical to those of the TM modes (7.133). For $n = 0$ or $m = 0$ all fields vanish. Hence, for the TM modes $n \geq 1$ and $m \geq 1$. The TM mode with the lowest cut-off frequency is the TM_{11} mode. Note that for $a \to \infty$ and with $b = d$ the results of the parallel-plate waveguide are recovered. Fig. 7.13 gives some examples of transversal field distribution for the rectangular waveguide.

7.9 The coaxial line

Fig. 7.14 shows a typical coaxial cable consisting of a solid inner conductor, a copper braid outer conductor, the dielectric filling material (e.g. teflon) and a protective jacket. Coaxial cables are also standardised. One example is the RG-142/U cable with a characteristic impedance of 50Ω. Its dimensions typically are: inner conductor diameter 0.94 mm; outer diameter of the dielectric core 2.95 mm; outer diameter of the outer conductor 4.34 mm; diameter of the protective jacket 4.95 mm and teflon filling ($\epsilon_r = 2.041$). We restrict the analysis to the TEM mode. This mode follows from the solution of the following potential problem

$$\nabla_t^2\phi(\boldsymbol{\rho}) = 0, \qquad \phi(\rho = a) = 0, \qquad \phi(\rho = b) = V, \tag{7.146}$$

with V the potential difference between the inner and the outer conductor. From the geometry of the problem and the boundary conditions it follows that the solution only depends on the radial co-ordinate ρ. Expressing the Laplacian in polar co-ordinates then yields

$$\rho\frac{d}{d\rho}[\rho\frac{d}{d\rho}\phi(\rho)] = 0, \tag{7.147}$$

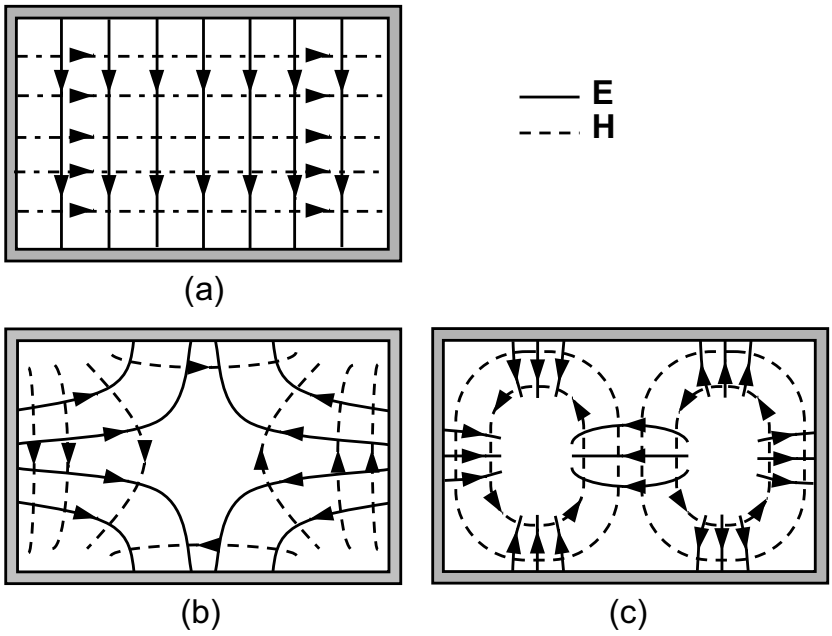


Figure 7.13: Examples of field lines of a few rectangular waveguide modes: TE_{10} mode (a), TE_{11} mode (b) and TM_{21} mode (c).

the general solution of which is $\phi(\rho) = A \ln \rho + B$. Enforcing the boundary conditions finally shows that

$$\phi(\rho) = V \frac{\ln \frac{\rho}{a}}{\ln \frac{b}{a}}. \tag{7.148}$$

The transversal fields are

$$\mathbf{E}_t(\rho) = -\nabla_t \phi(\rho) = -\frac{d}{d\rho}\phi(\rho)\mathbf{u}_\rho = -\frac{V}{\rho \ln \frac{b}{a}}\mathbf{u}_\rho, \tag{7.149}$$

$$\mathbf{H}_t(\rho) = -\frac{V}{Z_c \rho \ln \frac{b}{a}}\mathbf{u}_\phi. \tag{7.150}$$

These are the fields already obtained in Section 3.7 and Section 4.7 when solving the electrostatic and magnetostatic problem for the coaxial cable. It is clear that the solution in the non-static case is identical to that in the static case with the exception of a multiplicative factor $e^{-j\gamma z} = e^{-jkz} = e^{-j\omega\sqrt{\epsilon\mu}z}$ expressing the fact that the static field patterns now propagate along the cable with frequency independent speed $v = 1/\sqrt{\epsilon\mu}$.

The capacitance and inductance per unit of length are

$$C_{coax} = \frac{2\pi\epsilon}{\ln \frac{b}{a}}, \tag{7.151}$$

$$L_{coax} = \frac{\mu \ln \frac{b}{a}}{2\pi}. \tag{7.152}$$

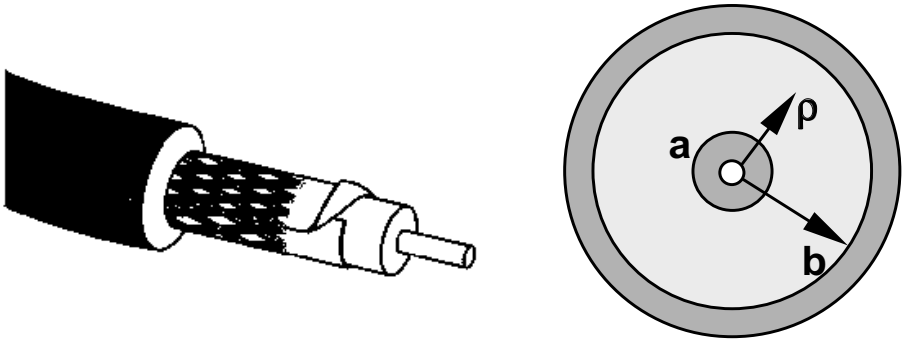


Figure 7.14: Coaxial cable.

The characteristic impedance of the equivalent transmission line is

$$R_{c,coax} = \sqrt{\frac{L}{C}} = \sqrt{\frac{\mu}{\epsilon}} \frac{\ln \frac{b}{a}}{2\pi}. \tag{7.153}$$

If a coaxial cable such as the RG-142/U cable mentioned above, is said to be a 50Ω cable, this means that $R_{c,coax} = 50\Omega$. The characteristic impedance is the ratio of the voltage to the current propagating along the line. For a TEM mode we have proved that the ratio of the amplitude of the electric field to the amplitude of the magnetic field is always equal to $Z_c = Z_{TEM} = \sqrt{\frac{\mu}{\epsilon}}$. One has to be careful to distinguish between the modal impedance Z_{TEM} and the characteristic impedance $R_{c,coax}$ of the equivalent transmission line.

The effect of losses can again be determined by applying the techniques put forward in Section 7.7.2. The results are

$$G_{coax} = \frac{2\pi\sigma}{\ln \frac{b}{a}}, \tag{7.154}$$

$$R_{coax} = \frac{R_s}{2\pi}(\frac{1}{a} + \frac{1}{b}). \tag{7.155}$$

To obtain (7.155) the following results were used

$$P_{prop} = |V|^2 \frac{\pi}{Z_c \ln \frac{b}{a}} = \frac{|V|^2}{2R_{c,coax}}, \tag{7.156}$$

$$P_{losses} = |V|^2 \frac{\pi}{Z_c^2 (\ln \frac{b}{a})^2} R_s (\frac{1}{a} + \frac{1}{b}). \tag{7.157}$$

The attenuation factor α becomes

$$\alpha_{coax} = \alpha_{cond} + \alpha_{diel} = \frac{R_s}{2Z_c \ln \frac{b}{a}}(\frac{1}{a} + \frac{1}{b}) + \frac{\sigma Z_c}{2}. \tag{7.158}$$

The coaxial cable also supports TE and TM modes. To know up to which frequency a coaxial cable remains monomode, it is important to know the cut-off frequency of the next mode. This mode is the TE_{11} mode and its cut-off frequency is $f_c = k_c/(2\pi\sqrt{\epsilon\mu})$. The wavenumber k_c is the smallest solution of

$$\frac{Y_1'(k_c a)}{J_1'(k_c a)} = \frac{Y_1'(k_c b)}{J_1'(k_c b)}, \tag{7.159}$$

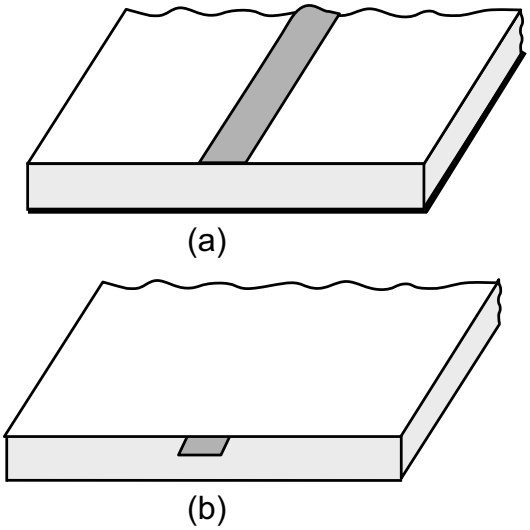


Figure 7.15: Microstrip (a) and stripline (b).

with Y_1' and J_1' the derivatives of the Neumann and the Bessel function of order one. For the RG-142/U cable the cut-off frequency is about 15GHz.

7.10 The microstrip and the stripline

In Section 6.2 we have already discussed the importance of the microstrip and the stripline (see Fig. 7.15) as interconnection structures. They are not only used in Printed Circuit Boards but also in *M*icrowave and *M*illimetre wave *I*ntegrated *C*ircuits (MMIC's). In these MMIC's microstrip lines and striplines are used to fabricate passive components such as inductors, capacitors, filters, couplers, ...

A microstrip line is a hybrid open waveguide. Hence, no pure TEM, TE or TM modes exist. For sufficiently low frequencies however, the quasi-TEM analysis of Section 1.3 applies and the microstrip can be characterised by a capacitance C_{micro} and an inductance L_{micro} per unit of length. The characteristic impedance $R_{c,micro}$ of the equivalent transmission line representation and the propagation coefficient β_{micro} are

$$R_{c,micro} = \sqrt{\frac{L_{micro}}{C_{micro}}}, \tag{7.160}$$

$$\beta_{micro} = \omega\sqrt{L_{micro}C_{micro}}. \tag{7.161}$$

For microstrips, the propagation coefficient is conventionally rewritten as

$$\beta_{micro} = k_0\sqrt{\epsilon_{r,eff}}, \tag{7.162}$$

with $k_0 = \omega/c$ the free space wave number and $\epsilon_{r,eff}$ the *effective relative dielectric permittivity*. This effective permittivity expresses the fact that fields partly propagate in the dielectric substrate and partly in the air. This results

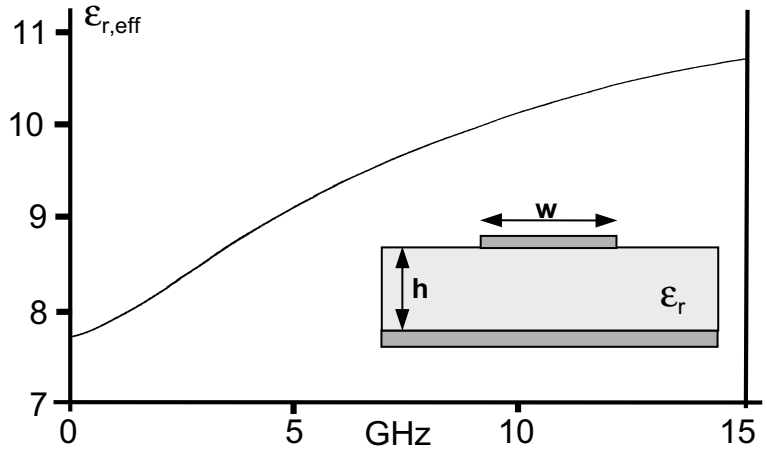


Figure 7.16: Effective dielectric permittivity as a function of frequency.

in a signal speed $v = c/\sqrt{\epsilon_{r,eff}}$ between the speed of light in the dielectric and the speed of light in the air c.

For the stripline (provided the dielectric is homogeneous) the analysis is much more simple as the stripline supports a TEM mode. It suffices to calculate its capacitance C_{strip} per unit of length. The corresponding inductance L_{strip} and characteristic impedance $R_{c,strip}$ then follow from

$$L_{strip} = \frac{\epsilon_r}{C_{strip}c^2}, \tag{7.163}$$

$$R_{c,strip} = \frac{\sqrt{\epsilon_r}}{C_{strip}c}, \tag{7.164}$$

with c the velocity of light in free space. The signal speed is of course $c/\sqrt{\epsilon_r}$.

When the frequency increases, the quasi-TEM analysis of the microstrip will no longer remain valid and a so called *full-wave* analysis becomes necessary. We will not go into detail at this point. At higher frequencies the quasi-TEM mode still exists but the longitudinal field components E_z and H_z become important, i.e. the mode loses its quasi-TEM character. At the same time the mode becomes dispersive: the signal speed changes and depends on frequency. This can be expressed by also making the effective dielectric permittivity frequency dependent. Fig. 7.16 shows the evolution of this effective dielectric permittivity for the microstrip depicted in the same figure. The metal strip is infinitely thin, $w = 3.04mm$, $h = 3.17mm$ and $\epsilon_r = 11.7$ The value of $\epsilon_{r,eff}$ at $\omega = 0$ is 7.7. This is the quasi-TEM value defined by (7.162). When the frequency increases the effective dielectric permittivity also increases. This is a consequence of the fact that at higher frequencies field lines tend to concentrate in the dielectric. For very high frequencies no field lines are found in the air. Hence $\epsilon_{r,eff}$ must asymptotically tend towards $\epsilon_{r,eff} = \epsilon_r = 11.7$. Outside the validity range of the quasi-TEM analysis, a transmission line equivalent with frequency dependent L and C can be derived. The analysis of this problem, to be found in the books cited in the Introduction, is quite complicated and falls outside the scope of this introductory course.

Chapter 8

Antennas and Radiation

8.1 Introduction

Most of us are familiar with antennas. Although in the new generation of cell phones, the antenna is no longer visible, base station antennas are now part of our landscape. Dish antennas as receivers for satellite broadcasting are equally omnipresent. A radio-telescope is yet another example that perhaps does not readily come to mind. The Concise Oxford Dictionary defines *antenna* or *aerial* (still preferred over antenna in Great-Britain) as *a metal rod, wire or other structure by which signals are transmitted or received as part of a radio transmission or receiving system.* More generally, the term antenna refers to a vary large variety of structures that can be used to either generate or receive electromagnetic fields.

In this introductory course it is impossible to cover a large number of antenna types and their design and operation principles. This would typically be the subject matter for a Master's course on "Antennas and Propagation". In this chapter we will restrict ourselves to so-called wire and patch antennas. In this case, the sources of the fields are the conductor currents on the metal wires or the metal patches. When transmitting, these currents are injected into the antenna at the antenna terminals, i.e. the connection points to the transmit circuit. When receiving, currents are induced in the metal parts of the antenna and fed into the receive circuit through these terminals.

To avoid unnecessary complications, discussions will be restricted to antennas radiating and receiving in homogeneous and infinite free space. As a first step to understanding the behaviour of an antenna, we will consider the fields generated by a known source current density and more particularly the so-called far fields. The introduction of the radiation impedance of the antenna allows to define the equivalent circuit of a receiving antenna. Next, the fundamental reciprocity that exists between receiving and transmitting will be highlighted. The open circuit voltage at the terminals of the antenna, when immersed in an incident plane wave, is the key to the equivalent circuit representation of the antenna in receive mode. Finally, we are in a position to discuss the Friis formula which elegantly describes the power transmitted between two antennas.

