



Photonics. An introductory course →

Jose Trull



iniciativa
digital politècnica
Publicacions Acadèmiques UPC

→ **UPCGRAU**

Photonics. An introductory course →

Jose Trull

Primera edición: febrero de 2021

- © Jose Trull, 2021
- © Iniciativa Digital Politècnica, 2021
Oficina de Publicacions Acadèmiques Digitals de la UPC
Jordi Girona 31,
Edifici Torre Girona, Plant 1, 08034 Barcelona
Tel.: 934 015 885
www.upc.edu/idp
E-mail: info.idp@upc.edu

DL: B 3256-2021
ISBN:978-84-9880-891-9

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede realizarse con la autorización de sus titulares, salvo excepción prevista en la ley.



Index

Index	5
Preface	9
1. Fundamental properties of light.....	13
Introduction	13
1.1 Classical description of light waves.....	14
1.2 Basic solutions to the scalar wave equation	16
1.3 Vector representation of light waves. Polarization	20
1.4 Conservation of energy. Intensity	27
1.5 Momentum relations and radiation pressure	31
1.6 Solved problems.....	34
1.7 Proposed problems	47
2. Basic models for optical beams.....	53
Introduction	53
2.1 Helmholtz equation for optical beams.....	54
2.2 Paraxial wave equation for optical beams	55
2.3 Basic solutions to the paraxial wave equation.....	57
2.4 Basic properties of Gaussian beams	61
2.5 The effect of a lens on a Gaussian beam	68
2.6 General solutions to the paraxial wave equation.....	73
2.7 Solved Problems.....	76
2.8 Proposed problems	89
3. Light emission and light-matter interaction	93
Introduction	93
3.1 Light emission from a dipole	95
3.2 Radiation properties of harmonic dipoles	101
3.3 Interaction of a single atom with an incident light field.....	109
3.4 Maxwell equations in macroscopic media	117
3.5 Energy relations in material media	132



3.6	Solved problems	133
3.7	Proposed problems	147
4.	Crystal Optics.....	151
	Introduction	151
4.1	Normal modes of propagation in birefringent media.....	152
4.2	Polarization state of the normal modes.....	157
4.3	Ray propagation in uniaxial crystals.....	161
4.4	Applications of wave propagation in uniaxial crystals	165
4.5	Solved problems.....	171
4.6	Proposed problems	187
5.	Light pulse propagation	189
	Introduction	189
5.1	Mathematical description of light pulses	190
5.2	Pulse duration and bandwidth	193
5.3	Instantaneous frequency and chirp	195
5.4	Optical pulse propagation in dispersive media.....	197
5.5	Solved problems.....	204
5.6	Proposed problems	217
6.	Basic diffraction phenomena	221
	Introduction	221
6.1	Angular spectrum representation of optical beams	223
6.2	The concept of spatial frequency	227
6.3	The connection between the angular spectrum representation and the scalar Rayleigh–Sommerfeld diffraction theory.....	229
6.4	Diffraction theory in the Fresnel and Fraunhofer approximation.....	230
6.5	Fourier transforms by lenses: Introduction to image processing.....	235
6.6	Resolution limit in optical systems.....	240
6.7	Diffraction from an array of periodic elements.....	241
6.8	Solved problems.....	245
6.9	Proposed problems	256
Appendix A.	The Laser	259
Appendix B.	Classical radiation from elementary dipoles.....	265
Appendix C.	Lorentz model of electron oscillation.....	271
Appendix D.	Basic properties of Fourier transforms.....	275
Appendix E.	Diffraction pattern from a circular aperture.....	279
Appendix F.	Fourier transforms by lenses.....	283
Bibliography		287
	General reference books in optics and photonics.....	287
	Advanced books on specialized topics.....	287
	Handbooks of Optics.....	288







Preface

What is light, we would be compelled to confess that nothing definite by itself exists to which this name could be given. If this name cannot be given to the external agent, which we must call radiation, and if we do not accept that true light is the operation performed in photometry for experimental and technical purposes... Nothing exists which flows between psyche and phantom to which the name of light could be given.

To the word light only one meaning remains, namely, absence of darkness, the very same meaning attributed to it by philosophers two thousand years ago.

Ronchi. The nature of light

Photonics is a science that deals with applications of light in science and technology. Growing from a common root, we can include a vast number of different branches of study, such as optical communications, material processing, optical sensing, metrology, imaging, microscopy, biophotonics, nonlinear optics, nanophotonics, plasmonics, quantum electronics, laser theory, and many others. A variety of photonics applications can be implemented in various overarching fields, from engineering to telecommunications, to medicine, computing, aviation, construction, and on and on.

The foundations for our knowledge of optics date back many centuries. The main reason for this field's relatively long history is likely that nearly our entire sensory perception of the world around us comes directly through the mechanism of vision. The history of optics is a testimony to the huge efforts of many eminent scientists (Alhazen, Newton, Fermat, Fresnel, Young, Rayleigh, Kirchhoff, Huygens, and many others), all of whom "shined some light" (and also cast some shadows) onto the nature of light. Their contributions consolidated the fields of geometrical optics, optical instrumentation, and wave optics during the eighteenth and nineteenth centuries.



Electricity, magnetism, and optics were once viewed as separate scientific disciplines over a long period of time, until the works of Maxwell and Hertz, among others, laid the foundations for studying the emission of electromagnetic radiation by oscillating charges, thus unifying the fields of electromagnetics and optics by recognizing that light forms a part of the electromagnetic wave spectrum. Subsequent attempts to describe the process of blackbody radiation using the standard laws of physics remained unsuccessful until Planck formulated his hypothesis of the emission and absorption of light quanta while, in parallel, Einstein introduced the basic processes of light–matter interaction, thus leading to the inception of quantum optics.

A major step forward in the development of the field of photonics occurred with the discovery of the laser in 1960, which sparked an increased number of this light source’s potential uses in terms of directionality, coherence, spectral bandwidth, and intensity. Since then, the number of proposed applications has grown unceasingly and thus given birth to new topics and ideas that were unthinkable just a few years ago.

This book originated with the idea of providing support for students taking the photonics course for the Bachelor’s Degree in Physical Engineering at UPC. Any attempt to cover in a single semester—even succinctly—all the different components that comprise the current field of photonics would be an insurmountable task. This is particularly true when bearing in mind that actual research is expanding continuously almost every specific branch in the vast field of photonics, requiring that the student has a solid background in optics and electromagnetism. While this book does not cover the fundamental techniques of photonics applications, it nevertheless constitutes a solid point of departure for becoming familiar with the basic concepts and models that are used widely in the field.

This book assumes that the student has a basic knowledge of electromagnetic theory, and it covers different topics related to the properties of the coherent interaction of light with matter in the frame of classical electrodynamics. Some interesting topics that have been omitted from the program but may be introduced in a second part of this course include extending these concepts to incoherent phenomena, interference effects, and the basic postulates of quantum optics, among others. To help the students in their study of the theoretical concepts, a list of solved problems is included at the end of each chapter. The bibliography at the end covers both the basic and more specialized literature for those students who wish to delve more deeply into the fascinating ideas of this field.

Finally, I would like to thank the people who helped me and stood by my side throughout the realization of this project. To my colleagues and friends, Ramon Vilaseca, Crina Cojocaru, and Ramon Herrero, I am grateful for their advice, useful comments, and continuous encouragement in transforming an original idea into reality. To my students, their feedback and useful comments on the course contents were indispensable. To Miguel Angel Trull, who helped me design and create most of the figures, I am eternally indebted. Finally, to my family, whose unconditional support and infinite patience made all of this possible.

Terrassa, November 2020



→ 1



Fundamental properties of light

Introduction

Photonics essentially deals with the theory of light–matter interaction. Studying this interaction involves describing different degrees of approach, and any accurate description is contingent on the particular problem under study. The various approaches can be classified according to the way each model treats light and matter.

Light	Matter	Model
Classical Maxwell's equations	Classical Material equations	Classical electrodynamics
Classical Maxwell's equations	Quantum Density matrix	Semi-classical model
Quantum Quantization of EM field	Quantum Density matrix	Quantum electrodynamics

In this first lesson, we will briefly review the main properties of light from a classical viewpoint. Within this framework, light is considered an electromagnetic (EM) wave generated by oscillating charges that can be described by Maxwell's equations. Although some properties are generally common to EM waves across the EM spectrum, some specific properties also exist, and these become more relevant at particular regions. Some examples:

1) The emission and absorption characteristics of radiation differ drastically. For very long wavelengths in the radio-wave region, single antennas can radiate EM waves that very closely resemble the emissions of a single dipole. As the wavelength becomes shorter and we move into the infrared, visible, and ultraviolet regions, emissions convert into a many-body problem where any radiation emitted is a result of the superposition of millions of emitters, which significantly alter the properties of the resulting optical radiation.



2) The quantum aspects of radiation become significant at these high frequency ranges, and they must be considered in order to properly describe the processes of light emission and absorption.

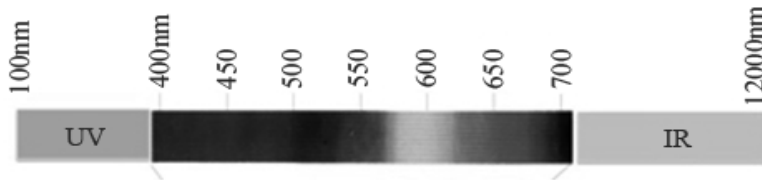
3) At even higher frequencies in the X-ray region and beyond, relevant new mechanisms emerge as the wavelength value shrinks to a comparable atomic size. These regimes call into question the suitability of using either the mean field model from classical electrodynamics (Chapter 3) or the dipolar electric approximation.

4) Waves are generated in radically different ways. A broad distinction can be made between those that are artificially generated by electronic devices (such as radio, TV, and microwaves) and naturally generated waves (such as those emitted by matter).

This course covers the study of the EM spectrum, specifically the range comprising the near infrared (IR), the visible (VIS) and the ultraviolet (UV) regions, all of which share some common properties. These commonalities correspond to atomic material oscillations, coherence properties, and the way in which matter absorbs this form of radiation.

The wavelength values in this region of the spectrum are summarized in Figure 1.

Figure 1.1
The EM spectrum in the
UV, visible and IR region



1.1 Classical description of light waves

Oscillating charges radiate electromagnetic fields. The microscopic Maxwell equations provide the mathematical model that describes electric and magnetic fields radiating from a set of elementary charges (ρ), currents ($\vec{j}$), their propagation, and their interaction with matter. These equations are:

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad [1.1A] \qquad \vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad [1.1C]$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad [1.1B] \qquad \vec{\nabla} \wedge \vec{B} = \mu_0 \vec{j} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \quad [1.1D]$$

where ϵ_0 is the dielectric permittivity and μ_0 is the magnetic permeability of vacuum.

♦ *Comment*

We will use the rationalized MKS unit system in this course. The field units and related quantities are:



Electric field: $[E] = \text{V/m}$

Polarization density: $[P] = \text{C/m}^2$

Displacement vector: $[D] = \text{C/m}^2$

Dielectric permittivity: $[\epsilon_0] = \text{C}^2/\text{Nm}^2$

$\epsilon_0 = 1/(4\pi \cdot 9 \cdot 10^9) \text{ C}^2/\text{Nm}^2$

Magnetic field: $[B] = \text{T}$

Magnetic intensity: $[H] = \text{A/m}$

Magnetic moment density: $[M] = \text{A/m}$

Magnetic permeability: $[\mu_0] = \text{Tm/A}$

$\mu_0 = 4\pi \cdot 10^{-7} \text{ Tm/A}$

For a detailed discussion of alternative unit systems for photonics (mainly Gaussian units), refer to the book Jackson, J.D. (1962). *Classical Electrodynamics*. John Wiley & Sons.

From Equations [1.1A] and [1.1D], we can readily obtain the continuity equation, which expresses the charge conservation:

$$\vec{\nabla} \cdot \vec{j} + \frac{\partial \rho}{\partial t} = 0 \quad [1.2]$$

These equations are completed by the Newton–Lorentz equation for describing the dynamics of particles with mass m_i and charge q_i that are under the influence of the EM forces exerted by the fields

$$m_i \frac{d^2 \vec{r}_i(t)}{dt^2} = q_i [\vec{E}(\vec{r}_i(t), t) + \vec{v}_i(t) \wedge \vec{B}(\vec{r}_i(t), t)] \quad [1.3]$$

(!) This chapter will consider the basic solutions for EM waves in vacuum.

As the fields radiate from the source region, they can propagate in vacuum (where ρ and $\vec{j}$ equal zero) while transferring energy and momentum.

The field propagation is governed by the wave equations. These are derived from Equation [1.1] by taking the curl of Equations [1.1C] and [1.1D]:

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) = \vec{\nabla} \wedge \left(-\frac{\partial \vec{B}}{\partial t} \right) = -\frac{\partial}{\partial t} (\vec{\nabla} \wedge \vec{B}) = -\mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} \quad [1.4A]$$

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{B}) = \vec{\nabla} \wedge \left(\mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \right) = \mu_0 \epsilon_0 \frac{\partial}{\partial t} (\vec{\nabla} \wedge \vec{E}) = -\mu_0 \epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} \quad [1.4B]$$

Under the assumed condition of propagation in vacuum, the divergence of the electric and magnetic fields is zero, and we can write the wave equations for the electric and magnetic fields in vacuum as:

$$\boxed{\begin{aligned} \nabla^2 \vec{E} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} &= 0 \\ \nabla^2 \vec{B} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} &= 0 \end{aligned}} \quad [1.4C]$$



From these equations, we can infer that the electric and magnetic fields propagate out of the source regions into vacuum with a velocity given by $c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \approx 3 \cdot 10^8 \text{ m/s}$

Since these are vector equations, any of the field components (and any linear combination of them) will satisfy the scalar wave equation

$$\nabla^2 U(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 U(\vec{r}, t)}{\partial t^2} = 0 \quad [1.5]$$

Optics must deal with many problems that preserve polarization conditions or use unpolarized radiation. Under these conditions, the vector character of the radiation is not relevant for studying the propagation and, instead, a scalar approximation of the wave equation is used.

The scalar function U can represent any of the components of the electric and magnetic fields. It is customary in many situations to define a scalar function that will be directly related to the optical intensity (a measurable quantity) through the relation $I = |U|^2$

For instance, the relation between the scalar wave function and the modulus of a plane wave's electric field should be $|E|^2 = 2|U|^2/\epsilon_0 c$ (as we will see later in this chapter).

1.2 Basic solutions to the scalar wave equation

1.2.1 1D solutions: Plane waves

Let us consider waves propagating along the z -axis as we seek solutions to the form $U(z, t)$. In this case, $\partial U/\partial x = \partial U/\partial y = 0$, and the wave equation simplifies to

$$\frac{\partial^2 U(z, t)}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 U(z, t)}{\partial t^2} = 0 \quad [1.6A]$$

A general solution to this equation (see Solved problems) is written as:

$$U(z, t) = f_+(z - ct) + f_-(z + ct) \quad [1.6B]$$

where f_+ and f_- correspond to arbitrary functions propagating, respectively, to the right or left, with speed c .

Solution [1.6B] corresponds to what is called a **progressive wave**. Other solutions to the wave equation are possible, such as the stationary wave solutions, expressed in the form $U(z, t) = p(z)q(t)$, which are separable.

At this point, we can be more specific regarding the function f as we consider the way light radiates away from atoms. As we will see in Chapter 3, the classical model describes light emission as oscillating dipoles radiating at frequencies of the order of 10^{15} Hz. This radiation is damped with damping constants of the order of 10^{-7} seconds (damping rate of 10^7 Hz). Thus, a typical emitted light burst will consist of a damped wave with many oscillations under its decaying envelope.



Under such conditions, **we can describe light emission using harmonic functions.**

The **plane wave** solution is written as

$$U(z, t) = U_o \cos(k(z - ct) + \phi) = U_o \cos(kz - kct + \phi) \quad [1.7]$$

where U_o is a constant amplitude, k is called the **wave vector**, and ϕ is the initial phase that allows setting the initial conditions.

From this expression, the following properties follow:

Spatial periodicity

The spatial period is called the **wavelength** and is related to the wave vector k in [1.7] by

$$U(z + \lambda, t) = U(z, t) \rightarrow \boxed{k = \frac{2\pi}{\lambda}}$$

Temporal periodicity

The condition of periodicity in the time domain gives the **period** T :

$$U(z, t + T) = U(z, t) \rightarrow \boxed{T = \frac{2\pi}{kc}}$$

Since the term kc has units of frequency, we define the **temporal frequency** ω by the relation

$$\boxed{\omega = kc}$$

This equation is called the **dispersion relation**, and it plays a very significant role in photonics (mainly when studying how light is modified by the presence of matter).

Solution [1.7] represents a wave propagating along the z -axis with constant amplitude over planes perpendicular to the propagation direction. Although it does not represent a realistic approximation of a light beam in most situations, it will play a leading role in the theory for conveniently extending other solutions.

Although the physically relevant magnitude is always the real-field representation, it is much more convenient in many situations to work within the **complex representation of the plane wave**, which is obtained directly from [1.7]:

$$U(z, t) = U_o \cos(kz - \omega t + \phi) = \frac{U_o}{2} \left[e^{i(kz - \omega t + \phi)} + e^{-i(kz - \omega t + \phi)} \right]$$

This expression can be rewritten as

$$\begin{aligned} U(z, t) &= \frac{U_o}{2} \left[e^{i\phi} e^{i(kz - \omega t)} + e^{-i\phi} e^{-i(kz - \omega t)} \right] = \\ &= \frac{1}{2} \left[U_o^{(+)} e^{i(kz - \omega t)} + U_o^{(-)} e^{-i(kz - \omega t)} \right] = \frac{1}{2} \left[U^{(+)} + U^{(-)} \right] \end{aligned}$$



thereby defining the variables below.

Complex amplitude: $U_o^{(+)} = U_o e^{i\phi}$ and $U_o^{(-)} = U_o e^{-i\phi}$

Phase: $kz - \omega t$

Complex analytic signal: $U^{(+)} = U_o^{(+)} e^{i(kz - \omega t)}$ and $U^{(-)} = U_o^{(-)} e^{-i(kz - \omega t)}$

The variables $U^{(+)}$ and $U^{(-)}$ are not independent, since the reality of the fields imposes the condition $U^{(+)} = [U^{(-)}]^*$ (the symbol * stands for complex conjugation).

The complex analytic signal is the basic expression of our plane wave solution, but we must remember that the physical solution corresponding to the real fields can be obtained from the analytic signal through the relation

$$U(z, t) = \text{Re}[U^{(+)}] = \frac{1}{2}[U^{(+)} + U^{(-)}] \quad [1.8]$$

1.2.2 3D solutions

Plane waves

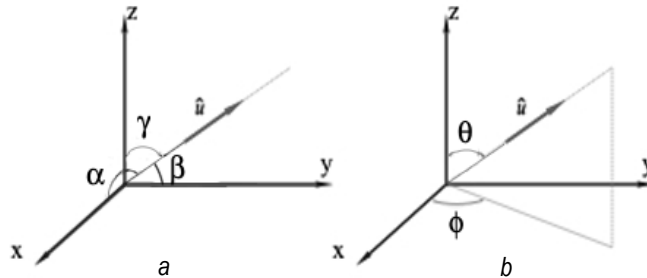
When dealing with single waves, one can always define a coordinate system in which the z -axis corresponds to the propagation direction. However, many situations involve the interaction of several waves propagating in different directions. The concept of a plane wave can be extended in a straightforward manner to include propagation in arbitrary directions through space, as specified through the unitary vector $\hat{u}$, by defining a vector representation of the wave vector:

$$\vec{k} = k \cdot \hat{u} = \frac{2\pi}{\lambda} \cdot \hat{u} = (k_x \hat{i} + k_y \hat{j} + k_z \hat{k})$$

where $\hat{u}$ can be defined either in terms of the direction cosines or using spherical coordinates:

$$\hat{u} = (\cos \alpha \hat{i} + \cos \beta \hat{j} + \cos \gamma \hat{k}) = (\cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k})$$

Figure 1.2
Representation of a unitary vector in terms of
a) direction cosines
b) spherical coordinates





The components of the wave vector can be expressed as:

$$\begin{aligned}k_x &= \frac{2\pi}{\lambda} \cos \alpha = \frac{2\pi}{\lambda_x} = 2\pi f_x \\k_y &= \frac{2\pi}{\lambda} \cos \beta = \frac{2\pi}{\lambda_y} = 2\pi f_y \\k_z &= \frac{2\pi}{\lambda} \cos \gamma = \frac{2\pi}{\lambda_z} = 2\pi f_z\end{aligned}$$

where f_x , f_y and f_z are called the spatial frequencies. These frequencies will be considered in further detail within the context of diffraction problems in Chapter 6.

Spherical waves

The solution corresponding to spherical waves is obtained by writing the Laplacian in spherical coordinates and assuming a solution with rotational symmetry (with $\partial/\partial\phi = \partial/\partial\theta = 0$) (see Solved Problems for details).

$$\frac{\partial^2(rU)}{\partial r^2} - \frac{1}{c^2} \frac{\partial^2(rU)}{\partial t^2} = 0 \quad [1.9]$$

This equation has solutions that are identical to the one found in Equation [1.6B]:

$$rU(r, t) = f_+(r - ct) + f_-(r + ct) \rightarrow U(r, t) = \frac{f_+(r - ct)}{r} + \frac{f_-(r + ct)}{r}$$

The harmonic spherical solution in the complex representation is obtained in the same way as described for plane waves:

$$U(r, t) = \frac{U_0}{r} \cos(kr - \omega t + \phi) = \frac{U_0}{2r} \left[e^{i(kr - \omega t + \phi)} + e^{-i(kr - \omega t + \phi)} \right]$$

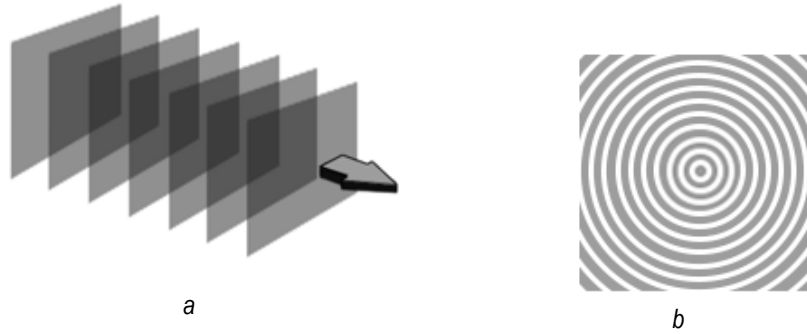
This expression can be rewritten as

$$\begin{aligned}U(r, t) &= \frac{1}{2} \left[\frac{U_0 e^{i\phi}}{r} e^{i(kr - \omega t)} + \frac{U_0 e^{-i\phi}}{r} e^{-i(kr - \omega t)} \right] = \frac{1}{2} \left[\frac{U_0^{(+)}}{r} e^{i(kr - \omega t)} + \frac{U_0^{(-)}}{r} e^{-i(kr - \omega t)} \right] \\U(r, t) &= \text{Re}[U^{(+)}(r) e^{i(kr - \omega t)}]\end{aligned}$$

As we will see, neither plane waves nor spherical waves constitute adequate realistic solutions, since plane waves extend infinitely in space and time, and spherical waves do not properly describe light emitted by atoms. However, plane waves constitute an appropriate basis for describing more complex optical beams, and spherical waves are good for describing the asymptotic limit at large distances from the sources.



Figure 1.3
Schematic representation of
a) scalar plane wave
b) spherical wave



1.3 Vector representation of light waves. Polarization

In the previous section, we defined the basic scalar solutions to the wave equation. However, this approximation is not adequate on many occasions, and we need to consider the full vector character of the wave solutions for the electric and magnetic fields. We should remember that these vector fields are not independent, since they are coupled through Maxwell's equations.

To study the basic **vector plane wave solutions for light waves**, let us consider a plane wave solution for an electric field that is oriented along a direction specified by the unitary vector $\hat{e}$:

$$\vec{E}(\vec{r}, t) = \text{Re}[\vec{E}^{(+)}(\vec{r}, t)] \quad \text{with} \quad \vec{E}^{(+)}(\vec{r}, t) = \hat{e}E^{(+)}(\vec{r}, t)$$

This can be written in terms of its Cartesian components as

$$\vec{E}^{(+)}(\vec{r}, t) = (E_{0x}^{(+)}\hat{i} + E_{0y}^{(+)}\hat{j} + E_{0z}^{(+)}\hat{k})e^{i(\vec{k}\vec{r} - \omega t)} \quad [1.10A]$$

Equivalent expressions hold for the magnetic field, which is oriented along a direction specified by the unitary vector $\hat{b}$:

$$\begin{aligned} \vec{B}(\vec{r}, t) &= \text{Re}[\vec{B}^{(+)}(\vec{r}, t)] \quad \text{with} \quad \vec{B}^{(+)}(\vec{r}, t) = \hat{b}B^{(+)}(\vec{r}, t) \\ \vec{B}^{(+)}(\vec{r}, t) &= (B_{0x}^{(+)}\hat{i} + B_{0y}^{(+)}\hat{j} + B_{0z}^{(+)}\hat{k})e^{i(\vec{k}\vec{r} - \omega t)} \end{aligned} \quad [1.10B]$$

Substituting in Maxwell's Equations [1.1A] (with $\rho = 0$) and [1.1B] gives

$$\begin{aligned} \vec{\nabla} \cdot \vec{E}^{(+)} &= 0 & \rightarrow & \quad i\vec{k} \cdot \vec{E}^{(+)} = 0 \\ \vec{\nabla} \cdot \vec{B}^{(+)} &= 0 & \rightarrow & \quad i\vec{k} \cdot \vec{B}^{(+)} = 0 \end{aligned}$$

These equations define a geometric property of the vector plane waves, indicating that both the electric and magnetic fields are perpendicular to the propagation direction and thus reflecting the transverse character of these solutions.



♦ *Comment*

The evolution of EM fields in the spatio-temporal domain is governed by Maxwell's equations, which are first-order differential equations. The electric field at a given point is related to the value of the fields at nearby points. A Fourier transformation of the fields in the spatial and temporal domain allows us to express the fields in terms of their spatial and temporal frequencies. In the reciprocal space, Maxwell's equations become local relations showing that each Fourier component of the field evolves independently in linear media.

The relations between the fields in space and reciprocal space are given by:

$$\vec{E}(\vec{r}, t) = \frac{1}{(2\pi)^{3/2}} \int d^3k \vec{E}(\vec{k}, t) e^{i\vec{k}\vec{r}} \quad \Leftrightarrow \quad \vec{E}(\vec{k}, t) = \int d^3r \vec{E}(\vec{r}, t) e^{i\vec{k}\vec{r}}$$

Substituting the fields given by [1.10] into Maxwell's Equation [1.1C] provides a relation between the electric and the magnetic field:

$$\vec{\nabla} \wedge \vec{E}^{(+)} = -\frac{\partial \vec{B}^{(+)}}{\partial t} \rightarrow i\vec{k} \wedge \vec{E}^{(+)} = i\omega \vec{B}^{(+)} \\ kE_0^{(+)} (\hat{u} \wedge \hat{e}) = \omega B_0^{(+)} \hat{b}$$

From this equation, two conclusions are readily obtained:

$$\begin{aligned} 1) \quad B_0^{(+)} &= \frac{k}{\omega} E_0^{(+)} = \frac{E_0^{(+)}}{c} \\ 2) \quad \hat{u} \wedge \hat{e} &= \hat{b} \rightarrow \hat{e} \wedge \hat{b} = \hat{u} \end{aligned} \quad [1.11]$$

Relation 1) provides the ratio between the magnitude of the electric and magnetic fields in a plane wave; while relation 2) fixes the relative orientation of the electric field, the magnetic field, and the wave propagation vectors. In addition to the transverse character of the plane EM wave, we can see that the electric and magnetic fields are perpendicular to each other.

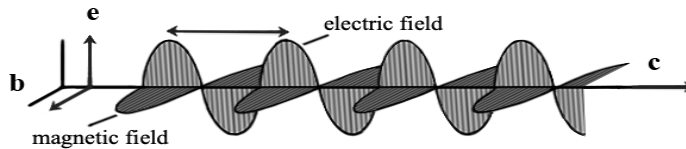


Figure 1.4
Representation
of an EM wave

However, the obtained solutions still allow an extra degree of freedom, since we have not fixed the direction of the electric field oscillation within the plane that is perpendicular to the wave propagation direction. We will call the plane containing the electric and magnetic fields *the plane of oscillation*. The evolution of the electric field oscillation within this plane gives rise to a new property of the wave solution: the **polarization state**.



In order to define the polarization state of a wave, we need to consider the possible oscillations of the electric field vector within the plane of oscillation. Any state of oscillation in a plane can be obtained as a superposition of two perpendicular harmonic motions (in this case, with identical frequencies). Without losing generality, we define a coordinate system where the propagation direction takes place along the z -axis and the plane of oscillation is defined by the xy -plane.

Expression [1.10 A] is particularized in this case to

$$\vec{E}^{(+)}(z, t) = [E_{0x} e^{i\phi_x} \hat{i} + E_{0y} e^{i\phi_y} \hat{j}] e^{i(kz - \omega t)} \quad [1.12A]$$

A similar equation can also be written for the magnetic field. However, since both fields are totally linked by the expressions derived previously, we can work only with the electric field and obtain the resulting magnetic field at the end.

Equation [1.12A] can be rewritten in the form

$$\vec{E}^{(+)}(z, t) = E_0 e^{i\phi_x} \left[\frac{E_{0x}}{E_0} \hat{i} + \frac{E_{0y}}{E_0} e^{i\delta} \hat{j} \right] e^{i(kz - \omega t)} \quad [1.12B]$$

where the electric field modulus $E_0 = \sqrt{E_{0x}^2 + E_{0y}^2}$ and the relative phase $\delta = \phi_y - \phi_x$ have been introduced.

Using the variable θ to define the angle between the direction of the electric field oscillation and the x -axis, we write the electric field finally as

$$\boxed{\vec{E}^{(+)}(z, t) = E_0^{(+)} [\cos \theta \hat{i} + \sin \theta e^{i\delta} \hat{j}] e^{i(kz - \omega t)}} \quad [1.12C]$$

where the term $E_0^{(+)}$ incorporates the arbitrary phase term $e^{i\phi_x}$. Any state of polarization of the plane EM wave can be described by the expression [1.12C]. The **state of polarization** is usually defined in terms of a vector known as the **Jones vector**:

$$[J] = \begin{bmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{bmatrix}$$

To properly classify the different states of light polarization, we define the **vibration plane** as the plane containing the vectors $\hat{e}$ and $\hat{u}$ at any instant of time.

In the most general case, light emitted by thermal and spectral sources and by light-emitting diodes is not polarized in the sense that the vibration plane changes randomly in time over periods of the order of 10^{-7} s. This is a consequence of the way light is generated by such sources, namely that it is emitted by many independent radiating dipoles. In this case, the light polarization state is not defined and we are at the limit where the scalar function approach (defined in the previous section) is valid.



This kind of emission stands in contrast to the EM wave emitted by, for instance, an antenna, as this type of wave can be considered a very good approximation of a polarized plane wave at a large distance.

If by some means we can establish a relative phase difference δ constant in time, we obtain a well-defined oscillation pattern. In this case, we say that the emitted light is **polarized**.

1.3.1 Polarization states of light

Linearly polarized light

Light is defined as linearly polarized when the electric field oscillates along a line defined by a fixed angle θ (the vibration plane stays constant during propagation). To obtain a linearly polarized wave, we have to fix the relative phase to the values

$$\delta = 0 \quad \text{or} \quad \delta = \pi$$

The corresponding Jones vectors in this case are real and can be written as

$$[L_0(\theta)] = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad \text{and} \quad [L_\pi(\theta)] = \begin{bmatrix} \cos \theta \\ -\sin \theta \end{bmatrix}$$

The relation $[L_\pi(\theta)] = [L_0(-\theta)]$ holds.

The expression for the real electric field is

$$\vec{E}(z, t) = \text{Re}[\vec{E}^{(+)}(z, t)] = E_0(\cos \theta \hat{i} \pm \sin \theta \hat{j}) \cos(kz - \omega t + \phi_x) \quad [1.13]$$

where the signs $+$ and $-$ refer to $\delta = 0$ and $\delta = \pi$.

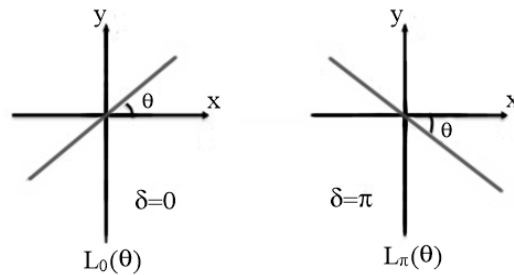


Figure 1.5
Linear polarization states
in the oscillation plane

Circularly polarized light

A circular state of polarization is obtained for the particular case with

$$E_{0x} = E_{0y} = E_0; \quad \delta = \pm\pi/2; \quad \theta = \pi/4 \text{ rad.}$$



The corresponding Jones vectors are

$$[C_R] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad [C_L] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}$$

and they define, respectively, right circular and left circular polarization states. One can immediately identify the complex character of these vectors, which in principle lack any physical interpretation. However, one should always keep in mind that the complex field representation is a useful tool for calculations, but it is always necessary to recover the physically measurable magnitude corresponding to the real electric field.

Under these conditions, the electric field reads

$$\vec{E}^{(+)}(z, t) = E_0 e^{i\phi_x} (\hat{i} \mp i \hat{j}) e^{i(kz - \omega t)}$$

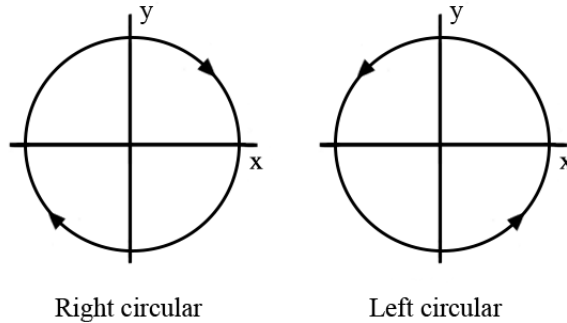
and

$$\vec{E}(z, t) = \text{Re}[\vec{E}^{(+)}(z, t)] = E_0 (\cos(kz - \omega t + \phi_x) \hat{i} \pm \sin(kz - \omega t + \phi_x) \hat{j}) \quad [1.14]$$

A close inspection of this expression shows that, for circular polarization, the electric field propagates while maintaining constant amplitude E_0 and the vibration plane performs a circular motion around the z -axis.

For right circular polarization, the electric vector tip rotates in a clockwise direction as a function of time when looked at by an observer facing the light wave at a fixed z position; while rotation in the opposite direction defines the left circular polarization state. The rotation direction is a function of time (at a fixed position), and it is shown in Figure 6.

Figure 1.6
Circular polarization states.
Rotation as a function of
time at fixed z position



Elliptically Polarized Light

In the general case of arbitrary (but constant) relative phase and angle, the tip of the electric field performs an elliptical motion during propagation as the amplitude and oscillation directions change simultaneously.



The general expression for the elliptical motion can be found with Equation [1.12A]. The real field can be written as below.

$$\vec{E} = \text{Re}[\vec{E}^{(+)}] = E_x \hat{i} + E_y \hat{j} = E_{ox} \cos(kz - \omega t + \phi_x) \hat{i} + E_{oy} \cos(kz - \omega t + \phi_y) \hat{j}$$

Field components:

$$\begin{aligned} E_x &= E_{ox} (\cos(kz - \omega t) \cos \phi_x - \sin(kz - \omega t) \sin \phi_x) \\ E_y &= E_{oy} (\cos(kz - \omega t) \cos \phi_y - \sin(kz - \omega t) \sin \phi_y) \end{aligned}$$

Rearranging terms:

$$\begin{aligned} \frac{E_x}{E_{ox}} \sin \phi_y - \frac{E_y}{E_{oy}} \sin \phi_x &= \cos(kz - \omega t) \sin(\phi_y - \phi_x) \\ \frac{E_x}{E_{ox}} \cos \phi_y - \frac{E_y}{E_{oy}} \cos \phi_x &= \sin(kz - \omega t) \sin(\phi_y - \phi_x) \end{aligned}$$

These relations can be written as:

$$\frac{E_x^2}{E_{ox}^2} + \frac{E_y^2}{E_{oy}^2} - 2 \frac{E_x E_y}{E_{ox} E_{oy}} \cos(\delta) = \sin^2 \delta$$

This expression represents the equation for an ellipse inscribed in a rectangle with sides parallel to the x -axis and y -axis, and dimensions $2E_{ox}$ and $2E_{oy}$.