In the second part of this chapter, the properties of a simple dipole antenna are discussed. From this it can easily be understood that the exact treatment of

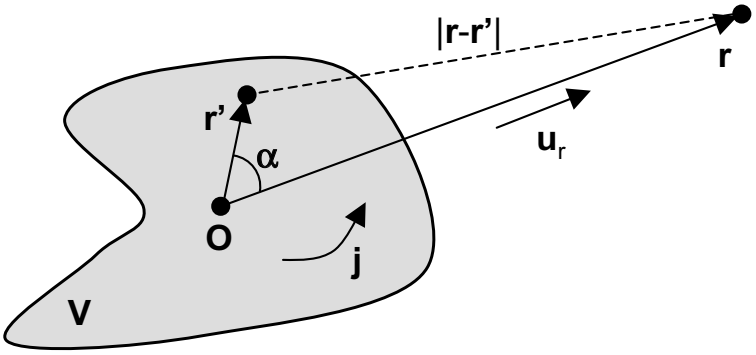


Figure 8.1: Relevant to the calculation of the far field.

the general wire antenna problem requires the solution of an integral equation for the currents on the metal wires or patches of the antenna. No in depth treatment of such an integral equation will be presented though. The concept of an antenna array is also very briefly outlined and the chapter concludes with an overview of various types of more advanced wire and patch antennas.

8.2 The far field of a current source in free space

In Section 2.6 the fields generated by a current density $\mathbf{j}(\mathbf{r})$ in a homogeneous and infinite space have been expressed in terms of the vector potential $\mathbf{a}(\mathbf{r})$ and the scalar potential $\phi(\mathbf{r})$. Now consider a source $\mathbf{j}(\mathbf{r})$ defined in volume V as depicted in Fig. 8.1. The origin O must be chosen inside volume V or at least close to that volume (ask yourself why this is necessary for the reasoning that follows). The vector potential generated by the source is given by (2.105)

$$\mathbf{a}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_V \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \mathbf{j}(\mathbf{r}')dV'. \tag{8.1}$$

We have replaced μ by μ_0 as we restrict ourselves to free space. The wavenumber k is the free space wave number $k = \omega\sqrt{\epsilon_0\mu_0}$; $\mathbf{r}$ is the position vector of the observation point where we want to know the fields and $\mathbf{r}'$ is the position vector of the variable integration point inside V. Now, and this is essential, suppose that the observation point $\mathbf{r} = r\mathbf{u}_r$ is located at a large distance of the source in the direction specified by the unit vector $\mathbf{u}_r$. When α represents the angle between $\mathbf{r}$ and $\mathbf{r}'$ the following approximation for the distance $|\mathbf{r}-\mathbf{r}'|$

$$|\mathbf{r}-\mathbf{r}'| = \sqrt{r^2 + r'^2 - 2rr'\cos\alpha} \approx r - r'\cos\alpha = r - \mathbf{r}'\cdot\mathbf{u}_r \tag{8.2}$$

holds, provided $r = |\mathbf{r}|$ is much larger than the maximum value of $r' = |\mathbf{r}'|$. The next step is to substitute this approximation for the distance in (8.1). As the denominator only influences the magnitude of $\mathbf{a}$, we only retain the first term in the right hand member of (8.2). For the numerator however, both terms are important as the numerator represents a phase factor, small changes of which

can have an important impact on the overall result. The result is

$$\lim_{kr\to\infty} \mathbf{a}(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{e^{-jkr}}{r}\int_V e^{jk\mathbf{u}_r\cdot\mathbf{r}'}\mathbf{j}(\mathbf{r}')dV' = \frac{e^{-jkr}}{r}\mathbf{N}(\theta,\phi). \tag{8.3}$$

The vector $\mathbf{N}(\theta,\phi)$ expresses the (θ,ϕ) angle dependence of the vector potential. The complete distance dependence is given by the factor $\frac{e^{-jkr}}{r}$ which is characteristic for a three dimensional free space problem. The vector potential directly leads to the expression of the magnetic field through $\mathbf{b} = \nabla\times\mathbf{a}$. Hence,

$$\begin{aligned}\mathbf{h}(\mathbf{r}) &= \frac{1}{\mu_0}\{\frac{1}{r\sin\theta}[\frac{\partial}{\partial\theta}(\sin\theta a_\phi(\mathbf{r})) - \frac{\partial}{\partial\phi}a_\theta(\mathbf{r})]\mathbf{u}_r \\ &+ \frac{1}{r}[\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}a_r(\mathbf{r}) - \frac{\partial}{\partial r}(ra_\phi(\mathbf{r}))]\mathbf{u}_\theta \\ &+ \frac{1}{r}[\frac{\partial}{\partial r}(ra_\theta(\mathbf{r})) - \frac{\partial}{\partial\theta}a_r(\mathbf{r})]\mathbf{u}_\phi\}. \end{aligned}\tag{8.4}$$

To obtain the above result, the curl operator has been expressed in spherical co-ordinates. Let us now substitute (8.3) into (8.4), but in doing so we will only retain the most important terms, i.e. those of order $1/r$ while neglecting all higher order terms (this approximation is consistent with the approximation that has already been introduced to obtain (8.3)). We arrive at

$$\begin{aligned}\lim_{kr\to\infty}\mathbf{h}(\mathbf{r}) &= \frac{1}{\mu_0}[-\frac{\partial}{\partial r}a_\phi(\mathbf{r})\mathbf{u}_\theta + \frac{\partial}{\partial r}a_\theta(\mathbf{r})\mathbf{u}_\phi] \approx -\frac{jk}{\mu_0}[-a_\phi(\mathbf{r})\mathbf{u}_\theta + a_\theta(\mathbf{r})\mathbf{u}_\phi] \\ &= -\frac{jk}{\mu_0}\mathbf{u}_r\times[a_\phi(\mathbf{r})\mathbf{u}_\phi + a_\theta(\mathbf{r})\mathbf{u}_\theta] = -\frac{jk}{\mu_0}[\mathbf{u}_r\times\mathbf{N}(\theta,\phi)]\frac{e^{-jkr}}{r}. \end{aligned}\tag{8.5}$$

As in the far field no further sources have to be taken into account, the electric field is easily obtained through

$$\mathbf{e}(\mathbf{r}) = \frac{1}{j\omega\epsilon_0}\nabla\times\mathbf{h}(\mathbf{r}). \tag{8.6}$$

The curl must be approximated in the same way as above for the magnetic field which readily leads to the final expression for the electric field, i.e.

$$\lim_{kr\to\infty}\mathbf{e}(\mathbf{r}) = -\frac{k^2}{j\omega\epsilon_0\mu_0}\mathbf{u}_r\times[\mathbf{u}_r\times\mathbf{N}(\theta,\phi)]\frac{e^{-jkr}}{r} = \mathbf{F}(\theta,\phi)\frac{e^{-jkr}}{r}, \tag{8.7}$$

where we have introduced the radiation vector $\mathbf{F}(\theta,\phi) = j\omega\mathbf{u}_r\times[\mathbf{u}_r\times\mathbf{N}(\theta,\phi)]$. Taking into account that $\mathbf{u}_r\times\mathbf{F}(\theta,\phi) = -j\omega[\mathbf{u}_r\times\mathbf{N}(\theta,\phi)]$, the magnetic field can also be expressed in terms of the radiation vector as

$$\lim_{kr\to\infty}\mathbf{h}(\mathbf{r}) = \frac{1}{Z_c}[\mathbf{u}_r\times\mathbf{F}(\theta,\phi)]\frac{e^{-jkr}}{r}, \tag{8.8}$$

with Z_c the free space characteristic impedance $Z_c = \sqrt{\frac{\mu_0}{\epsilon_0}}$, or approximately $120\pi\Omega$ or 377Ω. Expressions (8.7) and (8.8) define a spherical wave with its typical e^{-jkr}/r dependence. Locally, i.e. in any given direction $\mathbf{u}_r$, expressions (8.7) and (8.8) can be thought of as a plane wave propagating in that direction.

It is indeed easily verified that the electric field is perpendicular to the magnetic field, that both fields are perpendicular to their propagation direction $\mathbf{u}_r$ and that the ratio of their amplitudes is Z_c. In the far field approximation, the radial field components are negligible (i.e. they are of order $1/r^2$).

We however still have to determine under which condition the above far field approximations hold. To this end consider the following approximation of $|\mathbf{r}-\mathbf{r}'|$ which includes an extra term

$$|\mathbf{r}-\mathbf{r}'| \approx r - r'\cos\alpha + \frac{r'^2}{2r}\sin^2\alpha. \tag{8.9}$$

Introducing this extra term in the phase of (8.1) yields

$$\lim_{kr\to\infty}\mathbf{a}(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{e^{-jkr}}{r}\int_V e^{jk\mathbf{u}_r\cdot\mathbf{r}'}e^{-jkr'^2\sin^2\alpha/(2r)}\mathbf{j}(\mathbf{r}')dV'. \tag{8.10}$$

The far field approximation (8.3) will remain valid provided the extra phase term which now figures in the expression for the vector potential is much smaller than 2π, i.e. when

$$\frac{kr'^2}{2r}\sin^2\alpha \ll 2\pi. \tag{8.11}$$

As this must hold for all source points and for all directions $\mathbf{u}_r$, we replace $\sin^2\alpha$ by 1 and r' by its maximum value in the source volume, further noted as d (the reader should note at this point that the choice of the origin O plays a role here). Further replacing k by $2\pi/\lambda$ yields

$$r \gg \frac{d^2}{2\lambda}. \tag{8.12}$$

Traditionally, the far field boundary is defined as starting at a distance corresponding to 4 times the value of the right hand side, i.e. for a phase error smaller than $\pi/2$ or

$$r \geq \frac{2d^2}{\lambda}. \tag{8.13}$$

Fields points closer to the source are said to be in the *near field*.

8.3 Directivity

Let us now take a closer look at the energy radiated by the sources. To this end we first determine Poynting's vector in the far field

$$\mathbf{p}(\mathbf{r}) = \frac{|\mathbf{F}(\theta,\phi)|^2}{2r^2Z_c}\mathbf{u}_r. \tag{8.14}$$

The total radiated power P_{tot} is found by integrating the real part of $\mathbf{p}$ over a sphere with origin O. For all points of the surface of this sphere to be in the far field, its radius R must be chosen sufficiently large. Prove that the final expression

$$P_{tot} = \frac{1}{2Z_c}\int_0^{2\pi}\int_0^{\pi}|\mathbf{F}(\theta,\phi)|^2\sin\theta d\theta d\phi. \tag{8.15}$$

is independent of that radius. We can now ask the question whether this power is equally radiated in all directions. Such omnidirectional radiation would imply that $\frac{P_{tot}}{4\pi}$ watt are found in each elementary space angle $\sin\theta d\theta d\phi$ further denoted as $d\Omega$ (remark that $d\Omega$ integrated over all directions, i.e. the total space angle for a sphere, amounts to 4π). We have already pointed out that the far fields behave as locally plane waves. This implies that $\Re[\mathbf{p}(\mathbf{r})]r^2\sin\theta d\theta d\phi.\mathbf{u}_r = |\mathbf{F}(\theta,\phi)|^2/2Z_c\sin\theta d\theta d\phi$ can be interpreted as the power radiated in the elementary space angle $d\Omega$ associated with the direction $\mathbf{u}_r$. It is clear that this power can indeed differ for each direction $\mathbf{u}_r$. To quantify this, the directivity of the radiating source is introduced, defined as the ratio of the power radiated in $d\Omega$ to its mean value

$$D(\theta,\phi) = \frac{\frac{|\mathbf{F}(\theta,\phi)|^2}{2Z_c}}{\frac{P_{tot}}{4\pi}} = \frac{4\pi|\mathbf{F}(\theta,\phi)|^2}{\int\limits_{\Omega}|\mathbf{F}(\theta,\phi)|^2 d\Omega}. \tag{8.16}$$

Plotting the directivity as a function of the angles θ and ϕ yields the radiation pattern of the source. Fig. 8.2a shows a typical two-dimensional cross-section of a radiation pattern. The main lobe of the radiation pattern is associated with the direction of maximum directivity. In a particular cross-section, the 3 dB beamwidth (associated with the main lobe) is defined as the total angle over which the directivity does not drop below 3 dB of its maximum value. Side lobes are associated with other, local, maxima. Directions of minimal directivity are indicated as shadow regions and result from the destructive interference of the source contributions. As a result of this destructive interference, the directivity of many antennas will even exhibit "nulls", directions in which no power is at all radiated. Fig. 8.2b and 8.2c are examples of the complete radiation pattern or radiation body of an antenna.
Designing an antenna with a predefined directivity or at least satisfying a number of constraints such as desired direction of maximum directivity and 3 dB beamwidth or nulling in particular directions, is a very challenging task with a large number of important industrial applications. We refer the reader to the large body of literature on this topic.

8.4 Antennas

8.4.1 Radiation impedance of a transmitting antenna

As already stated in the Introduction, we will restrict ourselves to antennas consisting of perfect conductors embedded in free space. When the antenna acts as a transmitter, current is typically injected through a pair of terminals. This generic situation is depicted in Fig. 8.3. The transmitter circuit is modelled by its Thevenin equivalent with generator V_g and internal impedance Z_g. The current entering the antenna terminals is I and V represents the voltage over these terminals with $V = V_g - Z_g I$. Now consider the volume $\mathcal{V}$ bounded by the outer surface S_a of the antenna conductors (and as such excluding the conductors from $\mathcal{V}$), the surface S_c which encompasses the transmitter circuit and a spherical surface S_r with radius r which encloses the complete antenna. The overall dimensions of the volume enclosed by S_c are small with respect to the wavelength such that Kirchoff's laws will suffice to describe the transmitter. In

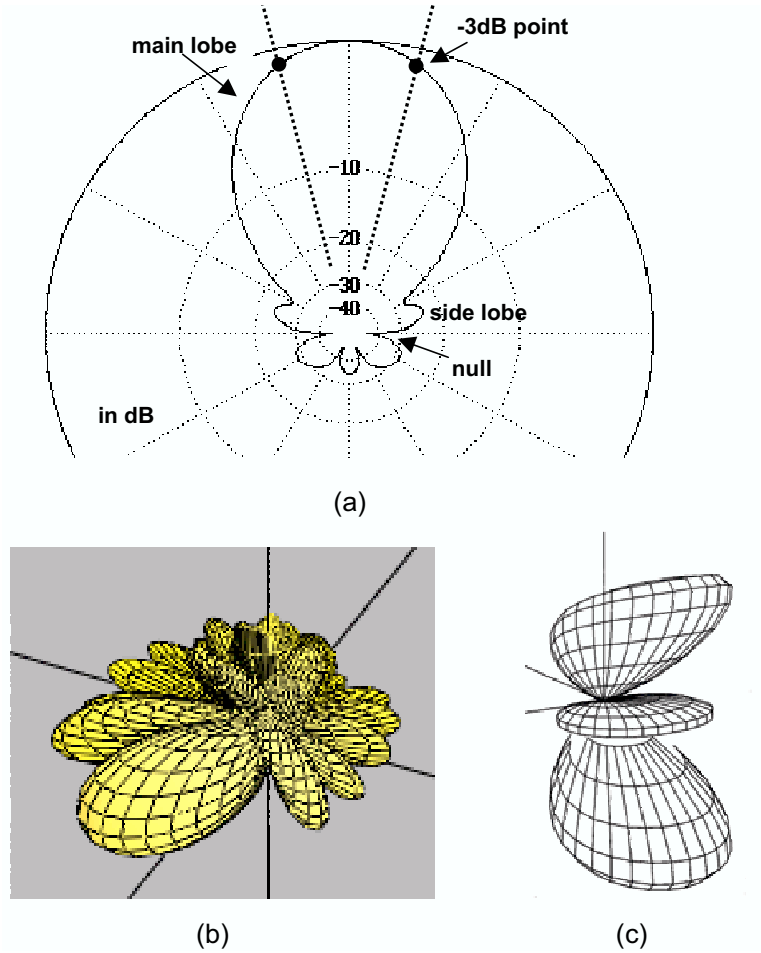


Figure 8.2: Typical two-dimensional cross-section of a radiation pattern (a), examples of radiation bodies (b) and (c).