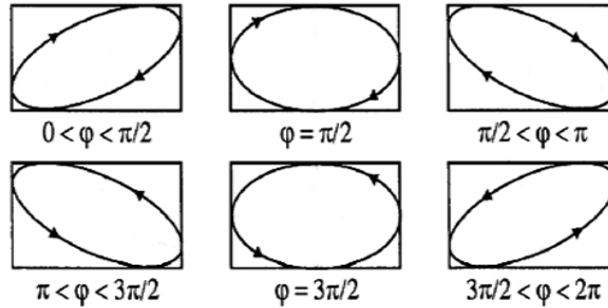


Figure 1.7
Elliptically polarized
states

1.3.2 Polarizers

A polarizer is any device used to obtain polarized light or to change the state of light polarization. We can analyse the effect of polarizers on incident light by looking at how the corresponding Jones vector is transformed by the action of the polarizer. The effect of the polarizer is described in terms of a characteristic matrix (the Jones matrix).

Linear polarizer matrix

A linear polarizer is defined by the orientation of the polarization state at the output. This output state should be proportional to the unitary Jones vector:



$$\hat{e}_{out} = \begin{pmatrix} \cos \theta_p \\ \sin \theta_p \end{pmatrix}$$

where θ_p is the angle of the polarizer axis with respect to the x -axis. To construct the characteristic matrix, we assume that the action of the polarizer on a given arbitrary incident state characterized by the Jones vector

$$\bar{e}_{in} = \begin{pmatrix} e_{in\ x} \\ e_{in\ y} \end{pmatrix}$$

is to obtain the projection of this initial state into the output state:

$$\bar{e}_{out} = (\bar{e}_{in} \cdot \hat{e}_{out}) \cdot \hat{e}_{out} = (e_{in\ x} \cdot \cos \theta_p + e_{in\ y} \cdot \sin \theta_p) \begin{pmatrix} \cos \theta_p \\ \sin \theta_p \end{pmatrix} = \begin{pmatrix} e_{in\ x} \cos^2 \theta_p + e_{in\ y} \cos \theta_p \sin \theta_p \\ e_{in\ x} \cos \theta_p \sin \theta_p + e_{in\ y} \sin^2 \theta_p \end{pmatrix}$$

This result can be expressed in the equivalent form:

$$\bar{e}_{out} = \begin{pmatrix} \cos^2 \theta_p & \cos \theta_p \sin \theta_p \\ \cos \theta_p \sin \theta_p & \sin^2 \theta_p \end{pmatrix} \begin{pmatrix} e_{in\ x} \\ e_{in\ y} \end{pmatrix} = L(\theta_p) \cdot \bar{e}_{in} \quad [1.15]$$

Here, L refers to the characteristic matrix we were searching for. (Note that the output state does not need to be normalized, since the intensity of the incident light beam will generally be different from that of the output beam.)

Example 1: Consider the action of the linear polarizer defined by $L(\theta_p)$ on an incident light beam that is linearly polarized at an angle θ .

The input state is $\hat{e}_{in} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$

The output plane wave after the linear polarizer will be described by the Jones vector:

$$\begin{aligned} \bar{e}_{out} &= L(\theta_p) \cdot \hat{e}_{in} = \begin{pmatrix} \cos^2 \theta_p \cos \theta + \cos \theta_p \sin \theta_p \sin \theta \\ \cos \theta_p \sin \theta_p \cos \theta + \sin^2 \theta_p \sin \theta \end{pmatrix} = \\ &= (\cos \theta_p \cos \theta + \sin \theta_p \sin \theta) \cdot \begin{pmatrix} \cos \theta_p \\ \sin \theta_p \end{pmatrix} \end{aligned}$$

$$\boxed{\bar{e}_{out} = \cos(\theta_p - \theta) \cdot \hat{e}_{out}}$$

The output beam intensity will be proportional to $\bar{e}_{out} \cdot \bar{e}_{out}^*$

$$I_{out} = \cos^2(\theta_p - \theta) I_{in}. \quad \text{This result is known as **Malus's Law**.}$$



Example 2: Action of a linear polarizer on incident circularly polarized light.

In this case, the incident beam is described by the Jones vector: $\vec{e}_{in} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mp i \end{pmatrix}$

Applying Equation [1.14] results in

$$\vec{e}_{out} = \frac{\cos \theta_p \pm i \sin \theta_p}{\sqrt{2}} \begin{pmatrix} \cos \theta_p \\ \sin \theta_p \end{pmatrix} = \frac{e^{\pm i \theta_p}}{\sqrt{2}} \begin{pmatrix} \cos \theta_p \\ \sin \theta_p \end{pmatrix}$$

1.4 Conservation of energy. Intensity

Many applications in photonics exploit the fact that light waves can transfer energy, linear momentum, and angular momentum. In this section, we will introduce the basic relations between these physical magnitudes and the electromagnetic field.

The energy conservation relations are obtained from Maxwell's Equations [1.1C] and [1.1D]. Let us multiply (scalar product) both sides of Equation [1.1C] by the magnetic field and of Equation [1.1D] by the electric field:

$$\begin{aligned} \vec{B} \cdot [1.1C] &\Rightarrow \vec{B} \cdot (\vec{\nabla} \wedge \vec{E}) = \vec{B} \cdot \left(-\frac{\partial \vec{B}}{\partial t} \right) \\ \vec{E} \cdot [1.1D] &\Rightarrow \vec{E} \cdot (\vec{\nabla} \wedge \vec{B}) = \vec{E} \cdot (\mu_0 \vec{j} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}) \end{aligned}$$

Using the vector identity $\vec{\nabla} \cdot (\vec{E} \wedge \vec{B}) = \vec{B} \cdot (\vec{\nabla} \wedge \vec{E}) - \vec{E} \cdot (\vec{\nabla} \wedge \vec{B})$ leads to the expression

$$\vec{\nabla} \cdot (\vec{E} \wedge \vec{B}) = -\vec{B} \cdot \frac{\partial \vec{B}}{\partial t} - \mu_0 \epsilon_0 \vec{E} \cdot \frac{\partial \vec{E}}{\partial t} - \mu_0 \vec{j} \cdot \vec{E}$$

which is easily rewritten as

$$\vec{\nabla} \cdot \left(\frac{\vec{E} \wedge \vec{B}}{\mu_0} \right) = -\frac{1}{2\mu_0} \frac{\partial B^2}{\partial t} - \frac{\epsilon_0}{2} \frac{\partial E^2}{\partial t} - \vec{j} \cdot \vec{E} \quad [1.15A]$$

$$\boxed{\vec{\nabla} \cdot \left(\frac{\vec{E} \wedge \vec{B}}{\mu_0} \right) = -\frac{\partial}{\partial t} \left(\frac{1}{2} \left(\frac{B^2}{\mu_0} + \epsilon_0 E^2 \right) \right) - \vec{j} \cdot \vec{E}}$$

In order to gain physical insight into this relation, let us take a look at the dimensions of each of the terms:



$$\begin{aligned}
\left[\frac{\vec{E} \wedge \vec{B}}{\mu_0} \right] &\rightarrow \frac{W}{m^2} \quad \text{Irradiance} \\
\left[\frac{B^2}{\mu_0} + \epsilon_0 E^2 \right] &\rightarrow \frac{J}{m^3} \quad \text{Energy Density} \\
[\vec{j} \cdot \vec{E}] &\rightarrow \frac{W}{m^3} \quad \text{Power Density}
\end{aligned}$$

By defining the following magnitudes:

Poynting Vector: $\vec{S} = \frac{\vec{E} \wedge \vec{B}}{\mu_0}$

Energy density of the EM field: $u = \frac{1}{2} \left(\frac{B^2}{\mu_0} + \epsilon_0 E^2 \right)$

Relation [15A] can be expressed as

$$\vec{\nabla} \cdot \vec{S} = -\frac{\partial u}{\partial t} - \vec{j} \cdot \vec{E}$$

By integrating this over the volume of space under consideration and making use of the divergence theorem, we obtain

$$\int \vec{\nabla} \cdot \vec{S} \, dV = -\int \frac{\partial u}{\partial t} dV - \int \vec{j} \cdot \vec{E} \, dV \quad \rightarrow \quad \boxed{\int \vec{S} \cdot d\vec{a} = -\frac{\partial U}{\partial t} - P} \quad [1.16A]$$

where $d\vec{a}$ stands for the surface element on the integrating volume contour pointing out of the surface, U is the total energy of the EM field enclosed within the volume, and P corresponds to the total power exerted on the charges by the fields within the volume. This expression constitutes the **energy conservation relation of the EM field**. In the absence of charges, we can simplify this relation:

$$\boxed{\int \vec{S} \cdot d\vec{a} + \frac{\partial U}{\partial t} = 0} \quad [1.16B]$$

The energy flowing out from a given volume through the enclosing surface results in a decrease in the total EM energy stored inside the volume.

Through this interpretation, the Poynting vector represents the irradiance of EM energy out of the surface. In other words, it is the energy propagated in a given direction per unit time and per unit area.



◆ *Comment*

The power generated in a system of charges by the EM field can be calculated in the following way:

$$\text{Elementary work: } dW = \vec{F} \cdot d\vec{r} = [q\vec{E} + (q\vec{v} \wedge \vec{B})] \cdot d\vec{r} = q\vec{E} \cdot d\vec{r} = \int \rho \vec{E} \cdot d\vec{r} dV$$

$$\text{Power: } \frac{dW}{dt} = \int \rho \vec{E} \cdot \frac{d\vec{r}}{dt} dV = \int \rho \vec{E} \cdot \vec{v} dV = \int \vec{j} \cdot \vec{E} dV$$

1.4.1 The Poynting vector of harmonic waves. Optical intensity

Let us calculate an explicit expression for the Poynting vector of a plane wave. To this end, we will use the expressions deduced in Section 1.2 to perform the vector product of the fields.

(!) This being a bilinear relationship involving the product of the electric and magnetic fields, some care is needed to take the real value of the fields when making the product. This is because we cannot directly use the complex representation and take the real part at the end due to the fact that $\text{Re}[A \cdot B] \neq \text{Re}[A] \cdot \text{Re}[B]$.

$$\vec{S} = \frac{\vec{E} \wedge \vec{B}}{\mu_0} = \frac{\text{Re}[\vec{E}^{(+)}] \wedge \text{Re}[\vec{B}^{(+)}]}{\mu_0}$$

For harmonic waves:

$$\begin{aligned} \vec{E}(\vec{r}, t) &= \hat{e} E_0 \cos(\vec{k}\vec{r} - \omega t + \phi) \\ \vec{B}(\vec{r}, t) &= \hat{b} B_0 \cos(\vec{k}\vec{r} - \omega t + \phi) \end{aligned} \quad \Rightarrow \quad \vec{S}(\vec{r}, t) = \frac{E_0 B_0}{\mu_0} \cos^2(\vec{k}\vec{r} - \omega t + \phi) [\hat{e} \wedge \hat{b}]$$

Using [1.11]:

$$\vec{S}(\vec{r}, t) = \frac{E_0^2}{c\mu_0} \cos^2(\vec{k}\vec{r} - \omega t + \phi) \hat{u} = \epsilon_0 c E_0^2 \cos^2(\vec{k}\vec{r} - \omega t + \phi) \hat{u}$$

From this expression, two important consequences may be derived:

1) For plane waves in isotropic media, the Poynting vector and consequently the EM energy flows in the wave vector's direction of propagation. The phase wavefronts (constant phase) are parallel to the energy wavefronts (constant energy).

(!) This is generally not true when light propagates in anisotropic media, as we will see it in Chapter 4.



2) The Poynting vector value oscillates with the optical frequency $\omega \approx 10^{15}$ Hz. These oscillations are too fast to be resolved by any electronic detector, and thus this is not a direct measurable quantity. The result of any real measurement by optical detectors will be the integrated signal over a period of time much longer than the period of a single oscillation. We define the **light irradiance or optical intensity** (units W/m²) as the temporal average of the modulus of the Poynting vector

$$I = \left\langle \left| \vec{S}(\vec{r}, t) \right| \right\rangle \quad [1.17A]$$

For linearly polarized light:

$$I = \left\langle \left| \vec{S} \right| \right\rangle = \frac{1}{T} \int_0^T \epsilon_0 c E_0^2 \cos^2(\vec{k}\vec{r} - \omega t + \phi) dt$$

Using the identity $\cos^2 a = \frac{1}{2}[1 + \cos(2a)]$, this results in

$$I = \frac{\epsilon_0 c E_0^2}{T} \int_0^T \frac{1}{2} [1 + \cos(2\vec{k}\vec{r} - 2\omega t + 2\phi)] dt = \frac{1}{2} \epsilon_0 c E_0^2 \quad [1.17B]$$

$$I = \frac{1}{2} \epsilon_0 c E_0^2$$

For fields with harmonic dependence, an alternative expression can be obtained by directly using the complex representation (analytic signal) (see Solved problems):

$$I = \frac{1}{2} \operatorname{Re} \left(\frac{E^{(+)} \wedge B^{(-)}}{\mu_0} \right) = \frac{1}{2} \epsilon_0 c |E_0^{(+)}|^2$$

Example 3: *Light intensity of circular polarized light*

For circular polarization, the electric field is written as (Equation [1.14]):

$$\vec{E}(\vec{r}, t) = E_0 [\cos(kz - \omega t + \phi) \hat{i} + \sin(kz - \omega t + \phi) \hat{j}]$$

with $\hat{e} = (\cos(kz - \omega t + \phi) \hat{i} + \sin(kz - \omega t + \phi) \hat{j})$

The magnetic field is:

$$\vec{B}(\vec{r}, t) = B_0 \hat{b} = \frac{E_0}{c} (\hat{k} \wedge \hat{e}) = \frac{E_0}{c} [-\sin(kz - \omega t + \phi) \hat{i} + \cos(kz - \omega t + \phi) \hat{j}]$$

The Poynting vector is



$$\vec{S} = \frac{\vec{E} \wedge \vec{B}}{\mu_0} = \frac{E_0^2}{\mu_0 c} \hat{k} \quad \Rightarrow \quad I = \langle |\vec{S}| \rangle = \frac{E_0^2}{\mu_0 c} = \epsilon_0 c E_0^2$$

As can be seen here for circular polarized light, the Poynting vector has constant magnitude, since in this case the temporal dependence is included in the unitary direction vectors of the fields.

For unpolarized light, the physically relevant magnitude is the wave intensity. In these situations, the scalar wave function is usually identified directly with the light intensity through the relation $I = |U|^2$, as discussed previously. The units of the scalar wave function are then $(\text{W/m}^2)^{1/2}$.

For monochromatic wave solutions, the electric field spans in time from $-\infty$ to $+\infty$, and the intensity is constant in time. These are called continuous wave (CW) sources. For real sources, the emitted intensity is not constant but fluctuates around a given average value due to the typical character of light emitted by atoms (discussed in much detail in Chapter 3). Actual lasers, however, are capable of delivering light pulses with temporal durations ranging from the nanosecond to the femtosecond regimes. In this case, the intensity profile of the pulse shows a temporal dependence at temporal scales that are larger than the optical period T . This temporal dependence is not washed out by the temporal average obtained when calculating the intensity, which removes the fast oscillations of the electric field under the pulse envelope. These questions will be discussed in further detail in Chapter 5. When considering spatially localized solutions in Chapter 2, we will see that the light distribution of a light wave packet in space and time can be expressed in terms of an intensity function $I(x, y, z, t)$, which represents the light energy per unit time and unit area. Incidentally, in radiometry, the magnitude defined by this definition corresponds to the irradiance, the term intensity being reserved for the energy distribution per unit solid angle. However, in the theory of coherent optical radiation, it is customary to refer to intensity, just as we have introduced the term.

By knowing the intensity distribution, we can determine the optical power:

$$P(t) = \int I(\vec{r}, t) dS$$

and total pulse energy

$$Q = \int P(t) dt$$

1.5 Momentum relations and radiation pressure

The linear momentum conservation equation for a system of particles and fields can be derived from Maxwell's equations, starting with the Lorentz force equation:

$$\vec{F} = \int (\rho \vec{E} + \vec{j} \wedge \vec{B}) dV$$



By writing ρ and $\mathbf{j}$ in terms of the fields using Equations [1.1A] and [1.1D], the following relation is found after some algebraic steps:

$$\frac{d\vec{G}_{mec}}{dt} + \frac{d\vec{G}_{EM}}{dt} = \int \vec{T} d\vec{a}$$

where

$\vec{G}_{mec} = \sum_i m_i \vec{v}_i$ is the mechanical momentum

$\vec{G}_{EM} = \epsilon_0 \int (\vec{E} \wedge \vec{B}) dV$ is the momentum of the EM field and $\vec{T}$ is the Energy stress tensor

We can also define the **angular momentum** of the electromagnetic field:

$$\vec{J}_{EM} = \epsilon_0 \int \vec{r} \wedge (\vec{E} \wedge \vec{B}) dV$$

Demonstrably, the total angular momentum can be decomposed into the sum of two different terms: the orbital angular momentum, related to the evolution of the Poynting vector in time; and the spin angular momentum, related to the polarization state of light.

◆ *Comment*

Radiation can exert forces on macroscopic objects. Determining these forces is something of a complicated issue and beyond the scope of this course. In essence, this complication is because the expression for the energy-stress tensor should use the total field (considering reflection from the object) and not just the incident field. This is quite an interesting topic and very relevant when considering the interaction of light with microscopic particles, as in the field of nanophotonics. Furthermore, it forms the basis of current topics in photonics, such as light tweezers and solar sails.

In the simplified situation of a plane wave incident on a given surface, we can obtain a simple expression for the pressure exerted by light on the object. We will introduce the basic expression with no further derivation.

Radiation pressure; P_{rad} :

$$P_{rad} = (1 + R) \frac{I}{c}$$

where R is the surface reflectivity, I the light intensity incident on the surface and c the speed of light.



Radiation force; F_{rad} :

The radiation pressure acting on the surface exerts a force,

$$F_{rad} = P_{rad} \cdot A = (1 + R) \frac{P}{c}$$

where P is the optical power of the light beam.

Momentum transfer; Δp :

The net momentum transfer is

$$\Delta p = F_{rad} \cdot \Delta t = (1 + R) \frac{Q}{c}$$

where Q corresponds to the light energy incident on the surface.

Classically, the effect of the radiation pressure is related to the magnetic field acting on the moving electrons accelerated by the electric field. This should not be confused with other forces acting on particles such as radiometric forces due to thermal gradients. When these gradients are created by laser light illumination, the effect is called photophoresis.

The radiation pressure acting on a medium with an extended interface when illuminated by light can be used to accelerate objects if other external forces can be neglected. This effect forms the basis of some ideas for spacecraft acceleration by means of solar sails (see, for instance, the IKAROS project of the Japanese Aerospace Exploration Agency).

For smaller particles, the forces of radiation pressure from CW visible light can provide stable optical potential wells, where particles can be trapped and accelerated (A. Ashkin, *Physical Review Letters*, 24, 156 (1970)). Ashkin and coworkers later demonstrated (Ashkin et al., *Optics Letters*, 11, 288 (1986)) that, when using tightly focused laser beams, the presence of an electric field gradient can trap microscopic particles near the beam focus. This forms the basis for the concept of **optical tweezers**, which are widely used nowadays to trap and manipulate dielectric spheres, DNA, bacteria, metallic particles, and others.

Particles smaller than the wavelength can be described by dipoles. Forces on these dipoles from monochromatic radiation have two contributions: a dipole force trapping the particle in the region of an electric field's larger gradient; and a scattering force that moves the particle in the direction of the beam propagation, which occurs due to momentum transfer.

The dipole forces, for instance, can be used to trap atoms in the highest intensity regions of optical traps (either in 1D or 2D configurations) if the light frequency is tuned slightly lower to one of the atom's resonance frequencies. This is due to the shift in their energy levels at the gradient.

The scattering forces transfer momentum to the atom in the direction of propagation when slightly detuned with respect to a resonant atomic frequency. Efficient absorption is obtained thanks to the Doppler effect acting on the moving particle. Since the spontaneous emission occurring after the absorption has a random directional emission, the total process leads to a net momentum transfer to the atom, thus decreasing its velocity. This effect forms the basis of atom laser cooling applications.



1.6 Solved problems

S1.1 Starting with the scalar wave equation, obtain the 1D plane wave and the spherical wave solutions.

Solution

The scalar 1D wave equation: $\left(\frac{\partial^2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right)U(z,t) = 0$

can be rewritten as

$$\left(\frac{\partial}{\partial z} - \frac{1}{c} \frac{\partial}{\partial t}\right) \cdot \left(\frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t}\right) U(z,t) = 0$$

By introducing the new variables $\eta = z - ct$ $\Rightarrow$ $z = \frac{1}{2}(\xi + \eta)$
 $\xi = z + ct$ $t = \frac{1}{2c}(\xi - \eta)$

the wave equation in the new variables $\frac{\partial^2 U(\eta, \xi)}{\partial \eta \partial \xi} = 0$

has the general solution

$$U(\eta, \xi) = f(\eta) + g(\xi) \rightarrow U(z, t) = f(z - ct) + g(z + ct)$$

When dealing with a spherical wave, the Laplacian for spherically symmetric solutions ($\partial/\partial\phi = \partial/\partial\theta = 0$) can be expressed in spherical coordinates as

$$\nabla^2 U(r, t) = \frac{1}{r^2} \cdot \frac{\partial}{\partial r} \left(r^2 \left(\frac{\partial U(r, t)}{\partial r} \right) \right) = \dots = \frac{1}{r} \frac{\partial^2}{\partial r^2} (rU(r, t))$$

The wave equation

$$\begin{aligned} \nabla^2 U(r, t) - \frac{1}{c^2} \frac{\partial^2 U(r, t)}{\partial t^2} &= 0 \rightarrow \frac{\partial^2 (rU(r, t))}{\partial r^2} - \frac{1}{c^2} \frac{\partial^2 (rU(r, t))}{\partial t^2} = 0 \\ \rightarrow \left(\frac{\partial^2}{\partial r^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) (rU(r, t)) &= 0 \end{aligned}$$

can be solved in the same way as for the plane wave:

$$\begin{aligned} \left[\begin{array}{l} \eta = r - ct \\ \xi = r + ct \end{array} \right] \rightarrow \frac{\partial^2 (rU(\eta, \xi))}{\partial \eta \partial \xi} = 0 \rightarrow rU(\eta, \xi) &= f(\eta) + g(\xi) \\ U(z, t) &= \frac{f(r - ct)}{r} + \frac{g(r + ct)}{r} \end{aligned}$$

■



S1.2 Write the equation of an electromagnetic wave with wavelength $\lambda = 532 \text{ nm}$ propagating in the xz -plane along the direction forming an angle of $\theta = 75^\circ$ with respect to the z -axis. The electric field is polarized along the y -axis, and its amplitude is $E_0 = 1000 \text{ V/cm}$.

Solution

The equation of the plane electromagnetic wave:

$$\begin{cases} \vec{E}(\vec{r}, t) = \hat{e} E_0 \cos(\vec{k} \cdot \vec{r} - \omega t) \\ \vec{B}(\vec{r}, t) = \hat{b} B_0 \cos(\vec{k} \cdot \vec{r} - \omega t) \end{cases}$$

The known parameters of the wave allow us to determine the following unknown parameters:

$$\lambda = 500 \text{ nm} \rightarrow k = \frac{2\pi}{\lambda} = 1.18 \cdot 10^7 \text{ m}^{-1}$$

$$\text{Direction of propagation: } \hat{u} = (\cos \alpha, \cos \beta, \cos \gamma) = (\cos 15^\circ, \cos 90^\circ, \cos 75^\circ)$$

$$\omega = \frac{2\pi c}{\lambda} = 3.54 \cdot 10^{15} \text{ rad / s}$$

$$E_0 = 10^5 \text{ V / m} \rightarrow B_0 = E_0 / c = 10^5 / 3 \cdot 10^8$$

$$\hat{e} = (0, 1, 0) \rightarrow \hat{b} = \hat{u} \wedge \hat{e} = (-\cos 75^\circ, 0, \cos 15^\circ)$$

$$\vec{E}(\vec{r}, t) = \hat{j} 10^5 \cos(1.18 \cdot 10^7 (x \cdot \cos 15^\circ \hat{i} + z \cdot \cos 75^\circ \hat{k}) - 3.54 \cdot 10^{15} t)$$

$$\vec{B}(\vec{r}, t) = 0.33 \cdot 10^{-3} (-\cos 75^\circ \hat{i} + \cos 15^\circ \hat{k}) \cos(1.18 \cdot 10^7 (x \cdot \cos 15^\circ \hat{i} + z \cdot \cos 75^\circ \hat{k}) - 3.54 \cdot 10^{15} t)$$

■

S1.3 A point source emits light of $\lambda = 600 \text{ nm}$ with a total power of 100 W . Write the expression for the scalar spherical wave representing the electric field at a distance of $r = 10 \text{ m}$ from the source.

Solution

The equation of the scalar spherical electromagnetic wave:

$$\begin{cases} E(r, t) = \frac{E_0}{r} \cos(k \cdot r - \omega t) \\ B(r, t) = \frac{B_0}{r} \cos(k \cdot r - \omega t) \end{cases}$$

The total power of the spherical wave is related to the amplitude of the electric field at a distance of $r = 1 \text{ m}$, E_0 , specifically by means of the equation

$$I = \frac{1}{2} \epsilon_0 c \frac{|E_0|^2}{r^2} = \frac{P}{4\pi r^2} \rightarrow E_0 = \sqrt{\frac{P}{2\pi \epsilon_0 c}} = 77.46 \text{ V / m}$$

$$\lambda = 600 \text{ nm} \rightarrow k = \frac{2\pi}{\lambda} = 1.05 \cdot 10^7 \text{ m}^{-1} \rightarrow \omega = \frac{2\pi c}{\lambda} = 3.14 \cdot 10^{15} \text{ rad / s}$$



$$\begin{cases} E(r, t) = \frac{77.46}{r} \cos(1.05 \cdot 10^7 \cdot r - 3.14 \cdot 10^{15} t) \\ B(r, t) = \frac{2.58 \cdot 10^{-7}}{r} \cos(1.05 \cdot 10^7 \cdot r - 3.14 \cdot 10^{15} t) \end{cases}$$

At $r = 10$ m the electric field is

$$E(10, t) = 7,746 \cos(10.5 \cdot 10^7 - 3.14 \cdot 10^{15} t)$$

■

S1.4 A light ray corresponding to a linearly polarized plane wave of wavelength λ propagates in an arbitrary direction specified by the unitary vector, which is expressed in spherical coordinates as $\hat{u} = \cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k}$.

The direction of the electric field remains constant during propagation, and it is perpendicular to the plane containing the light ray and the OZ axis. Write the general expression for the resulting electromagnetic wave.

Solution

Since the direction of the electric field is perpendicular to the plane containing the unitary vectors $\hat{u}$ and $\hat{k}$, its direction can be found by:

$$\hat{e} = \frac{\hat{u} \wedge \hat{k}}{|\hat{u} \wedge \hat{k}|} = \frac{\begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \cos \phi \sin \theta & \sin \phi \sin \theta & \cos \theta \\ 0 & 0 & 1 \end{vmatrix}}{\sin \theta} = \frac{(\sin \phi \sin \theta \hat{i} - \cos \phi \sin \theta \hat{j})}{\sin \theta} = (\sin \phi \hat{i} - \cos \phi \hat{j})$$

(!) We divided the vector product by its modulus, since the vectors $\hat{u}$ and $\hat{k}$ are in general not mutually perpendicular.

The direction of the magnetic field can be obtained immediately:

$$\hat{b} = \hat{u} \wedge \hat{e} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \cos \phi \sin \theta & \sin \phi \sin \theta & \cos \theta \\ \sin \phi & -\cos \phi & 0 \end{vmatrix} = \cos \phi \cos \theta \hat{i} + \sin \phi \cos \theta \hat{j} - \sin \theta \hat{k}$$

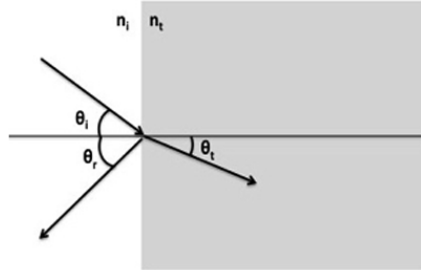
Once we have found the unitary vectors of the polarization states, we can write the corresponding fields:

$$\begin{aligned} \vec{E}^{(+)}(\vec{r}, t) &= E_o (\sin \phi \hat{i} - \cos \phi \hat{j}) \cos \left(\frac{2\pi}{\lambda} (\cos \phi \sin \theta x + \sin \phi \sin \theta y + \cos \theta z) - \frac{2\pi c t}{\lambda} \right) \\ \vec{B}^{(+)}(\vec{r}, t) &= \frac{E_o}{c} (\cos \phi \cos \theta \hat{i} + \sin \phi \cos \theta \hat{j} - \sin \theta \hat{k}) \cos \left(\frac{2\pi}{\lambda} (\cos \phi \sin \theta x + \sin \phi \sin \theta y + \cos \theta z) - \frac{2\pi c t}{\lambda} \right) \end{aligned}$$

■



S1.5 A plane electromagnetic wave is incident on a surface separating two media of refractive indices n_i and n_t , as shown in the figure below.



The wave propagates in the xz -plane, forming an angle θ_i with respect to the z -axis with TE polarization. Applying the boundary conditions at the interface, obtain an expression for the reflection and transmission coefficients and determine the energy reflected and transmitted at the boundary.

Solution

The electric field incident at the interface $\vec{E}_i^{(+)} = E_{oi}^{(+)} \hat{e}_i \cdot e^{i(\vec{k}_i \cdot \vec{r} - \omega t)}$ will separate into 1) a reflected beam propagating back into the incident medium of refractive index n_1 , and 2) a transmitted beam propagating in the medium of refractive index n_2 . All these beams are contained in a common plane known as the incidence plane. The amplitudes and directions of those beams are defined by the boundary conditions at the interface, and they depend strongly on the polarization state of the incident wave.

We define two different polarization orientations: TE polarization (electric field perpendicular to the incidence plane); and TM polarization (magnetic field perpendicular to the incidence plane).

The corresponding wave vectors can be written in terms of their propagating angles and index of refraction, in accordance with the figure.

$$\text{Incident wave: } \vec{k}_i = n_i k_o (-\sin \theta_i \hat{i} + \cos \theta_i \hat{k})$$

$$\text{Reflected wave: } \vec{k}_r = n_i k_o (-\sin \theta_r \hat{i} - \cos \theta_r \hat{k})$$

$$\text{Transmitted wave: } \vec{k}_t = n_t k_o (-\sin \theta_t \hat{i} + \cos \theta_t \hat{k})$$

Independently of the incident polarization state, the phase of the oscillating dipoles at the surface ($z = 0$) should be the same for all waves at each point and instant of time, since these dipoles are set into motion by the action of the incident field and radiate the wave that is reflected in medium 1 and the wave transmitted in medium 2.

This gives rise to the following conditions:



$$\vec{k}_i \cdot \vec{r} - \omega_i t = \vec{k}_r \cdot \vec{r} - \omega_r t = \vec{k}_t \cdot \vec{r} - \omega_t t \rightarrow$$

$$\omega_i = \omega_r = \omega_t = \omega \quad (\text{energy conservation})$$

$$n_i \sin \theta_i = n_i \sin \theta_r = n_t \sin \theta_t \quad (\text{transverse momentum conservation})$$

From the second condition, we extract two important laws (Snell's laws):

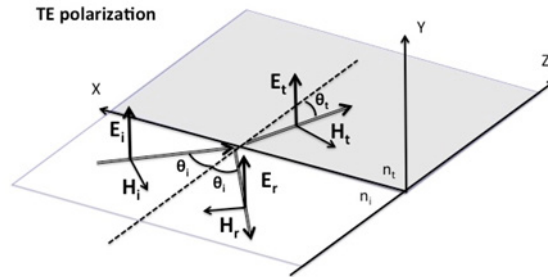
$$\text{Law of reflection: } \theta_i = \theta_r,$$

$$\text{Law of transmission: } n_i \sin \theta_i = n_t \sin \theta_t$$

Next, the boundary conditions for each polarization state are applied to calculate the amplitude of the corresponding reflected and transmitted waves as a function of the amplitude of the incident wave.

TE modes

The expressions for the complex amplitudes of the electric field and magnetic intensity $H = B/\mu_o$ (a magnitude to be detailed in Chapter 3) for this polarization mode can be obtained from the figure. The magnetic field is calculated from the electric field using Maxwell's Equation [1.3]



$$\vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t} \rightarrow i\vec{k} \wedge \vec{E} = i\omega \vec{B}$$

$$\vec{B} = \frac{nk_o E_o}{\omega} (\hat{u} \wedge \hat{e})$$

Using these relations, we obtain

$$\begin{aligned} \vec{E}_{oi} &= E_{oi} \hat{j}; & \vec{E}_{or} &= E_{or} \hat{j}; & \vec{E}_{ot} &= E_{ot} \hat{j} \\ \vec{H}_{oi} &= n_i \beta_i (-\cos \theta_i \hat{i} - \cos \theta_i \hat{k}); & \vec{H}_{or} &= n_i \beta_r (\cos \theta_i \hat{i} - \cos \theta_i \hat{k}); & \vec{H}_{ot} &= n_t \beta_t (-\cos \theta_t \hat{i} - \cos \theta_t \hat{k}) \end{aligned}$$

where we define $\beta_j = \frac{k_o E_{oj}}{\mu_o \omega}$; and $j = i, r, t$.

Boundary conditions: The tangential components of E and H are continuous across the boundary



$$E_{oi} + E_{or} = E_{ot}$$

$$-n_i \cos \theta_i E_{oi} + n_i \cos \theta_i E_{or} = -n_t \cos \theta_t E_{ot}$$

Defining the reflection (r) and transmission (t) coefficients as $r = E_{or}/E_{oi}$ and $t = E_{ot}/E_{oi}$, we can write

$$1 + r = t$$

$$1 - r = t \frac{n_t \cos \theta_t}{n_i \cos \theta_i} \rightarrow \boxed{\begin{aligned} r_{TE} &= \frac{n_t \cos \theta_t - n_i \cos \theta_i}{n_i \cos \theta_i + n_t \cos \theta_t} \\ t_{TE} &= \frac{2n_i \cos \theta_i}{n_i \cos \theta_i + n_t \cos \theta_t} \end{aligned}}$$

The obtained amplitude coefficients are generally complex. The reflected and transmitted energy can be obtained by defining the reflectance, R , and transmittance, T , at the boundary. When the medium is lossless, energy conservation requires that

$$R + T = 1.$$

In order to determine these quantities, let us assume that our incident beam has incident power P_i and an area transverse to the propagation direction equal to S_i . When incident at an angle θ_i , the projected area at the plane of the boundary will be $A = S_i / \cos \theta_i$. The reflected and transmitted beams will have transverse areas S_r and S_t , respectively, where

$$A = S_r / \cos \theta_r = S_t / \cos \theta_t.$$

Conservation of energy requires that the sum of the reflected and transmitted powers should be equal to the incident power:

$$P_i = P_r + P_t \rightarrow I_i \cdot S_i = I_r \cdot S_r + I_t \cdot S_t$$

$$1 = \frac{I_r}{I_i} + \frac{I_t \cdot S_t}{I_i \cdot S_i} = R + T$$

From this expression, we can relate the reflectance and transmittance to the previously obtained reflection and transmission coefficients:

$$R = \frac{I_r}{I_i} = \frac{\frac{1}{2} \epsilon_o n_i c |E_r|^2}{\frac{1}{2} \epsilon_o n_i c |E_i|^2} = \frac{|r E_i|^2}{|E_i|^2} \rightarrow \boxed{R = |r|^2}$$

$$T = \frac{P_t}{P_i} = \frac{I_t S_t}{I_i S_i} = \frac{\frac{1}{2} \epsilon_o n_t c |t E_i|^2 \cos \theta_t}{\frac{1}{2} \epsilon_o n_i c |E_i|^2 \cos \theta_i} \rightarrow \boxed{T = \frac{n_t \cos \theta_t}{n_i \cos \theta_i} |t|^2}$$

■



S1.6 We need to rotate the polarization direction of a linearly polarized wave by 90° using linear polarizers with minimum losses.

a) If we have two polarizers, describe which is the best choice for the orientation of the polarizers with respect to the incident polarization.

b) Which is the best choice if we have three polarizers?

Solution

a) If the final state must have linear polarization perpendicular to the incident state, we cannot use a single linear polarizer because, in this case, we would obtain zero transmission. Using two polarizers, the last one should be oriented along the direction corresponding to the desired output polarization.

Let us suppose that the initial state is polarized in the horizontal direction

$$\hat{e}_{in} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

such that the output state will be oriented along the vertical. Placing the first polarizer at an angle θ_p with respect to the horizontal, the second must be oriented at $\pi/2$ rad. We can write the expression for the transmitted field as

$$\begin{aligned} \vec{e}_{out} &= P_2 \cdot P_1 \cdot \hat{e}_{in} = \\ &= \begin{pmatrix} \cos^2\left(\frac{\pi}{2}\right) & \cos\left(\frac{\pi}{2}\right)\sin\left(\frac{\pi}{2}\right) \\ \cos\left(\frac{\pi}{2}\right)\sin\left(\frac{\pi}{2}\right) & \sin^2\left(\frac{\pi}{2}\right) \end{pmatrix} \cdot \begin{pmatrix} \cos^2\theta_p & \cos\theta_p \sin\theta_p \\ \cos\theta_p \sin\theta_p & \sin^2\theta_p \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \vec{e}_{out} &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos^2\theta_p \\ \cos\theta_p \sin\theta_p \end{pmatrix} = \begin{pmatrix} 0 \\ \cos\theta_p \sin\theta_p \end{pmatrix} = \cos\theta_p \sin\theta_p \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

The output intensity can be expressed in terms of the incident intensity through the relation

$$I_{out} = I_{in} \cdot \vec{e}_{out} \cdot \vec{e}_{out}^* = I_{in} \cos^2\theta_p \sin^2\theta_p.$$

To obtain the angle for which the output intensity will be maximum, the derivative must be zero:

$$\frac{d(\cos^2\theta_p \sin^2\theta_p)}{d\theta_p} = 0 \rightarrow 2\cos^3\theta_p \sin\theta_p - 2\cos\theta_p \sin^3\theta_p = 0 \rightarrow \cos^2\theta_p = \sin^2\theta_p$$

$$\tan^2\theta_p = 1 \rightarrow \theta_p = \pm \frac{\pi}{4} \text{ rad}$$



b) For the case of three polarizers, we must insert two polarizers at arbitrary angles θ_{p1} and θ_{p2} , and the last one at $\pi/2$ rad with respect to the polarization direction of the incident beam:

$$\vec{e}_{out} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos^2 \theta_{p2} & \cos \theta_{p2} \sin \theta_{p2} \\ \cos \theta_{p2} \sin \theta_{p2} & \sin^2 \theta_{p2} \end{pmatrix} \cdot \begin{pmatrix} \cos^2 \theta_{p1} & \cos \theta_{p1} \sin \theta_{p1} \\ \cos \theta_{p1} \sin \theta_{p1} & \sin^2 \theta_{p1} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\vec{e}_{out} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos^2 \theta_{p2} & \cos \theta_{p2} \sin \theta_{p2} \\ \cos \theta_{p2} \sin \theta_{p2} & \sin^2 \theta_{p2} \end{pmatrix} \cdot \begin{pmatrix} \cos^2 \theta_{p1} \\ \cos \theta_{p1} \sin \theta_{p1} \end{pmatrix}$$

$$\vec{e}_{out} = \cos \theta_{p1} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos^2 \theta_{p2} \cos \theta_{p1} + \cos \theta_{p2} \sin \theta_{p2} \sin \theta_{p1} \\ \cos \theta_{p2} \sin \theta_{p2} \cos \theta_{p1} + \sin^2 \theta_{p2} \sin \theta_{p1} \end{pmatrix}$$

$$\vec{e}_{out} = \cos \theta_{p1} \cos(\theta_{p2} - \theta_{p1}) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos \theta_{p2} \\ \sin \theta_{p2} \end{pmatrix} = \cos \theta_{p1} \sin \theta_{p2} \cos(\theta_{p2} - \theta_{p1}) \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The output intensity in this case is

$$I_T = I_{inc} \cos^2 \theta_{p1} \sin^2 \theta_{p2} \cos^2(\theta_{p2} - \theta_{p1}).$$

The condition of maximum transmitted intensity is obtained when the partial derivatives with respect to both angles simultaneously vanish:

$$\frac{\partial I_T}{\partial \theta_{p1}} = 0 \rightarrow \cos \theta_{p1} \sin(\theta_{p2} - \theta_{p1}) - \sin \theta_{p1} \cos(\theta_{p2} - \theta_{p1}) = 0 \rightarrow \sin(\theta_{p2} - 2\theta_{p1}) = 0$$

$$\frac{\partial I_T}{\partial \theta_{p2}} = 0 \rightarrow \cos \theta_{p2} \cos(\theta_{p2} - \theta_{p1}) - \sin \theta_{p2} \sin(\theta_{p2} - \theta_{p1}) = 0 \rightarrow \cos(2\theta_{p2} - \theta_{p1}) = 0$$

Since the angles θ_{p1} and θ_{p2} are smaller than $\pi/2$:

$$\left. \begin{array}{l} 2\theta_{p1} = \theta_{p2} \\ 2\theta_{p2} - \theta_{p1} = \pi/2 \end{array} \right\} \rightarrow 3\theta_{p1} = \pi/2$$

$$\boxed{\theta_{p1} = \frac{\pi}{6}; \quad \theta_{p2} = \frac{\pi}{3}; \quad \theta_{p3} = \frac{\pi}{2}}$$

■

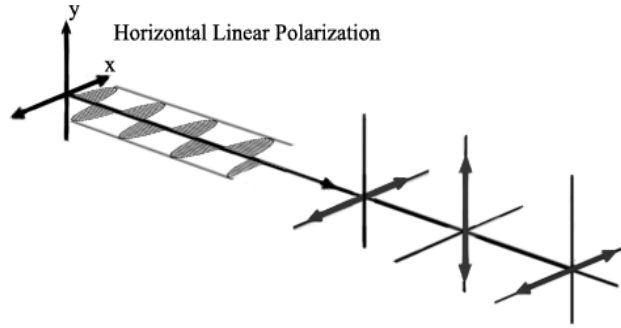
S1.7 An electromagnetic wave, which is horizontally linearly polarized (as shown in the figure), is incident on a device made of three distinct linear polarizers, which are initially prepared as in the configuration below.

a) At a given instant we start rotating the first and third polarizers simultaneously with constant angular velocity $\omega = 0.5$ rad/s, but in opposite directions, while keeping the second polarizer fixed in its initial orientation. Determine the transmitted intensity of the device as a function of the incident intensity, I_o , and time. Then, calculate both the first



temporal value for which the light transmission is maximum and the corresponding intensity value.

b) Repeat a) if the second polarizer is now rotating with identical angular velocity while the first and third polarizers are fixed in their initial orientations.



Solution

a) The action of the system of polarizers on the incident wave can be expressed using the Jones vector representation as the matrix product:

$$\vec{e}_{out} = P_3 \cdot P_2 \cdot P_1 \cdot \hat{e}_{in}$$

where the incident Jones vector is $\hat{e}_{in} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

All matrices correspond to linear polarizers and can be expressed as

$$P_{\theta_p} = \begin{pmatrix} \cos^2 \theta_p & \cos \theta_p \sin \theta_p \\ \cos \theta_p \sin \theta_p & \sin^2 \theta_p \end{pmatrix}$$

The angle of the second polarizer is fixed to $\theta_{p2} = \pi/2$, while the angles of the first and third polarizers vary with time according to the expressions

$$\theta_{p1}(t) = \omega t \quad \text{and} \quad \theta_{p3}(t) = -\omega t$$

By directly substituting these values into the corresponding matrices

$$\vec{e}_{out}(t) = \begin{pmatrix} \cos^2 \omega t & -\cos \omega t \sin \omega t \\ -\cos \omega t \sin \omega t & \sin^2 \omega t \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos^2 \omega t & \cos \omega t \sin \omega t \\ \cos \omega t \sin \omega t & \sin^2 \omega t \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\vec{e}_{out} = -\cos(\omega t) \sin^2(\omega t) \begin{pmatrix} \cos(\omega t) \\ -\sin(\omega t) \end{pmatrix} = -\cos(\omega t) \sin^2(\omega t) \begin{pmatrix} \cos(-\omega t) \\ \sin(-\omega t) \end{pmatrix}$$



we have a result showing that the output polarization state is linear and rotates counter clockwise with angular velocity ω .

The transmitted intensity is

$$I_T(t) = I_o (\cos(\omega t) \sin^2(\omega t))^2.$$

The maximum transmitted intensity will occur at

$$\frac{\partial I_T}{\partial t} = 0 \rightarrow \dots \rightarrow -\sin^2(\omega t) + 2 \cos^2(\omega t) = 0$$

$$\tan(\omega t) = \sqrt{2} \rightarrow t = \tan^{-1}(\sqrt{2}) / \omega = 2 \tan^{-1}(\sqrt{2})$$

$$t = 1.91s$$

$$I_T(1.91) = 0.148 I_o$$

b) If the second polarizer rotates but the other two stay fixed, we have

$$\theta_{p1} = 0; \quad \theta_{p2}(t) = \frac{\pi}{2} + \omega t; \quad \theta_{p3} = 0$$

and

$$\vec{e}_{out}(t) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \sin^2 \omega t & \sin \omega t \cos \omega t \\ \sin \omega t \cos \omega t & \cos^2 \omega t \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\vec{e}_{out}(t) = \sin^2(\omega t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow I_T = I_o \sin^4(\omega t)$$

The maximum transmitted intensity occurs first at $t = \pi s$, and its value is I_o

■

S1.8 A plane wave with amplitude E_0 propagates along the z -axis with linear horizontal polarization.

a) Express this wave as a superposition of right and left circular polarization states, writing the expression for the real electric field as a function of z and t .

b) The wave enters with normal incidence into an optical active material of length L . Being optically active means that we can define different refractive indices for the right and left circular polarizations. Considering that these indices are called n_R and n_L respectively, demonstrate that the resulting electric field at the output of the crystal of length L is linearly polarized and determine the orientation of this polarization state. (Consider that there is no reflection at the interfaces.)

**Solution**

a) The general expression for a plane wave propagating along the z -axis and that is linearly polarized along the horizontal axis is

$$\vec{E}^{(+)}(z, t) = \hat{e}_H E_o e^{i(kz - \omega t + \phi_x)} \quad \text{where the Jones vector is given by } \hat{e}_H = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \hat{i}.$$

The Jones vectors for right and left circular polarized light are given by

$$\hat{e}_R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{1}{\sqrt{2}} (\hat{i} - i\hat{j}) \quad \text{and} \quad \hat{e}_L = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{1}{\sqrt{2}} (\hat{i} + i\hat{j})$$

To express the incident wave as a superposition of circular waves:

$$\hat{e}_H = A\hat{e}_R + B\hat{e}_L \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{A}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} + \frac{B}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \rightarrow \begin{cases} \sqrt{2} = A + B \\ 0 = -A + B \end{cases} \rightarrow A = B = \frac{1}{\sqrt{2}}$$

We can then write the incident field as a superposition of right circular and left circular polarized waves:

$$\vec{E}^{(+)}(z, t) = \left(\frac{1}{\sqrt{2}} \hat{e}_R + \frac{1}{\sqrt{2}} \hat{e}_L \right) E_o e^{i(kz - \omega t + \phi_x)} = \frac{E_o}{2} \left[(\hat{i} - i\hat{j}) e^{i(kz - \omega t + \phi_x)} + (\hat{i} + i\hat{j}) e^{i(kz - \omega t + \phi_x)} \right]$$

$$\vec{E}(z, t) = \text{Re} \left[\vec{E}^{(+)}(z, t) \right] = \frac{E_o}{2} \left(\cos(\beta) \hat{i} + \sin(\beta) \hat{j} \right) + \frac{E_o}{2} \left(\cos(\beta) \hat{i} - \sin(\beta) \hat{j} \right)$$

$$\text{with } \beta = kz - \omega t + \phi_x$$

b) If this wave enters a medium where the index of refraction is different for right and left circular polarizations, we can define propagation wave vectors for each polarization:

$$k_R = n_R k_o \quad \text{and} \quad k_L = n_L k_o,$$

and then write the electric field at the output as

$$\vec{E}(L, t) = \frac{E_o}{2} \left[\left(\cos(\beta_R) \hat{i} + \sin(\beta_R) \hat{j} \right) + \left(\cos(\beta_L) \hat{i} - \sin(\beta_L) \hat{j} \right) \right]$$

$$\vec{E}(L, t) = \frac{E_o}{2} \left[\left(\cos(\beta_R) + \cos(\beta_L) \right) \hat{i} + \left(\sin(\beta_R) - \sin(\beta_L) \right) \hat{j} \right]$$

$$\text{with } \beta_R = n_R k_o L - \omega t + \phi_x \quad \text{and} \quad \beta_L = n_L k_o L - \omega t + \phi_x$$

Using the expressions



$$\cos(a) + \cos(b) = 2 \cos((a+b)/2) \cos((a-b)/2)$$

and

$$\sin(a) - \sin(b) = 2 \cos((a+b)/2) \sin((a-b)/2)$$

the electric field is rewritten as

$$\vec{E}(L, t) = E_o \left[\cos(\beta_+) \cos\left(\frac{n_R - n_L}{2} k_o L\right) \hat{i} + \cos(\beta_+) \sin\left(\frac{n_R - n_L}{2} k_o L\right) \hat{j} \right]$$

$$\vec{E}(L, t) = E_o \left(\cos\left(\frac{n_R - n_L}{2} k_o L\right) \hat{i} + \sin\left(\frac{n_R - n_L}{2} k_o L\right) \hat{j} \right) \cos(\beta_+)$$

$$\text{with } \beta_+ = \frac{n_R + n_L}{2} k_o L - \omega t + \phi_x$$

From this expression, we can see that the output field is linearly polarized with a Jones vector:

$$\hat{e}_{out} = \begin{pmatrix} \cos\left(\frac{n_R - n_L}{2} k_o L\right) \\ \sin\left(\frac{n_R - n_L}{2} k_o L\right) \end{pmatrix}$$

The angle of the linearly polarized wave at the exit of the crystal is

$$\theta = \frac{n_R - n_L}{2} k_o L = \frac{(n_R - n_L) L \pi}{\lambda}$$

which depends on the index difference and the length of the crystal.

■

S1.9 At the highest regions of the atmosphere, the radiation intensity emitted by the Sun (known as the solar constant) has been measured at $I = 1.35 \text{ kW/m}^2$.

- Calculate the amplitudes of the electric and magnetic fields in this region.
- The power radiated by the Sun
- The intensity and radiation pressure at the Sun's surface.

Data: The average distance between Earth and Sun is $1.5 \cdot 10^{11} \text{ m}$. The radius of the Sun is $6.96 \cdot 10^8 \text{ m}$.

Solution

a) The angle subtended by the Earth from the Sun is so small that we can consider the light reaching Earth to be formed by a parallel bundle of rays and, hence, it can be seen as a plane wave. Under these conditions, the electric and magnetic fields are given by the expressions



$$I = \frac{1}{2} \epsilon_o c |E_o|^2 \rightarrow E_o = \sqrt{\frac{2I}{\epsilon_o c}} = 1008.5 \text{ V / m}$$

$$B_o = \frac{E_o}{c} = 3.36 \cdot 10^{-6} \text{ T}$$

b) The power radiated by the Sun can be estimated by assuming that the Sun emits a spherical wave:

$$P_{rad} = I 4\pi R^2 = 3.8 \cdot 10^{26} \text{ W}$$

c) At the Sun's surface, the intensity is given by

$$I_{surf} = \frac{P_{rad}}{4\pi R_{Sun}^2} = 6.27 \cdot 10^7 \text{ W/m}^2$$

The radiation pressure at the Sun's surface is

$$p_{rad} = \frac{I_{surf}}{c} = 0.209 \text{ N/m}^2$$

■

S1.10 A pulsed laser emits a 200 ns pulse of 10 MW. This pulse hits an object of mass $m = 0.1 \text{ mg}$ suspended on a very thin filament of 40 cm length. If the object reflects 60% of the incident radiation, calculate the maximum angle of deviation for this pendulum.

Solution

The energy per pulse is calculated from the pulse duration and pulse peak power:

$$E = P \cdot \tau = 2 \text{ J} .$$

The momentum transfer from the laser to the mass is

$$\Delta p = (1 + R) \frac{E}{c} = 1.07 \cdot 10^{-8} \text{ Kgm / s}$$

This momentum transfer induces a change in velocity of $\Delta v = \frac{\Delta p}{m} = 0.107 \text{ m / s} .$

The mass acquires a kinetic energy, which will be converted into potential energy at the maximum angle of deviation:

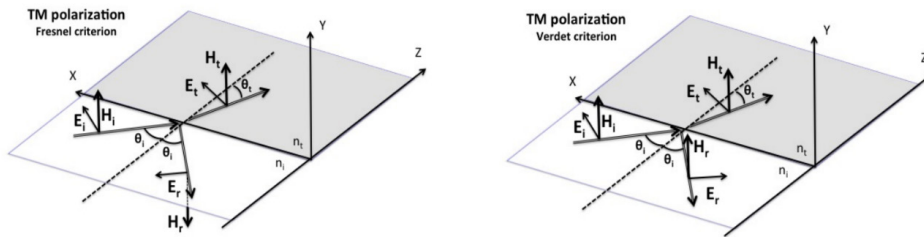
$$E_c = U \rightarrow \frac{1}{2} m \Delta v^2 = mgh \rightarrow h = \frac{\Delta v^2}{2g} = 5.8 \cdot 10^{-4} \text{ m}$$

$$h = l(1 - \cos \theta) \rightarrow \theta = 3.1^\circ$$

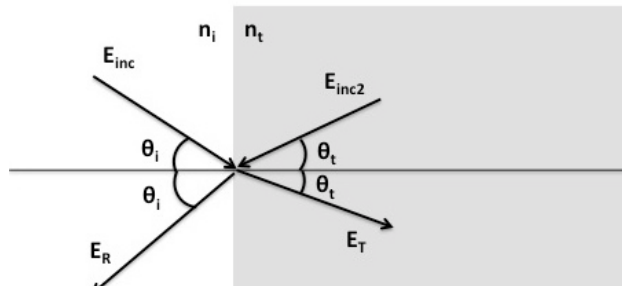


1.7 Proposed problems

P1.1 A plane EM wave is incident on a surface separating two media of refractive indices n_1 and n_2 , as shown in the figure. The wave propagates in the xz -plane, forming an angle θ with respect to the z -axis with TM polarization. Applying the boundary conditions at the interface, obtain an expression for the reflection and transmission coefficients, assuming the Fresnel criterion for the field orientations upon reflection from the surface. Describe what result should be obtained in the case where Verdet's criterion is applied.



P1.2 A plane electromagnetic wave is incident on the boundary separating air ($n_i = 1$) and on a medium with refractive index n_t at an incidence angle θ_i . The fields on each side can be written as the superposition of an incident and a reflected wave, as indicated in the figure. The polarization of the electromagnetic waves is transverse electric (TE mode), and the corresponding angle of propagation in the medium of index n_t is θ_t .



Obtain the scattering matrix relating the amplitude of the fields that are incident on the structure with the fields leaving the boundary:

$$\begin{pmatrix} E_R \\ E_T \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E_{inc} \\ E_{inc2} \end{pmatrix}$$



P1.3 An EM plane wave is given by the expression

$$\vec{E} = (-2\hat{i} + 2\sqrt{3}\hat{j}) \exp(-i(\sqrt{3}x + y + 6 \cdot 10^8 t))$$

Find the direction of polarization, the direction of propagation, the phase velocity, the amplitude, the frequency, and the wavelength

P1.4 A plane monochromatic wave propagating in vacuum has an electric field given by the expression

$$\vec{E}^+(\vec{r}, t) = E_o^+ (\sin \phi \hat{x} - \cos \phi \hat{y}) \exp[i(k(\cos \phi x + \sin \phi y) - \omega t)] .$$

Find: *a)* Magnetic field, *b)* Poynting vector, *c)* Wave intensity

P1.5 The polarization state of an EM wave can be written in vector notation, known as Jones vectors, in the following way:

$\hat{e} = \begin{pmatrix} E_x \\ E_y e^{i\delta} \end{pmatrix}$ where E_x and E_y correspond to the projections of the E field along two orthogonal directions contained in the plane transverse to the propagation direction, and $\delta = \phi_y - \phi_x$ is the relative phase difference.

Find the Jones vector expression for *a)* linear polarization states of arbitrary angle and *b)* right and left circular polarized light. Then, show that linear polarization can be obtained as the sum of two circularly polarized states

P1.6 The action of any polarizing element on the polarization state of the incident light can be represented in matrix form. The matrix characterizing a linear polarizer with its axis forming an angle θ with respect to the x -axis is given by

$$P(\theta) = \begin{pmatrix} \cos^2 \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \sin^2 \theta \end{pmatrix} .$$

A plane wave is linearly polarized in the Y direction and passes through a linear polarizer oriented at an angle θ_1 with respect to the x -axis. The outgoing wave from the first polarizer is transmitted through a second polarizer oriented at an angle θ_2 with respect to the x -axis. Determine the polarization state and intensity of the resulting transmitted wave.

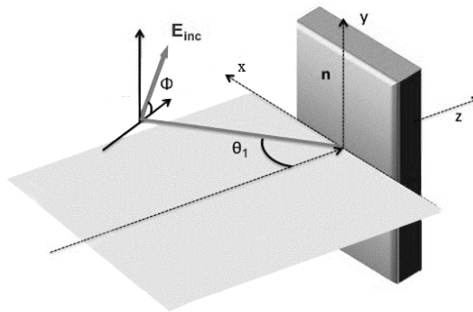
P1.7 A plane EM wave is incident on the surface separating air ($n_{\text{air}} = 1$) and glass ($n = 1.5$) with an incidence angle θ_1 . The polarization state of the wave is linear, forming an angle ϕ with the incidence plane, as shown in the figure.

a) Write the expression for the complex representation of the electric field incident on the glass at the surface plane ($z = 0$).



b) Applying the proper boundary conditions at the surface, calculate the expression for the reflected field for an arbitrary value of the incidence angle, expressed as a function of θ_1 and ϕ .

c) Obtain the value of the reflectivity for the particular incidence angles $\theta_1 = 0^\circ$ and for $\tan(\theta_1) = n$, then describe the resulting polarization state in each situation.



P1.8 The intensity of a plane and monochromatic electromagnetic wave is 100 W/m^2 . Calculate:

- The amplitude of the electric and magnetic fields associated with this wave.
- The radiation pressure.

This wave falls at normal incidence onto a piece of black cardboard with dimensions $20 \times 30 \text{ cm}$, which completely absorbs the radiation.

- Calculate the force exerted by the radiation pressure on the cardboard.
- Repeat the previous calculations, assuming that all the radiation is reflected.

P1.9 A Ti:sapphire laser emits pulses of duration 150 fs with a frequency of 76 MHz at a wavelength of 800 nm . If the average power measured by a slow detector is 1.5 W , determine the power and energy of a single pulse emitted by the laser. If the laser spot is focused to a diameter of 5λ , determine the intensity at this point.

P1.10 The amplitude of a monochromatic plane EM wave is $E_0 = 400 \text{ V/m}$. Calculate a) the amplitude of its magnetic field, b) its intensity, and c) the radiation pressure acting on an object that reflects 20% of the incident radiation.

P1.11 At the Earth's surface, we have measured the intensity of the solar radiation to be 0.75 kW/m^2 . We want to install solar cells to generate 25 kW of electric power. If the solar cells' efficiency is 30%, determine the total area of solar cells we need to mount in our installation.



P1.12 A laser emits a beam with 0.9 mW power over a circular section of 1.2 mm diameter. Determine the force exerted by the radiation on *a)* a black surface with 100% absorption and *b)* a 100% reflecting surface.

P1.13 A spacecraft moves in outer space propelled by the action of the electromagnetic radiation from the Sun (Intensity: 1.4 KW/m^2), which is incident normal to its perfect reflecting sail of 2500 m^2 surface and 7 kg mass.

a) Find the spacecraft's acceleration.

b) Find its speed after 24 hours, starting from a resting position.

c) Neglecting the action of gravity, find the time needed to reach the Moon from the Earth's surface.



→2



Basic models for optical beams

Introduction

In this chapter, our starting point for constructing more realistic light beam solutions will rely on the basic solutions derived in the previous chapter, i.e., the plane wave and the spherical wave. We do not intend to fully develop here the characteristics of optical beams, but to instead introduce a few optical models that are essential in the field of photonics.

A plane wave infinitely extended in space cannot properly represent a real optical beam, which must necessarily be confined to a finite region of space in the plane transverse to the beam propagation. However, in some situations it defines a convenient approximation or can be used as a convenient basis for the representation of more complex fields, just as we will do for instance in the Fourier expansion in reciprocal space. The spherical wave solutions are most usually employed as convenient asymptotic solutions that are far from real sources, or they are applied in scattering problems. The expansion of fields in terms of spherical harmonics, for instance, is the key point in developing angular momentum.

The most general solution with spatial and temporal dependence is called an **optical wave packet**, and it is represented by any function satisfying the wave equations $\mathbf{E}(\mathbf{r}, t) = \text{Re}[E_0(\mathbf{r}, t)e^{i\phi(\mathbf{r}, t)}]$. However, these models are used only to study situations where light is strongly localized in space and time. The difficulties found in mathematically treating such solutions justifies its use only in those situations where strong space–time coupling mechanisms appear. These treatments are out of the scope of this introductory course. Most of the situations can be adequately described by a much simpler approach considering factorized solutions along the lines of $\mathbf{E}(\mathbf{r}, t) = p(\mathbf{r})q(t)$.

In the field of photonics, we find mostly that light propagates in some given direction in the form of spatially localized waves (pointing toward some detector or probing some limited region of space, for instance). Any limitation in the transverse spatial profile will impose a spatial dependence on the complex amplitude of the beam, thereby leading to wavefront curvature and broadening of the transverse profile during propagation. This is



governed by the process of **diffraction**. Diffraction effects modify the light wavefront when we try to spatially confine light beams or when light is scattered from arbitrary objects.

A general **optical beam** represents a solution to the wave equation, in which its complex amplitude has spatial dependence. These kinds of beams modify their profiles upon propagation. In this chapter, we consider optical beams with a single temporal frequency, known as **monochromatic beams**. Thus, we will leave temporal-dependent propagation for a later chapter.

2.1 Helmholtz equation for optical beams

We learned in the previous chapter that a plane wave in a homogeneous medium has constant amplitude and undistorted wavefronts, acquiring a phase term upon propagation. In a more realistic case, the medium cannot be considered perfectly homogeneous and the beam will interact with different elements, such as objects or surfaces, and this will modify both the wavefronts and the amplitude.

The expression of a general monochromatic optical beam is

$$\vec{E}(\vec{r}, t) = \text{Re}[\vec{E}_0^{(+)}(\vec{r})e^{-i\omega t}] \quad [2.1]$$

where the complex amplitude is now considered to be dependent on the spatial coordinates.

Introducing this expression in the general wave Equation [1.4A], we obtain

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad \rightarrow \quad \vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) - \frac{\omega^2}{c^2} \vec{E} = 0$$

Considering the dispersion relation, it can be expressed in the form

$$\boxed{\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}_0^{(+)}(\vec{r})) - k^2 \vec{E}_0^{(+)}(\vec{r}) = 0} \quad [2.2]$$

This equation is called the **homogeneous Helmholtz equation**, and it constitutes the basic equation for studying monochromatic optical beams propagating in vacuum.

♦ *Comment*

Light rays and the eikonal equation

For arbitrary beams, the complex amplitude can be written as

$$E_o^{(+)}(\vec{r}) = E_o(\vec{r})e^{ik_o S(\vec{r})}$$

where $S(\vec{r})$ corresponds to the optical path length. The surfaces with constant $S(\vec{r})$ correspond to the beam wavefronts. We can obtain an equation for the evolution of the optical



wavefronts by inserting this expression of the fields into the Helmholtz equation or by directly using Maxwell's equations.

In directly applying this formalism for a medium of varying refractive index, $n(r)$, we arrive at an equation for $S(r)$. In the limiting case of $\lambda \rightarrow 0$ (i.e., $k_0 \rightarrow \infty$), this is called the eikonal equation:

$$\left[\nabla \bar{S}(\vec{r}) \right]^2 = n^2(\vec{r})$$

This is the basic equation for geometrical optics, and it is fundamental to the concept of optical rays, which by definition have a trajectory that is always perpendicular to wavefronts. These wavefronts are surfaces of equal optical path length.

The Helmholtz equation is also the starting point for developing a theory of diffraction of optical beams, as we will see in Chapter 6.

In this chapter, we will focus our attention on studying the propagation properties of highly directional beams, such as those produced by lasers, since this kind of radiation is highly relevant to the field of photonics.

2.2 Paraxial wave equation for optical beams

The general problem of beam propagation out of a region containing optical sources can be solved using the Helmholtz equation. Suppose we know the solution of a beam, generated by an optical source, in a region of space out of the source. If we are interested in studying the beam only in the vicinity of a specific direction, we could seek simpler solutions that adequately approximate the real solution within the region of interest. When a region of interest is close to a specified propagation direction, we call it the **paraxial region**, and the expressions describing the optical beam within this region are the **paraxial solutions**.

Let us consider that our direction of interest is the z -axis. A plane wave solution has an electric field transverse to the propagation direction, as described in Chapter 1. For a paraxial beam, the electric field will be almost perpendicular to the z -axis and, thus, the wave vector can be approximated by its z component. Under the **paraxial wave approximation**, we assume the following:

$$\begin{aligned} \text{Paraxial Wave approximation: } (k_x^2 + k_y^2 \ll k_z^2) \\ k = \sqrt{k_x^2 + k_y^2 + k_z^2} \approx k_z \left(1 + \frac{k_x^2 + k_y^2}{2k_z^2} \right) \approx k_z \rightarrow \vec{k} \approx k_z \hat{k} \end{aligned}$$

The expression of a paraxial beam is written as a plane wave propagating along a z -axis, with complex amplitude that is modulated in space. This kind of solution is also called a quasi-plane-wave solution.



The mathematical expression for a paraxial beam is

$$\vec{E}_p^{(+)}(\vec{r}, t) = \hat{e} E_0^{(+)}(x, y, z) e^{i(kz - \omega t)} \quad [2.3]$$

These beams constitute the solution to the beam propagation problem in regions close to the z -axis. They are also used to describe highly localized solutions along this axis, such as for a laser beam.

By considering that the divergence of the electric field is zero, we can write the Helmholtz equation as

$$\nabla^2 \vec{E}_p^{(+)} + k^2 \vec{E}_p^{(+)} = 0 \quad \rightarrow \quad \nabla^2 (E_0^{(+)}(x, y, z) e^{i(kz - \omega t)}) + k^2 E_0^{(+)}(x, y, z) e^{i(kz - \omega t)} = 0$$

$$\frac{\partial^2}{\partial z^2} (E_0^{(+)}(x, y, z) e^{i(kz - \omega t)}) + \nabla_{\perp}^2 E_0^{(+)}(x, y, z) e^{i(kz - \omega t)} + k^2 E_0^{(+)}(x, y, z) e^{i(kz - \omega t)} = 0$$

where the transverse Laplacian is $\nabla_{\perp}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$.

By performing the derivatives on z , we can easily obtain an equation for the complex amplitude

$$\frac{\partial^2 E_0^{(+)}}{\partial z^2} + 2ik \frac{\partial E_0^{(+)}}{\partial z} + \nabla_{\perp}^2 E_0^{(+)} = 0 \quad \rightarrow \quad \frac{\partial}{\partial z} \left(\frac{\partial E_0^{(+)}}{\partial z} + 2ik E_0^{(+)} \right) + \nabla_{\perp}^2 E_0^{(+)} = 0 \quad [2.4A]$$

(the explicit dependence of the complex amplitude on (x, y, z) will be omitted in the following in order to simplify the notation).

In many situations, this equation can be simplified further by considering the **slowly varying envelope approximation (SVEA)**, in which the beam's amplitude variations are considered to be much smaller than the amplitude over distances that are of the order of the wavelength.

$$\begin{aligned} \text{SVEA} \Rightarrow \frac{\Delta E_0^{(+)}}{E_0^{(+)}} \ll 1 \quad \text{for } \Delta z \approx \lambda &\rightarrow \frac{\frac{\partial E_0^{(+)}}{\partial z} \Delta z}{E_0^{(+)}} \ll 1 \quad \text{for } \Delta z \approx \lambda \\ \frac{\partial E_0^{(+)}}{\partial z} \lambda \ll E_0^{(+)} &\rightarrow \boxed{\frac{\partial E_0^{(+)}}{\partial z} \ll \frac{E_0^{(+)}}{\lambda} < 2k E_0^{(+)}} \end{aligned}$$

Within this approximation, the wave equation can be simplified to obtain the **paraxial wave equation (PWE)** for optical beams:

$$\boxed{\frac{\partial E_0^{(+)}}{\partial z} - \frac{i}{2k} \nabla_{\perp}^2 E_0^{(+)} = 0} \quad [2.4B]$$

This is the basic equation for studying the paraxial solutions to the wave equation, which describe the properties of optical beams in the paraxial region.

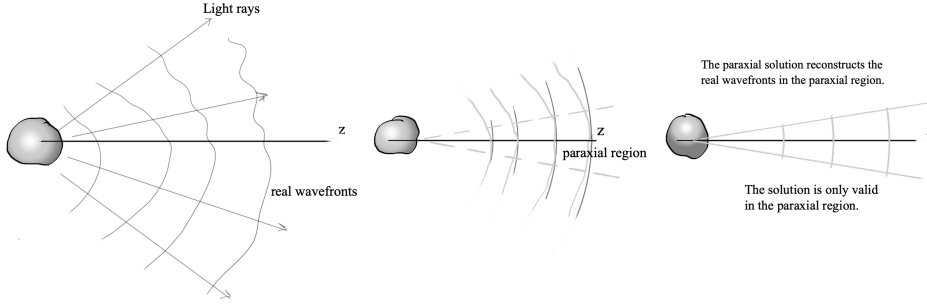


Figure 2.1
The concept of
paraxial solutions
and
paraxial regions

2.3 Basic solutions to the paraxial wave equation

We will describe in this section two important solutions to the PWE used in photonics. The first is the Fresnel approximation of the spherical wave, which is fundamental for developing the scalar theory of diffraction in the Fresnel limit. The second is the Gaussian beam, which accounts for the basic emission characteristics of most lasers.

2.3.1 Fresnel's approximation of spherical waves

Consider the expression of a scalar spherical wave, as defined in Chapter 1:

$$E^{(+)} = \frac{E_0^{(+)}}{r} e^{i(kr - \omega t)} \quad \text{with} \quad r^2 = x^2 + y^2 + z^2 = \rho^2 + z^2$$

where the transverse radius $\rho^2 = x^2 + y^2$ is defined.