the more general case, the antenna terminals are connected to the transmitter circuit through a transmission line, e.g. a coaxial cable. The calculations presented below can be extended to this more general case. The final result for the radiation impedance remains unaltered. The radius of the spherical surface S_r is chosen sufficiently large for all points on the surface to be in the far field region. As volume $\mathcal{V}$ does not contain sources, the energy balance equation (2.44) (with $\mathbf{j}_e = 0$) states that

$$\int\limits_{S_a+S_c+S_r} (\mathbf{e} \times \mathbf{h}^*) \cdot \mathbf{u}_n dS = \int\limits_{V} (-j\omega\mu|\mathbf{h}|^2 + j\omega\epsilon|\mathbf{e}|^2) dV. \tag{8.17}$$

The contribution of S_a to the left hand side vanishes as the tangential electric field is zero on the perfect conductors. To obtain the contribution of S_r, the far

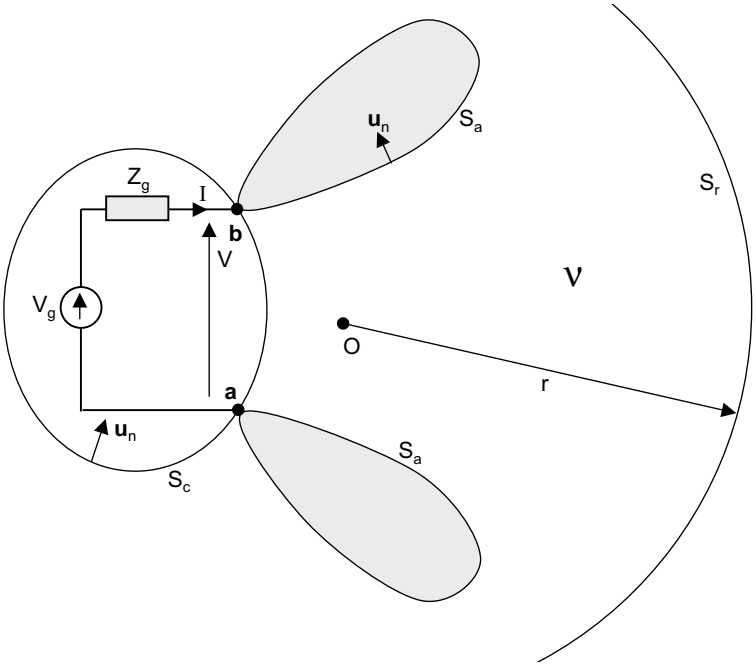


Figure 8.3: Antenna in transmission mode.

field expressions for the electric and magnetic field can be inserted, yielding

$$\int_{S_r} (\mathbf{e} \times \mathbf{h}^*) \cdot \mathbf{u}_n dS = \int_{\Omega} (\frac{e^{-jkr}}{r} \mathbf{F} \times \frac{1}{Z_c^*} \frac{e^{jkr}}{r} \mathbf{u}_r \times \mathbf{F}^*) \cdot \mathbf{u}_r r^2 d\Omega = \frac{1}{Z_c} \int_{\Omega} |\mathbf{F}|^2 d\Omega. \tag{8.18}$$

We have tacitly assumed that the origin for the calculation of $\mathbf{F}$ coincides with the origin of the spherical surface S_r.

We draw the attention of the reader to the calculation of the contribution of S_c as it requires a subtle combination of Maxwell's equations and circuit theory. As we required the dimensions of the volume enclosing the transmitter circuit to be small with respect to wavelength, the static approximation of Maxwell's equations applies, i.e.

$$\nabla \times \mathbf{e}(\mathbf{r}) = 0, \tag{8.19}$$

$$\nabla \times \mathbf{h}(\mathbf{r}) = \mathbf{j}(\mathbf{r}). \tag{8.20}$$

The first equation implies that the electric field can be derived from a scalar potential $\phi(\mathbf{r})$ through $\mathbf{e}(\mathbf{r}) = -\nabla\phi(\mathbf{r})$. As the reference point for the potential we choose one of the terminals of the antenna, say a in the case of Fig. 8.3. Consequently, ϕ takes the value V on terminal b. The integral over S_c can now be transformed as follows

$$\int_{S_c} (\mathbf{e} \times \mathbf{h}^*) \cdot \mathbf{u}_n dS = -\int_{S_c} (\nabla\phi \times \mathbf{h}^*) \cdot \mathbf{u}_n dS = -\int_{S_c} \nabla\phi \cdot (\mathbf{h}^* \times \mathbf{u}_n) dS$$

$$= -\int_{S_c} \nabla \cdot [\phi(\mathbf{h}^* \times \mathbf{u}_n)] dS + \int_{S_c} \phi \nabla \cdot (\mathbf{h}^* \times \mathbf{u}_n) dS$$

$$= \int_{S_c} \phi(\nabla \times \mathbf{h}^*) \cdot \mathbf{u}_n dS = \int_{S_c} \phi \mathbf{j}^* \cdot \mathbf{u}_n dS = -VI^*. \tag{8.21}$$

To obtain the above result we used the fact that the conduction current only penetrates through the surface S_c at the terminals a and b and the fact that

$$\int_{S_c} \nabla \cdot [\phi(\mathbf{h}^* \times \mathbf{u}_n)] dS = \int_{S_c} \nabla_s \cdot [\phi(\mathbf{h}^* \times \mathbf{u}_n)] dS = 0, \tag{8.22}$$

whereby ∇_s represents the surface divergence and taking into account that $\phi(\mathbf{h}^* \times \mathbf{u}_n)$ is tangential to the surface S_c. Also remember that the integral of the surface divergence of a function, taken over a closed surface, vanishes.

Finally, collecting all contributions, (8.17) can be rewritten as

$$\frac{1}{Z_c} \int_{\Omega} |\mathbf{F}|^2 d\Omega - VI^* = j\omega \int_V (\epsilon |\mathbf{e}|^2 - \mu |\mathbf{h}|^2) dV. \tag{8.23}$$

Now, linearity dictates that doubling the terminal current I will also double the terminal voltage V. Hence, it is logical to introduce an impedance Z_a expressing the ratio of V to I. $Z_a = R_a + jX_a$ is an inherent characteristic of the antenna, known as its *radiation impedance*, which from (8.23) turns out to be

$$R_a = \frac{1}{Z_c |I|^2} \int_{\Omega} |\mathbf{F}|^2 d\Omega, \tag{8.24}$$

$$X_a = -\frac{\omega}{|I|^2} \int_V (\epsilon |\mathbf{e}|^2 - \mu |\mathbf{h}|^2) dV, \tag{8.25}$$

R_a is called the radiation resistance, while X_a is the radiation reactance. One could worry about the fact that (8.24) and (8.25) still seem to depend on I. This is however not the case as linearity further dictates that $\mathbf{F}$, $\mathbf{e}$ en $\mathbf{h}$ must also be proportional to I. Remark that the radiation reactance R_a is completely determined by the near fields. Indeed, as in the far field $\epsilon |\mathbf{e}|^2 = \mu |\mathbf{h}|^2$, the integration in (8.25) is restricted to the near field. The radiation resistance solely depends on the far fields. The near field, and hence the antenna reactance, depends on the detailed field distribution near the antenna terminals and will be influenced by small details of the antenna. This is not the case for the radiation resistance as the far fields are much less susceptible to these details.

8.5 Equivalent circuit of a transmitting antenna

From the knowledge of the radiation impedance we can now easily derive the equivalent circuit of the antenna when acting as a transmitter. Fig. 8.4 shows the complete equivalent circuit of the transmitter and the antenna connected to its terminals. As the antenna itself consisted of perfect conductors, the power "dissipated" in Z_a completely corresponds to the total power radiated by the

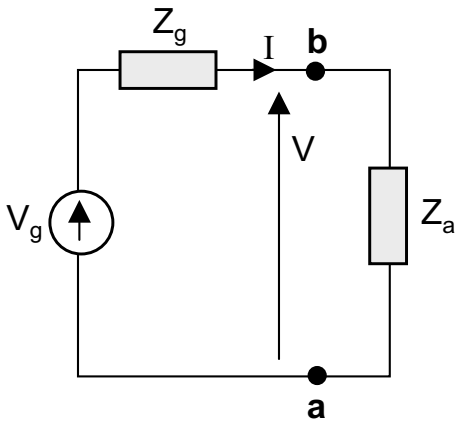


Figure 8.4: Equivalent circuit of a transmitting antenna.

antenna. One can easily check that the source delivers maximum power to the antenna provided $Z_g^* = Z_a$. This maximum power is given by

$$P_{rad,max} = \frac{R_a|I|^2}{2} = \frac{|V|^2}{2R_a}. \tag{8.26}$$

8.6 Open circuit voltage of a receiving antenna

Above, it was proved that, with respect to its terminals, a transmitting antenna behaves as an impedance, the radiation impedance. We would now like to know how this antenna behaves when used as a receiving antenna. The question that immediately comes to mind is the following: what is the relationship, if any, between an antenna used to transmit and the same antenna used as a receiver. To this end consider Fig. 8.5. The terminals of the antenna are left open and the incident field is denoted as $(\mathbf{e}^i(\mathbf{r}),\mathbf{h}^i(\mathbf{r}))$. Without loss of generality we will suppose that this incident field is a plane wave with its electric field given by

$$\mathbf{e}^i(\mathbf{r}) = \mathbf{E}_0 e^{-jk\mathbf{u}^i\cdot\mathbf{r}}. \tag{8.27}$$

Indeed, any field can be decomposed into a set of plane waves (this can e.g. be proved by spatially Fourier transforming Maxwell's equation with respect to two co-ordinates say x and y). Moreover, supposing that the sources of the incident field are at a large distance of the receiving antenna, we know that this field can be considered to be locally plane. The incident field is the field that exists and would remain undisturbed in the absence of the receiving antenna. It satisfies Maxwell's equation. However, in general, it will not satisfy the boundary conditions at the perfect metal conductors that constitute the antenna. One of these boundary conditions specifies that the tangential component of the electric field must be zero i.e. $\mathbf{e}^i(\mathbf{r})_{tan}$ must vanish on the antenna. As a consequence of not satisfying this boundary condition, surface currents will be induced on the antenna. These surface currents, as any currents, are themselves sources of electromagnetic fields which we will denote as $(\mathbf{e}^s(\mathbf{r}),\mathbf{h}^s(\mathbf{r}))$. This secondary field is known as the *scattered field*. The problem of an antenna immersed in an incident

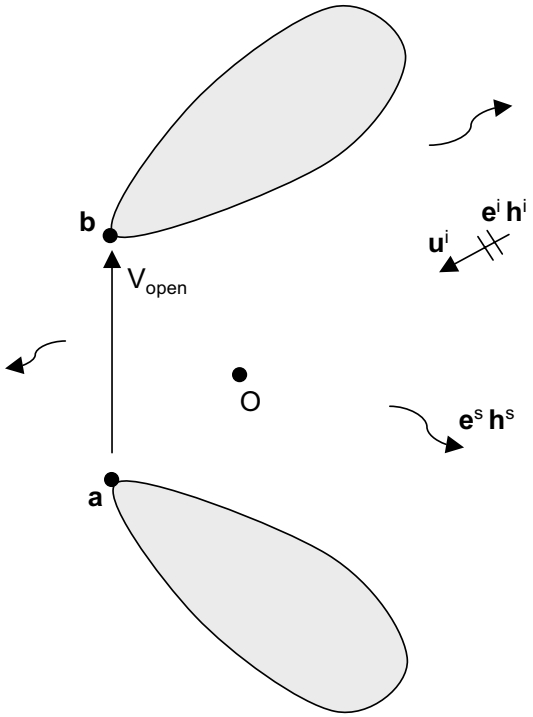


Figure 8.5: Antenna immersed in an incident plane wave.

plane wave is a special case of the more general scattering problem in which a metallic or dielectric object is immersed in an incident wave and generates a scattered field. A traditional radar uses this scattered field to determine the position of an airplane. A ground penetrating radar uses the scattered signal to e.g. detect the presence of mines. The value of the scattered field is such that the total field ($\mathbf{e}^t = \mathbf{e}^i + \mathbf{e}^s$, $\mathbf{h}^t = \mathbf{h}^i + \mathbf{h}^s$) does satisfy the zero tangential electric field boundary condition, i.e. $\mathbf{e}^t_{tan} = 0$ everywhere on the antenna.

The induced currents are not the only effect of the scattering process. At the same time a voltage difference between the antenna terminals will be generated. This voltage difference is called the *open circuit voltage* V_{open}. One could expect the complete solution of the scattering problem to be required to be able to determine V_{open}. Fortunately, this turns out not to be the case. It suffices to apply Lorentz reciprocity theorem in a clever way to relate V_{open} to the properties of the antenna in its transmission mode. We will not go into detail here but simply state the final result, i.e.

$$V_{open} = -\frac{2j\lambda}{Z_c I}\mathbf{E}_0 \cdot \mathbf{F}(-\mathbf{u}^i). \tag{8.28}$$

In (8.28) λ is the wavelength and Z_c is the characteristic impedance of free space. Of major importance is the fact that V_{open} is intimately linked with the properties of the antenna in its transmission mode. Indeed, in (8.28) $\mathbf{F}$ is the radiation vector of the antenna as determined in Section 8.2. Moreover, the required value of $\mathbf{F}$ is its value in the direction $-\mathbf{u}^i$ i.e. the receiving behaviour for a plane wave coming from a certain direction directly depends

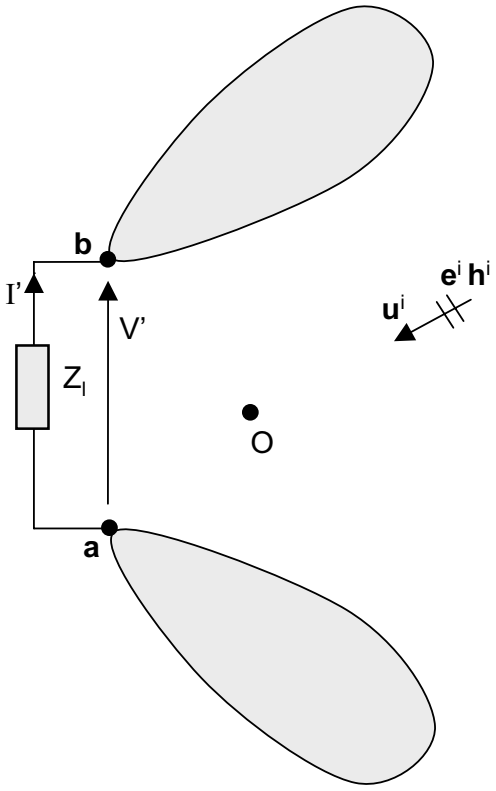


Figure 8.6: Receiving antenna with load impedance.

on the transmitting behaviour in that very same direction. Finally, V_{open} also depends upon the value of the incident electric field at the origin O, which is also the origin used in the calculation of the radiation vector $\mathbf{F}$. Remark that (8.28) does not depend on $\mathbf{E}_0$ and $\mathbf{F}$ separately but on their scalar product. The meaning and consequences of this will be discussed below.