This solution describes light beams with spherical wavefronts propagating away from a focus located at the origin. We can obtain a paraxial solution representing the spherical wavefronts close to the z -axis by means of the following approximations:

$$r^2 = z^2 \left(1 + \frac{\rho^2}{z^2} \right) \rightarrow r = z \sqrt{1 + \frac{\rho^2}{z^2}}$$

$$\text{In the paraxial region: } \rho \ll z \rightarrow r \approx z \left(1 + \frac{\rho^2}{2z^2} + \dots \right) \approx z + \frac{\rho^2}{2z} + \dots$$

We substitute this approximated value of r in the expression of the spherical wave while retaining the first-order term when expanding the denominator but keeping the first two terms in the phase term. This is necessary because the phase term is multiplied by k , which is a large number and thus the phase introduced by the second term could be non-negligible.

$$E_p^{(+)} = \frac{E_0^{(+)}}{z} e^{i(kz + \frac{k\rho^2}{2z} - \omega t)} = \frac{E_0^{(+)}}{z} e^{i\frac{k\rho^2}{2z}} e^{i(kz - \omega t)}$$

Comparing this expression with Equation [2.3], we define the Fresnel approximation solution for the complex amplitude as

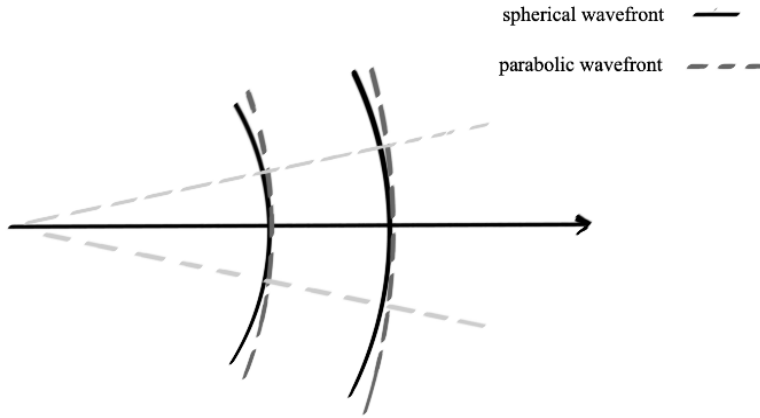


$$E_0^{(+)}(\rho, z) = \frac{E_0^{(+)}}{z} e^{i \frac{k \rho^2}{2z}} \quad [2.5A]$$

where $E_0^{(+)}$ is a constant. This solution represents a beam with a parabolic wavefront diverging from a focus located to the left at a distance z . As shown in the figure, this solution is a good approximation for spherical wavefronts close to the z -axis.

(!) Remember that these solutions can be used to represent the real spherical beam only in the region where the approximation is valid.

Figure 2.2
Fresnel or parabolic solutions to the PWE are useful approximations of spherical waves in the paraxial region



An equivalent solution representing a parabolic wavefront converging to a focus located to the right at a distance z can be expressed as

$$E_0^{(+)}(\rho, z) = \frac{E_0^{(+)}}{z} e^{-i \frac{k \rho^2}{2z}} \quad [2.5B]$$

These solutions are used in the following distinct situations:

- 1) A beam transmitted by a lens in the paraxial region introduces a phase retardation that is described by such solutions.
- 2) In diffraction theory, any beam is studied as a superposition of these solutions in the Fresnel region, which is called a wavelet decomposition.

2.3.2 The Gaussian beam

Another important solution to the PWE can be obtained as a shifted version of the Fresnel approximation of the spherical wave, which was obtained previously. It is immediately clear from the PWE [2.4B] that if $E_0^{(+)}(x, y, z)$ is a solution, then shifting this solution in z by a constant will give another valid solution $E_0^{(+)}(x, y, z+z_0)$. In particular, we construct a new solution by directly substituting the propagation variable z in Equation [2.5A] with



the shifted complex variable $q(z) = z - iz_R$ ($z_0 = -iz_R$). The meaning of the constant z_R , called the **Rayleigh range**, will be defined later in this section.

The obtained solution is called the **Gaussian beam**, and it can be expressed as

$$E_0^{(+)}(\rho, z) = \frac{E_0^{(+)}}{q(z)} e^{i \frac{k\rho^2}{2q(z)}} \quad [2.6]$$

The presence of the imaginary term $q(z)$ results in the solution behaving completely differently during propagation, as we will see. We can check that this is indeed a solution of the PWE by directly substituting [2.6] into the PWE [2.4B].

In order to more simply describe the properties of the Gaussian beam, we will rewrite expression [2.6] in the form $E_0^{(+)}(\rho, z) = |E_0^{(+)}(\rho, z)| e^{i\phi(\rho, z)}$

$$\begin{aligned} E_0^{(+)}(\rho, z) &= \frac{E_0^{(+)}}{z - iz_R} e^{i \frac{k\rho^2}{2(z - iz_R)}} = \frac{E_0^{(+)}(z + iz_R)}{z^2 + z_R^2} e^{i \frac{k\rho^2(z + iz_R)}{2(z^2 + z_R^2)}} = \\ &= \frac{iE_0^{(+)}(z_R - iz)}{z^2 + z_R^2} e^{-\frac{k\rho^2 z_R}{2(z^2 + z_R^2)}} \cdot e^{i \frac{k\rho^2 z}{2(z^2 + z_R^2)}} \end{aligned}$$

Next, we write the term $(z_R - iz)$ in terms of its modulus and phase, then use the identity $k = 2\pi/\lambda$ to obtain

$$E_0^{(+)}(\rho, z) = \frac{iE_0^{(+)} \sqrt{z_R^2 + z^2} e^{-iatg(z/z_R)}}{(z^2 + z_R^2)} \cdot e^{-\frac{\pi z_R \rho^2}{\lambda z_R^2 (1 + \frac{z^2}{z_R^2})}} \cdot e^{i \frac{k\rho^2 z}{2z^2 (1 + \frac{z^2}{z_R^2})}}$$

This expression can be simplified to

$$E_0^{(+)}(\rho, z) = \frac{iE_0^{(+)}}{\sqrt{z^2 + z_R^2}} e^{-\frac{\lambda z_R}{\pi} \frac{\rho^2}{(1 + \frac{z^2}{z_R^2})}} \cdot e^{i \frac{k\rho^2}{2z(1 + \frac{z^2}{z_R^2})}} \cdot e^{-iatg(z/z_R)}$$

In order to further simplify this expression, we define the following variables:

$$\text{Gaussian beam radius: } w(z) = \sqrt{\frac{\lambda z_R}{\pi} \left(1 + \frac{z^2}{z_R^2}\right)}$$

From this definition we obtain the beam waist radius as the beam radius at $z = 0$:

$$w_0 = w(0) = \sqrt{\frac{\lambda z_R}{\pi}}$$



The Rayleigh range, z_R , is then obtained as a function of the beam waist: $z_R = \frac{\pi w_0^2}{\lambda}$

Wavefront radius of curvature: $R(z) = z(1 + \frac{z_R^2}{z^2})$

Gouy phase shift: $\xi(z) = \text{atg}(z/z_R)$

With these definitions, the Gaussian beam expression is

$$E_0^{(+)}(\rho, z) = \frac{iE_0^{(+)}}{z_R \sqrt{1 + \frac{z^2}{z_R^2}}} e^{-\frac{\rho^2}{w^2(z)}} e^{i\left(\frac{k\rho^2}{2R(z)} - \xi(z)\right)}$$

Finally, redefining a new complex amplitude as $U_0^{(+)} = iE_0^{(+)}/z_R$, we obtain the final expression for the **Gaussian beam** and its related magnitudes:

q parameter $q(z) = z - iz_R$

Beam waist radius w_0

Rayleigh range $z_R = \frac{\pi w_0^2}{\lambda}$

Beam radius: $w(z) = w_0 \sqrt{1 + \frac{z^2}{z_R^2}}$

Wavefront radius of curvature $R(z) = z(1 + \frac{z_R^2}{z^2})$

Gouy phase shift $\xi(z) = \text{atg}(z/z_R)$

$$E_0^{(+)}(\rho, z) = \frac{U_0^{(+)} w_0}{w(z)} e^{-\frac{\rho^2}{w^2(z)}} \cdot e^{i\left(\frac{k\rho^2}{2R(z)} - \xi(z)\right)}$$

Gaussian beam [2.7]



2.4 Basic properties of Gaussian beams

2.4.1 Amplitude and intensity

Amplitude

The Gaussian beam solution [2.7] corresponds to propagation along the z -axis. **The position of the plane $z = 0$ is not arbitrary but is instead located at the plane where the beam radius takes its minimum value, w_0 .** This value, together with the wavelength, allow us to define the two most relevant parameters: the Rayleigh range, z_R ; and the q parameter, $q(z) = z - iz_R$.

At the plane $z = 0$, the beam transverse profile in the (x, y) plane has a Gaussian shape, with amplitude equal to $|U_0^{(+)}|$. The radius of the beam, also called the $1/e$ in amplitude radius, is defined as the value ρ , for which the field amplitude is equal to $|E_0^{(+)}| = |U_0^{(+)}|/e$. At the plane $z = 0$, this corresponds to $\rho = w_0$. These properties are represented in Figure 3.

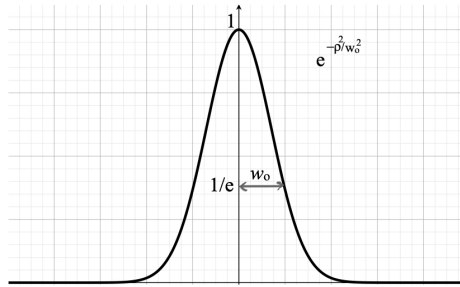


Figure 2.3
Gaussian dependence of the field amplitude and definition of beam radius

At any other plane z , the beam maintains the same gaussian profile in the transverse plane, but with an increased value of the radius and with decreased amplitude. Thus, the Gaussian beam propagates in space while maintaining a constant shape, but with an increasing radius during propagation. The transverse spatial broadening of optical beams is an intrinsic characteristic of any transversally localized light distribution, and it is linked to the physical process called **diffraction**.

The change in beam radius during propagation, $w(z)$, corresponds to the expression defined previously when deriving Equation [2.7]. Figure 4 plots this dependence as a function of z . From this definition, we can interpret the meaning of the Rayleigh range as the distance travelled by the beam, after which its radius increases by a factor of $\sqrt{2}$:

$$w(z_R) = w_0 \sqrt{2}.$$

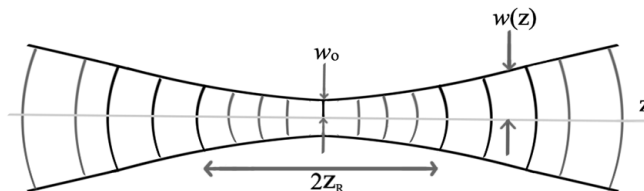


Figure 2.4
Dependence of the beam radius of a Gaussian beam, with distance



The **confocal parameter, b** , of a Gaussian beam is defined as the distance between the planes $z = \pm z_R$, and it is the length along the z -axis that the beam propagates while highly localized in the transverse plane.

$$\text{Confocal parameter} \rightarrow b = 2z_R = \frac{2\pi w_0^2}{\lambda}$$

As the beam propagates farther from its beam waist, the dependence of its radius on z becomes approximately linear:

$$\text{At } z \gg z_R \quad w(z) = w_0 \sqrt{1 + \frac{z^2}{z_R^2}} \approx w_0 \frac{z}{z_R} = \frac{\lambda}{\pi w_0} z$$

In this region of linear transverse expansion, we define the **divergence angle θ** . One should remember at this point that we are studying beam solutions to the paraxial wave equation, for which we define optical beams in the region around the propagation axis. At this limit, the angles should not be very large, and we can define this angle as

$$\theta \approx \frac{w(z)}{z} = \frac{w_0}{z_R} = \frac{\lambda}{\pi w_0} \quad [2.8]$$

From these definitions, one can immediately infer a basic property of Gaussian beam propagation: **the smaller the beam waist radius, w_0 , the shorter the confocal beam parameter and the larger the diverging angle.**

Some typical values are shown in the table below.

$\lambda = 500 \text{ nm}$			
w_0	z_R	b	θ
1 mm	6.3 m	12.6 m	0.16 mrad
100 μm	6.3 cm	12.6 cm	1.6 mrad
5 μm	157 μm	314 μm	32 mrad

As we transversally localize the optical beam, it will diffract more strongly. This has a profound impact on the field of photonics, because making devices with smaller and more compact dimensions requires that we overcome diffraction in order to propagate the beams in a tightly localized fashion along the device.

Different solutions involving waveguides and photonic crystals have been proposed in the last decade, giving rise to the birth of new lines of research in the study of photonics.

Intensity

The expression of the intensity of the beam depends on the specific conditions of our problem, since the Gaussian beam solution could represent a component of the electric field, the magnetic field, or be expressed in terms of the scalar wave function. For the case of the scalar wave function, we obtain:



$$I(\rho, z) = |U_0^{(+)}(\rho, z)|^2 = \frac{|U_0^{(+)}|^2 w_0^2}{w^2(z)} e^{-\frac{2\rho^2}{w^2(z)}} \quad [2.9]$$

$$I(\rho, z) = I_0(z) e^{-\frac{2\rho^2}{w^2(z)}}$$

where we define the Gaussian beam's **peak intensity** as the value of the intensity at $\rho = 0$

$$I_0(z) = \frac{|U_0^{(+)}|^2 w_0^2}{w^2(z)} = \frac{I_0 w_0^2}{w^2(z)} = \frac{|U_0^{(+)}|^2}{1 + \frac{z^2}{z_R^2}}$$

The units of $|U_0^{(+)}|$ are $\sqrt{W/m^2}$ in this case.

When we plot the transverse dependence of the intensity, we see that it has a gaussian shape with a $1/e$ radius in intensity, $w^I(z)$, given by

$$w^I(z) = \frac{w(z)}{\sqrt{2}} \quad \Rightarrow \quad I(w^I(z), z) = I_0(z) / e$$

At a transverse distance equal to the Gaussian beam radius $\rho = w(z)$, the intensity takes the value

$$I(w(z), z) = I_0(z) / e^2$$

The dependence of the beam intensity on propagation distance z shows that the intensity is maximum at the plane $z = 0$ and decreases when we move away from this plane. At a distance corresponding to the Rayleigh range, the peak intensity is half the value at the beam waist plane:

$$I_0(\pm z_R) = I_0 / 2$$

Below, we summarize the properties of the amplitude and intensity of Gaussian beams, based on our previous considerations.

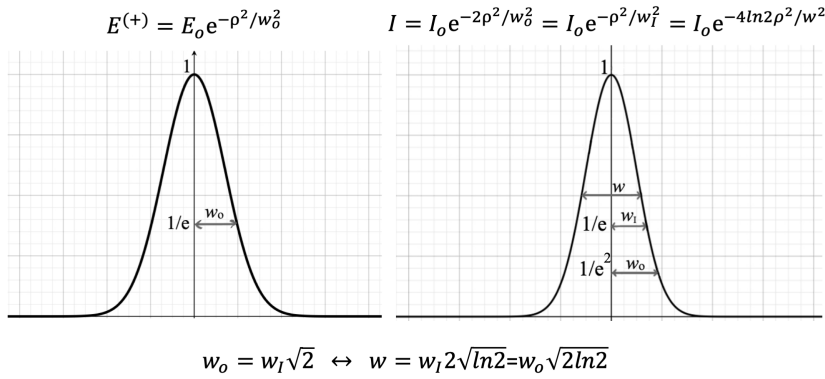
- The transverse profile of Gaussian beams maintains its gaussian shape during propagation.
- The minimum radius, called the beam waist (w_0), and the maximum peak intensity, I_0 , occur at the plane $z = 0$.
- At the particular distance called the Rayleigh range, z_R , $w(z_R) = \sqrt{2} \cdot w_0$ and $I_0(z_R) = I_0 / 2$.
- Far from the confocal region (defined by the confocal parameter $b = 2z_R$), the beam diverges linearly with a divergence angle $\theta = \lambda / (\pi w_0)$.



♦ **Comment**

The definition of the beam radius is in some ways arbitrary, since the mathematical gaussian solution extends to infinity in the transverse direction. In the theoretical calculations, the most widely used definition is the one we have given that corresponds to $1/e$ in amplitude (w_o). The definition $1/e$ in intensity (w^I) is also used and was derived previously. However, the experimental results often use a beam diameter called the full width half maximum in intensity, w . Due to the arbitrary nature of this choice, the Gaussian beam expressions found in the literature are not unique. In order to interpret the obtained results in the correct way, it is essential to pay attention to the convention used in each specific case. The relation between the different definitions can easily be obtained, and the most frequently used ones are the following.

Conventions in beam radius definitions



2.4.2 Power

The energy of the Gaussian beam is distributed in the transverse plane. The total power contained in the beam is calculated by integrating the beam intensity over its transverse area. Since the Gaussian beam solution has rotational symmetry along the z -axis, we obtain

$$P_{TOT} = \iint I(\rho, z) dS = \int_0^{2\pi} \int_0^\infty I(z) \rho d\phi d\rho = \int_0^\infty I(z) 2\pi \rho d\rho$$

After substituting the expression for $I(z)$, we have

$$P_{TOT} = \int_0^\infty \frac{I_o w_o^2}{w^2(z)} e^{-\frac{2\rho^2}{w^2(z)}} 2\pi \rho d\rho = -\frac{I_o w_o^2 \pi}{2} \int_0^\infty -\frac{4\rho}{w^2(z)} e^{-\frac{2\rho^2}{w^2(z)}} d\rho = \frac{\pi w_o^2}{2} I_o$$



This expression provides us with a specific way to obtain the maximum peak intensity of the Gaussian beam (I_0) from directly measurable physical quantities: the total beam power (which can be measured using power meters) and the beam waist radius (which can be measured in the way explained in problem 2.4 or by using a beam profiler):

$$I_0 = \frac{2P_{TOT}}{\pi w_0^2} \quad [2.10]$$

The relation we have just derived is obtained through integration in the transverse plane, going from zero to infinity. This is clearly an unphysical situation, and it describes the fact that the mathematical gaussian function has tails expanding out to infinity. Any realistic beam will be truncated by a physical aperture with a finite radius, so we should consider how the power is distributed with the transverse radius, ρ . This can be readily obtained by considering the integration of the intensity over a finite area in the transverse plane:

$$P(\rho, z) = \int_0^\rho I(z) 2\pi \rho d\rho = \int_0^\rho \frac{I_0 w_0^2}{w^2(z)} e^{-\frac{2\rho^2}{w^2(z)}} 2\pi \rho d\rho = \frac{\pi w_0^2}{2} I_0 \left[1 - e^{-\frac{2\rho^2}{w^2(z)}} \right]$$

$$P(\rho, z) = P_{TOT} \left[1 - e^{-\frac{2\rho^2}{w^2(z)}} \right]$$

From this expression, we can explore the value of the transmitted power across two particular apertures:

$$\begin{aligned} \rho = w(z) & \quad P(w(z), z) = P_{TOT} [1 - e^{-2}] = 0.86 P_{TOT} \\ \rho = 2w(z) & \quad P(2w(z), z) = P_{TOT} [1 - e^{-8}] = 0.999 P_{TOT} \end{aligned}$$

An aperture of radius equal to $2w(z)$ used in plane z will transmit 99.9% of the total beam power. Smaller apertures could be used to transmit most of the beam power. However, due to other considerations related to beam distortion after transmission through finite apertures, an optimum value must be imposed for the apertures to be used when working with Gaussian beams. A general rule of thumb is: any aperture with a radius smaller than $2w(z)$ will produce distortions in the beam profile and should not be used in an experimental configuration.

2.4.3 Phase

The total expression of the Gaussian beam solution reads:

$$U(\rho, z, t) = U_0^{(+)}(\rho, z) e^{i(kz - \omega t)} = |U_0^{(+)}| e^{i\phi(\rho, z)} e^{i(kz - \omega t)}$$

$$\text{with } \phi(\rho, z) = \frac{k\rho^2}{2R(z)} - \text{atg}(z/z_R)$$

And $U_0^{(+)}$ corresponds to the Gaussian solution of the PWE.



The first term describes the parabolic wavefronts with radius of curvature $R(z) = z(1 + z_R^2/z^2)$.

For positive z , they represent the wavefronts diverging from a Gaussian source whose beam waist is placed to the left at the distance z .

For negative z , this same solution represents the wavefronts converging towards a beam waist located at a distance z to the right of the actual plane.

For $z = 0$, we obtain $R(0) = \infty$. This means that the wavefronts are flat, as what happens in a plane wave. Thus, when the Gaussian beam is close to the beam waist, it resembles a plane wave but with a finite transverse dimension.

At large distances, $z \gg z_R$, the radius of curvature increases proportional to z becoming close to a spherical wavefront in the paraxial region.

At $z = z_R$, the radius of curvature takes its minimum value $R(z_R) = 2z_R$, as one can easily check. At this point, we thus obtain the maximum curvature of the wavefronts.

The second term, called the **Gouy phase**, changes from the value $-\pi/2$ for $z \rightarrow -\infty$ to the value $\pi/2$ for $z \rightarrow \infty$. There is a total shift of π when passing across the focused spot of a Gaussian beam.

This phase shift represents a displacement of the wavefronts of the Gaussian beam with respect to the phase fronts of a spherical wave (or plane wave) located at the beam waist. It is important to account for this phase factor in some applications, such as nonlinear conversion using focused beams. However, this course will not consider these special situations.

2.4.4 M^2 factor

The beam waist radius and the divergence angle are closely related, since we obtain a more divergent beam for a smaller beam waist (w_0). From the previously derived equations, we can easily check that the following relation holds for Gaussian beams:

$$\theta \cdot w_0 = \frac{\lambda}{\pi} \quad [2.11]$$

We define the product of the beam waist diameter (minimum diameter of the GB) times the total divergence angle of the beam in the far field:

$$2w_0 \cdot 2\theta = \frac{4\lambda}{\pi}$$

When we generate other kinds of transverse solutions to the PWE, we always observe similar behaviour, which is dictated by diffraction. Any real beam generated from a real laser will not be an exact Gaussian beam, which is an ideal mathematical solution of the PWE. We can always measure or calculate the corresponding product of minimum diameter ($2W$) times the total divergence angle (2Θ) for any real beam, then compare it to that of an ideal Gaussian beam. This ratio defines the M^2 factor of the beam:



$$M^2 = \frac{2\Theta \cdot 2W}{2\Theta \cdot 2w_0} = \frac{\pi}{4\lambda} [2\Theta \cdot 2W]$$

It can be demonstrated that the Gaussian beam profile gives the smallest value of the product (beam diameter)·(total divergence). Consequently, for any real beam, we find that $M^2 > 1$. Values of M^2 can be as low as 1.2–1.3 for gas lasers, such as He–Ne laser, reach as high as $M^2 > 100$ for some semiconductor lasers.

The way to obtain the measurement of the beam waist radius and the divergence angle of a real beam is far from evident in most situations. The most used techniques in photonics are beam profilers and the knife-edge technique (see the Proposed problems section for a brief outline of this method).

2.4.5 q parameter

As we have seen, the Gaussian beam solution is characterized by the q parameter. Knowledge of this parameter allows us to completely establish the evolution of the beam in the region of interest. The q parameter is a complex number that can be specified anywhere on plane z along the propagation direction of the Gaussian beam. The imaginary part is the minus Rayleigh range while the real part corresponds to the value of the distance of the actual plane to the beam waist radius position. Thus, a positive real value implies that we are in a plane located to the right of the beam waist plane and negative real value tell us we are at the left of the beam waist plane.

As an example, consider a Gaussian beam with Rayleigh range $z_R = 1$ m.

The q parameter at the $z = 0$ plane, where the beam waist is located is $q(0) = -i$.

The q parameter in the plane located at a distance $z = 10$ meters to the right of this plane is $q(10) = 10-i$. If we have $q = -2-i$, this means that we are in the plane $z = -2$ m, located to the left of the beam waist position.

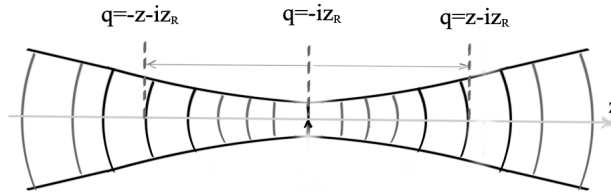


Figure 2.5
 q parameter expression at different planes

Sometimes we can also use the inverse of the q parameter, which can also be related to the wavefront radius of curvature $R(z)$ and to the beam radius $w(z)$. To see how this works, let us assume we are observing our Gaussian beam in plane z with

$q(z) = z - iz_R$. Then:

$$\frac{1}{q(z)} = \frac{1}{z - iz_R} = \frac{z + iz_R}{z^2 + z_R^2} = \frac{z}{z^2 + z_R^2} + i \frac{z_R}{z^2 + z_R^2}$$



This expression can be rewritten as

$$\frac{1}{q(z)} = \frac{1}{z \left(1 + \frac{z_R^2}{z^2}\right)} + i \frac{1}{z_R \left(1 + \frac{z^2}{z_R^2}\right)} = \frac{1}{z \left(1 + \frac{z_R^2}{z^2}\right)} + i \frac{w_0^2}{z_R w^2(z)}$$

Using the definitions for Equation [2.7], we have

$$\boxed{\frac{1}{q(z)} = \frac{1}{R(z)} + i \frac{\lambda}{\pi w^2(z)}} \quad [2.12]$$

Thus, the inverse of the q parameter gives us direct information for the beam parameters R and w in the desired plane.

-Basic Parameters defining a Gaussian beam-

- 1) Direction of propagation, defining the z -axis
 - 2) Position of its beam waist (minimum diameter) along the z -axis, locates the plane $z = 0$
 - 3) Beam waist radius w_0
 - 4) Wavelength
 - 5) Total power
-

2.5 The effect of a lens on a Gaussian beam

Every photonics optical system uses lenses and mirrors to control the propagation properties of the beam (i.e., its diameter and divergence) in order to satisfy the required specifications. Sometimes we need to focus the beam tightly to obtain high intensities in the region of interest; other times we need a beam propagating over large distance with no significant divergence. To understand how a Gaussian beam can be tailored according to our needs, we will study in this section a lens's effect on a Gaussian beam.

The first concept we should keep in mind is: **in a paraxial regime, a Gaussian beam remains Gaussian after passing through a lens. The lens will modify the beam waist radius, the beam waist position plane and the beam divergence.**

For each region between consecutive lenses, we can write the expression of the Gaussian beam in the same way as that derived in Equation [2.7], where the plane $z = 0$ corresponds to the plane where the beam waist is located, as shown in figure 5.

When a field distribution is incident on a lens, the effect is always that a phase factor and loss of intensity will be simultaneously introduced. The transmitted complex field amplitude is related to the incident field through a complex transmittance function:

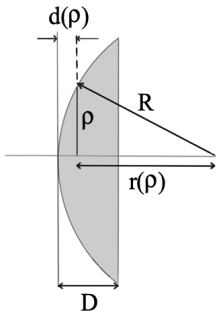


$$E_{out}^{(+)}(\rho) = t(\rho)E_{in}^{(+)}(\rho) = |t(\rho)|e^{i\phi(\rho)}E_{in}^{(+)}(\rho)$$

For our study, we can assume for simplicity's sake that the modulus of the transmittance function is 1. This implies that no light reflects from the lens surface (by means of, say, antireflective coating) and that the glass absorbs no light (i.e., a completely transparent lens at the wavelengths of interest). The phase factor of the transmittance function at any distance from the axis can be obtained as:

$$\phi(\rho) = kL(\rho) = \frac{2\pi}{\lambda}L(\rho)$$

where $L(\rho)$ is the optical path length at a height ρ from the axis.



In order to obtain an expression for the phase factor, we will assume we have a plano-convex lens of thickness D , as shown in the figure (this assumption simplifies the derivation, although the particular shape of the lens does not affect the final result as long as we are in the paraxial region).

The optical path can be written in this case as

$$L(\rho) = d(\rho) \cdot 1 + [D - d(\rho)] \cdot n = nD - d(\rho)[n - 1]$$

where $d(\rho)$ is the thickness of air at height ρ , and n is the index of refraction of the lens at the particular wavelength of interest. Evaluating the function $d(\rho)$ is straightforward from the figure:

$$d(\rho) = R - r(\rho) = R - \sqrt{R^2 - \rho^2}$$

We can simplify this expression further by assuming that we are in the paraxial region (where the Gaussian beam solution of the PWE is applicable). In this region, the condition $\rho \ll R$ holds. Therefore:

$$d(\rho) = R - R\sqrt{1 - \frac{\rho^2}{R^2}} \approx R - R\left(1 - \frac{\rho^2}{2R^2} + \dots\right) \approx \frac{\rho^2}{2R}$$

Substituting this result into the expression for the optical path gives

$$L(\rho) = nD - \frac{\rho^2(n-1)}{2R}$$

and

$$\phi(\rho) = knD - \frac{k\rho^2(n-1)}{2R} = knD - \frac{k\rho^2}{2f}$$

where we use the definition of the focal length of a thin lens, f , which is defined in geometrical optics as



$$\frac{1}{f} = (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$

Plano-convex lens: $R_1 = R$; $R_2 = \infty$ and $f = \frac{R}{(n-1)}$.

(A derivation of this result can be found in any general optics book, for instance Hecht, E. (2016). *Optics*. Pearson).

The first term introduces a constant phase factor with no relevant significance, which can be omitted at the limit of a thin lens when $D \rightarrow 0$, thus giving the final result:

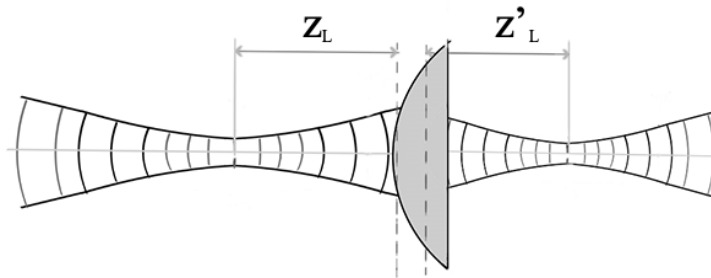
$$\boxed{\phi(\rho) = -\frac{k\rho^2}{2f}} \quad [2.13]$$

To understand the significance of this term, consider a plane wave incident on the lens with $E_0^{(+)} = 1$. Then the field transmitted by the lens will be

$$E_{out}^{(+)} = e^{-i\frac{k\rho^2}{2f}}$$

which, as discussed in Section 2.4.4, corresponds to a parabolic wavefront converging to a point located at distance $z = f$. This is precisely the effect of any lens acting on a plane wave: it focuses light on a spot in the focal plane. In the same way, the action of the lens over any field that is incident on it can be interpreted as a modification of the incident wavefront, which is done to make it more convergent (or divergent, if the focal length is negative).

Once we have derived the main result of this section, let us see how this affects the properties of the Gaussian beam incident on the lens. Consider a Gaussian beam propagating along the z -axis with a q parameter defining the beam in this region $q(z) = z - iz_R$. After setting the lens of focal length f , the new Gaussian beam will possess a new q parameter $q'(z') = z' - iz'_R$, where the variable z' is now measured from the position of the beam waist of the Gaussian beam after the lens.



If the lens is placed at a distance z_L from the beam waist, the field that is incident on the lens can be written as



$$E_{in}^{(+)}(z_L, \rho) = \frac{E_o^{(+)} w_0}{w(z_L)} e^{-\frac{\rho^2}{w^2(z_L)}} e^{i \frac{k \rho^2}{2R(z_L)}} e^{-i \xi(z_L)} \rightarrow$$

$$I = |E_{in}^{(+)}|^2 = \frac{I_0 w_0^2}{w^2(z_L)} e^{-\frac{2\rho^2}{w^2(z_L)}} = \frac{2P_{TOT}}{\pi w^2(z_L)} e^{-\frac{2\rho^2}{w^2(z_L)}}$$

where the intensity is expressed in terms of the total power of the beam. After the lens, the beam will be modified according to Equation [2.13]. For a thin lens, we additionally can assume that the lens will not modify the beam radius. These assumptions lead to the conditions:

$$E_{out}^{(+)}(\rho, z'_L) = e^{-i \frac{k \rho^2}{2f}} E_{in}^{(+)}(\rho, z_L) = \sqrt{\frac{2P_{TOT}}{\pi w^2(z_L)}} e^{-\frac{\rho^2}{w^2(z_L)}} e^{i \frac{k \rho^2}{2} \left(\frac{1}{R(z_L)} - \frac{1}{f} \right)} e^{-i \xi(z_L)}$$

$$w(z_L) = w(z'_L)$$

where $z'_L = -|z'_L|$ corresponds to the distance from the lens to the plane on which the new beam waist is located. From these assumptions, we can relate the fields at both sides of the lens. It holds true that:

1/ The intensity of the beam coming out of the lens will be the same as the intensity incident on the lens, since $w(z_L) = w(z'_L)$.

2/ We can define the wavefront radius of curvature after the lens as

$$\frac{1}{R(z'_L)} = \frac{1}{R(z_L)} - \frac{1}{f}$$

Using these relations together with Equation [2.12] allows formulating the lens's effect in terms of the q parameters:

$$\boxed{\frac{1}{q_{out}} = \frac{1}{q_{in}} - \frac{1}{f} \quad \Leftrightarrow \quad q_{out} = z'_L - i z'_R = \frac{q_{in} \cdot f}{f - q_{in}} = \frac{(z_L - i z_R) \cdot f}{f - z_L + i z_R}} \quad [2.14]$$

The Gaussian beam after the lens is written in terms of the propagation variable z' :

$$E^{(+)}(\rho, z') = \sqrt{\frac{2P_{TOT}}{\pi w^2(z')}} e^{-\frac{\rho^2}{w^2(z')}} e^{i \left(\frac{k \rho^2}{2R(z')} - \xi(z') \right)}$$

with

$$z'_R = \frac{\pi w_0^2}{\lambda} \quad w(z') = w'_0 \sqrt{1 + \frac{z'^2}{z_R'^2}} \quad R(z') = z' \left(1 + \frac{z_R'^2}{z'^2} \right)$$

**Example**

A lens of focal length f is placed at a distance $z_L = f$ from the plane of the beam waist of a Gaussian beam of wavelength λ with Rayleigh range z_R . Find the position of the new beam waist after the lens and the corresponding beam waist radius.

The q parameter incident on the lens is

$$q_{in} = z_L - iz_R = z_L - i \frac{\pi w_0^2}{\lambda}$$

After the lens, the new q parameter is given by Equation [2.14]:

$$q_{out} = \frac{z_L f - iz_R f}{f - z_L + iz_R} = \frac{[z_L f(f - z_L) - z_R^2 f] - i[z_L z_R f + z_R f(f - z_L)]}{(f - z_L)^2 + z_R^2}$$

$$q_{out} = z'_L - iz'_R = \frac{z_L f(f - z_L) - z_R^2 f}{(f - z_L)^2 + z_R^2} - i \frac{z_R f^2}{(f - z_L)^2 + z_R^2}$$

In this case, the lens is placed at a distance $z_L = f$ from the beam waist, and we obtain

$$q_{out} = z'_L - iz'_R = -f - i \frac{f^2}{z_R}$$

From this expression, we can infer directly that the new beam waist will be placed at a distance $|z'_L| = f$ to the right of the lens. The new beam waist radius is obtained directly from the expressions of the Rayleigh range parameter:

$$z'_R = \frac{f^2}{z_R} \Rightarrow \frac{\pi w_0'^2}{\lambda} = \frac{f^2 \lambda}{\pi w_0^2} \Rightarrow \boxed{z'_L = -f \quad w'_0 = \frac{f \lambda}{\pi w_0}}$$

(!) Note that this result is quite unexpected if we think in terms of geometrical optics, where an object placed at the focal plane of a lens is imaged at infinity. In the case of the Gaussian beam, we consider the beam waist position to be an object and obtain a new beam waist (i.e., an image) that appears at a distance equal to the focal length. In fact, the usual lens equation used in geometrical optics is not valid for Gaussian beams (see Solved problems for an alternative expression).

The difference lies essentially in how geometrical optics defines a light ray by presuming that it propagates as a plane wave. A ray bundle is treated as a set of plane waves propagating in slightly different directions.



2.6 General solutions to the paraxial wave equation

The two basic solutions described in the previous sections are not the only way to solve the PWE of interest. The Gaussian beam is the simplest of a complete set of solutions, which can be derived in systems with rectangular (Hermite–Gauss) or cylindrical (Laguerre–Gauss) symmetry. They are used extensively in laser optics and photonics.

Hermite–Gauss modes are obtained as solutions to the PWE with rectangular symmetries. These solutions can be found by starting with a separable ansatz of the form

$$u_{nl}^{HG}(x, y, z) = u_n^{HG}(x, z) \cdot u_m^{HG}(y, z)$$

where each of these functions satisfy the paraxial wave equation in one transverse dimension. A solution to this equation is written as a Gaussian beam modulated by a product of Hermite polynomials, H (n and m indicate the order of the corresponding Hermite polynomial):

$$u_{n,m}^{HG}(x, y, z) = \frac{A \cdot w_O}{w(z)} e^{\frac{-\rho^2}{w^2(z)}} e^{i \left(\frac{k\rho^2}{2R(z)} - (n+m+1)\xi(z) \right)} \cdot H_n \left(\frac{\sqrt{2}x}{w(z)} \right) \cdot H_m \left(\frac{\sqrt{2}y}{w(z)} \right)$$

Figure 6 shows some examples corresponding to this type of solution.

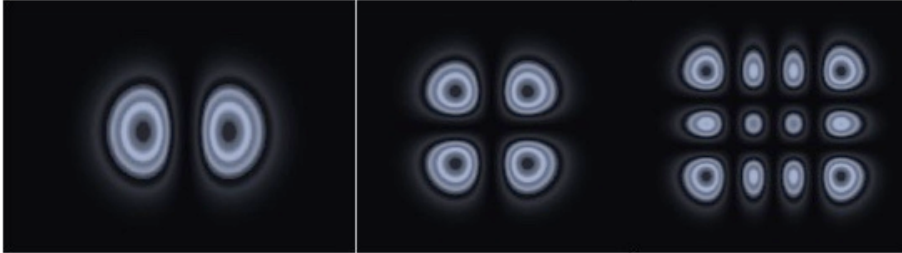


Figure 2.6

Hermite–Gaussian beams:
 $u_{1,0}^{HG}$, $u_{1,1}^{HG}$ and $u_{2,3}^{HG}$

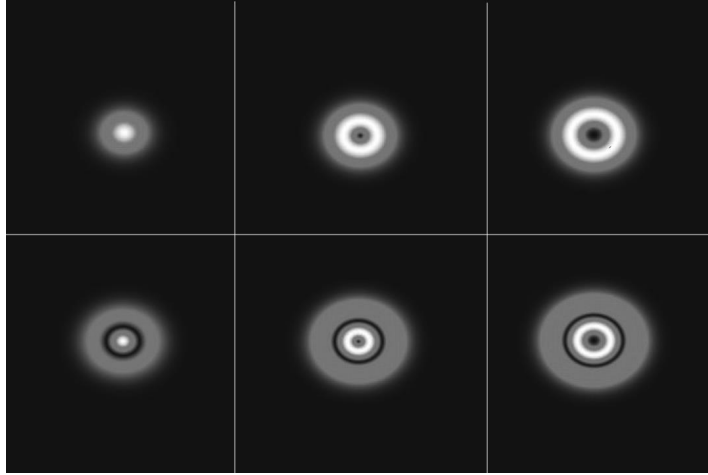
Another complete set of solutions are the Laguerre–Gaussian modes with rotational symmetry. These are solutions to the PWE with cylindrical coordinates:

$$u_{p,l}^{LG}(\rho, \phi, z) = \frac{A \cdot w_O}{w(z)} \cdot \left[\frac{\rho}{w(z)} \right]^{|l|} e^{\frac{-\rho^2}{w^2(z)}} e^{i \left(\frac{k\rho^2}{2R(z)} - (2p+|l|+1)\xi(z) \right)} \cdot e^{il\phi} \cdot L_p^{|l|} \left(\frac{2\rho^2}{w^2(z)} \right)$$

where $L_p^{|m|}$ are the generalised Laguerre polynomials. The profiles of these modes show concentric rings whose number is determined by the mode index p . The mode index $|m|$ is contained in the azimuthal phase term $\exp(im\phi)$. It is where helical wavefronts originate and it causes the orbital angular momentum associated with these beams (vortex beams). Some examples are shown in Figure 7.



Figure 2.7
Laguerre–Gaussian
beams.
Top:
 $u_{0,0}^{LG}, u_{0,1}^{LG}, u_{0,2}^{LG}$
Bottom:
 $u_{1,0}^{LG}, u_{1,1}^{LG}, u_{1,2}^{LG}$



Applying the theory of linear systems as a general approach allows constructing other paraxial solutions expressed in terms of the transfer function or impulse response function.

The simplest way to construct the general solution is to consider the Fourier transform of the transverse field distribution in a given plane. Let us consider a general beam profile $E^{(+)}(x, y, z)$ in plane z and assume that it has a 2D Fourier transform:

$$E^{(+)}(x, y, z) = \frac{1}{(2\pi)^2} \iint E^{(+)}(k_x, k_y, z) e^{i(k_x x + k_y y)} dk_x dk_y \quad \Leftrightarrow$$

$$E^{(+)}(k_x, k_y, z) = \iint E^{(+)}(x, y, z) e^{-i(k_x x + k_y y)} dx dy$$

where $E^{(+)}(k_x, k_y, z)$ is the 2D Fourier transform of the transverse field distribution in plane z .

The PWE [2.4B] applied to this generic field distribution can be solved more easily in the Fourier space, since the transverse Laplacian is easily calculated in this case, thus leading to an equation for the Fourier transform field:

$$\frac{\partial}{\partial z} \left[\frac{1}{(2\pi)^2} \iint E(k_x, k_y, z) e^{i(k_x x + k_y y)} dk_x dk_y \right] -$$

$$-\frac{i}{2k} \nabla_{\perp}^2 \left[\frac{1}{(2\pi)^2} \iint E(k_x, k_y, z) e^{i(k_x x + k_y y)} dk_x dk_y \right] = 0$$

$$\frac{1}{(2\pi)^2} \iint \left[\frac{\partial E^{(+)}(k_x, k_y, z)}{\partial z} + i \frac{(k_x^2 + k_y^2)}{2k} E^{(+)}(k_x, k_y, z) \right] \cdot e^{i(k_x x + k_y y)} dk_x dk_y = 0$$

Since this relation is satisfied for any arbitrary field distribution, the term in brackets must vanish, thereby leading to the equation



$$\frac{\partial E^{(+)}(k_x, k_y, z)}{\partial z} + i \frac{(k_x^2 + k_y^2)}{2k} E^{(+)}(k_x, k_y, z) = 0$$

which is easily solved, thus giving

$$E^{(+)}(k_x, k_y, z) = E^{(+)}(k_x, k_y, 0) e^{-i \frac{(k_x^2 + k_y^2)z}{2k}} \quad [2.15]$$

In linear system theory, the function relating the Fourier transform of the field in plane z to the Fourier transform of the field in plane $z = 0$ is generally denoted as the **transfer function** $H(k_x, k_y, z)$:

$$E^{(+)}(k_x, k_y, z) = E^{(+)}(k_x, k_y, 0) H(k_x, k_y, z) \quad \Rightarrow \quad H(k_x, k_y, z) = e^{-i \frac{(k_x^2 + k_y^2)z}{2k}}$$

The solution in real space can be obtained by performing the inverse Fourier transform:

$$E^{(+)}(x, y, z) = \frac{1}{(2\pi)^2} \iint E^{(+)}(k_x, k_y, 0) e^{-i \frac{(k_x^2 + k_y^2)z}{2k}} e^{i(k_x x + k_y y)} dk_x dk_y$$

and can be expressed in terms of the field distribution at $z = 0$ by replacing $E^{(+)}(k_x, k_y, 0)$ in terms of $E^{(+)}(x, y, 0)$:

$$E^{(+)}(x, y, z) = \frac{1}{(2\pi)^2} \iint \left[\iint E^{(+)}(x', y', 0) e^{-i(k_x x' + k_y y')} dx' dy' \right] \cdot e^{-i \frac{(k_x^2 + k_y^2)z}{2k}} e^{i(k_x x + k_y y)} dk_x dk_y$$

By interchanging the order of integration, we obtain

$$E^{(+)}(x, y, z) = \iint \left[\frac{1}{(2\pi)^2} \iint e^{-i \frac{(k_x^2 + k_y^2)z}{2k}} e^{i(k_x(x-x') + k_y(y-y'))} dk_x dk_y \right] E^{(+)}(x', y', 0) dx' dy'$$

$$E^{(+)}(x, y, z) = \iint E^{(+)}(x', y', 0) h(x-x', y-y', z) dx' dy'$$

[2.16A]

The solution in real space in plane z can be written as a convolution of the field distribution at $z = 0$ with the **impulse response function** $h(x, y, z)$:

$$h(x, y, z) = \frac{1}{(2\pi)^2} \iint e^{-i \frac{(k_x^2 + k_y^2)z}{2k}} e^{i(k_x x + k_y y)} dk_x dk_y \quad [2.16B]$$

Close inspection of Equation [2.16B] shows that the transfer function is in fact the Fourier transform of the impulse response function, so $h(x, y, z)$ and $H(k_x, k_y, z)$ constitute a Fourier transform pair.



2.7 Solved Problems

S2.1 Derive the relation between the beam radius w and the full width half maximum diameter of the beam intensity, w_{FWHM} . Write the expression of the Rayleigh range in terms of this parameter.

Solution

The intensity of the Gaussian beam at a given plane as a function of the beam radius at this plane is written as

$$I = I_o e^{-2\rho^2/w^2}$$

According to this expression, the beam radius w corresponds to the radius of the gaussian profile at the level I_o/e^2 .

The radius of the beam at the level where the intensity takes half its maximum value is calculated by taking $I = I_o/2$:

$$\frac{I_o}{2} = I_o e^{-2\rho_{1/2}^2/w^2} \rightarrow \ln 2 = 2 \frac{\rho_{1/2}^2}{w^2}$$

The diameter we are searching for is twice the obtained radius:

$$\phi_{FWHM} = 2\rho_{1/2} = 2w\sqrt{\frac{\ln 2}{2}} = w\sqrt{2 \ln 2}$$

The Rayleigh range of the beam can be expressed in terms of the FWHM diameter as

$$z_R = \frac{\pi w_o^2}{\lambda} = \frac{\pi \phi_o^2}{\lambda 2 \ln 2}$$

■

S2.2 Obtain the expression for the intensity of a Gaussian beam in terms of the total power. Calculate the diameter and area of a beam with constant circular intensity profile (top-hat beam) with the same peak intensity and total power of the Gaussian beam.

Solution

The intensity of the Gaussian beam at an arbitrary plane z is

$$I(\rho, z) = \frac{I_o w_o^2}{w^2(z)} e^{-2\rho^2/w^2(z)},$$

where I_o is the peak intensity of the Gaussian beam.

The total power is obtained by integrating over the whole area of the beam:



$$P_{tot} = \int_0^\infty \int_0^{2\pi} I(\rho, z) \rho d\rho d\phi = \int_0^\infty I(\rho, z) 2\pi \rho d\rho = \int_0^\infty \frac{I_o w_o^2}{w^2(z)} 2\pi \rho e^{-2\rho^2/w^2(z)} d\rho$$

$$P_{tot} = -\frac{I_o \pi w_o^2}{2} \int_0^\infty \frac{4\rho}{w^2(z)} e^{-2\rho^2/w^2(z)} d\rho = -\frac{I_o \pi w_o^2}{2} \left[e^{-2\rho^2/w^2(z)} \right]_0^\infty$$

$$P_{tot} = \frac{\pi w_o^2}{2} I_o$$

The intensity of the Gaussian beam in terms of the total power is obtained from this expression:

$$I(\rho, z) = \frac{2P_{tot}}{\pi w^2(z)} e^{-2\rho^2/w^2(z)}$$

Our top-hat beam has constant intensity distribution, I_o , over the whole beam equal to the peak intensity of the Gaussian beam. In this case, the relationship between the power and the intensity is

$$P_{TH} = I_o \pi \rho_{TH}^2 = \frac{I_o \pi \phi_{TH}^2}{4}$$

where ρ_{TH} is the beam radius and ϕ_{TH} is the beam diameter. By equating the total power of both beams, we obtain

$$\frac{\pi w_o^2}{2} I_o = I_o \frac{\pi \phi_{TH}^2}{4} \rightarrow \phi_{TH} = \sqrt{2} \cdot w_o$$

The top-hat beam area is

$$S_{TH} = \frac{\pi \phi_{TH}^2}{4} = \frac{\pi w_o^2}{2}$$

■

S2.3 The light emitted from a Nd:YAG laser at a wavelength of 1064 nm is a Gaussian beam of 1 W optical power and beam divergence $2\theta = 1$ mrad. Determine:

- the beam waist radius
- the depth of focus (confocal parameter)
- the peak intensity and the axial intensity at the planes $z = 1$ m and $z = 100$ m
- the q parameter of the beam at the planes $z = 1$ m and $z = 100$ m

**Solution**

a) The beam waist radius can be obtained from the beam divergence, which is known by means of Equation [2.8]:

$$2\theta = \frac{2\lambda}{\pi w_o} \rightarrow w_o = \frac{2\lambda}{\pi(2\theta)}$$

Directly substituting the laser beam parameters gives $w_o = 677.4 \mu\text{m}$.

b) The depth of focus is twice the Rayleigh range, which can be calculated by knowing the beam waist radius:

$$b = 2z_R = \frac{2\pi w_o^2}{\lambda} = 2.71\text{m}$$

c) The peak intensity is obtained using the relation derived in problem 2 as a function of the total power:

$$I_o = \frac{2P_{\text{tot}}}{\pi w_o^2} = 1.387\text{MW} / \text{m}^2 = 138.7 \text{ W} / \text{cm}^2$$

(The typical units used in laser optics for measuring highly intense ultrashort laser pulses are GW/cm^2 .)

The axial intensity at a given distance z from the beam waist plane:

$$I(0, z) = \frac{I_o w_o^2}{w^2(z)} = \frac{I_o}{1 + z^2/z_R^2} = \frac{I_o}{1 + (2z/b)^2}$$

Using the values of peak intensity and depth of focus obtained previously, we have the following:

At $z = 1 \text{ m}$, the on-axis intensity is $I(0, z) = 89.79 \text{ W}/\text{cm}^2$.

And at $z = 100 \text{ m}$, we obtain $I(0, z) = 2.54 \cdot 10^{-2} \text{ W}/\text{cm}^2$.

d) The q parameter at a given plane is $q = z - iz_R = z - i(b/2)$.

At the beam waist plane, $z = 0$ and $q = -1.355i \text{ m}$.

At $z = 1 \text{ m}$, $q = 1 - i = 1 - 1.355i \text{ m}$.

At $z = 100 \text{ m}$, $q = 100 - 1.355i \text{ m}$.

■

S2.4 Consider a Gaussian beam whose width, w , and radius of curvature, R , are known for a particular plane. Determine the position of the beam waist and the beam waist radius of the beam.



Solution

The q parameter of a Gaussian beam may be determined directly if the position and beam waist radius are known. Alternatively, knowing the q parameter allows finding out the values of the beam waist radius and its position for a given beam. We will use the latter in this problem, where our initial knowledge of the system is the wavefront radius of curvature and beam radius at a given plane. As we demonstrated in the discussion preceding Equation [2.12], the inverse of the q parameter is related directly to these beam parameters, which are initially known for this problem in the form of

$$\frac{1}{q(z)} = \frac{1}{R(z)} + i \frac{\lambda}{\pi w^2(z)}$$

Inverting this expression, we obtain the q parameter directly:

$$\frac{1}{q} = \frac{\pi w^2 + i\lambda R}{R\pi w^2} \rightarrow q = R\pi w^2 \cdot \frac{\pi w^2 - i\lambda R}{(\pi w^2)^2 + (\lambda R)^2} = \frac{R\pi^2 w^4 - i\lambda R^2 \pi w^2}{\pi^2 w^4 + \lambda^2 R^2}$$

Here, we omitted the dependence on z for clarity. Since $q = z - iz_R$, we equate the real and imaginary parts:

$$z = R \frac{\pi^2 w^4}{\pi^2 w^4 + \lambda^2 R^2} = \frac{R}{1 + \left(\frac{\lambda R}{\pi w^2}\right)^2}$$

$$z_R = \frac{\lambda R^2 \pi w^2}{\pi^2 w^4 + \lambda^2 R^2}$$

From the Rayleigh range, we can obtain an expression for the beam waist radius:

$$z_R = \frac{\pi w_o^2}{\lambda} = \frac{\pi}{\lambda} \cdot \frac{\lambda^2 R^2 w^2}{\pi^2 w^4 + \lambda^2 R^2} = \frac{\pi}{\lambda} \cdot \frac{w^2}{\left(\frac{\pi w^2}{\lambda R}\right)^2 + 1}$$

$$w_o = \frac{w}{\sqrt{1 + \left(\frac{\pi w^2}{\lambda R}\right)^2}}$$

■

S2.5 A Gaussian beam is incident on a plano-convex lens of focal length f , which is located at the beam waist plane. Determine the parameters of the transformed Gaussian beam after the lens.

Solution

If we consider the lens to be perfectly transparent, the effect of the beam is that a quadratic phase factor is introduced, thus inducing a curvature of the wavefront. As we have



seen in this chapter, this phase transforms the incident Gaussian beam into another Gaussian beam after the lens. The characteristics of the output beam are easily obtained from the expression:

$$\frac{1}{q_{out}} = \frac{1}{q_{in}} - \frac{1}{f}$$

where, for a thin lens, q_{in} corresponds to the q parameter of the beam just at the entrance of the lens; and q_{out} is the q parameter of the transformed Gaussian beam at the plane just after the lens. From this expression:

$$q_{out} = \frac{q_{in} \cdot f}{f - q_{in}}$$

In this problem, the beam waist of the incident beam is located at the lens plane and, hence, $q_{in} = -iz_R$:

$$q_{out} = \frac{-iz_R f}{f + iz_R} = -iz_R f \frac{f - iz_R}{f^2 + z_R^2} = -\frac{z_R^2 f}{f^2 + z_R^2} - i \frac{z_R f^2}{f^2 + z_R^2}$$

Rearranging the terms:

$$q_{out} = -\frac{f}{1 + (f/z_R)^2} - i \frac{z_R}{1 + (z_R/f)^2} = -z' - iz_R'$$

The imaginary term corresponds to the new Rayleigh range, from which the new beam waist radius is obtained ($z_R' = \pi w_o'^2 / \lambda$):

$$w_o' = \frac{w_o}{\sqrt{1 + (\pi w_o^2 / \lambda f)^2}}$$

In the limit of a Rayleigh range that is much larger than the focal length, this situation is similar to that of a plane wave illuminating a lens, for which we take a geometrical optics approach:

$$z_R \gg f \rightarrow z' = -\frac{f}{1 + \left(\frac{f}{z_R}\right)^2} \approx -f$$

The position of the beam waist after the lens is at the back focal plane of the lens:

$$z_R \gg f \rightarrow w_o' = \frac{w_o}{\sqrt{1 + (\pi w_o^2 / \lambda f)^2}} \approx \frac{\lambda f}{\pi w_o}$$

In terms of the beam diameters, this expression becomes



$$2w_o' = \frac{4\lambda f}{\pi 2w_o} \rightarrow D' = \frac{4\lambda f}{\pi D_o} \approx 1.2 \frac{\lambda f}{D_o}$$

This formula corresponds to the diameter of the focused spot of a plane wave using a lens of focal length f and diameter D_o . ■

S2.6 a) A lens of focal length f is placed at an arbitrary distance, z , from the beam waist plane of a Gaussian beam. Determine the beam radius at the back focal plane of the lens, $w'(f)$.

b) A laser is emitting a Gaussian beam, and we need to determine the position of the beam waist plane, located at a distance A inside the laser cage (measured from the laser output plane). In order to obtain this value, we take three experimental measurements of the beam radius: The first value is measured at a distance L_1 from the laser aperture; a second value is measured at a distance $L_2 = L_1 + D$, and the third value is measured at the back focal plane of a lens, placed at an arbitrary distance larger than L_2 . The focal length of the lens is $f = 50$ cm, and the corresponding results of our measurements are:

Plane	Distance	Beam radius
1	$L_1 = 150$ cm	$w(L_1) = 745$ μ m
2	$L_2 = L_1 + D = 250$ cm	$w(L_2) = 1005$ μ m
3	At distance f from the lens	$w'(f) = 160$ μ m

Find the value of A .

Solution

To obtain the beam radius at the plane located at a distance f after the lens, we will propagate the q parameter up to the desired plane, starting from the beam waist position of the entrance beam.

Initial q parameter: $q_o = -iz_R$

At the entrance of the lens, the q parameter will be $q_{in} = z - iz_R$

And at the output of the lens, we get

$$q_{out} = \frac{q_{in} \cdot f}{f - q_{in}} = \frac{f(z - iz_R)}{f - z + iz_R}$$

At the desired plane, the q parameter becomes $q = q_{out} + f = \frac{q_{in} \cdot f}{f - q_{in}} + f = \frac{f^2}{f - q_{in}}$



By inverting this expression: $\frac{1}{q} = \frac{f - q_{in}}{f^2} = \frac{f - (z - iz_R)}{f^2} = \frac{f - z}{f^2} + i \frac{z_R}{f^2}$

We can obtain information about the beam radius using the general relation [2.12]. From the imaginary part, we get

$$\frac{1}{q} = \frac{1}{R} + i \frac{\lambda}{\pi w'^2} \rightarrow \frac{\lambda}{\pi w'^2} = \frac{z_R}{f^2} = \frac{\pi w_o^2}{\lambda f^2}$$

The final result: $w' = \frac{\lambda f}{\pi w_o}$.

b) To obtain the value of A , we will use the information included in the table.

The beam radius at L_1 and L_2 are related to the Rayleigh range and beam waist radius by means of the usual relation:

$$z_1 = L_1 + A \rightarrow w^2(z_1) = w_o^2 \left(1 + \frac{z_1^2}{z_R^2} \right)$$

$$z_2 = z_1 + D \rightarrow w^2(z_2) = w_o^2 \left(1 + \frac{z_2^2}{z_R^2} \right)$$

Subtracting the second from the first, we have

$$\frac{w^2(z_2) - w^2(z_1)}{w_o^2} = \frac{z_2^2 - z_1^2}{z_R^2} = \frac{(z_1 + D)^2 - z_1^2}{z_R^2}$$

After a few algebraic steps, we obtain an expression for z_1 :

$$z_1 = \frac{\frac{z_R^2}{w_o^2} (w^2(z_2) - w^2(z_1)) - D^2}{2D} = \frac{\frac{\pi^2 w_o^2}{\lambda^2} (w^2(z_2) - w^2(z_1)) - D^2}{2D}$$

Now, using the result obtained in a), we can express z_1 in terms of our measured quantities:

$$z_1 = \frac{f^2 (w^2(z_2) - w^2(z_1))}{2D w'^2(f)} - \frac{D}{2}$$

Directly substituting all the values gives $z_1 = L_1 + A = 1.722 \text{ m}$

And the final result is $A = 22.2 \text{ cm}$.

■

S2.7 A Gaussian beam of Rayleigh range z_R is incident on an optical system formed by 2 lenses of focal lengths f_1 and $f_2 = 2f_1$. The separation between lenses is $D = 2f_1$, and the



beam waist of the incident Gaussian beam is placed at a distance $z = f_1$ from the first lens. Calculate a) the value of the new Rayleigh range after the second lens, and b) the distance from the second lens to the plane where the new beam waist will be located. Then, particularize the expressions obtained for the case $f_1 = z_R$.

Solution

The q parameter just before the first lens is $q_{in1} = f_1 - iz_R$.

$$\text{And at the output of the first lens: } q_{out1} = \frac{f_1 \cdot q_{in1}}{f_1 - q_{in1}} = \frac{f_1^2 - iz_R f_1}{iz_R} = -f_1 - i \frac{f_1^2}{z_R}.$$

After propagation of the Gaussian beam between lenses 1 and 2 at a distance $D = 2f_1$, the q parameter before the second lens will be

$$q_{in2} = q_{out1} + D = f_1 - i \frac{f_1^2}{z_R}$$

$$\text{And after the second lens } q_{out2} = \frac{f_2 \cdot q_{in2}}{f_2 - q_{in2}} = \frac{2f_1 \cdot q_{in2}}{2f_1 - q_{in2}}$$

After directly substituting q_{in2} and some straightforward algebra, we obtain the expression

$$q_{out2} = 2f_1 \left(\frac{1 - f_1^2/z_R^2}{1 + f_1^2/z_R^2} \right) - i \frac{4z_R}{1 + z_R^2/f_1^2}$$

$$\text{From this result, the beam waist after the second lens is located at } z' = -2f_1 \left(\frac{f_1^2 - z_R^2}{f_1^2 + z_R^2} \right).$$

(!) Note: from this result, when $f_1 > z_R$, $z' < 0$, the beam waist is located to the right of the lens; but if $f_1 < z_R$ then $z' > 0$, then the beam waist will be virtual (located to the left of the lens). Therefore, the Gaussian beam diverges from the second lens.

The beam waist radius is given by the imaginary part of q_{out2} :

$$z'_R = \frac{4z_R}{1 + z_R^2/f_1^2} \rightarrow w_o' = \frac{2w_o}{\sqrt{1 + \left(\frac{\pi w_o^2}{\lambda f_1} \right)^2}}$$

For the particular case $f_1 = z_R$, a direct substitution gives

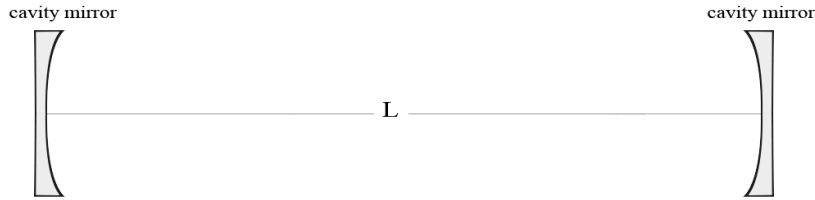
$$z' = 0 \text{ and } w_o' = \sqrt{2}w_o$$

The new beam waist will be located exactly at the second lens. ■

S2.8 Gaussian beams are used in interferometric gravitational wave detectors. Each arm of the interferometer consists of two mirrors separated by a distance L . For the light to



resonate in the cavity, the mirror's surface must coincide with the Gaussian beam wavefronts.



a) If the mirrors at both ends are to be equal, indicate the position of the beam waist of the Gaussian beam resonating inside the cavity and calculate the beam waist radius, w_o , which minimizes the value of the beam radius at the mirrors. Calculate the beam waist radius and Rayleigh range if our system has $L = 4 \text{ km}$ and the laser wavelength is $\lambda = 1064 \text{ nm}$.

b) Calculate the full width half maximum (FWHM) diameter in intensity, the q parameter, and the wavefront radius of curvature of the Gaussian beam at the mirror locations.

c) A new generation of interferometers uses larger spot sizes on the mirror to reduce thermal noise. In this new configuration, the radius of curvature of the mirrors is set to $R = 2.076 \text{ km}$. Find in this configuration the new Rayleigh range, the beam waist radius, and the beam radius at the mirrors. ($L = 4 \text{ km}$; $\lambda = 1064 \text{ nm}$)

Solution

a) Since the cavity is symmetric, the beam waist plane must be located in the middle of the cavity, and the position of the two mirrors will be $z = \pm L/2$.

Next, we must find the value of w_o by minimizing the value of $w(z)$ at $z = \pm L/2$.

The radius of the beam in plane z can be expressed in terms of the beam waist radius:

$$w^2(z) = w_o^2 \left(1 + \frac{z^2}{z_R^2} \right) = w_o^2 \left(1 + \frac{z^2 \lambda^2}{\pi^2 w_o^4} \right) \xrightarrow{\text{at } z = \pm L/2} w^2(\pm L/2) = w_o^2 + \frac{L^2 \lambda^2}{4\pi^2 w_o^4}.$$

The value of w_o minimizing $w(L/2)$,

$$\frac{d(w^2(L/2))}{dw_o} = 0 \rightarrow 2w_o + \frac{L^2 \lambda^2}{4\pi^2} \left(-\frac{2}{w_o^3} \right) = 0$$

$$\text{is given by } w_o = \sqrt{\frac{L\lambda}{2\pi}} \text{ and } z_R = \frac{\pi w_o^2}{\lambda} = \frac{L}{2}.$$

Thus, the total cavity length must be equal to the confocal parameter of the Gaussian beam.



For the values of the problem, $L = 4$ km and $\lambda = 1064$ nm, whereby one obtains

$$w_o = 2.60 \text{ cm and } z_R = 2 \text{ km.}$$

b) At the mirror location, the beam radius is

$$w(L/2) = w_o \sqrt{1 + \frac{L^2}{4z_R^2}}$$

And using the values obtained in a), we get $w(L/2) = 3.67 \text{ cm}$.

To obtain the FWHM diameter, we can directly use the result obtained in problem 1:

$$\phi_{FWHM}(L/2) = \sqrt{2 \ln 2} \cdot w(L/2) = 4.33 \text{ cm}$$

The q parameter of the beam at the mirror position:

$$q(\pm L/2) = \pm(L/2) - iz_R = (\pm 2000 - i2000) \text{ m}$$

Wavefront radius of curvature:

$$R(L/2) = \frac{L}{2} \left(1 + \frac{z_R^2}{(L/2)^2} \right) = L = 4 \text{ km}$$

c) By changing the mirror configuration, the properties of the Gaussian beam must also change to accommodate to the new cavity. In this new situation the radius of curvature of the mirrors is changed to $R' = 2.076$ km. From this value we obtain the new Rayleigh range:

$$R'(L/2) = \frac{L}{2} \left(1 + \frac{z_R'^2}{(L/2)^2} \right) \rightarrow z'_R = 389 \text{ m}$$

And the corresponding new beam waist radius is

$$w'_o = \sqrt{\frac{\lambda z'_R}{\pi}} = 1.15 \text{ cm}$$

Finally, the beam radius at the mirror's location is

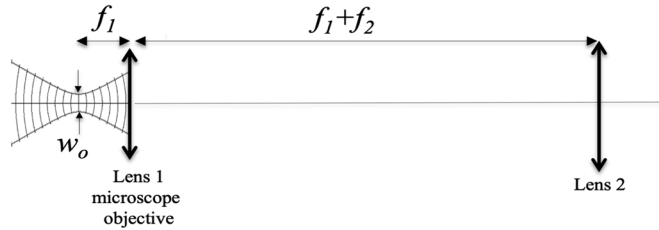
$$w(L/2) = w'_o \sqrt{1 + \frac{(L/2)^2}{z_R'^2}} = 6 \text{ cm}$$

S2.9 A tightly-focused Gaussian beam of wavelength 800 nm has a beam waist radius $w_o = 200 \mu\text{m}$. We need to acquire the image of the beam waist plane using a CCD camera



with a square sensor of 5 mm size. To obtain a good resolution image, we will use a magnifying system.

a) Our first setup is the (confocal) arrangement shown in the figure, which uses a microscope objective with focal length $f_1 = 9$ mm and a second lens with focal length $f_2 = 10$ cm.



a1) Determine the value of the beam waist radius of the Gaussian beam after lens 2. If our beam diameter has twice the beam radius at $1/e^2$ the value of its intensity, discuss if it will fit into the CCD sensor.

a2) At which distance from the original beam waist plane should the CCD be placed?

a3) If the beam waist radius, w_o , is changed, do we need to change the position of our CCD to record the magnified beam waist image?

b) In our second setup, we replace the confocal system with a single lens placed at a distance f from the beam waist plane.

b1) Find the value of the focal length we must use to obtain the same magnification at the CCD as the one obtained in case a).

b2) At which distance from the original beam waist plane should the CCD be placed?

b3) If the beam waist radius, w_o , is changed, do we need to change the position of our CCD to record the magnified beam waist image?

Solution

a1) To obtain the beam waist radius after lens 2, we will propagate the Gaussian beam through the optical system. Since the beam waist of the initial Gaussian beam is placed at the focal plane of the microscope objective, the q parameter on lens 1 is

$$q_{in1} = f_1 - iz_R, \text{ where } z_R = \frac{\pi w_o^2}{\lambda} = 15.71 \text{ cm}$$

The q parameter of the new Gaussian beam after the objective is

$$q_{out1} = \frac{f_1 \cdot q_{in1}}{f_1 - q_{in1}} = -f_1 - i \frac{f_1^2}{z_R}$$



After propagation of a distance $D = f_1 + f_2$, this becomes $q_{in2} = q_{out1} + D = f_2 - i \frac{f_1^2}{z_R}$.

And, after lens 2:

$$q_{out2} = \frac{f_2 \cdot q_{in2}}{f_2 - q_{in2}} = -f_2 - i \frac{f_2^2 z_R}{f_1^2}$$

From this result, we can conclude:

- The new beam waist after lens 2 is located at a distance f_2 from lens 2, in the back focal plane.
- The new Rayleigh range after lens 2: $z'_R = \left(\frac{f_2}{f_1}\right)^2 \cdot z_R \rightarrow w'_o = \frac{f_2}{f_1} w_o = 2.22 \text{ mm}$.

The new beam waist has a magnification of $M = f_2/f_1 = 11.1$.

w_o' corresponds to the $(1/e)$ radius in amplitude and also to the $(1/e^2)$ radius in intensity. Thus, the beam diameter at $(1/e^2)$ in intensity is

$$\phi_{1/e^2} = 2w_o' = 4.44 \text{ mm}$$

Hence, it will fit in the 5 mm-size sensor.

a2) The camera will record the desired image when its distance from the original beam waist is $D = f_1 + (f_1 + f_2) + f_2 = 218 \text{ mm}$, at the back focal plane of lens 2.

a3) From the results in Section a1), we can see that the position of the image is always the back focal plane of lens 2, irrespectively of the size of w_o . The size of the image will change if the dimensions of the original beam change, though always maintaining a constant magnification factor M .

b1) By replacing our system with a single lens placed at a distance f from the original beam waist, we obtain:

$$q_{in} = f - iz_R \rightarrow q_{out} = \frac{f^2 - ifz_R}{iz_R} = -f - i \frac{f^2}{z_R}$$

The new beam waist will appear at the focal plane of the lens with Rayleigh range

$$z'_R = \frac{f^2}{z_R} \text{ and beam waist radius } w'_o = \frac{f\lambda}{\pi w_o}.$$

If we want to the same magnification as in a), then $w'_o = M \cdot w_o$ and

$$f = \frac{M w_o^2 \pi}{\lambda} = 1.74 \text{ m}$$



b2) Now the camera should be placed at a distance of $2f = 3.48$ m from the original beam waist position, a distance almost 16 times greater than in the previous configuration.

b3) Additionally, if the beam waist of the Gaussian beam changes, the required focal length of the lens should also change, as seen in *b1*), and the position of the CCD should also be changed accordingly.

■



2.8 Proposed problems

P2.1 Check by direct substitution that the Fresnel approximation to a spherical wave is a solution to the paraxial wave equation. Find a limit of validity to this approximation.

P2.2 A Gaussian beam is transmitted through a thin lens of focal length f . Find an expression for the location of the waist of the transmitted beam as a function of the location of the waist of the incident beam.

P2.3 Considering that the beam waist corresponds to the position of the object or image in a single lens configuration, write the corresponding lens equation and compare it with the well-known thin lens formula of geometrical optics.

P2.4 An experimental way to measure the beam diameter is the knife-edge method. A very sharp edge is placed at the z plane where we want to measure the beam. A detector placed on the beam axis measures the beam energy. As we move the edge into the beam, the energy going into the detector decreases. Show that the beam radius corresponds to the edge displacement between the positions for which we measure 16% and 84 % of the total beam intensity.

P2.5 A Gaussian beam of wavelength 800 nm has a Rayleigh range $z_R = 1$ m. At the beam waist position, we place a lens with focal length $f_1 = 5$ cm. A second lens with $f_2 = 15$ cm is placed at a distance $D = f_1 + f_2$ to the right of the first lens.

- a) Find the q parameter of the Gaussian beam in the plane just after the second lens.
- b) Find the position of the beam waist and the beam waist radius after the second lens.
- c) Find the expression of the Gaussian beam amplitude at the output of the second lens if the total power of the beam is $P_{\text{tot}} = 10$ W

P2.6 A Gaussian beam of total power 10 W and wavelength 500 nm has a beam waist radius of 1 mm and is placed at 10 cm from a lens of focal length $f = 5$ cm. Find the beam radius, the radius of curvature, and the on-axis intensity at a distance $z = 2$ cm after the lens.

P2.7 A Gaussian beam of wavelength λ has a beam waist radius w_0 . The beam is transformed by a collimator formed by two lenses arranged in the following configuration. The first lens, of focal length f_1 , is placed at a distance $z = f_1$ from the beam waist plane of the initial Gaussian beam. A second lens of focal length f_2 is placed at a distance $f_1 + f_2$ from the first lens.