8.7 Equivalent circuit of a receiving antenna

For the antenna to act as a receiver, we have to connect its terminals to the receiver circuit. This is shown in Fig. 8.6. The receiver is represented by a simple Thevenin equivalent load Z_l (which can of course be frequency dependent). When immersed in an incident plane wave, the scattering process will now also result in a current I' flowing through the load Z_l. The voltage difference between terminals a and b is different from the open circuit voltage V_{open} and is given by V'. Again applying Lorentz reciprocity theorem (the details of which are not discussed here), the following value for I' is obtained

$$I' = -\frac{V_{open}}{Z_a + Z_l}. \tag{8.29}$$

This immediately leads to the equivalent circuit description of the receiving antenna displayed in Fig. 8.7. The fact that the radiation impedance of the

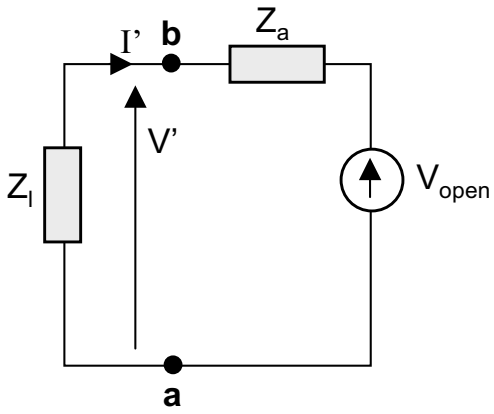


Figure 8.7: Equivalent circuit of a receiving antenna.

antenna is identical when sending or receiving again confirms the intimate relationship between the antenna as a receiver or as a transmitter. Note that the open circuit voltage V_{open} turns out to be the Thevenin source of the equivalent circuit of Fig. 8.7. If we remove Z_l ($Z_l = \infty$), V_{open} is indeed the voltage between terminals a and b when left unloaded.

8.8 Effective cross-section of a receiving antenna

We now focus our attention on the power P_l available at the load impedance $Z_l = R_l + jX_l$. This power must be sufficiently large as compared to the noise in order to be able to detect a signal. To state this in a more quantitative way, the receiver circuit itself must be studied in more detail.

From the equivalent circuit in Fig. 8.7 the power dissipated in Z_l is

$$P_l = \frac{R_l}{2|Z_l + Z_a|^2}|V_{open}|^2. \tag{8.30}$$

The maximum value of this quantity

$$P_{max} = \frac{|V_{open}|^2}{8R_a}, \tag{8.31}$$

is obtained for $Z_l = Z_a^*$. This allows to express P_l as

$$P_l = \frac{4R_lR_a}{|Z_l + Z_a|^2}P_{max} = MP_{max}, \tag{8.32}$$

introducing the mismatch factor M. Next, expression (8.28) for the open circuit voltage is substituted, leading to

$$P_{max} = \frac{\lambda^2}{2Z_c^2R_a}\frac{|\mathbf{F}(-\mathbf{u}^i)\cdot\mathbf{E}_0|^2}{|I|^2}. \tag{8.33}$$

The radiation resistance (8.24) can be expressed in terms of the directivity using (8.16)

$$R_a = \frac{1}{Z_c|I|^2}\int_\Omega |\mathbf{F}|^2 d\Omega = \frac{4\pi|\mathbf{F}(-\mathbf{u}^i)|^2}{Z_c|I|^2D(-\mathbf{u}^i)}. \tag{8.34}$$

This finally allows to eliminate $R_a|I|^2$ from P_{max}

$$P_{max} = \frac{\lambda^2 D(-\mathbf{u}^i)|\mathbf{F}(-\mathbf{u}^i)\cdot\mathbf{E}_0|^2}{8\pi Z_c|\mathbf{F}(-\mathbf{u}^i)|^2}. \tag{8.35}$$

The *effective cross-section* $\sigma_{eff}(\mathbf{u}^i)$ of the antenna is now defined as the ratio of the power dissipated in the load (i.e. the power available at the receiver) to the power density of the incident plane wave

$$\sigma_{eff}(\mathbf{u}^i) = \frac{P_l}{\frac{|\mathbf{E}_0|^2}{2Z_c}} = \frac{\lambda^2 D(-\mathbf{u}^i)M|\mathbf{F}(-\mathbf{u}^i)\cdot\mathbf{E}_0|^2}{4\pi|\mathbf{F}(-\mathbf{u}^i)|^2|\mathbf{E}_0|^2}. \tag{8.36}$$

The power of the incident wave $\frac{|\mathbf{E}_0|^2}{2Z_c}$ is expressed in Watt/m^2. The receiving antenna acts as if it consisted of a surface $\sigma_{eff}(\mathbf{u}^i)$ (in m^2), perpendicular to the direction of the incident plane wave, which is able to capture the complete power that penetrates that surface.

To account for the fact that the metal of the antenna will not be perfectly conducting and hence exhibit Joule losses and/or for the fact that some of the incident power will be dissipated in non-perfect dielectrics in the immediate neighbourhood of the antenna, an efficiency coefficient $\eta(\theta,\phi)$ is introduced together with the power gain $G(\theta,\phi)$

$$G(\theta,\phi) = \eta(\theta,\phi)D(\theta,\phi), \tag{8.37}$$

defined as the product of the directivity and the efficiency coefficient. Finally, a compact description of the effective cross-section is obtained by further defining the polarisation factor Q as

$$Q(-\mathbf{u}^i) = \frac{|\mathbf{F}(-\mathbf{u}^i)\cdot\mathbf{E}_0|^2}{|\mathbf{F}(-\mathbf{u}^i)|^2|\mathbf{E}_0|^2}, \tag{8.38}$$

which allows to rewrite σ_{eff} as

$$\sigma_{eff}(\mathbf{u}^i) = \frac{\lambda^2 G(-\mathbf{u}^i)}{4\pi} MQ(-\mathbf{u}^i). \tag{8.39}$$

The effective cross-section turns out to be proportional to the square of the wavelength, to the power gain (which itself is directly proportional to the directivity of the antenna, defined for the antenna acting as a transmitter), to the mismatch factor (which expresses the matching, or rather lack of it, between the load impedance and the radiation impedance) and the polarisation factor. It can easily be seen from (8.38) that this polarisation factor does neither depend upon the amplitude of the incident wave nor on the amplitude of the radiation vector but only upon the "angle" (or rather the cosine of that angle) between the complex vectors $\mathbf{F}(-\mathbf{u}_i)$ and $\mathbf{E}_0$.

Let us take a closer look as to what this angle means. Consider a plane perpendicular to the direction of incidence $\mathbf{u}_i$ as shown in Fig. 8.8. In this plane $\mathbf{F}(-\mathbf{u}^i)$ and $\mathbf{E}_0$ can be represented as

$$\mathbf{F}(-\mathbf{u}^i) = A(\mathbf{u}_x - j\tau\mathbf{u}_y), \tag{8.40}$$

$$\mathbf{E}_0 = B(\mathbf{u}'_x - j\tau'\mathbf{u}'_y), \tag{8.41}$$

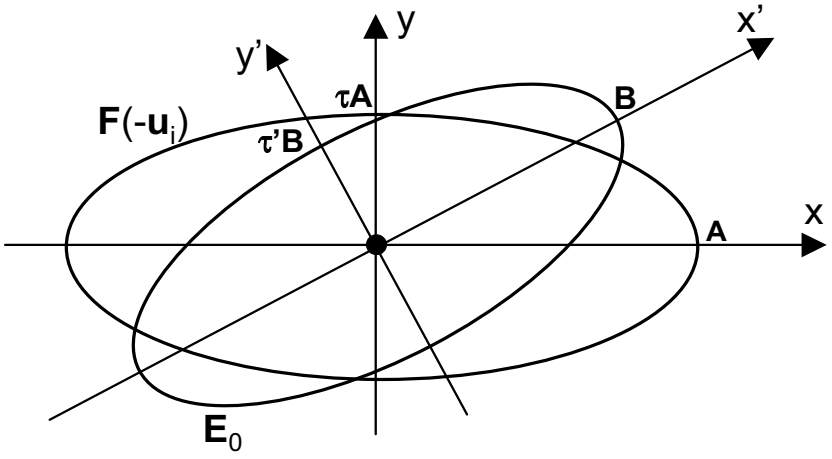


Figure 8.8: Polarisation ellipses of the incident plane wave and of the radiation vector.

with $\mathbf{u}_x$ en $\mathbf{u}_y$ along the main axes of the polarisation ellipse of $\mathbf{F}(-\mathbf{u}^i)$ and with $\mathbf{u}'_x$ en $\mathbf{u}'_y$ along the main axes of the polarisation ellipse of $\mathbf{E}_0$. The eccentricities of the respective ellipses are given by τ and τ'. With α representing the angle between $\mathbf{u}_x$ and $\mathbf{u}'_x$, $Q(-\mathbf{u}^i)$ can be expressed as

$$Q(-\mathbf{u}^i) = \frac{(\tau' - \tau)^2}{(1+\tau'^2)(1+\tau^2)} + \cos^2\alpha \frac{(1-\tau^2)(1-\tau'^2)}{(1+\tau^2)(1+\tau'^2)}. \tag{8.42}$$

Q reaches its maximum for $\alpha = 0$, i.e. when the main axes of both polarisation ellipses coincide (and this regardless of the values of τ and τ')

$$Q_{max} = \frac{(1-\tau'\tau)^2}{(1+\tau'^2)(1+\tau^2)}. \tag{8.43}$$

The minimum value is obtained when these axes are perpendicular, i.e. for $\alpha = \pi/2$

$$Q_{min} = \frac{(\tau' - \tau)^2}{(1+\tau'^2)(1+\tau^2)}. \tag{8.44}$$

The maximum maximorum $Q = 1$ is obtained for $\tau = -\tau'$ en $\alpha = 0$. In that case the ellipses are identical in shape and orientation, but seen from the antenna, the field vectors rotate in opposite sense. The minimum minimorum $Q = 0$ is obtained for $\tau = \tau'$ en $\alpha = \pi/2$. The ellipses are again identical in shape but are oriented perpendicularly and the field vectors rotate in the same sense. It must by now be clear to the reader that the cosine of the angle between two complex vectors expresses the quite intricate phenomenon of how the polarisation of these vectors (i.e their orientation in space as a function of time) influences their interaction.

The maximum value $D(-\mathbf{u}^i)\lambda^2/(4\pi)$ of the effective cross-section $\sigma_{eff}(\mathbf{u}^i)$ is obtained in the absence of impedance mismatch ($M = 1$), in the absence of losses in or near the antenna ($\eta = 1$) and provided the polarisation ellipses of the radiation vector $\mathbf{F}(-\mathbf{u}^i)$ and of the incident field $\mathbf{E}_0$ are identical in shape

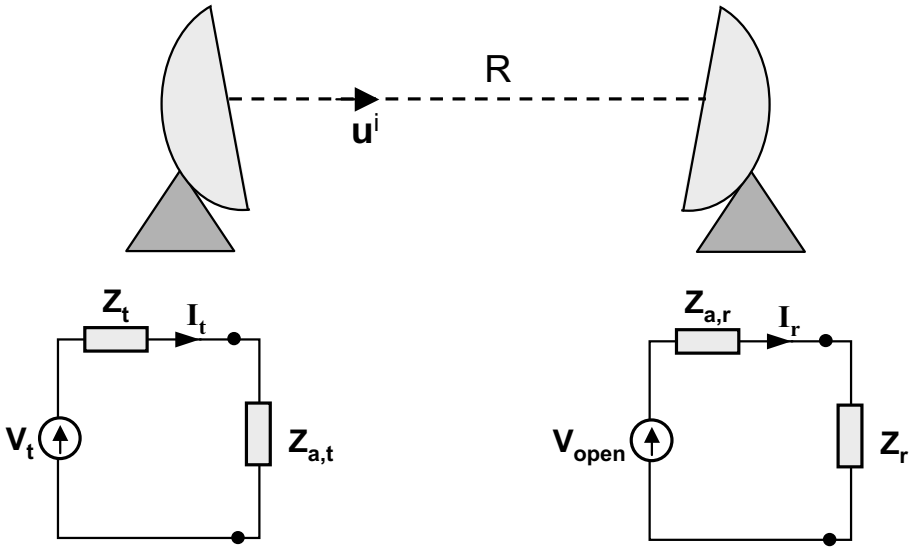


Figure 8.9: Communication between two antennas.

and orientation with the respective field vectors rotating in opposite sense as viewed from the antenna ($Q = 1$). Note that rotating in the opposite sense as viewed from the antenna implies that the rotation sense is identical when the rotation is observed from the point of view of the antenna for $\mathbf{F}(-\mathbf{u}^i)$ and from the direction the incident plane wave emerges from for $\mathbf{E}_0$.

8.8.1 Antenna link: Friis formula

Wireless systems use a communication link between two antennas. It is important to know how much power will be available at the receiver for a given total radiated power at the transmitter. This problem can be solved with the results obtained in previous sections. Consider the communication between a transmitter antenna and a receiver antenna as depicted in Fig. 8.9. The transmitter antenna, at the left side, is fed by a Thevenin source V_t with internal impedance Z_t. Its radiation impedance and radiation vector respectively are $Z_{a,t}$ and $\mathbf{F}_t(\theta, \phi)$. I_t is the current at the terminals of the transmitter antenna. The radiation impedance and radiation vector of the receiver antenna are given by $Z_{a,r}$ and $\mathbf{F}_r(\theta, \phi)$. In defining the radiation vectors, we are confronted with the problem that their value depends upon the complex amplitude of the current at the antenna terminals. This point was already mentioned when discussing the radiation impedance (see (8.24) and (8.25)). Here, we assume that the radiation vectors introduced above were obtained for a unit terminal input current of $1A$. The reader is urged to verify that the final result of this section is not influenced by this assumption. The receiver antenna is connected to the load Z_r. The distance R between transmitter and receiver is sufficiently large for each antenna to be in the far field of the other one and the receiving antenna is located in the direction $\mathbf{u}^i$ as seen from the transmitter. We want to calculate the power dissipated in Z_r and express this power as a fraction of the total power radiated by the transmitter.

The current I_t flowing through the terminals of the transmitter antenna is

$$I_t = \frac{V_t}{Z_t + Z_{a,t}}. \tag{8.45}$$

As the receiver is in the far field of the transmitter, the incident field at the receiver is locally plane with its electric field given by

$$\mathbf{E}^i = I_t \mathbf{F}_t(\mathbf{u}^i) \frac{e^{-jkR}}{R}. \tag{8.46}$$

The presence of I_t as a multiplicative factor is due to the fact that $\mathbf{F}_t$ has been defined for a unit terminal input current. The open circuit voltage of the receiver due to the incident electric field is

$$V_{open} = -\frac{2j\lambda}{Z_c} \mathbf{E}^i \cdot \mathbf{F}_r(-\mathbf{u}^i). \tag{8.47}$$

The value of I in the general expression (8.28) is 1 because $\mathbf{F}_r$ was also defined for a unit terminal input current. From this we easily derive the power dissipated in the load as

$$P_r = \frac{1}{2} R_r |I_r|^2 = \frac{1}{2} R_r \frac{|V_{open}|^2}{|Z_{a,r} + Z_r|^2}, \tag{8.48}$$

with R_r the real part of Z_r. Next, we introduce the mismatch factors M_t and M_r of both antennas

$$M_t = \frac{4R_t R_{a,t}}{|Z_t + Z_{a,t}|^2}, \qquad M_r = \frac{4R_r R_{a,r}}{|Z_r + Z_{a,r}|^2}, \tag{8.49}$$

with R_t the real part of Z_t. As in the previous section, the radiation resistances can again be expressed in terms of the corresponding directivities

$$R_{a,t} = \frac{4\pi |\mathbf{F}_t(\mathbf{u}^i)|^2}{Z_c D_t(\mathbf{u}^i)}, \qquad R_{a,r} = \frac{4\pi |\mathbf{F}_r(-\mathbf{u}^i)|^2}{Z_c D_r(-\mathbf{u}^i)}. \tag{8.50}$$

Note again that the unit amplitude terminal current assumption was invoked. The directivities $D_t(\mathbf{u}^i)$ and $D_r(-\mathbf{u}^i)$ can be replaced by the corresponding power gains $G_t(\mathbf{u}^i)$ en $G_r(-\mathbf{u}^i)$ by taking into account the efficiency factors of both antennas.
Finally, note that the maximum power the transmitter source can deliver to the transmitter antenna is

$$P_{t,max} = \frac{|V_t|^2}{8R_t}. \tag{8.51}$$

Collecting the above results yields the final ratio of the received power P_r to the maximum available power at the transmitter

$$\frac{P_r}{P_{t,max}} = M_r G_r(-\mathbf{u}^i) \frac{\lambda^2}{(4\pi)^2 R^2} \frac{|\mathbf{F}_t(\mathbf{u}^i) \cdot \mathbf{F}_r(-\mathbf{u}^i)|^2}{|\mathbf{F}_t(\mathbf{u}^i)|^2 |\mathbf{F}_r(-\mathbf{u}^i)|^2} G_t(\mathbf{u}^i) M_t. \tag{8.52}$$

This result is known as the *Friis formula*, usually expressed in dB

$$\begin{aligned} [\tfrac{P_r}{P_{t,max}}]_{dB} = &-21.98dB - 20log_{10}(R/\lambda) + (M_t)_{dB} + (M_r)_{dB} \\ &+[G_r(-\mathbf{u}^i)]_{dB} + [G_t(\mathbf{u}^i)]_{dB} + (Q_{rt})_{dB}. \end{aligned} \tag{8.53}$$

The above formula shows that the power budget available for the communication between two antennas diminishes with $20dB$ when increasing the distance between transmitter and receiver by an extra wavelength. The factor $\frac{\lambda}{(4\pi R)}$ is known as the free space loss factor and is due to the spherical spreading of the waves. At 2.45GHz e.g. (i.e. for $\lambda \approx 12cm$) this free space loss factor is about 60dB for a distance R of 10m. The available power budget can further be reduced by the mismatch and by the radiation efficiency of both transmitter and receiver.