- a) Find the position of the beam waist after the second lens.
- b) Find the value of the new beam waist radius after the second lens.
- c) Calculate the value of the confocal parameter of the Gaussian beam after the second lens, expressing it in terms of the confocal parameter of the initial Gaussian beam and the focal lengths of the lenses.



P2.8 A Gaussian beam of wavelength 532 nm has a Rayleigh range of $z_R = 10$ cm and a total power of 10 mW. A divergent lens with focal length $f = -5$ cm is placed at a distance $z = z_R$ to the left of the beam waist. Calculate the transmitted Gaussian beam's beam radius and wavefront radius of curvature at the plane located at a distance $z = 5$ cm to the right of the lens. Then, write the expression for the intensity of the Gaussian beam at this plane.



→ 3



Light emission and light-matter interaction

Introduction

In the previous chapters, we studied the basic solutions for the propagation of EM waves in vacuum. We did so by establishing the basic equation that describes the propagation of optical beams in the paraxial region and some of their most significant solutions. Next, we turn to the problem of considering the presence of matter and how it interacts with the EM field. Two main problems will be discussed in this chapter: 1) light emission from radiating dipoles; and 2) the interaction of the emitted EM radiation with matter. We will find significant aspects in both these situations, specifically regarding the properties of optical radiation when it interacts with matter. Although a more rigorous treatment would at least require a density-matrix or time-dependent perturbation approach to a quantum treatment of matter, many situations can be properly understood by means of the classical electrodynamics approach. This chapter will deal with the main aspects of the classical theory related to light emission and interaction. The most relevant aspects are summarized without entering into all the details of the theory, which can be found in various textbooks on classical electrodynamics (such as Jackson, J.D. (1999). *Classical Electrodynamics*. John Wiley & Sons; and Oughstun, K.E. (2006). *Electromagnetic and Optical Pulse Propagation*. Springer.).

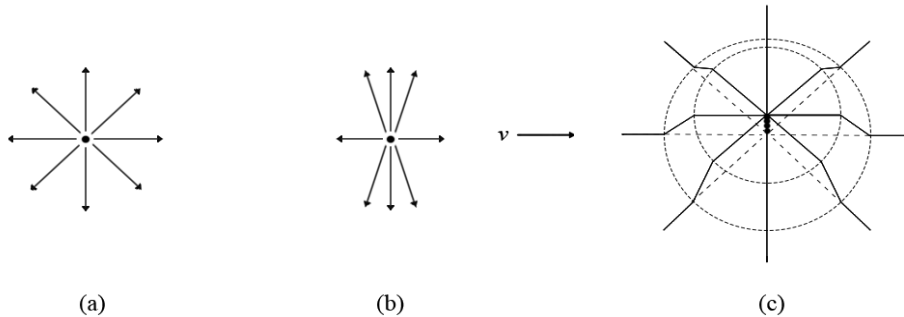
The first important aspect to be considered when studying the emission of EM radiation from matter is that **electromagnetic radiation is generated by accelerated charges**, while the light-matter interaction problem essentially refers to the radiated EM field's action on charges.

A point charge at rest generates a radial static electric field that points either away from (positive) or toward (negative) the charge. When the charge moves at constant velocity, we observe a displacement of the lines in the direction of the movement with no change in curvature (see Figure 1). The sudden acceleration and stop of a point charge generates a burst of EM radiation propagating out from the charge at speed c (in vacuum), as outlined in Figure 1. Note that the electric field lines are not changed in the direction of acceleration. When the charge oscillates with harmonic motion, electromagnetic waves are generated.



In the low frequency regime, any wavelength can be obtained using circuit oscillators driving electrons up and down along the antenna. A well-defined EM emission that closely resembles a spherical or plane wave can be obtained from individual antennas; and a more directional emission pattern can be obtained by the coherent superposition of a few oscillators. Although the basic properties of an electromagnetic wave are common to the entire range of the EM spectrum, the way they are generated differs drastically from region to region, thus imposing some peculiarities on the observed characteristics.

Figure 3.1
Electric field lines for
a) static charge,
b) charge with constant $v \approx c$, and
c) accelerated charge



The following summarizes the main aspects of the radiation problem in the EM spectrum ranging from near UV through the visible to IR.

- Light is generated by oscillation of charges at the microscopic level, such as
 - Electron transitions in atoms, molecules and solids;
 - vibrational and rotational transitions in molecules and solids; and
 - band transitions in solids.
- The basic aspect of this kind of emission is the **quantum nature** of the processes:
 - quantized frequency emission;
 - transitions allowed by the radiative process (dipolar, quadrupolar, Raman, etc.);
 - the interplay between spontaneous and stimulated emissions; and
 - resonance effects (discrete line emission).

Light emission from matter is most often a **many body emission** problem. In this sense, the statistical properties of the emitters must be carefully considered. However, the detailed treatment of coherence and the statistical properties of light are beyond the scope of this course and will not be covered in detail.

Based on the aspects considered, any light source can be classified according to its emission properties:

A. Spontaneous emission sources

Thermal sources

Solids at a given temperature emit a continuous spectrum of electromagnetic radiation that reaches a maximum at a particular wavelength, depending on temperature (Wien's law). The total energy emitted by a blackbody is proportional to T^4 , as described by the

Stefan–Boltzmann law. Max Planck’s discovery of the law governing the emission of radiation by a blackbody constitutes one of the most important milestones in the birth of quantum theory.

Spectral sources

At a given pressure, gas-discharge lamps generate radiation at a discrete set of wavelengths, which depend on the material.

Light-Emitting Diodes

Electron-hole recombination processes in semiconductors (p-n junctions) are used as efficient light sources. Combining LEDs at three basic wavelengths (red, blue, and green) allow generating any colour, thanks to the additive properties of light superposition.

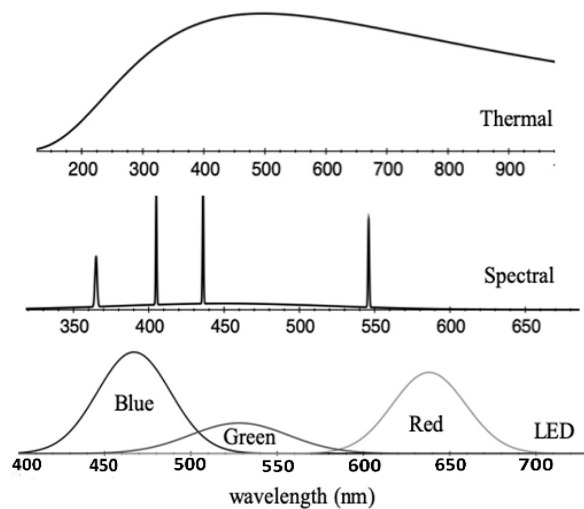


Figure 3.2
Typical spectrum of
a) thermal,
b) spectral, and
c) LED
sources

B. Stimulated emission sources

The possibility of using stimulated emission resulted in the invention of a coherent light source, the LASER, whose properties have made possible a revolution in the field of photonics. A brief description of the operation principle of a laser is given in Appendix A.

3.1 Light emission from a dipole

A mathematical study of the classical approach to radiation starts with Maxwell’s Equations [1.1], which must include the presence of charges and currents. A wave equation (for homogeneous media) in this case can be derived as



$$\begin{aligned}\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) &= -\frac{\partial}{\partial t}(\mu_0 \vec{j} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}) \\ \vec{\nabla} \cdot (\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} + \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} &= -\mu_0 \frac{\partial \vec{j}}{\partial t}\end{aligned}\quad [3.1A]$$

$$\boxed{\nabla^2 \vec{E} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} = \vec{\nabla} \left(\frac{\rho}{\epsilon_0} \right) + \mu_0 \frac{\partial \vec{j}}{\partial t}}$$

$$\begin{aligned}\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{B}) &= \mu_0 (\vec{\nabla} \wedge \vec{j}) + \mu_0 \epsilon_0 \frac{\partial}{\partial t} (\vec{\nabla} \wedge \vec{E}) \\ -\nabla^2 \vec{B} &= \mu_0 (\vec{\nabla} \wedge \vec{j}) - \mu_0 \epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2}\end{aligned}\quad [3.1B]$$

$$\boxed{\nabla^2 \vec{B} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} = -\mu_0 (\vec{\nabla} \wedge \vec{j})}$$

The solution to these equations for arbitrary charge and current distribution is a very complicated problem. One can see that the effect of sources appears in the equations under derivative operators, indicating a nonlocal relation between them and the fields. The solution for radiation problems can be solved more simply in terms of the electromagnetic potentials, which can be derived directly from Maxwell's equations:

Vector Potential: $\vec{A}$

Since the magnetic field is divergence-free [1.1B], it can always be expressed in terms of a new function called the vector potential, such as

$$\vec{B} = \vec{\nabla} \wedge \vec{A} \quad [3.2A]$$

Coulomb Potential: ϕ

Substituting [3.2A] in [1.1C] leads to

$$\vec{\nabla} \wedge \vec{E} = -\frac{\partial}{\partial t} (\vec{\nabla} \wedge \vec{A}) \quad \rightarrow \quad \vec{\nabla} \wedge \left(\vec{E} + \frac{\partial \vec{A}}{\partial t} \right) = 0$$

thus expressing the fact that the vector inside the bracket is irrotational. We define a new scalar variable, the Coulomb potential, ϕ , as

$$\vec{E} + \frac{\partial \vec{A}}{\partial t} = -\vec{\nabla} \phi \quad \rightarrow \quad \vec{E} = -\vec{\nabla} \phi - \frac{\partial \vec{A}}{\partial t} \quad [3.2B]$$

For stationary problems, the temporal derivative vanishes. This relation then simplifies to the electrostatic definition of Coulomb's potential, thus leading to Laplace or Poisson equations of electrostatics.

In terms of the defined potentials, Maxwell's equations are rewritten:



$$[1.1D] \rightarrow \quad \vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{A}) = \mu_0 \vec{j} - \mu_0 \epsilon_0 \vec{\nabla} \frac{\partial \phi}{\partial t} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{A}}{\partial t^2} \quad \text{and} \quad [3.3A]$$

$$\boxed{\nabla^2 \vec{A} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{A}}{\partial t^2} = -\mu_0 \vec{j} + \vec{\nabla} \left(\vec{\nabla} \cdot \vec{A} + \mu_0 \epsilon_0 \frac{\partial \phi}{\partial t} \right)}$$

$$[1.1A] \rightarrow \quad \vec{\nabla} \cdot \left(-\vec{\nabla} \phi - \frac{\partial \vec{A}}{\partial t} \right) = \frac{\rho}{\epsilon_0} \rightarrow \boxed{\nabla^2 \phi + \frac{\partial}{\partial t} (\vec{\nabla} \cdot \vec{A}) = -\frac{\rho}{\epsilon_0}} \quad [3.3B]$$

These equations couple the vector potential and Coulomb potential with the charge and current distributions. From the definitions in [3.2], we immediately see that the electromagnetic fields are invariant under a change in the potentials, which is given by a gauge transformation:

$$\vec{A}' = \vec{A} + \vec{\nabla} \cdot f \quad \text{and} \quad \phi' = \phi - \frac{\partial f}{\partial t}$$

with f being a generic function of r and t .

Thus, by selecting a particular gauge, one can choose the value of the divergence of the vector potential.

◆ *Comment*

Two important gauge transformations are used in electrodynamics:

1) *Coulomb gauge*: $\vec{\nabla} \cdot \vec{A} = 0$

This transformation is very useful for the quantization of the EM field, leading to a dynamical equation for the potential vector, while the temporal dependence for the Coulomb potential is eliminated.

2) *Lorenz gauge*: $\vec{\nabla} \cdot \vec{A} = -\mu_0 \epsilon_0 \frac{\partial \phi}{\partial t}$

The Lorenz gauge is used when studying dipole radiation, as we will see.¹ Another advantage is that it is relativistic-invariant.

By selecting a vector potential fulfilling the Lorenz gauge condition, we can write a set of decoupled wave equations for the vector and Coulomb potentials, which are driven by the density current and charge density, respectively:

¹ (*) For a detailed discussion of these aspects, refer to the book by Cohen-Tannoudji, C., Dupont-Roc, J., and Grynberg, G. (1989). *Photons and Atoms: Introduction to Quantum Electrodynamics*. John Wiley & Sons.



$$\begin{aligned}\bar{\nabla}^2 \bar{A} - \mu_0 \epsilon_0 \frac{\partial^2 \bar{A}}{\partial t^2} &= -\mu_0 \bar{J} \\ \bar{\nabla}^2 \phi - \mu_0 \epsilon_0 \frac{\partial^2 \phi}{\partial t^2} &= -\frac{\rho}{\epsilon_0}\end{aligned}\quad [3.4]$$

These equations are solved for given current and charge density distributions. The basic solutions to these equations are called the Liénard–Wiechert solutions, and they are expressed in terms of the retarded sources; i.e., the potentials at a particular instant of time, t , are defined in terms of the values of the sources at a previous time $t' = t - R/c$, in which $R = |\vec{r} - \vec{r}'|$ defines the distance between the point in space where we calculate the potential and the position of the source element:

$$\begin{aligned}\bar{A}(\vec{r}, t) &= \frac{\mu_0}{4\pi} \int \frac{\bar{J}(\vec{r}', t' = t - R/c)}{R} dV', \\ \phi(\vec{r}, t) &= \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\vec{r}', t' = t - R/c)}{R} dV',\end{aligned}\quad [3.5]$$

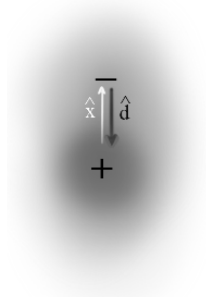
where V' corresponds to the volume that includes the charge and current density distributions.

These equations are usually used in antenna theory to find radiated solutions for different configurations. Although we could use these expressions to determine the emission of light, a simpler formalism can be adopted by considering the specific characteristics of the emission of radiation by atoms. As we have discussed, light is emitted by the oscillating charges in atoms, molecules, or solids, which are fundamentally neutral (ionization of matter takes place usually at high frequencies and intensities). The multipole expansion of any charge distribution shows that it can be considered as a superposition of a point charge (equal to the total charge of the distribution), an electric dipole, an electric quadrupole, or a magnetic dipole, and so on (refer to any standard book on electromagnetic theory for details).

For neutral charge distributions, the leading term is consequently a dipole, so radiation by atoms can be considered at first order as an emission from a classical electric dipole.

Electric dipole

An elementary electric dipole is formed by a positive (q) and negative ($-q$) charge of equal magnitude, separated by a given distance (x). For the case of atoms, we can consider that this distance is defined by the relative position of the center of mass of the electron cloud with respect to the nucleus when distorted by the presence of an external electric field. We define the **dipolar moment** of an elementary dipole as



$$\vec{p} = -q \cdot \vec{x} = qx \cdot \hat{d}$$



where the unitary vector $\hat{d}$ points from the negative to the positive charge. An electronic cloud oscillating in time will lead, then, to an oscillatory dipole, $\vec{p}(t)$ thus becoming an emission source of EM radiation (oscillatory motion is accelerated).

Although radiation by a dipole could be obtained from Equation [3.5], it can be studied more simply by considering the **electric Hertz potential, $\vec{\Pi}$** .

The Hertz potential is obtained from the vector and Coulomb potentials satisfying the relations

$$\vec{A} = \mu_0 \frac{\partial \vec{\Pi}}{\partial t} \quad \text{and} \quad \phi = -\frac{\vec{\nabla} \cdot \vec{\Pi}}{\epsilon_0} \quad [3.6A]$$

The Hertz potential satisfies the wave equation

$$\vec{\nabla}^2 \vec{\Pi} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{\Pi}}{\partial t^2} = -\vec{p} \quad [3.6B]$$

We can recognize in this equation the utility of this approach, since the **electric dipole becomes the source term in the wave equation for the Hertz potential**. For a given dipole distribution, the general solution to this wave equation is

$$\vec{\Pi}(\vec{r}, t) = \frac{1}{4\pi} \int \frac{\vec{p}(\vec{r}', t' = t - R/c)}{R} dV' \quad [3.6C]$$

The electric and magnetic fields are given by

$$\begin{aligned} \vec{E} &= \frac{\vec{\nabla} \cdot (\vec{\nabla} \cdot \vec{\Pi})}{\epsilon_0} - \mu_0 \frac{\partial^2 \vec{\Pi}}{\partial t^2} \\ \vec{B} &= \mu_0 \left(\vec{\nabla} \wedge \frac{\partial \vec{\Pi}}{\partial t} \right) \end{aligned} \quad [3.6D]$$

The formalism just derived is used in Appendix B to obtain the solution for the field radiated from an elementary dipole, in the form $\vec{p}(\vec{r}, t) = p(t)\delta(\vec{r} - \vec{r}_0)\hat{d}$.

The solution is quite complex and involves different terms that show different dependences with distance. We can distinguish the following different regimes when studying the radiated emission patterns.

1. Near field zone: $R \ll \lambda \approx l$

In this zone, we are observing the emission pattern at distances much smaller than two characteristic lengths of the problem: the wavelength, λ ; and the dipole size, l . The only significant terms in this region are those decaying, such as $(1/R^3)$, which corresponds to the electrostatic expression for an electric dipole.



2. Intermediate zone: $R \approx \lambda, l$

At this zone the evolution of the fields is quite complex. We will not study the radiation in this regime, since for optical waves we are interested in the emission at distances longer than the wavelength (typically of the order of microns).

3. Far field or radiation zone: $R \gg \lambda, l$

For studying light emission, we are mostly interested in this region. We can also apply the **dipolar approximation**, which assumes that the dipole dimensions are much smaller than the wavelength, so the electric field can be considered constant over the dimensions of a single dipole. This is justified because atomic dimensions (angstroms) are much smaller than optical wavelengths (microns). In this region, the terms of the solution decaying faster than $(1/R)$ are negligible at the observing distances, and the only surviving **radiation term** decays as $(1/R)$.

EM field emitted by an elementary dipole in the dipolar and radiation field approximation $R \gg \lambda \gg l$

As derived in Appendix B, the EM field radiated by an elementary dipole is written as

$$\begin{aligned} \vec{e}_{rad}(\vec{r}, t) &= \frac{\mu_0 \ddot{P}(\vec{r}_0, t') \sin \theta}{4\pi R} \hat{e} \\ \vec{b}_{rad}(\vec{r}, t) &= \frac{\mu_0 \ddot{P}(\vec{r}_0, t') \sin \theta}{4\pi R c} \hat{b} \end{aligned} \quad [3.7A]$$

where R is the distance from the observation point to the dipole $R = |\vec{r} - \vec{r}_0|$, and the temporal derivatives of the dipole are evaluated at the retarded time $t' = t - R/c$. The unitary vectors defining the relative orientation of the electric and magnetic field vectors with respect to the direction of propagation are given by the expressions

$$\begin{aligned} \hat{e} &= [\hat{R} \wedge (\hat{R} \wedge \hat{d})] / \sin \theta = (\cos \theta \cos \phi \hat{x} + \cos \theta \sin \phi \hat{y} - \sin \theta \hat{z}) \\ \hat{b} &= (\hat{d} \wedge \hat{R}) / \sin \theta = (-\sin \phi \hat{x} + \cos \phi \hat{y}) \\ \hat{R} &= (\sin \theta \cos \phi \hat{x} + \sin \theta \sin \phi \hat{y} + \cos \theta \hat{z}) \\ \hat{d} &= \hat{z} \end{aligned} \quad [3.7B]$$

where we define a right-handed coordinate system centred on the dipole position, with the dipole vector directed along the z -axis. The electric field is contained in the plane that includes the dipole and the observation direction while the magnetic field is perpendicular to that plane and satisfies the condition [1.11].

The angular dependence of the radiated pattern, proportional to $\sin^2 \theta$, indicates that the dipole does not radiate in the direction of the dipole. The typical emission pattern is shown in Figure 3.

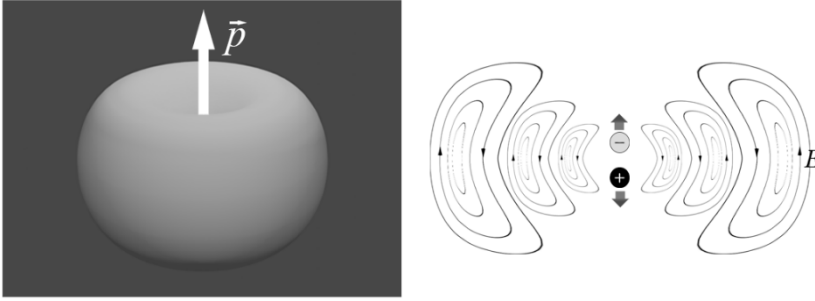


Figure 3.3
Radiation pattern of the
EM field generated by an
oscillating dipole

The Poynting vector is calculated directly from the expression of the fields:

$$\vec{S}_{rad} = \frac{\vec{e}_{rad} \wedge \vec{b}_{rad}}{\mu_0} = \frac{\mu_0 (\ddot{p})^2 \sin^2 \theta}{16\pi^2 R^2 c} \hat{R} = \frac{(\ddot{p})^2 \sin^2 \theta}{16\pi^2 \epsilon_0 c^3 R^2} \hat{R} \quad [3.8]$$

This is a general expression for any temporal dependence of the dipole moment.

A very important magnitude of radiation theory is the **total power radiated** by a dipole:

$$P_{rad} = \iint |\vec{S}_{rad}| ds = \int_0^\pi \int_0^{2\pi} \frac{\mu_0 (\ddot{p})^2 \sin^2 \theta}{16\pi^2 c R^2} R^2 \sin \theta d\theta d\phi = \frac{\mu_0 (\ddot{p})^2 2\pi}{16\pi^2 c} \int_0^\pi \sin^3 \theta d\theta$$

where we integrate the Poynting vector over all directions in space and use the differential surface element in spherical coordinates $ds = R^2 \sin \theta d\theta d\phi$. The integral on the azimuthal angle θ can be easily evaluated:

$$\int_0^\pi \sin^3 \theta d\theta = \int_0^\pi (1 - \cos^2 \theta) \sin \theta d\theta = \int_0^\pi \sin \theta d\theta - \int_0^\pi \cos^2 \theta \sin \theta d\theta = \left[-\cos \theta + \frac{\cos^3 \theta}{3} \right]_0^\pi = \frac{4}{3}$$

thus leading to the expression

$$P_{rad} = \frac{\mu_0 (\ddot{p})^2}{6\pi c} = \frac{(\ddot{p})^2}{6\pi \epsilon_0 c^3} \quad \text{Larmor's radiation formula} \quad [3.9]$$

3.2 Radiation properties of harmonic dipoles

Since dipoles oscillating with harmonic motion generate EM waves, we can assume an explicit temporal dependence of the dipole moment:

$$p(t) = p_0 \cos(\omega t + \phi)$$

which is expressed in complex notation as

$$p^{(+)}(t) = p_0^{(+)} e^{-i\omega t} \quad \text{and} \quad p(t) = \text{Re}[p^{(+)}(t)].$$



Using this temporal dependence and remembering that the dipole moment is evaluated at the retarded time t' , we obtain from [3.7A]:

$$\begin{aligned}\vec{e}_{rad}^{(+)}(\vec{r}, t) &= -\frac{\mu_0 \omega^2 p_0^{(+)} \sin \theta}{4\pi R} e^{-i\omega t'} \hat{e} = -\frac{\mu_0 \omega^2 p_0^{(+)} \sin \theta}{4\pi R} e^{-i\omega(t-R/c)} \hat{e} = \\ &= -\frac{\mu_0 \omega^2 p_0^{(+)} \sin \theta}{4\pi R} e^{i(kR-\omega t)} \hat{e} \\ \vec{b}_{rad}^{(+)}(\vec{r}, t) &= -\frac{\mu_0 \omega^2 p_0^{(+)} \sin \theta}{4\pi R c} e^{i(kR-\omega t)} \hat{b}\end{aligned}\quad [3.10]$$

The instantaneous Poynting vector is

$$\vec{S}(t) = \frac{\vec{e}_{rad} \wedge \vec{b}_{rad}}{\mu_0} = \frac{\mu_0 \omega^4 |p_0^{(+)}|^2 \sin^2 \theta}{16\pi^2 R^2 c} \cos^2(kR - \omega t) \hat{R}$$

which allows evaluating the intensity of the radiated emission (as discussed in Chapter 1):

$$I = \langle |\vec{S}(t)| \rangle = \frac{1}{2} \text{Re} \left[\frac{|\vec{e}^{(+)} \wedge \vec{b}^{(-)}|}{\mu_0} \right] = \frac{\mu_0 \omega^4 |p_0^{(+)}|^2 \sin^2 \theta}{32\pi^2 R^2 c}$$

where $\langle \dots \rangle$ denotes the temporal average.

By integrating the intensity in all directions in space, we obtain the **total averaged radiated power**:

$$\langle P_{rad} \rangle = \iint Ids = \int_0^{2\pi} \int_0^\pi IR^2 \sin \theta d\theta d\phi$$

Substituting I and integrating over the azimuthal angle, we have

$$\langle P_{rad} \rangle = \frac{\mu_0 \omega^4 |p_0^{(+)}|^2}{12\pi c} = \frac{\omega^4 |p_0^{(+)}|^2}{12\pi \epsilon_0 c^3} \quad [3.11]$$

3.2.1 Light emission from single atoms and molecules

The results obtained from the classical theory can be used to describe the light emitted by single atoms or molecules at distances where the dipolar and radiation zone approximation is valid.

The model of an atom has been changing over time as knowledge of atomic physics has developed. However, Bohr's model remains quite an accurate representation: electrons in an atom occupy electronic orbitals with well-defined energy levels, which are determined as a function of the basic quantum numbers (n, l, m, s) . Each energy level is defined by its energy E_i and can be occupied by a single electron, as stated by Pauli's Exclusion Principle. In the quantum mechanics framework, these energy levels constitute the eigenstates of the atom and have a well-defined energy value.



If we want to model molecules instead of considering a single isolated atom, we need to incorporate more energy levels representing the vibrational and rotational states of the molecule around its equilibrium positions. These levels can be modelled as generated from harmonic oscillation potentials.

The classical approach used to model atomic oscillations was defined by Lorentz, and it considers that a classical oscillator model can describe the electronic motion in the atom or molecule under excitation (see Appendix C for details). Radiation emitted by excited atoms, molecules, or solids corresponds to jumps between stationary states of the system, with well-defined frequencies characterized by the formula

$$E_i - E_f = \hbar\omega_o$$

with Planck's constant $\hbar = 1.054 \cdot 10^{-34} \text{ J} \cdot \text{s} = 6.582 \cdot 10^{-16} \text{ eV} \cdot \text{s}$, and E_i and E_f being the initial and final energy states of the system.

Let us consider an atom that is making a transition characterized by the emission of light at frequency ω_0 . In order to evaluate the magnitude of the dipole moment in the classical electrodynamics approach, we associate to this particular emission an oscillating dipole:

$$\vec{p}(t) = \text{Re}[\vec{p}^{(+)}(t)] = \text{Re}[qx^{(+)}(t)\hat{d}] \quad [3.12]$$

where q is the absolute value of the electron charge, x is the displacement of the negative charge with respect to the positive one, and $\hat{d}$ is a unitary vector that at any time is directed from the negative to the positive charge.

The expression for $x(t)$ is obtained directly from the harmonic oscillator equation of Lorentz's model:

$$\frac{d^2 x^{(+)}}{dt^2} + \omega_0^2 x^{(+)} = 0$$

whose solution is $x^{(+)}(t) = x_0^{(+)} e^{-i\omega_0 t}$,

thus creating an oscillating dipole moment:

$$\boxed{\vec{p}^{(+)} = qx_0^{(+)} e^{-i\omega_0 t} \hat{d}} \quad [3.13]$$

◆ *Comment*

Atoms exhibit sharp resonances both in their spontaneous radiation emission and in their stimulated response to applied signals, which are usually described in this model by simple harmonic resonances.

Most atoms respond primarily to the electric field of an applied signal rather than to the magnetic field. The strongest atomic transitions are usually of the type known as **electric dipole transitions** (other types of atomic transitions are classified as magnetic dipole,



electric quadrupole, or even higher-order). These transitions can be classically studied through the classical electron oscillator model, in which the natural oscillation frequencies of oscillation correspond to the atomic resonances. With some simple extensions, this simplified model can describe every significant feature of a real atomic quantum transition.

From this expression, we can directly obtain the radiated electromagnetic fields by substituting the second derivative of the dipole moment in expression [3.7A]:

$$\begin{aligned}\vec{e}_{rad}^{(+)} &= \frac{\mu_o \omega_o^2 q x_o^{(+)} \sin \theta}{4\pi R} e^{-i\omega_o(t-R/c)} \hat{e} \\ \vec{b}_{rad}^{(+)} &= \frac{\mu_o \omega_o^2 q x_o^{(+)} \sin \theta}{4\pi R c} e^{-i\omega_o(t-R/c)} \hat{b}\end{aligned}\quad [3.14]$$

To obtain this expression, the dipole moment is evaluated at the retarded time $t' = t - R/c$ and a phase factor of π is omitted, which does not affect the direction of oscillation of the dipole.

The total averaged radiated power is readily obtained from expression [3.11]:

$$\langle P_{rad} \rangle = \frac{\mu_o \omega_o^4 q^2 |x_o^{(+)}|^2}{12\pi c} \quad [3.15A]$$

In the case of a real atom, we have several possible transitions between energy levels, so a more realistic solution should include this possibility. The total radiated power from a given initial state can thus be obtained by summing over all these possible emission paths. This idea can be refined somewhat in order to obtain a quantitative agreement between this radiated power and the radiated power calculated through a quantum mechanical approach, namely by introducing a weight for each possible transition, called the strength factor: f_i with

$$\sum_i f_i = 1$$

The total radiated power from a given initial level is then given by the expression

$$\langle P_{rad}^{TOT} \rangle = \sum_i f_i \cdot \frac{\mu_o \omega_{oi}^4 q^2 |x_{oi}^{(+)}|^2}{12\pi c} \quad [3.15B]$$

Still, an additional term should be included in the model in order to obtain a good correspondence between this classical description and the more exact quantum mechanical calculation. Since we have not included losses in the preceding calculation, emission could take place over an infinite period of time. Experimentally, however, it is well established that no electronic population can be permanently maintained in an excited state and the excited state will be depleted at a given decay rate, which is called the **spontaneous emission decay rate**.



3.2.2 Emission from damped harmonic dipole oscillators

This mechanism of damping is derived from first principles only in a quantum mechanical approach. In the classical model, it must be incorporated in a phenomenological way. The two different approaches below can be stated in order to justify (although not rigorously) the presence of such damping in the atomic emission process.

A. Energy decay rate

Let us consider that a given amount of energy has been gained by the atom, so part of its population is transferred to excited states. In our Lorentz oscillator model, we can define an initial energy for the oscillators at frequency ω_0 as

$$E = \frac{1}{2} m \omega_0^2 x_0^2$$

In terms of this energy, we can write the total radiated power in the form

$$\langle P_{rad} \rangle = \frac{\mu_0 q^2 \omega_0^2}{12\pi c} \omega_0^2 x_0^2 = \frac{\mu_0 q^2 \omega_0^2}{6\pi c m} E \quad [3.16]$$

In the absence of other mechanisms of energy losses, we can consider that any reduction of the initial energy of the system must appear as energy radiated by the atom, leading to a decrease in the energy of the oscillating dipole. Thus, we can consider the total radiated power as a rate of energy loss per unit time:

$$\begin{aligned} \langle P_{rad} \rangle &= -\frac{dE}{dt} = \frac{\mu_0 q^2 \omega_0^2}{6\pi c m} E \\ \frac{dE}{E} &= -\frac{\mu_0 q^2 \omega_0^2}{6\pi c m} dt \quad \Rightarrow \quad \boxed{E(t) = E e^{-\gamma t}} \end{aligned}$$

where we define the classical expression of the spontaneous decay rate as

$$\gamma = \frac{\mu_0 \omega_0^2 q^2}{6\pi m c} = \frac{\omega_0^2 q^2}{6\pi m \epsilon_0 c^3} \quad [3.17]$$

◆ Comment

The emission of radiation is not the only possible mechanism for the loss of the excited electronic population in a material system. Other possibilities are, for instance, the loss of energy by collisions with other atoms in what are called **non-radiative transitions**. A total decay rate can be defined in this case as

$$\gamma_{tot} = \gamma_{rad} + \gamma_{non\ rad}$$

It is this total decay rate that is measured in, for instance, a typical fluorescence decay measurement. These measurements are very important in order to determine the lifetime of a given excited state.



B. Radiation reaction

An alternative way to obtain the damping coefficient of atomic oscillators considers that the harmonic dipole loses energy due to the presence of a fictitious friction force acting on the atom:

$$F_f = -bv = -\gamma mv$$

This fictitious friction force can be interpreted as a consequence of the reaction of the radiated field on the atom itself. We can quantify the magnitude of this loss with an energy conservation argument: the work performed by the friction force on the atom during an oscillation period will result in an energy loss of the oscillator; since radiation is the only way to lose energy in an isolated atom, the total radiated energy over a period should equal the work done on the dipole. This argument can be cast in mathematical form:

$$[Total\ work\ exerted\ by\ friction\ force = Total\ radiated\ energy]_{over\ a\ period}$$

$$\int_0^T F_f \cdot v dt = \int_0^T P_{rad} dt$$

Using [3.9]:

$$\int_0^T \gamma m \dot{x} \cdot \dot{x} dt = \int_0^T \frac{\mu_o q^2 (\ddot{x})^2}{6\pi c} dt$$

Integrating by parts the r.h.s. to obtain

$$\gamma m \int_0^T \dot{x} \cdot \dot{x} dt = \frac{\mu_o q^2}{6\pi c} \left([\dot{x} \cdot \ddot{x}]_0^T - \int_0^T \dot{x} \cdot \dddot{x} dt \right)$$

Considering that the atom is performing harmonic motion (very slowly damped), the expression in brackets cancels out when evaluated at between 0 and T , and the third order derivative can be written as

$$\ddot{x} = -\omega_o^2 x$$

leading to the final result

$$\gamma \int_0^T \dot{x}^2 dt = \frac{\mu_o q^2 \omega_o^2}{6\pi m c} \int_0^T \dot{x}^2 dt \quad \Rightarrow \quad \boxed{\gamma = \frac{\mu_o q^2 \omega_o^2}{6\pi m c}}$$

Substituting typical values for atomic transitions into this expression gives decay rate values of the order of $10^8\ s^{-1}$ (see Solved problems for details).

Once we have obtained an expression for the radiative emission decay rate, we can write an explicit expression for the emission of radiation by a damped harmonic oscillator, in which we include an extra term in the Lorentz model equation to include damping:



$$\frac{d^2 x^{(+)}}{dt^2} + \gamma \frac{dx^{(+)}}{dt} + \omega_o^2 x^{(+)} = 0$$

The solution of this equation reads

$$x^{(+)}(t) = x_o^{(+)} e^{-\frac{\gamma}{2}t} e^{-i\omega t} \quad \text{with} \quad \omega^2 = \omega_o^2 - \left(\frac{\gamma}{2}\right)^2$$

and

$$p^{(+)}(t) = qx_o^{(+)} e^{-\frac{\gamma}{2}t} e^{-i\omega t} \quad [3.18]$$

In order to calculate the radiation fields and the total radiated power, we perform the second derivative of the dipole moment. Considering the typical values of the damping constant and the optical frequencies in the visible spectrum, we use the approximation $\gamma \ll \omega \approx \omega_o$:

$$\ddot{p}^{(+)} = \left(-\frac{\gamma}{2} - i\omega\right)^2 p^{(+)} \simeq -\omega_o^2 p^{(+)}$$

Emission of radiation by damped harmonic oscillator

$$\boxed{\begin{aligned} \vec{e}_{rad}^{(+)} &= \frac{\mu_o q x_o^{(+)} \omega_o^2 \sin \theta}{4\pi R} e^{-\frac{\gamma t}{2}} e^{i(kR - \omega_o t)} \hat{e} \\ \vec{b}_{rad}^{(+)} &= \frac{\mu_o q x_o^{(+)} \omega_o^2 \sin \theta}{4\pi R c} e^{-\frac{\gamma t}{2}} e^{i(kR - \omega_o t)} \hat{b} \\ \langle P_{rad} \rangle &= \frac{\mu_o q^2 \omega_o^4 |x_o^{(+)}|^2}{12\pi c} e^{-\gamma t} \end{aligned}} \quad [3.19]$$

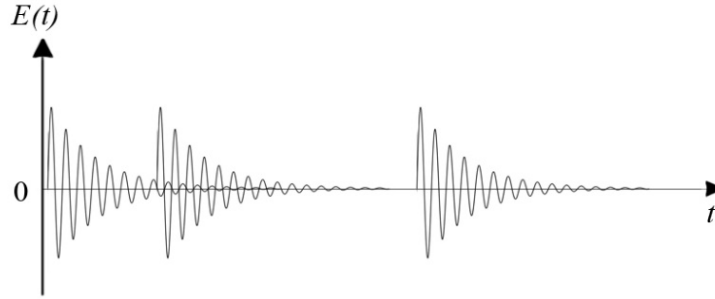
This is the central result of this section. Below, we summarize the main properties of the radiation from a unitary harmonic dipole oscillator.

- Radiated fields proportional to $1/R$.
- Angular emission field pattern proportional to $\sin\theta$. No radiation is emitted in the direction of the oscillation (compared to the scalar solution of a spherical wave, which emits in all directions).
- Radiated power proportional to the fourth power of the frequency (this dependence is typical of small dipoles at the limit named Rayleigh scattering, as we will see).
- In absence of other loss mechanisms, the radiated power decays at a constant rate given by the spontaneous emission decay rate.

If the atom is excited continuously by some means, the emitted light will generally consist of a set of damped bursts of EM radiation with typical durations in the range of tenths of a nanosecond, as schematically shown in the figure.



Figure 3.4
Schematic representation of light emitted by a damped harmonic oscillator



3.2.3 Natural linewidth of the emitted radiation

In the previous section, we derived the damping constant associated with the emission of light from oscillating dipoles. Aside from the temporal decay of the radiated power, another consequence of this damped emission is that the radiated light is not purely monochromatic but instead has a finite bandwidth.

The **bandwidth** of a signal is a measure of its frequency content. In the case we are interested in, it is related to the finite lifetime of any excited state. The bandwidth associated with a pure radiative transition is called the **natural linewidth**.

The bandwidth of a given temporal signal can be obtained by Fourier transformation of the temporal field:

$$e(\omega) = \int_{-\infty}^{\infty} e_{rad}(t) e^{i\omega t} dt = \frac{\mu_o q \omega_o^2 x_o^{(+)} \sin \theta}{4\pi R} \int_0^{\infty} e^{-\frac{\gamma}{2}t} \cos(\omega_o t) e^{i\omega t} dt$$

$$e(\omega) = \frac{\mu_o q \omega_o^2 x_o^{(+)} \sin \theta}{8\pi R} \left[\frac{1}{\frac{\gamma}{2} + i(\omega_o - \omega)} + \frac{1}{\frac{\gamma}{2} - i(\omega_o + \omega)} \right]$$

Since the emitted frequencies are close to the resonance value, the first term in brackets is much larger than the second one. It is customary to neglect the second term in what is called the **rotating wave approximation**.

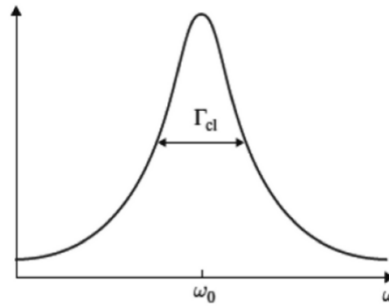
The spectrum of the radiated emission is defined as proportional to the modulus square of its amplitude $e(\omega)$, giving

$$I(\omega) = \frac{I_o \omega_o^4}{\frac{\gamma^2}{4} + (\omega_o - \omega)^2} \quad \text{with} \quad I_o = \left(\frac{\mu_o^2 q^2 |x_o^{(+)}|^2 \sin^2 \theta}{64\pi^2 R^2} \right)$$

The spectrum associated with the natural linewidth has a characteristic frequency dependence that is expressed as a **Lorentzian profile with a full width half maximum given by the damping constant**.

Since typical damping constant rates are of the order of a few GHz, we can estimate the relative bandwidths of a purely radiative transition:

$$\frac{\Delta\omega}{\omega} \approx \frac{10^9}{10^{15}} \approx 10^{-6}$$



Emission by a collection of n atoms

A detailed treatment of the emission from a dense number of emitters is out of the scope of an introductory course. We can point few general aspects

Any real source is made of a large number of emitters, each of them emitting in the way described. The superposition of the emission from all individual scatterers results in light emitted from natural sources, which in general has the following characteristics.

- Incoherence: The total intensity is the sum of the individual intensities.
- Not polarized: No fixed phase relation exists between the individual emitters.

Additional broadening mechanisms:

- Doppler broadening
- Collisional broadening

3.3 Interaction of a single atom with an incident light field

3.3.1 Scattering and absorption

After studying the emission properties of an oscillating dipole, we now consider the process that describes the interaction of a light wave with an atom. This fundamental process forms the basis of the mechanism for light-matter interaction, which is highly relevant in photonics.

The starting point of the classical treatment is again the Lorentz model, into which we now incorporate a forcing term provided by the action of the external electric field on the atom.



Consider an external electric field in the form of a plane wave with external frequency ω_e , linear polarization, and propagating along a direction given by the z -axis. When this field is incident on an atom (modelled as an oscillating dipole, in accordance with Lorentz's model), it induces the oscillation of the electronic cloud driven by the electrical force, described by the equation

$$\frac{d^2 \vec{x}}{dt^2} + \gamma \frac{d\vec{x}}{dt} + \omega_o^2 \vec{x} = -\frac{q\vec{E}_{ext}}{m} \quad [3.21A]$$

Since the dipole moment is given directly by the expression $\vec{p} = -q\vec{x}$, we directly obtain an equation for the dipole moment evolution:

$$\frac{d^2 p}{dt^2} + \gamma \frac{dp}{dt} + \omega_o^2 p = \frac{q^2 E_{ext}}{m} \quad [3.21B]$$

which is now an equation for the modulus, since the direction of p is always parallel to E . The expression for the external field is

$$E_{ext}^{(+)} = E_o^{(+)} e^{-i\omega_e t} \quad [3.21C]$$

This field will induce a harmonic-type solution for the dipole moment, which can be expressed as

$$p^{(+)}(t) = p_o^{(+)} e^{-i\omega t} \quad [3.21D]$$

The general solution to the forced harmonic oscillator equation consists of the sum of a solution to the homogeneous equation and a particular solution:

$$p^{(+)} = p_{Hom}^{(+)} + p_{Part}^{(+)}$$

The homogeneous solution, oscillating at the natural resonant frequency of the system ω_o , corresponds to [3.18]. As we have discussed previously, its decay over time will be of the order of $1/\gamma$. If we are interested in the stationary response of the atoms after this transient period, we have to consider the particular solution, which will oscillate at the external frequency imposed by the external field.

Substituting the ansatz for $p^{(+)}$ (Equation [3.21D]) and $E^{(+)}$ (Equation [3.21C]) in [3.21B] with $\omega = \omega_e$ gives

$$-\omega_e^2 p_o^{(+)} - i\gamma\omega_e p_o^{(+)} + \omega_o^2 p_o^{(+)} = \frac{q^2 E_o^{(+)}}{m} \Rightarrow \boxed{p_o^{(+)} = \frac{q^2 E_o^{(+)} / m}{(\omega_o^2 - \omega_e^2) - i\gamma\omega_e}} \quad [3.21E]$$

The dipole oscillation corresponding to this stationary solution is used to find the **scattered fields** from the single atom:

$$\begin{aligned} \vec{e}_{scatt}^{(+)} &= \frac{\mu_o \omega_e^2 \sin \theta}{4\pi R} \left[\frac{q^2 E_o^{(+)} / m}{(\omega_o^2 - \omega_e^2) + i\gamma\omega_e} \right] e^{i(kR - \omega_e t)} \hat{u}_e \\ \vec{b}_{scatt}^{(+)} &= \frac{\mu_o \omega_e^2 \sin \theta}{4\pi R c} \left[\frac{q^2 E_o^{(+)} / m}{(\omega_o^2 - \omega_e^2) + i\gamma\omega_e} \right] e^{i(kR - \omega_e t)} \hat{u}_b \end{aligned} \quad [3.22]$$



Although the expression corresponding to the scattered fields shares some similar behaviour with the radiated fields [3.19], there is a substantial difference in their frequency dependence, which is a fundamental aspect for differentiating light emission and light-matter interaction processes. Atoms acting as light sources can emit only those frequencies corresponding to jumps between their stationary energy levels; these jumps may correspond to very well-defined discrete levels (as in the case of atoms) or to some quasi-continuous bands (as in the case of solids).

When the atoms are interacting with an external EM wave of arbitrary frequency, we may distinguish the following two different basic processes.

Absorption: The external frequency corresponds to energy that is resonant with some jump between the stationary energy levels of the medium. In this case, the EM energy can be transferred from the EM wave to the atomic part of the electronic population that is being excited. The absorbed energy can be reemitted in the form of EM radiation, or it can be transferred to the lattice through phonon excitation and finally converted into a variation of the internal energy of the material.

For instance, it is well known that if we illuminate some gas with a light source that has a continuous spectrum, we will observe in the transmission a set of dark lines corresponding precisely to the frequencies absorbed by the gas, which correspond exactly to the frequencies emitted by the same gas when it is excited by some means, as for example in a discharge lamp.

Scattering: The external frequency is not resonant with any of the electronic transitions between levels. In these conditions, the effect of the external electric field is still an induced oscillation of the electronic orbitals, but in this case there is no net transfer of energy to the atom and no electronic transition is induced. The energy of the EM wave is re-emitted in all directions (except in the direction of the dipole oscillation) after a period of time that is on the order of the spontaneous emission decay rate.

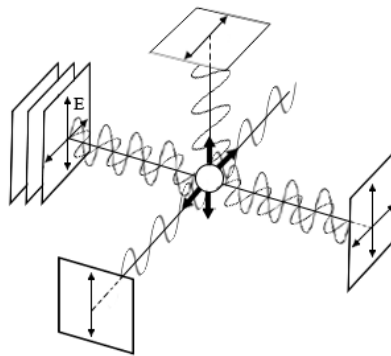


Figure 3.5
Scattering of light by an oscillating dipole

The total radiated power in a scattering event may be obtained as usual:

$$\langle P_{scatt} \rangle = \frac{\mu_0 \omega_e^4 |p_o^{(+)}|^2}{12\pi c} = \frac{\mu_0 \omega_e^4}{12\pi c} \left[\frac{q^4 |E_o^{(+)}|^2 / m^2}{(\omega_e^2 - \omega_o^2)^2 + (\gamma \omega_e)^2} \right]$$



and expressed in terms of the incident intensity. For a plane wave $I_{inc} = \frac{1}{2} \epsilon_0 c |E_o^{(+)}|^2$,

the **total scattered power of a single dipole** is written as

$$\langle P_{scatt} \rangle = \frac{\mu_0 q^4}{6\pi \epsilon_0 c^2 m^2} \left[\frac{\omega_e^4}{(\omega_e^2 - \omega_o^2)^2 + \gamma^2 \omega_e^2} \right] \cdot I_{inc} \quad [3.23A]$$

In scattering theory, it is usual to work with a related magnitude known as the **scattering cross section** (with units of area), which represents the ratio of the total scattered power to the incident intensity. This magnitude defines the response of an atom to the incident field in the dipolar classical approach:

$$\sigma_{scatt} = \frac{\langle P_{scatt} \rangle}{I_{inc}} = \frac{q^4}{6\pi \epsilon_0^2 c^4 m^2} \left[\frac{\omega_e^4}{(\omega_e^2 - \omega_o^2)^2 + \gamma^2 \omega_e^2} \right] \quad [3.23B]$$

The behaviour of the atomic response depends strongly on the frequency, and we usually distinguish the following three different regimes when comparing the external frequency with the resonant frequency of the dipole.

1. Thomson scattering regime ($\omega_e \gg \omega_o \gg \gamma$)

In this regime, the incident frequency is higher than both the resonant frequency and the decay rate of the dipole. The scattering cross section is approximately constant and is called the Thomson scattering cross section:

$$(\omega_e^2 - \omega_o^2)^2 + \gamma^2 \omega_e^2 \approx \omega_e^4 \left(1 + \frac{\gamma^2}{\omega_e^2} \right) \approx \omega_e^4 \quad [3.24A]$$

$$\sigma_{Thomson} = \frac{q^4}{6\pi \epsilon_0^2 c^4 m^2}$$

2. Resonant scattering ($\omega_e \approx \omega_o$)

In this regime, the effect of scattering increases strongly and, at resonance energy, it is effectively transferred to the medium in the process of absorption. We can use the following approximation:

$$(\omega_e^2 - \omega_o^2) \approx (\omega_e - \omega_o)(\omega_e + \omega_o) \approx 2\omega_e(\omega_e - \omega_o)$$

$$\frac{\omega_e^4}{(\omega_e^2 - \omega_o^2)^2 + \gamma^2 \omega_e^2} \approx \frac{\omega_e^4}{4\omega_e^2(\omega_e - \omega_o)^2 + \gamma^2 \omega_e^2} \approx \frac{\omega_e^2 / 4}{(\omega_e - \omega_o)^2 + (\gamma/2)^2}$$

The resulting scattering cross section acquires the profile of a Lorentzian:



$$\sigma_{abs} = \frac{q^4}{6\pi\epsilon_0^2 c^4 m^2} \left[\frac{(\omega_e / 2)^2}{(\omega_e - \omega_o)^2 + (\gamma / 2)^2} \right] \quad [3.24B]$$

3. Rayleigh scattering ($\omega_e, \gamma \ll \omega_o$)

At frequencies well below the resonant frequency, we are in the domain of Rayleigh scattering. In this regime:

$$\frac{\omega_e^4}{(\omega_e^2 - \omega_o^2)^2 + \gamma^2 \omega_e^2} \approx \frac{\omega_e^4}{\omega_o^4}$$

We write the corresponding Rayleigh scattering cross section as

$$\sigma_{Rayleigh} = \frac{q^4 \omega_e^4}{6\pi\epsilon_0^2 c^4 m^2 \omega_o^4} \quad [3.24C]$$

This regime is characterized by a scattering cross section proportional to the fourth power of the frequency, which for example is responsible for the blue colour of the sky.

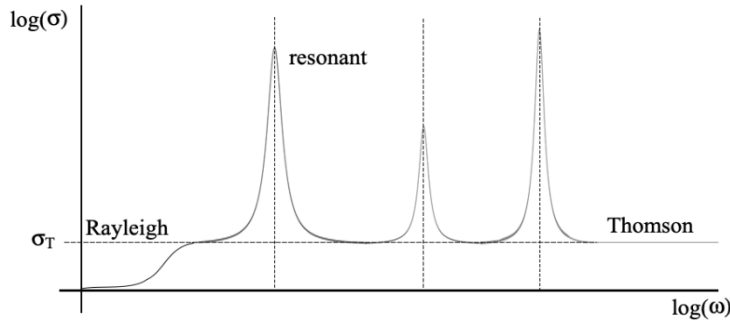


Figure 3.6
Different regimes of the scattering cross section

3.3.2 Power transfer from an external field

As the energy scatters out from the atom, the energy of the incident field decreases. We may calculate the power transfer from the electric field to the atom:

$$P_{transfer} = \vec{F}_{ext} \cdot \vec{v} = -q\vec{E}_{ext} \cdot \dot{\vec{x}} = E_{ext} \cdot \dot{p} \quad [3.25]$$

Since this is a bilinear product of the fields, real values should be used for their evaluation:

$$P_{transfer} = \frac{1}{4} \left[E_{ext}^{(+)} \cdot \dot{p}^{(+)} + E_{ext}^{(+)} \cdot \dot{p}^{(-)} + E_{ext}^{(-)} \cdot \dot{p}^{(+)} + E_{ext}^{(-)} \cdot \dot{p}^{(-)} \right]$$

The optical fields and dipole moment oscillate at very large frequencies of the order of 10^{15} Hz, so the instantaneous power is the averaged value and not a physically accessible magnitude. The first and fourth terms of this expression average out to zero, while the



second and third are complex conjugate of each other. The averaged power transfer is then written as

$$\langle P_{transfer} \rangle = \frac{1}{2} \text{Re} [E_{ext}^{(+)} \cdot \dot{p}^{(-)}] \quad [3.26]$$

From equations [3.21D] and [3.21E], we get

$$\dot{p}^{(-)} = i\omega_e \frac{q^2 E_o^{(-)} / m}{(\omega_o^2 - \omega_e^2) + i\gamma\omega_e}$$

$$\langle P_{transfer} \rangle = \frac{1}{2} \text{Re} \left[i \frac{\omega_e q^2 |E_o|^2}{m} \left(\frac{(\omega_o^2 - \omega_e^2) - i\gamma\omega_e}{(\omega_o^2 - \omega_e^2)^2 + \gamma^2\omega_e^2} \right) \right] = \frac{q^2 |E_o|^2}{2m} \left[\frac{\gamma\omega_e^2}{(\omega_o^2 - \omega_e^2)^2 + \gamma^2\omega_e^2} \right]$$

This can be expressed in terms of the incident intensity:

$$\langle P_{transfer} \rangle = \frac{q^2 \gamma \omega_e^2}{\epsilon_o c m [(\omega_o^2 - \omega_e^2)^2 + \gamma^2 \omega_e^2]} I_{inc} \quad [3.27]$$

The **extinction cross section** is defined as the ratio of the power transfer to the incident intensity:

$$\sigma_{ext} = \frac{\langle P_{transfer} \rangle}{I_{inc}} = \frac{q^2}{\epsilon_o c m} \left[\frac{\gamma \omega_e^2}{(\omega_o^2 - \omega_e^2)^2 + \gamma^2 \omega_e^2} \right] \quad [3.28]$$

We have discussed different properties related to the radiation of single dipoles, as well as the interaction between external fields and single dipoles. In this case, there are no other possible mechanisms for any energy transfer to the surroundings, so the total energy transferred to the dipole should ultimately radiate outward. By this argument, we can see that, for a single atom, the scattering and extinction cross section should be equal. Comparison of these two magnitudes gives an alternative way to obtain an expression for the spontaneous emission decay rate. Equating expressions [3.23B] and [3.28]:

$$\frac{q^4 \omega_e^4}{6\pi \epsilon_o^2 m^2 c^4} = \frac{q^2 \gamma \omega_e^2}{\epsilon_o c m} \Rightarrow \gamma = \frac{q^2 \omega_e^2}{6\pi \epsilon_o m c^3} = \frac{\mu_o q^2 \omega_e^2}{6\pi m c}$$

3.3.3 Interaction of light with N atoms

We will briefly consider some effects of an EM optical wave interacting with a collection of atoms. In this case, new mechanisms of energy released from the excited atoms can involve non-radiative transitions. What is more, different linewidth broadening mechanisms must be considered and local field effects could also be important, among many other effects.

Let us consider a light beam crossing a cell filled with a gas. Each one of the atoms will perform forced oscillations (as explained in the previous sections) while the intensity of the radiation will decrease along the initial direction of propagation. The net effect of the incident wave's intensity loss is called extinction and can be described by an extinction cross section in the same way as derived previously for a single atom. In order to derive



an expression for the dependence of light intensity as a function of distance within the medium, let us consider a thin slab of the cell, of length Δx and perpendicular to the propagation direction. If the atom density in the slab is N , the loss of electromagnetic radiation is written as

$$u(x + \Delta x) - u(x) \approx -NP_{transfer}\Delta t = -N\sigma_{ext}I(x)\Delta x / c$$

where σ_{ext} is the extinction coefficient of a single atom and Δt is the time taken to cross the slab. Dividing by Δx and taking the limit $\Delta x \rightarrow 0$, we obtain the differential equation

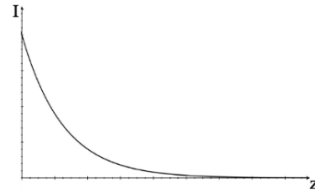
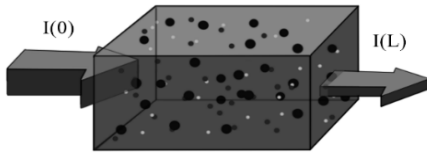
$$\frac{du}{dx} = -\frac{N\sigma_{ext}I(x)}{c}$$

This equation can be transformed to an equation for the intensity change by recognizing that

$$I(x) = c \cdot u(x)$$

$$\frac{dI}{dx} = -N\sigma_{ext}I \quad \rightarrow \quad \boxed{I(x) = I_0 e^{-\alpha_{ext}x}} \quad [3.29]$$

This relation is the **Beer–Lambert law**, which defines an optical wave's extinction with distance in a material medium.



The total extinction coefficient, α_{ext} , is defined as N times the extinction cross section of a single dipole, in units of m^{-1} .

The extinction of light in the direction of propagation is mainly due to two different processes: **scattering and absorption**:

$$\alpha_{ext} = \alpha_{scat} + \alpha_{abs}$$

Scattering

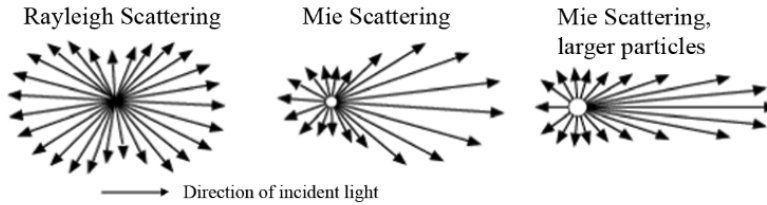
The process of scattering corresponds to the forced emission of radiation from each single dipole, as studied previously, and it leads to a re-distribution of the incident radiation. We distinguish the following different kinds of scattering.

Elastic scattering: The emitted radiation frequency is equal to the frequency of the incident wave.

Rayleigh scattering ($l \ll \lambda$): Emission proportional to ω^4 and isotropic.

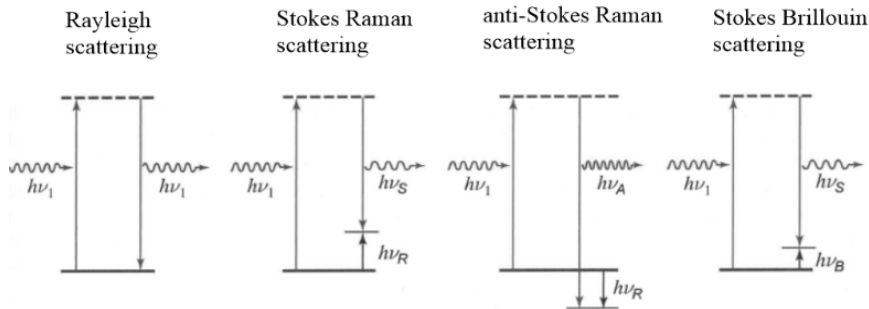


Mie scattering ($l \geq \lambda$): Emission independent of frequency and directed mainly in the forward direction.



Inelastic scattering: When vibrational states or phonons are present in the medium, the frequency of the scattered light can be different from that of the incident light. The two basic types are **Raman scattering** and **Brillouin scattering**.

In Raman scattering, the difference in energies between the initial and final states corresponds to some vibrational transition of the medium. Raman spectroscopy is one of the most sensitive techniques used nowadays in photonics technology, and it has many different applications. In Brillouin scattering, the difference in energies corresponds to phonon excitation in the lattice.



Absorption

The incident energy is transferred to the dipoles and is converted into thermal, acoustical, or internal vibrational energy within the medium. Absorption takes place at frequencies close to the medium resonances. Typical resonance transitions for materials that are transparent in the visible domain lie within the ultraviolet (electronic resonances) or infrared (vibrational and rotational) parts of the spectrum.

For dense media, the effects of absorption and scattering are much richer and more complicated, requiring in-depth study that is beyond the scope of this introductory course. However, a particular aspect of this interaction is worth commenting on. From the viewpoint of classical electrodynamics, a light wave that is incident on a medium forces the oscillation of the medium's elementary dipoles, which emit light. The resulting field in the medium is the sum of the radiation from all charges and the incident field. When the medium is homogeneous, the **Ewald–Oseen extinction theorem** applies: when an EM



wave is incident on a homogeneous medium with a sharp boundary, it is extinguished inside the medium through the interaction process and is replaced by a wave propagating in the medium at a velocity and in a direction that are different from those of the incident wave.

This impressive result makes it possible to define a particular trajectory for light beams in dense transparent media. When the medium is not completely homogeneous, the inhomogeneities cause some residual scattering and rectilinear propagation.

3.4 Maxwell equations in macroscopic media

After studying the light generation process and the interaction of single atoms with external fields, we now face the problem of an EM wave propagating in a medium composed of a much greater number of atoms. In this case, one should consider the interactions among all the charged particles present in the medium. The number of charges in a unit volume of matter (10^{22} – 10^{23} atoms/cm³ in dense media), even for diluted gases (usually $> 10^8$ atoms/cm³), makes it impractical to attempt following the individual interactions with each charge. The total microscopic field at any point in the medium is the superposition of all the microscopic fields generated by the atoms in response to the incident field. The resultant field shows large random fluctuations from point to point.

However, when we are interested in the field values in regions much larger than the typical atomic sizes, we can obtain a good representation of the resulting macroscopic fields by determining the averaged values of the microscopic fields over volumes that should be large in order to include a high enough number of charges, as this obtains a well-behaved field. The values must also be smaller than the wavelength so that we can define an effective electric field distribution as a continuous variable in space.

For light waves, the macroscopic fields are obtained by averaging over volumes of dimensions of the order of 1 nm³. In this sense, we can move to the macroscopic magnitudes obtained in this averaging procedure.

Macroscopic fields:

$$\vec{E}(\vec{r}, t) = \frac{1}{\Delta V} \int_{\Delta V} \vec{e}(\vec{r}', t) d^3 r' \quad \vec{B}(\vec{r}, t) = \frac{1}{\Delta V} \int_{\Delta V} \vec{b}(\vec{r}', t) d^3 r'$$

and we can define macroscopic charge density and currents in a similar way.

The **macroscopic Maxwell equations** are written for these averaged effective magnitudes as

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho_{macro}}{\epsilon_o} \quad [3.30A] \quad \vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad [3.30C]$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad [3.30B] \quad \vec{\nabla} \wedge \vec{B} = \mu_o \vec{j}_{macro} + \mu_o \epsilon_o \frac{\partial \vec{E}}{\partial t} \quad [3.30D]$$

where the macroscopic charge and current densities satisfy the continuity equation that expresses the charge conservation over a macroscopic volume:



$$\frac{\partial \rho_{macro}}{\partial t} + \vec{\nabla} \cdot \vec{j}_{macro} = 0$$

The total charge and current densities include the effects of bound and free charges. We can separate these two contributions to write

$$\begin{aligned}\rho_{macro} &= \rho_{free} + \rho_{bound} \\ \vec{j}_{macro} &= \vec{j}_{free} + \vec{j}_{bound}\end{aligned}$$

The effect of the bound charges can be adequately described through a polarization density existing in the medium, $\vec{P}$:

$$\rho_{bound} = -\vec{\nabla} \cdot \vec{P}$$

The bound currents can arise due to the temporal variation of the polarization density or from the existence of magnetic moments in the medium, resulting in a macroscopic magnetization density, $\vec{M}$:

$$\vec{j}_{bound} = \frac{\partial \vec{P}}{\partial t} + \vec{\nabla} \wedge \vec{M}$$

where we define

$$\vec{P} = N\vec{p} \quad \text{and} \quad \vec{M} = N\vec{m}$$

with N being the atomic density, p the atomic dipole moment, and m the atomic magnetic moment. These polarization and magnetization densities are created in the medium through the action of the incident external EM field.

♦ *Comment*

Although, the separation between free and bond charges is in some ways arbitrary and other approaches can be used in electromagnetic theory, it is usually adopted to describe light–matter interaction processes.

The effect of the currents generated by free electrons in the material, driven by the external fields, is generally expressed in terms of the material conductivity, σ :

$$\vec{j}_{free} = \sigma \vec{E}$$

Considering these definitions, we can rewrite Maxwell's equations by including these magnitudes as

$$\begin{aligned}\vec{\nabla} \cdot \vec{E} &= \frac{\rho_{free} - \vec{\nabla} \cdot \vec{P}}{\epsilon_0} & \vec{\nabla} \wedge \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \cdot \vec{B} &= 0 & \vec{\nabla} \wedge \vec{B} &= \mu_0 \sigma \vec{E} + \mu_0 \frac{\partial \vec{P}}{\partial t} + \mu_0 \vec{\nabla} \wedge \vec{M} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t}\end{aligned}$$



Regrouping some terms:

$$\begin{aligned}\vec{\nabla} \cdot (\epsilon_o \vec{E} + \vec{P}) &= \rho_{free} \\ \vec{\nabla} \wedge \left(\frac{\vec{B}}{\mu_o} - \vec{M} \right) &= \sigma \vec{E} + \frac{\partial(\epsilon_o \vec{E} + \vec{P})}{\partial t}\end{aligned}$$

We define below new variables describing the response of matter to light excitation.

Electric Displacement: $\vec{D} = \epsilon_o \vec{E} + \vec{P}$ (C/m²)

Magnetic Intensity: $\vec{H} = \frac{\vec{B}}{\mu_o} - \vec{M}$ (A/m)

In terms of the newly defined variables, we obtain the macroscopic Maxwell equations (in the absence of free charges; $\rho_{free} = 0$):

$$\boxed{\begin{aligned}\vec{\nabla} \cdot \vec{D} &= 0 & \vec{\nabla} \wedge \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \cdot \vec{B} &= 0 & \vec{\nabla} \wedge \vec{H} &= \sigma \cdot \vec{E} + \frac{\partial \vec{D}}{\partial t}\end{aligned}} \quad [3.31]$$

3.4.1 Wave equations for material media

From Maxwell's equations, we can now derive the corresponding wave equations that describe the propagation of the electric and magnetic fields in the medium. As we did with the corresponding microscopic equations, we start by taking the curl of Equation [3.30C]:

$$\begin{aligned}\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) &= -\vec{\nabla} \wedge \frac{\partial \vec{B}}{\partial t} = -\vec{\nabla} \wedge \frac{\partial(\mu_o \vec{H} + \mu_o \vec{M})}{\partial t} \\ \vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) &= -\mu_o \sigma \frac{\partial \vec{E}}{\partial t} - \mu_o \frac{\partial^2 \vec{D}}{\partial t^2} - \mu_o \frac{\partial(\vec{\nabla} \wedge \vec{M})}{\partial t} \\ \vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \mu_o \sigma \frac{\partial \vec{E}}{\partial t} + \mu_o \epsilon_o \frac{\partial^2 \vec{E}}{\partial t^2} &= -\mu_o \frac{\partial \left(\frac{\partial \vec{P}}{\partial t} + \vec{\nabla} \wedge \vec{M} \right)}{\partial t}\end{aligned}$$

This obtains a wave equation for the electric field:

$$\boxed{\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \mu_o \sigma \frac{\partial \vec{E}}{\partial t} + \mu_o \epsilon_o \frac{\partial^2 \vec{E}}{\partial t^2} = -\mu_o \frac{\partial \vec{j}_{bound}}{\partial t}} \quad [3.32A]$$



Comparing this expression with the wave equation in vacuum [1.4C], we can comment below on some aspects related to the presence of matter, aside from the trivial effect of substituting the microscopic field with its macroscopic averaged value.

- The expression $\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E})$ is not generally reduced to the Laplacian operation, since by virtue of previous results the divergence of the electric field is not vanishing in general (particularly when propagating in inhomogeneous or birefringent media).
- The presence of conductivity gives rise to conductive losses in the medium, which are expressed in terms of the first derivative of the field.
- The bound density current acts as an inhomogeneous source of the wave equation.

We can analogously derive a wave equation for the magnetic field by taking the curl of the magnetic intensity:

$$\begin{aligned}\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{H}) &= \vec{\nabla} \wedge \sigma \vec{E} + \vec{\nabla} \wedge \frac{\partial \vec{D}}{\partial t} \\ \vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{B}) &= \mu_o \sigma (\vec{\nabla} \wedge \vec{E}) + \mu_o \epsilon_o \frac{\partial (\vec{\nabla} \wedge \vec{E})}{\partial t} + \mu_o \vec{\nabla} \wedge \frac{\partial \vec{P}}{\partial t} + \mu_o \vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{M})\end{aligned}$$

The wave equation for the magnetic field is

$$\boxed{\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{B}) + \mu_o \sigma \frac{\partial \vec{B}}{\partial t} + \mu_o \epsilon_o \frac{\partial^2 \vec{B}}{\partial t^2} = \mu_o \vec{\nabla} \wedge \vec{j}_{bound}} \quad [3.32B]$$

There exist no general analytical solutions to the wave equations for arbitrary bound density currents. In order further understand the mechanisms of light–matter interaction, we need to consider more deeply the relationship between the medium’s incident fields, generated polarization, and magnetization density.

3.4.2 Material equations

Material equations define the relationship not only between the electric field acting on the material and the generated polarization density, but also between the magnetic intensity and the resulting magnetization of the medium.

The material polarization density is generated by the atomic oscillations due to the driving action of the incident electric field, as explained previously. The stationary response for incident monochromatic radiation is a function of the electric field acting on the dipole. If we apply a single impulse to the dipole (which, at the limit, can be assumed to be a delta impulse function) instead of a monochromatic plane wave, the response will be a decaying exponential oscillation. The induced polarization from a single excitation is called the **electric susceptibility**. The non-instantaneous response of the oscillating dipole to a single impulse means that the dipole response to an applied field is not local in time.



For an arbitrary electric field incident on the medium, the resulting polarization density can be written as a series expansion in terms of the powers of E :

$$P_i(\vec{r}, t) = \varepsilon_0 \int_{-\infty}^t \chi_{ij}^{(1)}(\vec{r}, t-t') E_j(\vec{r}, t') dt' + \varepsilon_0 \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \chi_{ijk}^{(2)}(\vec{r}, t-t'_1, t-t'_2) E_j(\vec{r}, t'_1) E_k(\vec{r}, t'_2) dt'_1 dt'_2 + \dots \quad [3.33A]$$

Here, the indices i, j , and k correspond to the Cartesian coordinates x, y , and z , and summation over repeated indices is assumed. The integration limits span from negative infinity to time t . The upper limit reflects the causal relationship between P and E .

This relation can be written in the frequency space by performing a Fourier transformation in time, thus leading to the expression

$$P_i(\vec{r}, \omega) = \varepsilon_0 \chi_{ij}^{(1)}(\vec{r}, \omega) E_j(\vec{r}, \omega) + \sum_{S,T} \varepsilon_0 \chi_{ijk}^{(2)}(\vec{r}, \omega = \omega_S + \omega_T) E_j(\vec{r}, \omega_S) E_k(\vec{r}, \omega_T) \quad [3.33B]$$

In writing this expression, we do not consider the more general case, since we are assuming a spatially local response. This assumption must be reconsidered in situations where the fields show very large variations at short distances. In these situations, a non-local response should be considered. In order to shorten the notation, we will omit the spatial dependence in the following expressions, whenever this locally spatial response is assumed.

In many real situations, this expression may be highly simplified, depending on the characteristics of the incident electric field or the type of medium we are considering.

We define the following specific situations.

1. Linear/non-linear response

When the incident fields are very intense (comparable to the atomic fields acting on the electrons), the response of the dipoles becomes non-linear. Thus, the second-order and higher terms can be significant when describing the light-matter interaction process. In this case, we describe the process as non-linear, and the full expression must be used. This can be the case when a very intense laser pulse is incident on the medium with intensities of the order of GW/cm^2 . Although this kind of interaction is very important in many actual photonics studies, we will not consider nonlinear effects in this introductory course.

When the incident electric field intensities are lower, we may assume a linear response of the system, and the corresponding polarization density becomes

$$P_i(t) = \varepsilon_0 \int_{-\infty}^t \chi_{ij}^{(1)}(t-t') E_j(t') dt' \quad [3.33C]$$

$$P_i(\omega) = \varepsilon_0 \chi_{ij}^{(1)}(\omega) E_j(\omega)$$



B. Isotropic/anisotropic media

The medium is isotropic when the resulting polarization density does not depend on the direction of propagation or the polarization state of the incident electric field. Examples of isotropic media are gases, liquids, amorphous solids, and crystals with cubic symmetry.