To enhance the communication it is important that the respective antenna gains (for the transmitter the gain in the direction $\mathbf{u}^i$ and for the receiver the gain in the direction $-\mathbf{u}^i$) are as high as possible. This of course explains why, when considering an antenna link between two dish antennas, these antennas are oriented such that their axes coincide. In a GSM system, the size of the antenna of the cell phone ultimately restricts the gain. Moreover, in that case omnidirectionality is important as communication must remain possible whatever the orientation of the cell phone. At the base station several fixed antennas are used which each cover part of space and hence antennas with higher gains can be used.
Finally, remark that here again the polarisation of receiver and transmitter play a crucial role. The polarisation factor Q_{rt}

$$\frac{|\mathbf{F}_t(\mathbf{u}^i) \cdot \mathbf{F}_r(-\mathbf{u}^i)|^2}{|\mathbf{F}_t(\mathbf{u}^i)|^2 |\mathbf{F}_r(-\mathbf{u}^i)|^2} \tag{8.54}$$

is completely analogous to the one encountered in (8.28). In the worst case, a polarisation mismatch between transmitter and receiver polarisation (e.g. a transmitter generating a vertically polarised electric field while receiver currents can only be induced by a horizontal electric field) makes communication impossible.

8.9 Thin wire antennas

The difficulty to determine the actual current distribution on an antenna poses a major problem when studying different antenna types. Generally speaking, (often complex) numerical techniques must be invoked to determine these currents. To that purpose, powerful CAD-tools are nowadays available, reducing the importance of approximate analytical techniques. In this section such an analytical approximation to the linear thin wire or dipole antenna will be discussed as it provides sound physical insight. At the same time, we will explain how an integral equation can be formulated for the unknown currents on the antenna, but the actual solution of that integral equation will not be pursued.

Consider the dipole antenna with length $2L$, located in free space, as shown on Fig. 8.10. The diameter of the perfectly conducting wires is very small with respect to the wavelength such that azimuthal current density variations can be neglected, allowing to idealise the antenna as an infinitely thin wire with a current $I(z)$ running along it. The electric field of this current is given by (2.85)

$$\mathbf{e}(\mathbf{r}) = -\frac{j\omega\mu_0}{4\pi} \int_{-L}^{+L} \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} I(z')\mathbf{u}_z dz' - \frac{1}{4\pi\epsilon_0} \nabla \int_{-L}^{+L} \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \rho(z') dz'. \tag{8.55}$$

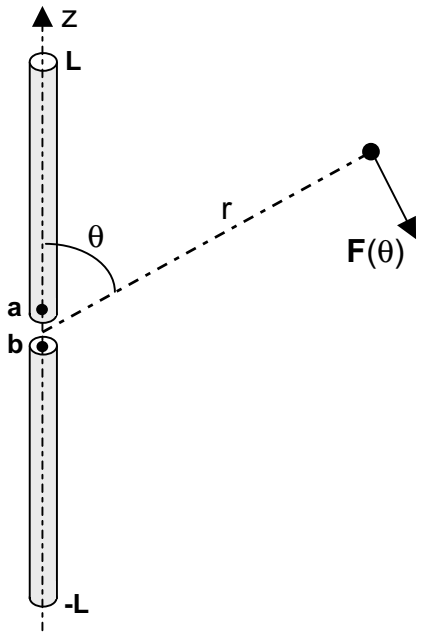


Figure 8.10: Dipole antenna with length $2L$.

with $\mathbf{r}' = z'\mathbf{u}_z$ and with $\rho(z')$ the line charge density corresponding with the current $I(z')$

$$\rho(z) = -\frac{1}{j\omega}\nabla \cdot \mathbf{j}(\mathbf{r}) = -\frac{1}{j\omega}\frac{d}{dz}I(z). \tag{8.56}$$

The antenna is directly excited by a voltage source such that a potential difference V_s is enforced between its terminals a and b. Provided the spacing between these terminals remains sufficiently small, the electric field along the line segment ab is given by $E_s = -V_s/\Delta$ with Δ the distance between a and b. As the dipole antenna is perfectly conducting, the tangential component of the electric field must be zero on its surface except near the source (on the segment ab) where the tangential electric field is forced to be E_s. Expression (8.55) is valid for any observation point $\mathbf{r}$. We now let $\mathbf{r}$ approach the wire. The tangential electric field boundary condition then leads to

$$\lim_{\mathbf{r}\to wire\ surface}[-\frac{j\omega\mu_0}{4\pi}\int_{-L}^{+L}\frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}I(z')dz'$$

$$-\frac{1}{4\pi\epsilon_0}\mathbf{u}_z \cdot \nabla\int_{-L}^{+L}\frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\rho(z')dz'] = \nu, \tag{8.57}$$

with $\nu = 0$ for $\mathbf{r}$ approaching the wire (the source segment ab excluded) and $\nu = E_s$ for $\mathbf{r}$ approaching the source segment ab. Equation (8.57) is an integral equation for the unknown current $I(z)$ along the dipole antenna. Taking the

limit in (8.57) by putting $\mathbf{r} = z\mathbf{u}_z$ simplifies (8.57) to

$$[-\frac{j\omega\mu_0}{4\pi}\int_{-L}^{+L}\frac{e^{-jk|z-z'|}}{|z-z'|}I(z')dz'$$

$$-\frac{1}{4\pi\epsilon_0}\frac{d}{dz}\int_{-L}^{+L}\frac{e^{-jk|z-z'|}}{|z-z'|}\rho(z')dz'] = \nu, \tag{8.58}$$

with $-L \leq z \leq +L$. The reader will immediately remark the presence of a non-integrable singularity in (8.58) at $z = z'$. This proves that the limiting procedure must be handled with care. The mathematical problem that arises is well-known. It is that of determining the so-called self-patch contribution. We will not go into further detail here, nor will we discuss the actual numerical solution of the integral equation (8.58).

Let us now concentrate on the far field of the dipole antenna. The integral in expression (8.3) for $\mathbf{N}(\theta, \phi)$ reduces to a line integral

$$\mathbf{N}(\theta,\phi) = \mathbf{u}_z\frac{\mu_0}{4\pi}\int_{-L}^{L}e^{jkz'\cos\theta}I(z')dz'. \tag{8.59}$$

Due to the rotational symmetry with respect to the z-axis, all fields are ϕ-independent. The far fields are

$$\mathbf{e}(\mathbf{r}) = \frac{e^{-jkr}}{r}\mathbf{F}(\theta), \tag{8.60}$$

$$\mathbf{h}(\mathbf{r}) = \frac{1}{Z_c}\frac{e^{-jkr}}{r}\mathbf{u}_r \times \mathbf{F}(\theta), \tag{8.61}$$

with

$$\mathbf{F}(\theta) = \frac{j\omega\mu_0}{4\pi}\sin\theta\int_{-L}^{+L}e^{jkz'\cos\theta}I(z')dz'\mathbf{u}_\theta. \tag{8.62}$$

This expression for $\mathbf{F}$ is proportional to the Fourier transform of the current with respect to the Fourier variable $k\cos\theta$. Only Fourier variables in the range $[-k, k]$ are relevant.

As emphasised before, the current $I(z')$ can only be determined numerically, e.g. using the integral equation derived above. It however turns out that a suitable *approximation* or *ansatz* for that current is the sinusoidal current

$$I(z) = I_m \sin k(L - |z|) \tag{8.63}$$

with the current entering the terminals given by $I(0) = I_m \sin kL$. This immediately leads to

$$\mathbf{F}(\theta) = \frac{jI_mZ_c}{2\pi}\frac{\cos(kL\cos\theta) - \cos kL}{\sin\theta}\mathbf{u}_\theta. \tag{8.64}$$

For small θ-values $\mathbf{F}$ is proportional to $\sin\theta$. Fig. 8.11 shows cross-sections of

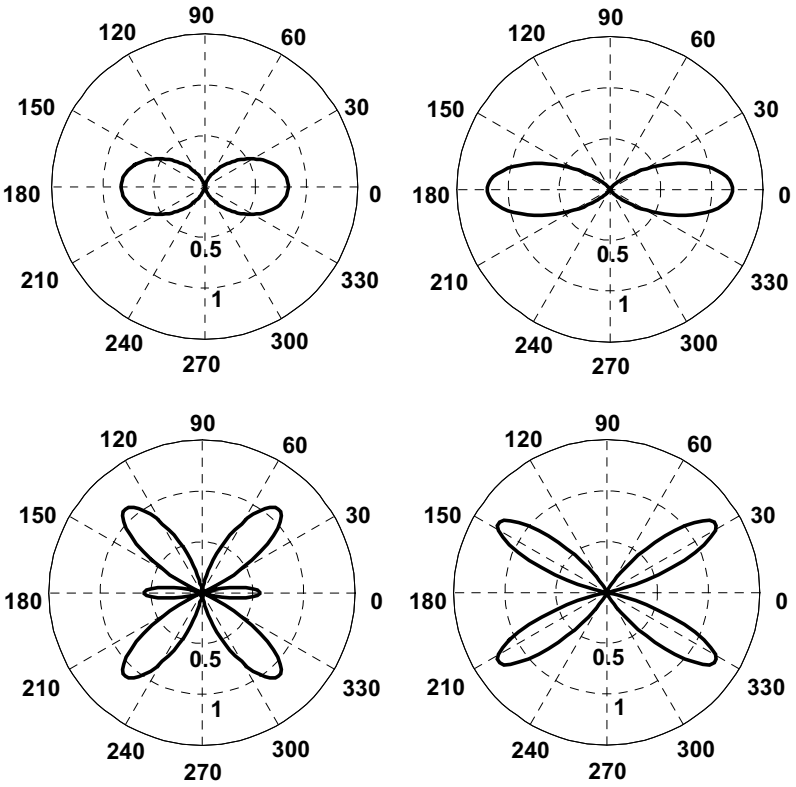


Figure 8.11: Directivity for thin wire antennas with $2L = \lambda/2$ (top left), $2L = \lambda$ (top right), $2L = 3\lambda/2$ (bottom left) and $2L = 2\lambda$ (bottom right).

the directivity for different lengths of the dipole. The complete radiation pattern is obtained by rotating these cross-sectional results around the axis. The pattern of the half-wavelength dipole ($2L = \lambda/2$) in Fig. 8.11 (top left) has a 3dB beamwidth of 78^o. The beamwidth of the full wavelength dipole in Fig. 8.11 (top right) is only 47^o. When still increasing the length of the dipole, as in Fig. 8.11 (bottom left) ($2L = 3\lambda/2$) and Fig. 8.11 (bottom right) ($2L = 2\lambda$), the radiation pattern starts exhibiting several lobes.

The approximate sinusoidal representation of the current (8.63) is inaccurate for antennas that are a multiple of the wavelength. In that case the terminal current $I(0)$ vanishes and the radiation impedance becomes infinite. A thorough comparison with the numerical solution of the integral equation shows that (8.63) becomes inaccurate for $2L > \lambda$. The difference between the exact and approximate solution also depends upon the diameter of the dipole. This difference is much less pronounced for the radiation pattern and the radiation resistance than for the near field and for the radiation reactance.

To conclude this section we turn to the example of a dipole antenna at 1GHz ($\lambda = 30cm$) with a diameter of $20\mu m$. Fig. 8.12 shows $|I(z)|$ (in A) for a dipole with length $2L = 10cm$ generated by 1V potential difference between the terminals. The distance on the horizontal axis is expressed in terms of the wavelength. The dashed curve (which almost completely coincides with the exact

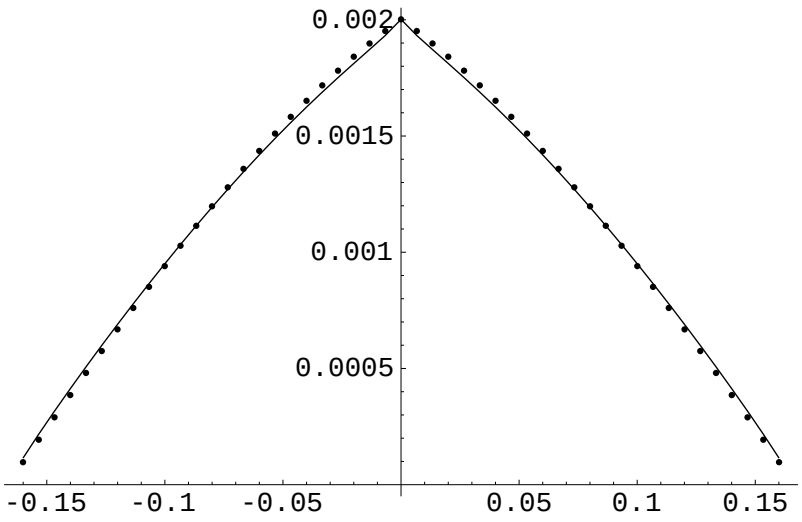


Figure 8.12: Amplitude of the current on a dipole with length $2L = \lambda/3$.

result) corresponds to the sinusoidal ansatz of (8.63). The radiation impedance turns out to be $Z_a = (25.6 - j499)\Omega$. Fig. 8.13 shows $|I(z)|$ (in A) for a dipole with length $2L = 2\lambda/3 = 20cm$, again generated by 1V potential difference between the terminals. The ansatz (8.63) is already much less accurate. The radiation impedance is $Z_a = (249 + j677)\Omega$.

8.9.1 Half-wavelength dipole

This important special case is obtained for $2L = \lambda/2$

$$I(z) = I_m \cos \frac{\pi z}{2L}, \tag{8.65}$$

$$\mathbf{F}(\theta) = \frac{jI_m Z_c}{2\pi} \frac{\cos(\frac{\pi}{2}\cos\theta)}{\sin\theta} \mathbf{u}_\theta, \tag{8.66}$$

$$R_a = \frac{Z_c}{2\pi} \int_0^\pi \frac{[\cos(\frac{\pi}{2}\cos\theta)]^2}{\sin\theta} d\theta = 0.1940 Z_c = 73.08\Omega. \tag{8.67}$$

As the length of the antenna is half a wavelength, the antenna is said to be "in resonance". To understand the meaning of this, Fig. 8.14 shows the radiation impedance as a function of frequency of a 15 cm dipole. The radius of the wire is 1 mm. The numerical results were obtained with the NEC code. NEC or *N*umerical *E*lectromagnetics *C*ode is one of the first numerical codes developed for antenna modelling. Many freeware version of it can be found on the web. The first resonance frequency is just below 1 GHz. For frequencies below the resonance frequency the dipole behaves capacitively ($X_a < 0$) to switch to an inductive behaviour ($X_a > 0$) above the resonance frequency. At resonance, capacitive and inductive effects cancel out. Three other resonance frequencies can also be observed. Fig. 8.15 shows a more detailed picture of the radiation impedance as a function of frequency of the same 15 cm dipole in the neighbourhood of its first resonance frequency. The radius of the wire is now only

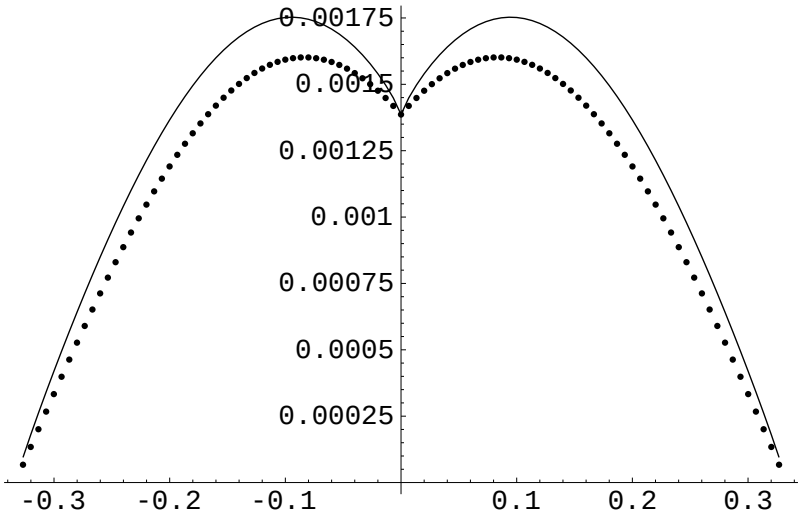


Figure 8.13: Amplitude of the current on a dipole with length $2L = 2\lambda/3$.

$0.1\mu m$. Numerical results were again obtained with NEC. At resonance (981.5 MHz), the radiation impedance is real with a value of 72.37Ω, which differs only by 1 percent from (8.67).

8.9.2 Very short dipole

Let us now investigate the radiation properties of a piece a wire which is small compared to the wavelength ($2L \ll \lambda$). The final results are

$$I(z) = I_m k(L - |z|), \tag{8.68}$$

$$\mathbf{F}(\theta) = \frac{jk^2L^2I_mZ_c}{4\pi}\sin\theta\mathbf{u}_\theta = \frac{j\omega\mu_0 I(0)L}{4\pi}\sin\theta\mathbf{u}_\theta, \tag{8.69}$$

$$R_a = \frac{Z_c}{6\pi}k^2L^2 = 19.99k^2L^2\Omega, \tag{8.70}$$

with $I(0) = I_m kL$. For $L = \lambda/20$ the radiation resistance amounts to 1.97Ω. Hence, for a total radiated power of $100W$, a terminal current of $10A$ is required. For a half-wavelength dipole the same radiated power is obtained for a current of about $1.65A$. The radiation patterns of the short dipole and of the half-wavelength dipole are hardly different as shown in Fig. 8.16. The half-wavelength dipole radiates much more efficiently than the short dipole, hence its practical importance. As in both cases the antenna is assumed to be perfectly conducting this has nothing to do with Joule losses: for the same radiated power, the currents on the short dipole are simply much larger than on the half-wavelength dipole. Of course, in the inevitable presence of conductor losses, these losses will be much higher for the short dipole reducing its efficiency. In many practical situations the use of a short dipole cannot be avoided e.g. a $80cm$ antenna for a radio receiver in the AM band for wavelengths between $200m$ and $600m$.

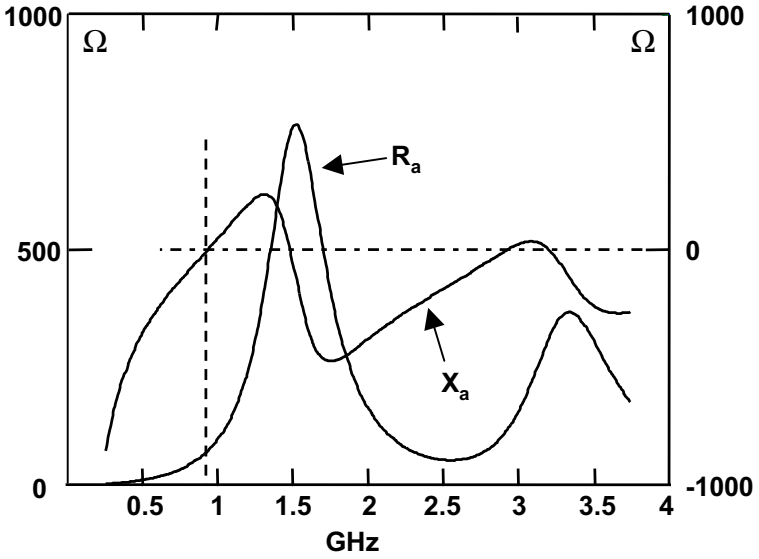


Figure 8.14: Radiation impedance $Z_a = R_a + jX_a$ (Ω) of a 15 cm dipole as a function of frequency.

8.10 Antenna arrays

To obtain a high gain antenna with a pencil like main beam, this antenna must be large with respect to the wavelength (e.g. a dish antenna). Another way to realise high gain is to combine a number of smaller antennas arranged in an array. One extreme example of such an array is a radio-telescope array used for radio-astronomy. Here, the individual dish antennas, which already themselves exhibit a high gain, are combined into one large antenna capable of capturing faint signals. An array of antennas offers a lot of flexibility as each of the constitutive antennas can be separately driven. This allows to design arrays with predefined radiation patterns or at least with radiation patterns satisfying a number of constraints. It is also possible to design an array with an electronically steerable beam.

In the sequel the basic principles of antenna arrays will be discussed together with a few examples. In this introductory course we again have to refer the reader to the vast body of literature on this topic for a more in-depth treatment. Consider a set of N antennas at positions $\mathbf{s}_i$, $(i = 1, 2, ..., N)$, each with their own radiation vector $\mathbf{F}_i(\theta, \phi)$ as depicted in Fig. 8.17. Each radiation vector is defined with respect to its own origin located at $\mathbf{r} = \mathbf{s}_i$. The results presented below are only correct provided the radiation vector of each antenna, as used in this section, is the radiation vector of that antenna in the presence of the other antennas and with the terminals of these antennas left open! Only when the spacing of the elements (as the individual antennas are also called) of the array is sufficiently large, such that the mutual influence becomes negligible, will these radiation vectors be identical to their counterparts for a single antenna placed in free space. However, in the vast majority of cases, the array elements will be closely spaced. Consequently, determining the correct radiation vectors

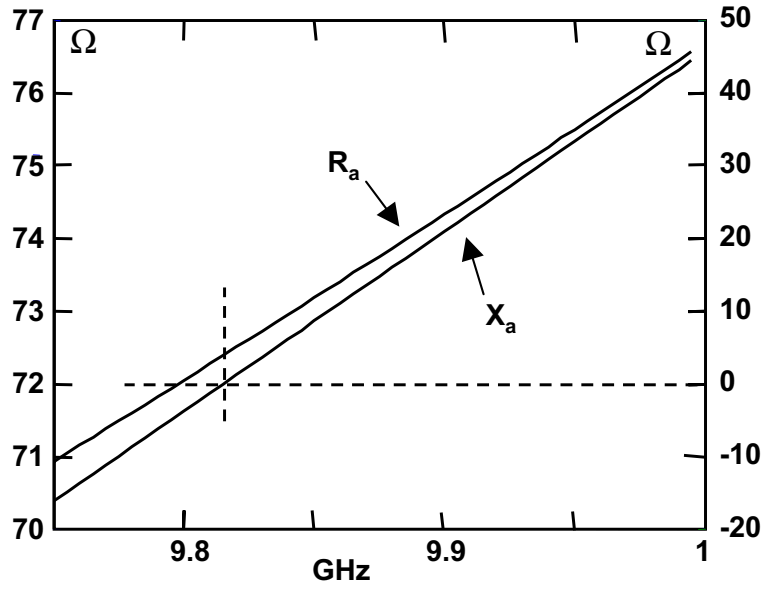


Figure 8.15: Radiation impedance $Z_a = R_a + jX_a$ (Ω) of a 15 cm dipole in the neighbourhood of its first resonance frequency.

often poses a real challenge.
The total electric far field of the array is given by

$$\mathbf{e}_{tot}(\mathbf{r}) = \sum_{i=1}^{N} \mathbf{F}_i(\theta, \phi) \frac{e^{-jk|\mathbf{r}-\mathbf{s}_i|}}{|\mathbf{r}-\mathbf{s}_i|}. \tag{8.71}$$

Next, we choose a common origin O sufficiently close to all the elements of the array. Following a similar reasoning as the one used in Section 8.2 to obtain the far field, the distance $|\mathbf{r}-\mathbf{s}_i|$ is approximated as

$$|\mathbf{r}-\mathbf{s}_i| \approx r - \mathbf{s}_i \cdot \mathbf{u}_r, \tag{8.72}$$

such that

$$\mathbf{e}_{tot}(\mathbf{r}) = \mathbf{F}(\theta, \phi) \frac{e^{-jkr}}{r}, \tag{8.73}$$

with

$$\mathbf{F}(\theta, \phi) = \sum_{i=1}^{N} \mathbf{F}_i(\theta, \phi) e^{jk\mathbf{s}_i \cdot \mathbf{u}_r}. \tag{8.74}$$

We have again carefully retained the phase information. $\mathbf{F}(\theta, \phi)$ is the radiation vector of the array. It is the weighted sum of the individual radiation vectors.

Now suppose that the array is formed by identical antennas (due to the mutual influence, this assumption will in practice only be met approximately). The radiation vector of each element is denoted as $\mathbf{F}_{norm}(\theta, \phi)$ with $\mathbf{F}_{norm}$ defined for a unit voltage over the antenna terminals. If, as in previous sections, the reader prefers the radiation vector definition for a unit current flowing through the antenna terminals, the theory presented here can be repeated, but the appropriate radiation vectors in this case are those obtained for each individual

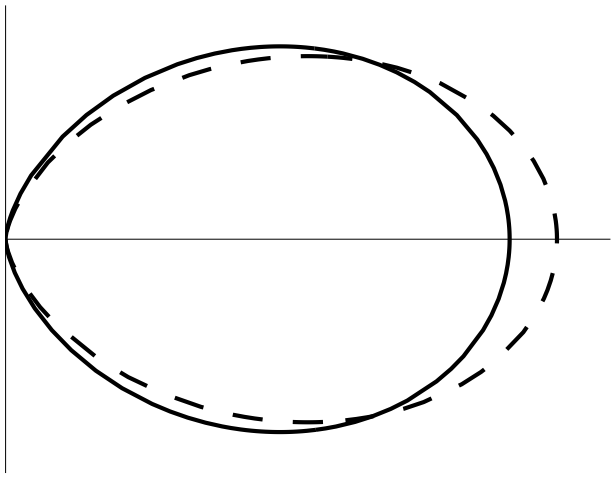


Figure 8.16: Radiation pattern of a very short dipole (full line) and of a half-wavelength dipole (dashed line).

antenna with the other ones short-circuited at their terminals. If we excite each antenna i with a voltage V_i, the radiation vector of the array becomes

$$\mathbf{F}(\theta,\phi) = \sum_{i=1}^{N} V_i e^{jk\mathbf{s}_i\cdot\mathbf{u}_r}\mathbf{F}_{norm}(\theta,\phi) = R(\theta,\phi)\mathbf{F}_{norm}(\theta,\phi), \tag{8.75}$$

where we have introduced the *array factor* $R(\theta,\phi)$,

$$R(\theta,\phi) = \sum_{i=1}^{N} V_i e^{jk\mathbf{s}_i\cdot\mathbf{u}_r}. \tag{8.76}$$

This result constitutes the simplest array theory and clearly shows the flexibility offered by an array. Although $\mathbf{F}_{norm}$ could e.g. exhibit the simple not very directive radiation pattern of a half-wavelength dipole, the combination of carefully positioning (choice of $\mathbf{s}_i$) and exciting (choice of V_i) each array element yet allows to obtain a desired radiation pattern through the multiplicative array factor $R(\theta,\phi)$. It is however overly optimistic to believe that $R(\theta,\phi)$ can completely be chosen at will especially for closely spaced array elements. The coupling between array elements and their feeding networks will prevent this.

To conclude this brief discussion of antenna arrays, consider the case of the linear array with equidistant elements positioned on the z-axis as depicted in Fig. 8.18. The distance between the elements is a. With the origin positioned at the phase centre of the first antenna element ($\mathbf{s}_i = i \times a\mathbf{u}_z$) the array factor becomes

$$R = \sum_{i=1}^{N} V_i e^{jka(i-1)\cos\theta}. \tag{8.77}$$

This polynomial is known as the Schelkunoff polynomial. Further suppose that all elements are steered with an identical voltage amplitude and with a constant

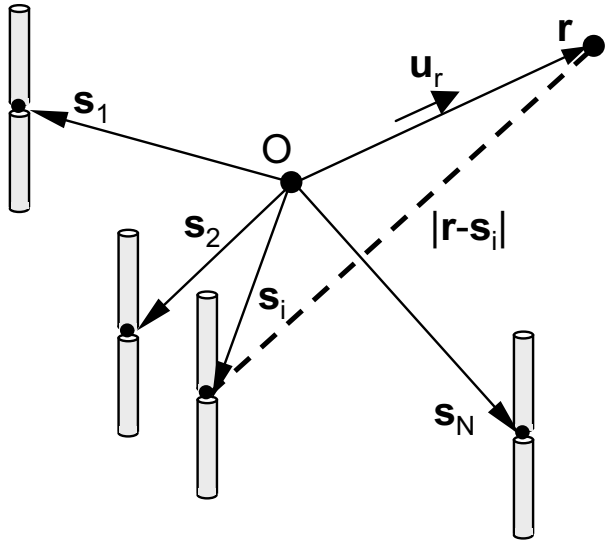


Figure 8.17: Antenna array.

phase difference α between consecutive elements

$$V_i = Ve^{j\alpha(i-1)}. \tag{8.78}$$

Such an array is known as a *phased array*. The array factor can now be expressed analytically

$$R = e^{j\frac{N-1}{2}\tau}\frac{\sin N\frac{\tau}{2}}{\sin\frac{\tau}{2}}, \tag{8.79}$$

with $\tau = \alpha + ka\cos\theta$. Fig. 8.19 displays

$$\frac{|R|}{N} = \frac{|\sin N\frac{\tau}{2}|}{N|\sin\frac{\tau}{2}|}. \tag{8.80}$$

The largest possible value of $|R|$ determines the direction of the main lobe of the array factor. In this direction all the contributions of the array elements interfere constructively. This maximum is obtained for $\tau_n = n2\pi$ and hence for angles θ_n satisfying

$$\cos\theta_n = \frac{n - \frac{\alpha}{2\pi}}{\frac{a}{\lambda}}. \tag{8.81}$$

As $-1 \leq \cos\theta_n \leq 1$, n must be in the range

$$-\frac{a}{\lambda} + \frac{\alpha}{2\pi} \leq n \leq \frac{a}{\lambda} + \frac{\alpha}{2\pi}. \tag{8.82}$$

The number of main lobes is hence determined by a/λ. For $a/\lambda < 1/2$ there is only a single direction in which constructive interference occurs. Shadow regions or zeros of the array factor correspond to directions of destructive interference determined by $\tau_n = n2\pi/N$. By changing α, the direction of the main lobe can be changed. This is called electronic beam steering as opposed to mechanical beam steering when the antenna itself is rotated. Also note that the main

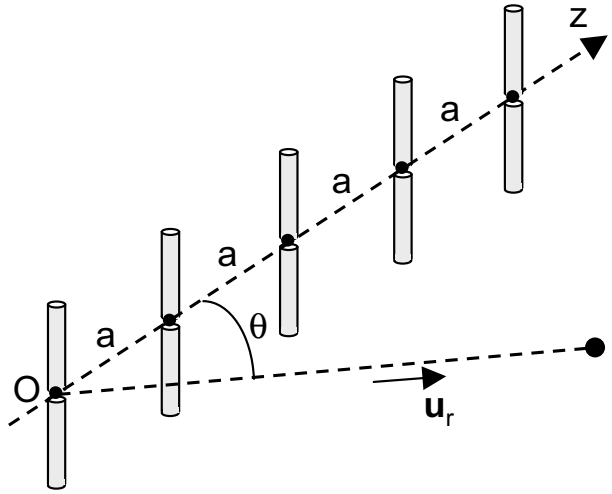


Figure 8.18: Linear antenna array.

lobe becomes more directive (more pencil like) when increasing the number of elements.
The array factor exhibits rotational symmetry with respect to the z-axis. This is not automatically the case for the radiation pattern of the array as the array factor must still be multiplied with the radiation vector of a single element.

Another interesting array is obtained for the very particular values of V_i given by

$$V_i = C_{N-1}^{i-1} V, \tag{8.83}$$

with C_{N-1}^{i-1} the binomial coefficients and for a half-wavelength spacing of the elements ($a = \lambda/2$). The array factor then turns out to be

$$R = V \sum_{i=1}^{N} C_{N-1}^{i-1} e^{j(i-1)\pi \cos\theta} = V(1 + e^{j\pi\cos\theta})^{N-1}, \tag{8.84}$$

or

$$|R| = |V||\cos(\frac{\pi}{2}\cos\theta)|^{N-1}. \tag{8.85}$$

This is an array factor with a single main lobe and no side lobes.

8.11 Overview of different types of antennas

This chapter concludes with a brief overview of a number of antenna types chosen among the vast variety of existing ones.

We already know the dipole antenna. Leaving out one arm of the dipole and introducing a metal plate to back the remaining arm, results in a monopole antenna (Fig. 8.20a) The metal plate acts as a mirror. In a car, this metal plate can e.g. be the roof of the car. This metal plate can also be left out as e.g. for the antenna on a radio receiver.
The radiation vector of a dipole antenna is zero along its axis. This can be avoided by folding the two arms to obtain the V-antenna depicted in Fig.

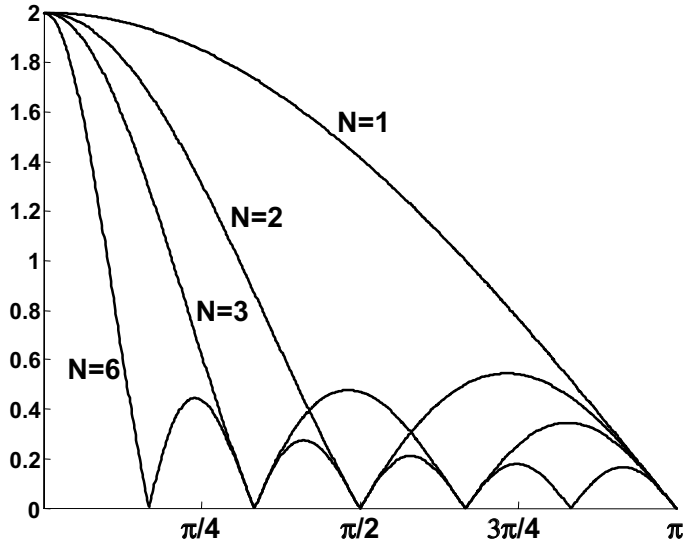


Figure 8.19: $|\sin N\frac{\tau}{2}|/(N|\sin\frac{\tau}{2}|)$ as a function of τ for $N = 1, 2, 3, 6$.

8.20b. Another type of wire antenna results when connecting the tips of a half-wavelength dipole with an additional wire. This is the folded dipole of Fig. 8.20c consisting of two parallel half-wavelength dipoles. For an identical input current, the radiated power is four times larger and the radiation resistance becomes $R_a = 4 \times 73.1\Omega = 292.4\Omega$. The radiated power can further be enhanced by backing this antenna by a metal plate placed at a distance of $\lambda/4$ (Fig. 8.20d). This metal plate prevents the antenna from radiating backwards. The metal plate can be replaced by a wire mesh.

Fig. 8.20e shows a Yagi-antenna, in the past typically used for reception of television broadcasting and a familiar sight on our roofs. The folded dipole is the actual receiving antenna. The other short-circuited dipoles in front of the folded dipole are called "directors" while the short-circuited dipole behind the folded dipole is the "reflector". The combination of directors, folded dipole and reflector results in a directive antenna with a large gain in the forward direction.

In the case of the monopole antenna, the wire can take the shape of a helix (Fig. 8.20f). This helix antenna generates a circularly polarised field.

The above wire antennas are narrow band antennas: they will only perform optimally close to the resonance frequency. Designing broadband antennas that operate over a large frequency range is difficult. One such an example is sketched in Fig. 8.21a. It is a spiral antenna the arms of which satisfy $r = e^{A\phi}$. For arms that extend to infinity, the properties of this antenna only depend upon angles and not upon distances. This implies that such an antenna must exhibit an infinite bandwidth. In practice, the spiral antenna will have finite dimensions and hence a finite bandwidth. The spiral antenna also generates a circularly polarised field. Another broadband antenna is the log-periodic antenna depicted in Fig. 8.21b. It consists of a number of driven folded dipole antennas. The law

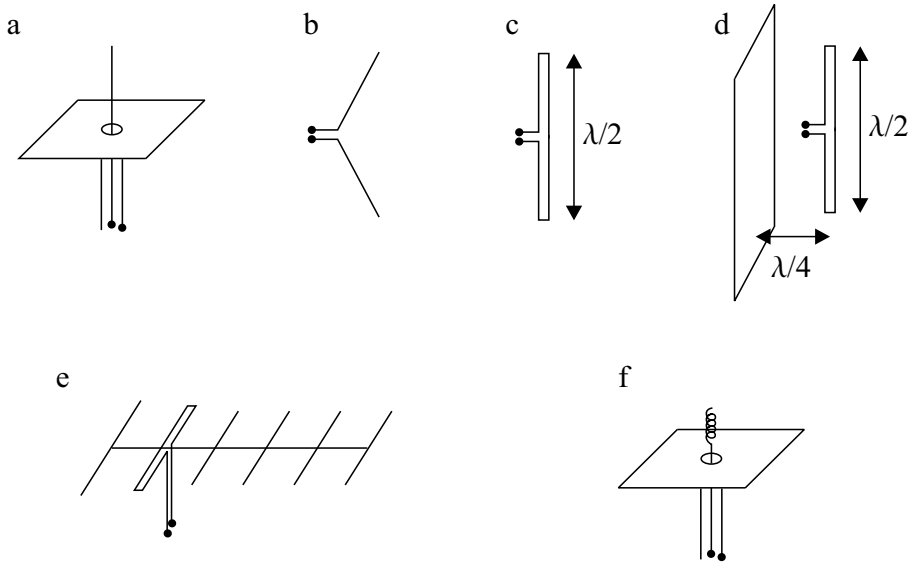


Figure 8.20: Some typical wire antennas.

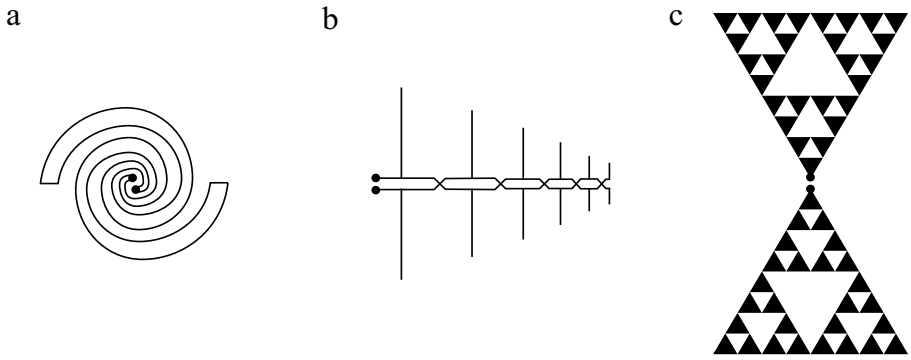


Figure 8.21: Some broadband wire antennas.

dictating the lengths of the dipoles and the distances between them is

$$\frac{l_i}{l_{i-1}} = \frac{d_i}{d_{i-1}} = \tau, \tag{8.86}$$

with τ an arbitrary constant. If the number of dipoles is infinite, the log-periodic antenna is electrically identical at frequency f and at frequency τf. This again yields a broadband behaviour that will of course depend upon the actual number of elements. The log-periodic antenna is an example of a fractal antenna which looks identical on each resolution level. Another example of such a fractal antenna is given in Fig. 8.21c. The black triangles are metal patches.

The bandwidth of a dipole antenna can also be increased by slowly increasing the wire diameter with increasing distance from the antenna terminals. This yields the biconical antenna of Fig. 8.22a. Another variation of the biconical antenna uses individual wires (Fig. 8.22b). By flattening the cones, we obtain

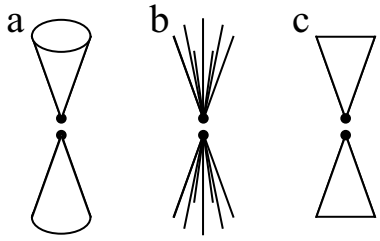


Figure 8.22: Biconical antennas (a) and (b) and bowtie antenna (c).

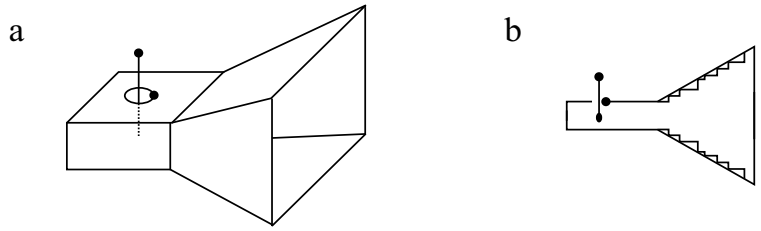


Figure 8.23: Horn antennas.

the bowtie antenna (Fig. 8.22c) consisting of two metal triangles.

A very different type of antenna is the rectangular horn antenna depicted in Fig. 8.23a. This antenna is fed by a rectangular waveguide (to the left). The horn gradually opens up and as such "adapts" the cross-section of the waveguide to free space. This assures minimal reflection of the eigenmode propagating in the waveguide and maximal power transmission to free space. The horn is said to be impedance matched with respect to the waveguide. A typical way to excite the waveguide is to insert a small wire through a hole or aperture in the waveguide wall. This wire acts as a monopole antenna. Circular horns are also used, in combination with circular waveguides. To increase the bandwidth of a horn antenna, its inner surface is shaped as shown on Fig. 8.23b. Such a horn is known as a "corrugated" horn. A horn antenna is often combined with a metal reflector e.g. a paraboloid yielding highly directive dish antennas for satellite television or for radio-astronomy purposes.

Planar antennas are realised using microstrip technology. Fig. 8.24a shows a microstrip patch antenna. The shaded area is metallised and rests on a dielectric substrate. This substrate is itself backed by a metal plate. At specific frequencies the current on the metal patch is in resonance and efficient radiation is obtained. This type of antenna is again inherently narrow banded. Many variations of this basic patch antenna exist (by changing the shape of the patch, by using a single or several feed points, by combining single patches into an array, by using a curved substrate, etc.). We must again refer the reader to the literature for more details on the present research in this domain.
An interesting variant of the patch antenna is the slot antenna of Fig. 8.24b. Here the radiation emerges from a slot in a completely metallised plane. The slot is fed by a microstrip line that passes under the slot as also depicted in Fig.

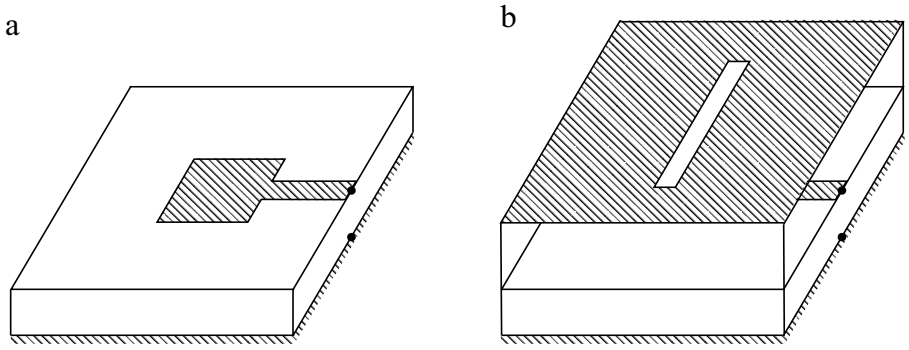
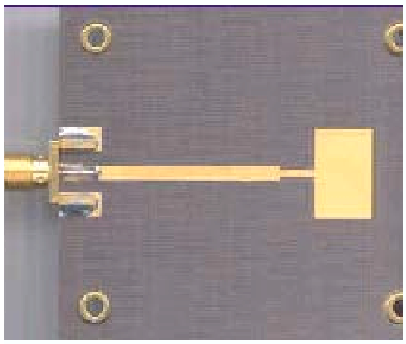
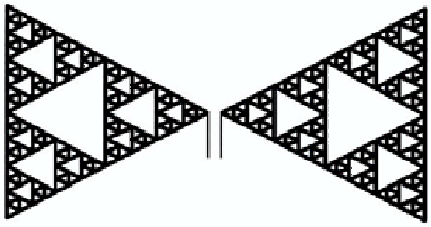


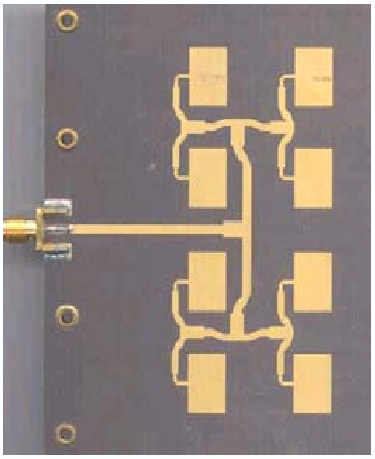
Figure 8.24: Microstrip patch antenna and slot antenna.

8.24b. Microstrip antennas are ideally suited for arrays as a large number of patches can easily be accommodated on a single substrate. Each element can then be excited using the additional signals layers of the multilayered substrate of which the array constitutes the outer layer.

It will be clear to the reader that predicting and optimising the behaviour of the above antennas with respect to the radiation pattern, the bandwidth, the input impedance, the polarisation, the size, etc. not only requires the antenna basics provided in this chapter, but also a lot of insight, experience and above all some of the powerful CAD-tools that have been put on the market in the past decade. This chapter concludes with a set of pictures of a variety of antennas selected among the large amount of examples to be found on the web.







Appendix A

Vector Analysis in Three Dimensions

A.1 Multiplicative relationships

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) \tag{A.1}$$

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c} = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{b}\cdot\mathbf{c})\mathbf{a} \tag{A.2}$$

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b}) \tag{A.3}$$

$$(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}) \tag{A.4}$$

A.2 Differential relationships

$$\nabla(ab) = a(\nabla b) + (\nabla a)b \tag{A.5}$$

$$\nabla(\mathbf{a}\cdot\mathbf{b}) = \mathbf{a}\times(\nabla\times\mathbf{b}) + \mathbf{b}\times(\nabla\times\mathbf{a}) + (\mathbf{b}\cdot\nabla)\mathbf{a} + (\mathbf{a}\cdot\nabla)\mathbf{b} \tag{A.6}$$

$$\nabla\cdot(a\mathbf{a}) = a(\nabla\cdot\mathbf{a}) + (\nabla a)\cdot\mathbf{a} \tag{A.7}$$

$$\nabla\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\cdot(\nabla\times\mathbf{a}) - \mathbf{a}\cdot(\nabla\times\mathbf{b}) \tag{A.8}$$

$$\nabla\times(a\mathbf{b}) = \nabla a\times\mathbf{b} + a\nabla\times\mathbf{b} \tag{A.9}$$

$$\nabla\times(\mathbf{a}\times\mathbf{b}) = \mathbf{a}\nabla\cdot\mathbf{b} - \mathbf{b}\nabla\cdot\mathbf{a} + (\mathbf{b}\cdot\nabla)\mathbf{a} - (\mathbf{a}\cdot\nabla)\mathbf{b} \tag{A.10}$$

$$\nabla\times\nabla a = 0 \tag{A.11}$$

$$\nabla\cdot(\nabla\times\mathbf{a}) = 0 \tag{A.12}$$

$$\nabla\times(\nabla\times\mathbf{a}) = \nabla(\nabla\cdot\mathbf{a}) - \nabla^2\mathbf{a} \tag{A.13}$$

$$\nabla\cdot(\nabla a) = \nabla^2 a \tag{A.14}$$

$$\nabla f(a) = f'(a)\nabla a \tag{A.15}$$

Formulas (A.5)–(A.15) also apply in two dimensions provided ∇, $a(\mathbf{r})$, $\mathbf{a}(\mathbf{r})$, $b(\mathbf{r})$ and $\mathbf{b}(\mathbf{r})$, with $\mathbf{r}$ the place vector which is the argument of these scalar and vector functions, are replaced by $\nabla_t = \frac{\partial}{\partial x}\mathbf{u}_x + \frac{\partial}{\partial y}\mathbf{u}_y$, $a_t(\boldsymbol{\rho})$, $\mathbf{a}_t(\boldsymbol{\rho})$, $b_t(\boldsymbol{\rho})$ and $\mathbf{b}_t(\boldsymbol{\rho})$ with $\boldsymbol{\rho} = x\mathbf{u}_x + y\mathbf{u}_y$ and where the subscript "t" stands for the transversal component of a vector or a scalar, i.e. its component in the xy-plane.

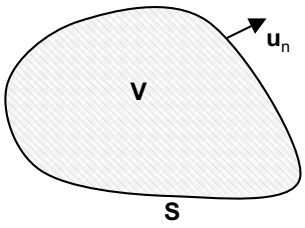


Figure A.1: Relevant to Gauss and Green's theorems for a volume.

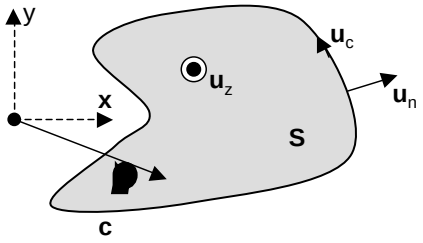


Figure A.2: Relevant to Gauss theorems for a surface.

A.3 Integral relationships

The following Gauss theorems hold for a volume V with boundary surface S and with $\mathbf{u}_n$ the outward unit normal to S (see Fig. A.1):

$$\int_V \nabla \cdot \mathbf{a}(\mathbf{r}) dV = \int_S \mathbf{a}(\mathbf{r}) \cdot \mathbf{u}_n \, dS, \tag{A.16}$$

$$\int_V \nabla \times \mathbf{a}(\mathbf{r}) dV = \int_S \mathbf{u}_n \times \mathbf{a}(\mathbf{r}) \, dS, \tag{A.17}$$

$$\int_V \nabla a(\mathbf{r}) dV = \int_S a(\mathbf{r}) \mathbf{u}_n \, dS, \tag{A.18}$$

The following Gauss theorems hold for a surface S in the xy-plane with circumference c and with $\mathbf{u}_n$ the outward unit normal to c (see Fig. A.2):

$$\int_S \nabla_t \cdot \mathbf{a}_t(\boldsymbol{\rho}) \, dS = \int_c \mathbf{a}_t(\boldsymbol{\rho}) \cdot \mathbf{u}_n \, dc = [\int_c \mathbf{a}_t(\boldsymbol{\rho}) \times \mathbf{dc}] \cdot \mathbf{u}_z, \tag{A.19}$$

$$\int_S \nabla_t \times \mathbf{a}_t(\boldsymbol{\rho}) \, dS = \int_c \mathbf{u}_n \times \mathbf{a}_t(\boldsymbol{\rho}) \, dc = \int_c [\mathbf{a}_t(\boldsymbol{\rho}) \cdot \mathbf{dc}] \, \mathbf{u}_z, \tag{A.20}$$

$$\int_S \nabla_t a(\boldsymbol{\rho}) \, dS = \int_c a(\boldsymbol{\rho}) \mathbf{u}_n \, dc = [\int_c a(\boldsymbol{\rho}) d\mathbf{l}] \times \mathbf{u}_z, \tag{A.21}$$

with $\mathbf{dc} = \mathbf{u}_z \times \mathbf{u}_n dc$.

The following Green's theorems hold for a volume V with boundary surface S and with $\mathbf{u}_n$ the outward unit normal to S:

$$\int_V (f\nabla^2 g + \nabla f \cdot \nabla g) dV = \int_S f \frac{\partial g}{\partial n} dS, \tag{A.22}$$

$$\int_V (f\nabla^2 g - g\nabla^2 f) dV = \int_S (f \frac{\partial g}{\partial n} - g \frac{\partial f}{\partial n}) dS, \tag{A.23}$$

$$\int_V ((\nabla \times \mathbf{f}) \cdot (\nabla \times \mathbf{g}) - \mathbf{f} \cdot \nabla \times \nabla \times \mathbf{g}) \; dV = \int_S (\mathbf{f} \times (\nabla \times \mathbf{g})) \cdot \mathbf{u}_n \; dS. \tag{A.24}$$

A.4 Co-ordinate systems

A.4.1 Cylindrical co-ordinates

$$\mathbf{u}_\rho = \mathbf{u}_x \cos\phi + \mathbf{u}_y \sin\phi. \tag{A.25}$$

$$\mathbf{u}_\phi = -\mathbf{u}_x \sin\phi + \mathbf{u}_y \cos\phi. \tag{A.26}$$

$$\mathbf{u}_x = \mathbf{u}_\rho \cos\phi - \mathbf{u}_\phi \sin\phi. \tag{A.27}$$

$$\mathbf{u}_y = \mathbf{u}_\rho \sin\phi + \mathbf{u}_\phi \cos\phi. \tag{A.28}$$

$$x = \rho\cos\phi, \qquad y = \rho\sin\phi. \tag{A.29}$$

$$\rho = \sqrt{x^2 + y^2}, \qquad \mathrm{tg}\phi = \frac{y}{x}. \tag{A.30}$$

$$a_\rho = a_x \cos\phi + a_y \sin\phi. \tag{A.31}$$

$$a_\phi = -a_x \sin\phi + a_y \cos\phi. \tag{A.32}$$

$$a_x = a_\rho \cos\phi - a_\phi \sin\phi. \tag{A.33}$$

$$a_y = a_\rho \sin\phi + a_\phi \cos\phi. \tag{A.34}$$

$$\mathbf{u}_\rho \times \mathbf{u}_\phi = \mathbf{u}_z. \tag{A.35}$$

$$\mathbf{u}_\phi \times \mathbf{u}_z = \mathbf{u}_\rho. \tag{A.36}$$

$$\mathbf{u}_z \times \mathbf{u}_\rho = \mathbf{u}_\phi. \tag{A.37}$$

$$\frac{\partial}{\partial\rho}\mathbf{a} = (\frac{\partial}{\partial\rho}a_\rho)\mathbf{u}_\rho + (\frac{\partial}{\partial\rho}a_\phi)\mathbf{u}_\phi + (\frac{\partial}{\partial\rho}a_z)\mathbf{u}_z. \tag{A.38}$$

$$\frac{\partial}{\partial\phi}\mathbf{a} = \mathbf{u}_z \times \mathbf{a} + (\frac{\partial}{\partial\phi}a_\rho)\mathbf{u}_\rho + (\frac{\partial}{\partial\phi}a_\phi)\mathbf{u}_\phi + (\frac{\partial}{\partial\phi}a_z)\mathbf{u}_z. \tag{A.39}$$

$$\frac{\partial}{\partial z}\mathbf{a} = (\frac{\partial}{\partial z}a_\rho)\mathbf{u}_\rho + (\frac{\partial}{\partial z}a_\phi)\mathbf{u}_\phi + (\frac{\partial}{\partial z}a_z)\mathbf{u}_z. \tag{A.40}$$

$$\nabla a = (\frac{\partial}{\partial\rho}a)\mathbf{u}_\rho + \frac{1}{\rho}(\frac{\partial}{\partial\phi}a)\mathbf{u}_\phi + (\frac{\partial}{\partial z}a)\mathbf{u}_z. \tag{A.41}$$

$$\nabla \cdot \mathbf{a} = \frac{1}{\rho}\frac{\partial}{\partial\rho}[\rho a_\rho] + \frac{1}{\rho}\frac{\partial}{\partial\phi}a_\phi + \frac{\partial}{\partial z}a_z = \frac{\partial}{\partial\rho}a_\rho + \frac{a_\rho}{\rho} + \frac{1}{\rho}\frac{\partial}{\partial\phi}a_\phi + \frac{\partial}{\partial z}a_z. \tag{A.42}$$

$$\nabla\times\mathbf{a} = [\frac{1}{\rho}\frac{\partial}{\partial\phi}a_z - \frac{\partial}{\partial z}a_\phi]\mathbf{u}_\rho + [\frac{\partial}{\partial z}a_\rho - \frac{\partial}{\partial\rho}a_z]\mathbf{u}_\phi + [\frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho a_\phi) - \frac{1}{\rho}\frac{\partial}{\partial\phi}a_\rho]\mathbf{u}_z. \tag{A.43}$$

$$\nabla^2 a = \frac{1}{\rho}\frac{\partial}{\partial\rho}[\rho\frac{\partial}{\partial\rho}a] + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2}a + \frac{\partial^2}{\partial z^2}a = \frac{\partial^2}{\partial\rho^2}a + \frac{1}{\rho}\frac{\partial}{\partial\rho}a + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2}a + \frac{\partial^2}{\partial z^2}a. \tag{A.44}$$

A.4.2 Spherical co-ordinates

$$\mathbf{u}_r = \mathbf{u}_x \sin\theta\cos\phi + \mathbf{u}_y \sin\theta\sin\phi + \mathbf{u}_z \cos\theta. \tag{A.45}$$

$$\mathbf{u}_\theta = \mathbf{u}_x \cos\theta\cos\phi + \mathbf{u}_y \cos\theta\sin\phi - \mathbf{u}_z \sin\theta. \tag{A.46}$$

$$\mathbf{u}_\phi = -\mathbf{u}_x \sin\phi + \mathbf{u}_y \cos\phi. \tag{A.47}$$

$$\mathbf{u}_x = \mathbf{u}_r \sin\theta\cos\phi + \mathbf{u}_\theta \cos\theta\cos\phi - \mathbf{u}_\phi \sin\phi. \tag{A.48}$$

$$\mathbf{u}_y = \mathbf{u}_r \sin\theta\sin\phi + \mathbf{u}_\theta \cos\theta\sin\phi + \mathbf{u}_\phi \cos\phi. \tag{A.49}$$

$$\mathbf{u}_z = \mathbf{u}_r \cos\theta - \mathbf{u}_\theta \sin\theta. \tag{A.50}$$

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta. \tag{A.51}$$

$$r = \sqrt{x^2+y^2+z^2}, \qquad \mathrm{tg}\theta = \frac{\sqrt{x^2+y^2}}{z}, \qquad \mathrm{tg}\phi = \frac{y}{x}. \tag{A.52}$$

$$a_r = a_x \sin\theta\cos\phi + a_y \sin\theta\sin\phi + a_z \cos\theta. \tag{A.53}$$

$$a_\theta = a_x \cos\theta\cos\phi + a_y \cos\theta\sin\phi - a_z \sin\theta. \tag{A.54}$$

$$a_\phi = -a_x \sin\phi + a_y \cos\phi. \tag{A.55}$$

$$a_x = a_r \sin\theta\cos\phi + a_\theta \cos\theta\cos\phi - a_\phi \sin\phi. \tag{A.56}$$

$$a_y = a_r \sin\theta\sin\phi + a_\theta \cos\theta\sin\phi + a_\phi \cos\phi. \tag{A.57}$$

$$a_z = a_r \cos\theta - a_\theta \sin\theta. \tag{A.58}$$

$$\mathbf{u}_r \times \mathbf{u}_\theta = \mathbf{u}_\phi. \tag{A.59}$$

$$\mathbf{u}_\phi \times \mathbf{u}_r = \mathbf{u}_\theta. \tag{A.60}$$

$$\mathbf{u}_\theta \times \mathbf{u}_\phi = \mathbf{u}_r. \tag{A.61}$$

$$\frac{\partial}{\partial r}\mathbf{a} = (\frac{\partial}{\partial r}a_r)\mathbf{u}_r + (\frac{\partial}{\partial r}a_\theta)\mathbf{u}_\theta + (\frac{\partial}{\partial r}a_\phi)\mathbf{u}_\phi. \tag{A.62}$$

$$\frac{\partial}{\partial \theta}\mathbf{a} = \mathbf{u}_\phi \times \mathbf{a} + (\frac{\partial}{\partial \theta}a_r)\mathbf{u}_r + (\frac{\partial}{\partial \theta}a_\theta)\mathbf{u}_\theta + (\frac{\partial}{\partial \theta}a_\phi)\mathbf{u}_\phi. \tag{A.63}$$

$$\frac{\partial}{\partial \phi}\mathbf{a} = \mathbf{u}_z \times \mathbf{a} + (\frac{\partial}{\partial \phi}a_r)\mathbf{u}_r + (\frac{\partial}{\partial \phi}a_\theta)\mathbf{u}_\theta + (\frac{\partial}{\partial \phi}a_\phi)\mathbf{u}_\phi. \tag{A.64}$$

$$\nabla a = (\frac{\partial}{\partial r}a)\mathbf{u}_r + \frac{1}{r}(\frac{\partial}{\partial \theta}a)\mathbf{u}_\theta + \frac{1}{r\sin\theta}(\frac{\partial}{\partial \phi}a)\mathbf{u}_\phi. \tag{A.65}$$

$$\nabla\cdot\mathbf{a} = \frac{1}{r^2}\frac{\partial}{\partial r}[r^2 a_r] + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}[\sin\theta a_\theta] + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}a_\phi. \tag{A.66}$$

$$\nabla\times\mathbf{a} = \frac{1}{r\sin\theta}[\frac{\partial}{\partial\theta}(\sin\theta a_\phi) - \frac{\partial}{\partial\phi}a_\theta]\mathbf{u}_r + \frac{1}{r}[\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}a_r - \frac{\partial}{\partial r}(r a_\phi)]\mathbf{u}_\theta$$
$$+\frac{1}{r}[\frac{\partial}{\partial r}(r a_\theta) - \frac{\partial}{\partial\theta}a_r]\mathbf{u}_\phi. \tag{A.67}$$

$$\nabla^2 a = \frac{1}{r^2}\frac{\partial}{\partial r}[r^2\frac{\partial}{\partial r}a] + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}[\sin\theta\frac{\partial}{\partial\theta}a] + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}a. \tag{A.68}$$