In these situations, the generated polarization density is parallel to the incident electric field, and the general tensor character of the susceptibility is simplified to a scalar:

$$\begin{aligned}\vec{P}(t) &= \epsilon_o \int_{-\infty}^t \chi^{(1)}(t-t') \vec{E}(t') dt' \\ \vec{P}(\omega) &= \epsilon_o \chi^{(1)}(\omega) \vec{E}(\omega)\end{aligned}\tag{3.33D}$$

When the medium is anisotropic, the polarization density is no longer parallel to the incident electric field, and the more general expression with the electric susceptibility tensor should be used. This occurs, for instance, when studying light propagation in crystals with no cubic symmetries.

C. Dispersive/non-dispersive media

The response of the medium to any optical field with a finite extension in time will not always be instantaneous and, consequently, the convolution should be used when working in the temporal domain.

However, in the case of a monochromatic electric field incident on the medium, the medium can be assumed to be non-dispersive, i.e., with instantaneous response. In this case, the expression in the temporal and frequency domain is identical:

$$\begin{aligned}P_i(t) &= \epsilon_o \chi_{ij}^{(1)} E_j(t) \\ P_i(\omega) &= \epsilon_o \chi_{ij}^{(1)}(\omega) E_j(\omega)\end{aligned}\tag{3.33E}$$

D. Homogeneous/inhomogeneous media

In an inhomogeneous medium, the electric susceptibility will have a spatial dependence and thus the polarization density will depend on the position. The presence of these inhomogeneities can be due to some impurities present in the medium giving rise to scattering, for instance; or they can be constructed artificially in order to purposely modify the optical properties of the material. The latter is the case with photonic crystals, where the dispersion relation can be strongly altered by the presence of the periodicity, thus leading to unprecedented control over light propagation properties in media upon which many current devices and sensors are based.

In this course, we focus on studying propagation in homogeneous media, thereby omitting the spatial dependence of the electric susceptibility.

The expression of the material equations, then, must be defined for each particular study case while leaving its more general form to situations where we have no previous knowledge of the nature of the interaction.



The simplest case of light-matter interaction occurs for light propagation in the linear regime in a homogeneous, non-dispersive (or using monochromatic radiation) and isotropic medium. In this case, the material dispersion reads:

$$\vec{P} = \epsilon_o \chi^{(1)} \vec{E} \quad [3.34A]$$

Regarding the material magnetization, a **magnetic susceptibility** is defined simply through the expression

$$\vec{M} = \chi_m \vec{H} \quad [3.34B]$$

We will not consider other more complex relations, since we generally assume in this course that our medium is not magnetic.

3.4.3 Index of refraction

For our case of interest, once we have established the material equation defining the increase in the polarization density of a medium in response to the electric field incident on it, we may define the interaction in terms of new properties of the medium: **index of refraction** and **dielectric constant**.

The index of refraction is the most important property for optically characterizing a given medium, and its expression depends strongly on the material properties. Let us derive the expression for the case of linear, homogeneous, non-dispersive, isotropic media.

With these assumptions, the material equations correspond to those expressed in [3.34A], which can be substituted in the wave Equation [3.32A] to give

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \mu_o \sigma \frac{\partial \vec{E}}{\partial t} + \mu_o \epsilon_o \frac{\partial^2 \vec{E}}{\partial t^2} = -\mu_o \frac{\partial}{\partial t} \left[\frac{\partial(\epsilon_o \chi^{(1)} \vec{E})}{\partial t} + \vec{\nabla} \wedge \chi_m \vec{H} \right]$$

Rearranging some terms and substituting the equation for the curl of the magnetic intensity, we have

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \mu_o \sigma \frac{\partial \vec{E}}{\partial t} + \mu_o \epsilon_o (1 + \chi^{(1)}) \frac{\partial^2 \vec{E}}{\partial t^2} = -\mu_o \chi_m \frac{\partial}{\partial t} \left[\sigma \vec{E} + \frac{\partial(\epsilon_o (1 + \chi^{(1)}) \vec{E})}{\partial t} \right]$$

This leads to the result

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \mu_o (1 + \chi_m) \sigma \frac{\partial \vec{E}}{\partial t} + \mu_o \epsilon_o (1 + \chi^{(1)}) (1 + \chi_m) \frac{\partial^2 \vec{E}}{\partial t^2} = 0$$

To define new physical magnitudes associated with a given material, we use:

Relative electric permittivity: $\epsilon_r = 1 + \chi^{(1)}$

Relative magnetic permeability: $\mu_r = 1 + \chi_m$

Dielectric constant: $\epsilon = \epsilon_o \epsilon_r = \epsilon_o (1 + \chi^{(1)})$, and

Magnetic permeability: $\mu = \mu_o \mu_r = \mu_o (1 + \chi_m)$



We may write alternative expressions for the electric displacement vector and magnetic field as:

$$\begin{aligned}\vec{D} &= \epsilon \vec{E} \\ \vec{B} &= \mu \vec{H}\end{aligned}\quad [3.35]$$

For the case of electric displacement, we should remember that we are considering the particular case of a linear, isotropic, non-dispersive and homogeneous medium. We define a more general expression for the displacement field in terms of the electric field with

$$D_i(t) = \int_{-\infty}^t \epsilon_{ij}(t-t') E_j(t') dt'$$

(we explore the important consequences of this relation in Chapter 5).

With the new definitions, the wave equation can be simplified to

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \mu \sigma \frac{\partial \vec{E}}{\partial t} + \mu \epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = 0$$

Considering that a monochromatic beam is incident on the medium:

$$\vec{E} = \text{Re}[\hat{e} E_o^{(+)}(\vec{r}) e^{-i\omega t}]$$

In this case, the temporal derivatives can be easily evaluated:

$$\begin{aligned}\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}_o^{(+)} - i\omega \mu \sigma \vec{E}_o^{(+)} - \omega^2 \mu \epsilon \vec{E}_o^{(+)} &= 0 \\ \vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}_o^{(+)} - \omega^2 \mu \epsilon_o \left[\epsilon_r + i \frac{\sigma}{\omega \epsilon_o} \right] \vec{E}_o^{(+)} &= 0\end{aligned}$$

Defining the **generalized relative dielectric constant** at frequency ω :

$$\epsilon_{eff} = \epsilon_r + i \frac{\sigma}{\omega \epsilon_o} = 1 + \chi^{(1)} + i \frac{\sigma}{\omega \epsilon_o} \quad [3.36A]$$

The **index of refraction** at frequency ω , $n(\omega)$:

$$n^2(\omega) = \mu_r \epsilon_{eff} = \mu_r (1 + \chi^{(1)} + i \frac{\sigma}{\omega \epsilon_o}) \quad [3.36B]$$

The index of refraction defined in this way is a complex quantity:

$$n(\omega) = n_{real}(\omega) + i n_{im}(\omega)$$

In terms of the index of refraction, the wave equation is

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}_o^{(+)} - \frac{n^2(\omega) \omega^2}{c^2} \vec{E}_o^{(+)} = 0$$



The **dispersion relation in the material** is defined as

$$k(\omega) = \frac{n(\omega)\omega}{c} = \frac{n_{real}(\omega)\omega}{c} + i \frac{n_{im}(\omega)\omega}{c} = k_{real}(\omega) + i k_{im}(\omega) \quad [3.37]$$

leading to the basic equation for studying the propagation of beams in material media, known as the **Helmholtz equation for material media**:

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}_O^{(+)} - k^2(\omega) \vec{E}_O^{(+)}) = 0 \quad [3.38]$$

By comparing this equation with the one obtained in Chapter 2, we can see that the effect of the medium is to replace the vacuum wave vector, k_o , with the medium wave vector $k = nk_o$. An immediate consequence of this is that the solutions we found in Chapter 2 for propagation in vacuum are also solutions for propagation in a homogeneous medium just by replacing k_o with k . In particular, the expression for the plane wave in a medium becomes

$$\vec{E}^{(+)}(\vec{r}, t) = \hat{e} E_O^{(+)} e^{i(\vec{k}\vec{r} - \omega t)} = \hat{e} E_O^{(+)} e^{i((n_{re} + i n_{im}) \vec{k}_o \vec{r} - \omega t)}$$

$$\vec{E}^{(+)}(\vec{r}, t) = \hat{e} E_O^{(+)} e^{-n_{im} \vec{k}_o \vec{r}} e^{i(n_{re} \vec{k}_o \vec{r} - \omega t)}$$

From this expression, we may outline the following different consequences.

1) The effect of the real part of the refractive index is to change the phase velocity of the wavefronts from its value in vacuum, c . The phase velocity can be expressed generally as

$$v_{phase}(\omega) = \frac{\omega}{k_{real}(\omega)} = \frac{c}{n_{real}(\omega)} \quad [3.39A]$$

2) The presence of the imaginary refractive index term results in an exponentially decreasing amplitude of the plane wave inside the medium. At the frequencies for which the medium is transparent, this imaginary term vanishes and the wave can propagate without absorption in the medium. Using the Beer–Lambert law defined in [3.29] for a direct comparison of the intensity of the plane wave, we can define the extinction coefficient in terms of the imaginary refractive index:

$$\alpha_{ext}(\omega) = 2n_{im}(\omega)k_o = \frac{2n_{im}(\omega)\omega}{c} \quad [3.39B]$$

If we consider the medium as perfectly homogeneous, the effect of the scattering is to generate the propagating wave inside the medium, according to the Ewald–Ossen theorem. Then, the loss of intensity will be due mainly to the effect of absorption. In these circumstances, the extinction coefficient equals the absorption coefficient. The term absorption coefficient is then mostly used.

As we have seen, the index of refraction of the medium is the main property describing the light–matter interaction process. It is crucial for determining the frequency dependence in order to properly account for the medium's response to light waves of changing



frequency. Although a detailed derivation must use a quantum mechanical treatment, the right frequency dependence is obtained from a simplified classical treatment based on the Lorentz's model (see Appendix C for details). In the next section, we will derive the expression for the index of refraction in specific situations.

3.4.4 Classical derivation of the refractive index

The classical derivation of the index of refraction through the Lorentz's oscillator model is based on a microscopic derivation of the generalized relative dielectric constant and the use of Equation [3.36B].

From the expression

$$[n_{real} + in_{im}]^2 = \mu_r [\epsilon_{eff}^{real} + i\epsilon_{eff}^{im}]$$

we get

$$\begin{aligned} n_{real}^2 - n_{im}^2 &= \epsilon_{eff}^{real} \\ 2n_{real}n_{im} &= \epsilon_{eff}^{im} \end{aligned} \quad [3.40A]$$

An expression for the real and imaginary refractive index is then obtained:

$$\begin{aligned} n_{real} &= \sqrt{\frac{\epsilon_{eff}^{real} + \sqrt{(\epsilon_{eff}^{real})^2 + (\epsilon_{eff}^{im})^2}}{2}} \\ n_{im} &= \sqrt{\frac{-\epsilon_{eff}^{real} + \sqrt{(\epsilon_{eff}^{real})^2 + (\epsilon_{eff}^{im})^2}}{2}} \end{aligned} \quad [3.40B]$$

The particular expression for the refractive index depends on the properties of the material, which are reflected in the relations [3.33]. In the following, we will consider a few specific situations.

Index of refraction in diluted dielectric media

In dielectric media, the material response is due to bound charges and, consequently, to the electric susceptibility of the material. No free charges are present in these materials, so the material conductivity vanishes and the effective relative dielectric constant is reduced to

$$\epsilon_{eff} = 1 + \chi^{(1)}$$

Additionally, if we consider that the medium is dilute, as for instance in a gas cell at low pressure, the index of refraction is close to 1 and we can derive a simplified expression for the refractive index:

$$n^2 = \mu_r (1 + \chi^{(1)}) \rightarrow n = \sqrt{\mu_r (1 + \chi^{(1)})} \approx \sqrt{\mu_r} \left(1 + \frac{\chi^{(1)}}{2}\right)$$



Considering a non-magnetic material ($\mu_r = 1$), we obtain

$$n_{real} \approx 1 + \frac{\chi_{real}^{(1)}}{2} \quad \text{and} \quad n_{im} = \frac{\chi_{im}^{(1)}}{2}$$

To obtain the expression for the electric susceptibility, we start with Lorentz model for the electronic oscillation of the atom. In the simplest situation, let us consider that we want to obtain the refractive index at frequencies close to a single atomic resonance (two-level atom approximation). For a dilute gas, we can consider the medium to be homogeneous and isotropic. The equation for the polarization density $P = Np$ can be obtained directly from the equation for the dipole moment [3.21B] and multiplying it by N . The solution is obtained directly as in [3.21E]:

$$P(\omega) = \frac{Nq^2 / m}{(\omega_o^2 - \omega^2) - i\gamma\omega} E(\omega)$$

The electric susceptibility is obtained directly using [3.33E]:

$$\chi^{(1)}(\omega) = \frac{P(\omega)}{\epsilon_o E(\omega)} = \frac{Nq^2 / \epsilon_o m}{(\omega_o^2 - \omega^2) - i\gamma\omega} = \frac{\omega_p^2}{(\omega_o^2 - \omega^2) - i\gamma\omega} \quad [3.41]$$

where the **plasma frequency**, ω_p , is defined as $\omega_p^2 = Nq^2 / \epsilon_o m$. Dividing this expression into its real and imaginary parts, we have

$$\chi_{real}^{(1)} + i\chi_{im}^{(1)} = \frac{\omega_p^2 (\omega_o^2 - \omega^2)}{(\omega_o^2 - \omega^2)^2 + \gamma^2 \omega^2} + i \frac{\gamma \omega \omega_p^2}{(\omega_o^2 - \omega^2)^2 + \gamma^2 \omega^2}$$

From this result, we can finally write the expression for the real index of refraction and absorption coefficient:

$$n_{real} \approx 1 + \frac{\omega_p^2 (\omega_o^2 - \omega^2)}{2[(\omega_o^2 - \omega^2)^2 + \gamma^2 \omega^2]} \quad \alpha = \frac{2n_{im}\omega}{c} = \frac{\gamma \omega^2 \omega_p^2}{c[(\omega_o^2 - \omega^2)^2 + \gamma^2 \omega^2]}$$

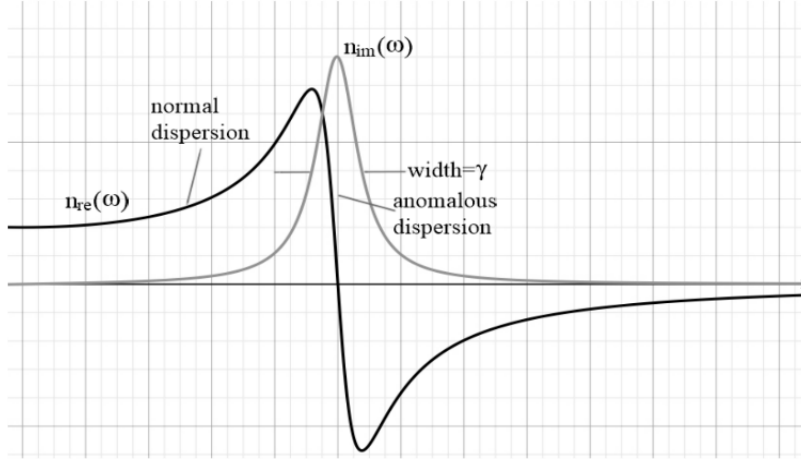
Figure 7 shows a plot of these functions. The absorption coefficient is different from zero in regions that are close to an atomic resonance. The slope of the real index of refraction increases as a function of frequency in the out-of-resonance region, which is called the normal dispersion region ($\partial n / \partial \omega > 0$). At frequencies close to resonance, the slope changes its sign in the region called the anomalous dispersion of the refractive index ($\partial n / \partial \omega < 0$).

If we want to obtain the index of refraction in a frequency region where more than one resonance is present, we can extend this formalism by introducing more resonances into the expression for susceptibility and by introducing the strength factors discussed in regard to Equation [3.15B]:

$$\chi^{(1)}(\omega) = \sum_i \frac{f_i \omega_p^2}{(\omega_{oi}^2 - \omega^2) - i\gamma_i \omega}$$



Figure 3.7
Real and imaginary parts of the refractive index



Index of refraction in dense media

When we consider the light–matter interaction in dense dielectric media (with zero conductivity), we cannot consider that the applied external field is the field acting on the atoms inside the material. From classical electrodynamics, the expression for the local field is obtained as a correction to the incident field:

$$\vec{E}_{local} = \vec{E} + \frac{\vec{P}}{3\epsilon_0}$$

In this case, Equation [3.21] changes to

$$\frac{d^2 P}{dt^2} + \gamma \frac{dP}{dt} + \omega_o^2 P = \frac{Nq^2}{m} E_{local}$$

$$(\omega_o^2 - \omega^2 - i\gamma\omega)P = \frac{Nq^2}{m} \left(E + \frac{P}{3\epsilon_0} \right)$$

Using the relations $P = \epsilon_o \chi^{(1)} E$, we have

$$(\omega_o^2 - \omega^2 - i\gamma\omega)\chi^{(1)} = \omega_p^2 \left(1 + \frac{\chi^{(1)}}{3} \right) \quad \rightarrow \quad \frac{3\chi^{(1)}}{3 + \chi^{(1)}} = \frac{\omega_p^2}{(\omega_o^2 - \omega^2) - i\gamma\omega}$$

Using $n^2 = 1 + \chi^{(1)}$:

$$\frac{n^2 - 1}{n^2 + 2} = \frac{\omega_p^2}{3[(\omega_o^2 - \omega^2) - i\gamma\omega]}$$

This expression is called the **Clausius–Mossotti** relation, and it can be extended straightforwardly to include more resonances, as we did previously.



The index of refraction of materials is usually expressed in terms of the wavelength instead of the frequency. For the visible region, the index of refraction in the normal dispersion region is usually written in the following form, which is called the **Sellmeier equation**:

$$n_{real}^2(\lambda) = A + \frac{B\lambda^2}{\lambda^2 - C} + \frac{D}{\lambda^2 - E}$$

Example: Ordinary index of refraction of the crystal BBO:

$$n_o^2 = 2.7405 + \frac{0.0184}{\lambda^2 - 0.0179} - 0.0155\lambda^2$$

Index of refraction in conductors

In this last example, we describe the index of refraction in a conductor characterized by a very large conductivity value of σ . To obtain the response of the medium to the incident EM wave, the action on the conduction electrons must be studied. This system can be considered a free-electron plasma, as described classically using Drude's model:

$$\frac{d^2\vec{x}}{dt^2} + \Gamma \frac{d\vec{x}}{dt} = \frac{-q\vec{E}}{m} \quad [3.42]$$

This model differs from the previously considered Lorentz model in two important aspects. First, the motion of the free charges is not restricted to the vicinity of a given atom by the action of the Coulomb potential, so they can move along the crystal while behaving as an electron gas. Second, the oscillatory motion of the electrons in response to the applied electromagnetic field is damped, and the damping constant is not related to any damped response of the electronic orbitals but to the energy lost due to collisions with the atomic lattice. This damping constant can also be expressed in terms of the relaxation time τ , which is typically on the order of 10^{-14} s at room temperature:

$$\Gamma = 1/\tau$$

Using this model, we can derive an expression for the medium conductivity. The motion of the electrons results in a conduction current that can be expressed as

$$\vec{j}_c = -N_f q \vec{v} = -N_f q \dot{\vec{x}} = \sigma \vec{E}$$

where N_f is the free electron density in the metal. Under periodic driving fields at external frequency ω , we expect that the particular stationary solution corresponding to the electrons displacement and velocity should also be periodic: $\vec{v}^{(+)} = \vec{v}_o^{(+)} e^{-i\omega t}$.

Writing Equation [3.41] in terms of the velocity as

$$\frac{d\vec{v}}{dt} + \Gamma \vec{v} = \frac{-q\vec{E}}{m} \rightarrow \vec{v}_o^{(+)} = \frac{-(q/m)}{\Gamma - i\omega} \vec{E}_o^{(+)}$$

we obtain an expression for the conduction current, which is



$$j_c^{(+)} = \frac{N_f q^2 / m}{\Gamma - i\omega} E_o^{(+)} = \left[\frac{\epsilon_o \omega_{pf}^2 \Gamma}{\omega^2 + \Gamma^2} + i \frac{\epsilon_o \omega_{pf}^2 \omega}{\omega^2 + \Gamma^2} \right] \cdot E_o^{(+)}$$

where we introduce the plasma frequency for the free charges: $\omega_{pf}^2 = N_f q^2 / \epsilon_o m$.

The conductivity is obtained directly from this relation:

$$\sigma(\omega) = \left[\frac{\epsilon_o \omega_{pf}^2 \Gamma}{\omega^2 + \Gamma^2} + i \frac{\epsilon_o \omega_{pf}^2 \omega}{\omega^2 + \Gamma^2} \right] \quad [3.43]$$

Finally, we can obtain the contribution made by the free electrons to the effective dielectric constant and refractive index, specifically by using [3.36A]:

$$\epsilon_{rc} = \frac{i\sigma}{\epsilon_o \omega} = -\frac{\omega_{pf}^2}{\omega^2 + \Gamma^2} + i \frac{\omega_{pf}^2 \Gamma}{\omega(\omega^2 + \Gamma^2)} = -\frac{\omega_{pf}^2 \tau^2}{1 + \omega^2 \tau^2} + i \frac{\omega_{pf}^2 \tau}{\omega(1 + \omega^2 \tau^2)}$$

In the case of very large conductivity values and assuming a negligible contribution from the bound charges, the effective dielectric constant is given by the expressions

$$\epsilon_{eff}^{real} = 1 - \frac{\omega_{pf}^2 \tau^2}{1 + \omega^2 \tau^2} \quad \text{and} \quad \epsilon_{eff}^{im} = \frac{\omega_{pf}^2 \tau}{\omega(1 + \omega^2 \tau^2)}$$

from which the index of refraction can be obtained.

The behaviour of conductors depends strongly on the value of the driving frequency of the exciting electromagnetic field with respect to the plasma frequency and the relaxation time. Some general properties based on the previous results are established below.

a) Large frequencies ($\omega \gg \Gamma \rightarrow \omega\tau \gg 1$)

In this regime, the condition $\epsilon_{eff}^{real} > \epsilon_{eff}^{im}$ holds, and $\epsilon_{eff} = 1 - \frac{\omega_{pf}^2}{\omega^2}$

The first important characteristic of conductors in this regime is that they change from being highly reflective to transparent when the driving frequency is, respectively, below or above the free-electron plasma frequency. At frequencies close to the plasma frequency, the wave vector becomes

$$k = \frac{n\omega}{c} = \frac{\sqrt{\epsilon_{eff}} \omega}{c} \approx \frac{\sqrt{\omega^2 - \omega_{pf}^2}}{c}$$

which becomes imaginary at frequencies below the plasma frequency, resulting in an evanescent wave inside the medium. At the same time, the dielectric constant becomes negative in this regime, by which the relations [3.40B] lead to the condition $n_{im} \gg n_{real}$.

Under these conditions, the reflectivity of the conductor (at normal incidence)



$$R = \left| \frac{n-1}{n+1} \right|^2 = \frac{(n_{real}-1)^2 + n_{im}^2}{(n_{real}+1)^2 + n_{im}^2} \approx 1$$

is close to 1, one of the characteristic properties of conductors.

b) Low frequencies ($\omega < \Gamma \rightarrow \omega\tau < 1$)

This regime is typical of frequencies in the infrared, and the imaginary component of the dielectric constant (proportional to the real part of the conductivity) dominates:

$$\epsilon_{eff}^{real} \ll \epsilon_{eff}^{im} \text{ and } \epsilon_{eff} \approx i\epsilon_{eff}^{im}$$

Under these conditions, the effective relative dielectric constant [3.36A] becomes purely imaginary:

$$\epsilon_{eff}(\omega) \approx i \frac{\sigma_{real}(\omega)}{\epsilon_0 \omega} \approx i \frac{\omega_{pf}^2 \tau}{\omega}$$

Using expressions [3.40B], we obtain

$$n_{real} \approx n_{im} \approx \sqrt{\frac{\epsilon_{eff}^{im}}{2}} = \sqrt{\frac{\sigma_{real}}{2\epsilon_0 \omega}} = \sqrt{\frac{\omega_{pf}^2 \tau}{2\omega}}$$

The large imaginary refractive index gives rise to a large absorption coefficient in the medium. Typically, optical radiation is strongly absorbed in metallic materials within distances of the order of $1/\alpha = c/2n_{im}\omega$. This distance is called the skin depth, and it is given by

$$\delta(\omega) = \frac{c}{2n_{im}(\omega)\omega} = \frac{c}{\sqrt{2\omega_{pf}^2 \tau \omega}}$$

with strong dependence on the frequency.

The plasma frequency value depends on the characteristics of the conductor. For instance, noble metals such as Au or Ag have plasma frequencies in the UV region and, consequently, become highly reflective and opaque in the visible region. At frequencies in the visible range close to the plasma frequency, their properties are strongly affected by the interband transitions resulting from the bound charges. To properly described the optical properties of these metals in the visible region, we should use the general relation [3.36B] to consider the action of both the bound (Lorentz) and the free (Drude) electron contributions.

In contrast, elements such as ITO (Indium Tin Oxide) have a plasma frequency in the IR region (around 1200 nm), thus making them transparent at optical frequencies. This important property favours using ITO in many applications requiring transparent electrodes.



3.5 Energy relations in material media

The conservation of energy in the presence of material media can be described in the same way that has been outlined in Chapter 1 for energy relations in vacuum. In this case, extra terms appear due to the energy needed to polarize and magnetize the medium. The detailed derivation is worked out in the collection of problems that give the equation for energy conservation:

$$\vec{\nabla} \cdot \vec{S} + \frac{\partial u_m}{\partial t} = -\frac{1}{2} \left[\mu_o \vec{H} \frac{\partial \vec{M}}{\partial t} - \mu_o \vec{M} \frac{\partial \vec{H}}{\partial t} + \vec{E} \frac{\partial \vec{P}}{\partial t} - \vec{P} \frac{\partial \vec{E}}{\partial t} \right] - \vec{j} \cdot \vec{E}$$

The following definitions apply.

Poynting vector in a material medium: $\vec{S} = \vec{E} \wedge \vec{H}$

Energy Density of the EM field in the medium: $u_m = \frac{1}{2} [\vec{D} \cdot \vec{E} + \vec{B} \cdot \vec{H}]$

The term in brackets represents the work done per unit volume and time to polarize and magnetize the medium.

From Maxwell's equations and the definition of the Poynting vector, we can obtain the expression for the intensity of a plane wave propagating in a medium of refractive index n (see Solved problems):

$$I = \frac{1}{2} \epsilon_o c n |E_o^{(+)}|^2$$



3.6 Solved problems

S3.1 Two point, linear, and harmonic dipoles oscillate with identical amplitude, p_o , and frequency, ω . Both are placed at the origin of our system of reference and oscillate along different orientations. The first dipole oscillates along the z -axis, while the second dipole oscillates along the y -axis. Using the general expression of the radiated fields of a generic dipole, $\vec{p} = p_o \hat{d}$, in the electric dipole (long wavelength) and radiation zone approximation, we have

$$\vec{e}_{rad} = \frac{\mu_o \ddot{p}}{4\pi R} [\hat{u} \wedge (\hat{u} \wedge \hat{d})]$$

$$\vec{b}_{rad} = \frac{\mu_o \ddot{p}}{4\pi R c} [\hat{d} \wedge \hat{u}]$$

Write the expression for the total radiated electric and magnetic fields and the total intensity along an arbitrary direction in the xy -plane (use spherical coordinates to define the observation direction, $\hat{u}$).

Solution

Any observation direction in space can be expressed in terms of the usual spherical components, such as $\hat{u} = (\cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k})$. In this problem, our observation plane is the xz -plane, so $\phi = 0$ (in general, the whole space is fully covered by letting $\phi \in [0, 2\pi]$ and $\theta \in [0, \pi]$). With this convention, our case should take both $\phi = 0$ and $\phi = \pi$ to cover the whole xz -plane. Instead, we will take $\phi = 0$ and let $\theta \in [0, 2\pi]$.

In the xz -plane, our propagation direction is defined by $\hat{u} = (\sin \theta \hat{i} + \cos \theta \hat{k})$.

Fields of dipole 1: For this dipole, $\hat{d}_1 = \hat{k}$, and performing the vectorial products gives us:

$$\vec{e}_{rad}^1 = \frac{\mu_o \ddot{p}}{4\pi R} [\hat{u} \wedge (\hat{u} \wedge \hat{d}_1)] = \frac{\mu_o \ddot{p}}{4\pi R} \sin \theta (\cos \theta \hat{i} - \sin \theta \hat{k})$$

$$\vec{b}_{rad}^1 = \frac{\mu_o \ddot{p}}{4\pi R c} [\hat{d}_1 \wedge \hat{u}] = \frac{\mu_o \ddot{p}}{4\pi R c} \sin \theta \hat{j}$$

Fields of dipole 2: In this case, $\hat{d}_2 = \hat{j}$, and

$$\vec{e}_{rad}^2 = \frac{\mu_o \ddot{p}}{4\pi R} [\hat{u} \wedge (\hat{u} \wedge \hat{d}_2)] = \frac{\mu_o \ddot{p}}{4\pi R} (-\hat{j})$$

$$\vec{b}_{rad}^2 = \frac{\mu_o \ddot{p}}{4\pi R c} [\hat{d}_2 \wedge \hat{u}] = \frac{\mu_o \ddot{p}}{4\pi R c} (\cos \theta \hat{i} - \sin \theta \hat{k})$$

The harmonic dipole can be expressed as $p^{(+)}(t') = p_o^{(+)} e^{-i\omega t'}$, where t' is the retarded time $t' = t - R/c$, and the derivatives are taken with respect to the retarded time. Thus:



$$\ddot{p} = -\omega^2 p_o^{(+)} e^{-i\omega t} = -\omega^2 p_o^{(+)} e^{i(\vec{k}\vec{R} - \omega t)}$$

The total radiated fields are obtained as the sum of the fields radiated by the two dipoles. Using the previous results, we obtain

$$\begin{aligned}\vec{e}_{tot}^{(+)} &= \frac{\mu_o \omega^2 p_o^{(+)}}{4\pi R} \left(-\sin \theta \cos \theta \hat{i} + \hat{j} + \sin^2 \theta \hat{k} \right) e^{i(k \sin \theta x + k \cos \theta z - \omega t)} \\ \vec{b}_{tot}^{(+)} &= \frac{\mu_o \omega^2 p_o^{(+)}}{4\pi R c} \left(-\cos \theta \hat{i} - \sin \theta \hat{j} + \sin \theta \hat{k} \right) e^{i(k \sin \theta x + k \cos \theta z - \omega t)}\end{aligned}$$

The total intensity is given by the expression

$$\begin{aligned}I_{tot} &= \frac{1}{2} \text{Re} \left[\left| \frac{\vec{e}_{tot}^{(+)} \wedge \vec{b}_{tot}^{(-)}}{\mu_o} \right|^2 \right] = \frac{1}{2} \cdot \frac{\mu_o |p_o^{(+)}|^2 \omega^4}{16\pi^2 R^2 c} \left\| \begin{array}{ccc} \hat{i} & \hat{j} & \hat{k} \\ -\sin \theta \cos \theta & 1 & \sin^2 \theta \\ -\cos \theta & -\sin \theta & \sin \theta \end{array} \right\| = \\ &= \frac{\mu_o |p_o^{(+)}|^2 \omega^4}{32\pi^2 R^2 c} \left| (\sin \theta + \sin^3 \theta) \hat{i} + (\sin^2 \theta \cos \theta + \cos \theta) \hat{k} \right| = \\ &= \frac{\mu_o |p_o^{(+)}|^2 \omega^4}{32\pi^2 R^2 c} (1 + \sin^2 \theta) \left| \sin \theta \hat{i} + \cos \theta \hat{k} \right| \\ I_{tot} &= \frac{\mu_o |p_o^{(+)}|^2 \omega^4}{32\pi^2 R^2 c} (1 + \sin^2 \theta)\end{aligned}$$

Alternative solution in the spherical coordinate system

The spherical coordinate system uses the unitary vectors $\hat{u}_r, \hat{u}_\phi$ and $\hat{u}_\theta$.

These are expressed in terms of the Cartesian basis as

$$\begin{aligned}\hat{u}_r &= \frac{\partial \vec{r} / \partial r}{|\partial \vec{r} / \partial r|} = \sin \theta \cos \phi \hat{i} + \sin \theta \sin \phi \hat{j} + \cos \theta \hat{k} \\ \hat{u}_\phi &= \frac{\partial \vec{r} / \partial \phi}{|\partial \vec{r} / \partial \phi|} = \cos \phi \cos \theta \hat{i} + \cos \phi \sin \theta \hat{j} - \sin \theta \hat{k} \\ \hat{u}_\theta &= \frac{\partial \vec{r} / \partial \theta}{|\partial \vec{r} / \partial \theta|} = -\sin \phi \hat{i} + \cos \phi \hat{j}\end{aligned}$$

with $\vec{r} = r(\cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k})$.

Inverting this system, we obtain the relation between the Cartesian and spherical basis:



$$\begin{aligned}
 \hat{i} &= \sin \theta \cos \phi \hat{u}_r - \sin \phi \hat{u}_\phi + \cos \theta \cos \phi \hat{u}_\theta & \hat{i} &= \sin \theta \hat{u}_r + \cos \theta \hat{u}_\theta \\
 \hat{j} &= \sin \theta \sin \phi \hat{u}_r + \cos \phi \hat{u}_\phi + \cos \theta \sin \phi \hat{u}_\theta & \text{In the } xz\text{-plane: } \hat{j} &= \hat{u}_\phi \\
 \hat{k} &= \cos \theta \hat{u}_r - \sin \theta \hat{u}_\theta & \hat{k} &= \cos \theta \hat{u}_r - \sin \theta \hat{u}_\theta
 \end{aligned}$$

Dipole 1: Direction of observation $\hat{u} = \hat{u}_r$ and dipole oscillation direction

$$\hat{d}_1 = \cos \theta \hat{u}_r - \sin \theta \hat{u}_\theta$$

$$\hat{d}_1 \wedge \hat{u} = \sin \theta \hat{u}_\phi \quad \text{and} \quad \hat{u} \wedge (\hat{u} \wedge \hat{d}_1) = \sin \theta \hat{u}_\theta$$

Dipole 2: Direction of observation $\hat{u} = \hat{u}_r$ and dipole oscillation direction $\hat{d}_2 = \hat{u}_\phi$

$$\hat{d}_2 \wedge \hat{u} = \hat{u}_\theta \quad \text{and} \quad \hat{u} \wedge (\hat{u} \wedge \hat{d}_2) = -\hat{u}_\phi$$

Including these results in the expressions of the fields, the total electric and magnetic fields can be expressed as

$$\begin{aligned}
 \vec{e}_{tot}^{(+)} &= \frac{\mu_o \omega^2 P_o^{(+)}}{4\pi R} (\hat{u}_\phi - \sin \theta \hat{u}_\theta) e^{i(kR - \omega t)} \\
 \vec{b}_{tot}^{(+)} &= \frac{\mu_o \omega^2 P_o^{(+)}}{4\pi R c} (-\sin \theta \hat{u}_\phi - \hat{u}_\theta) e^{i(kR - \omega t)}
 \end{aligned}$$

and the total intensity as

$$I_{tot} = \frac{\mu_o \omega^4 |P_o^{(+)}|^2}{32\pi^2 R^2 c} |(1 + \sin^2 \theta) \hat{u}_r| = \frac{\mu_o \omega^4 |P_o^{(+)}|^2}{32\pi^2 R^2 c} (1 + \sin^2 \theta)$$

■

S3.2 A hydrogen atom emits radiation of frequency $\omega = 4 \cdot 10^{15} \text{ rad / s}$. Assuming that the maximum displacement of the dipole moment is $x_o = 5 \cdot 10^{-11} \text{ m}$, calculate the damping factor due to radiation reaction and the total average power radiated by the oscillator.

Solution

The damping constant corresponding to the classical expression for the spontaneous decay rate is given by

$$\gamma = \frac{\mu_o q^2 \omega^2}{6\pi m c}$$

where m is the reduced mass of the atom, which is a good approximation of the electron mass, and q is the electron charge. Directly substituting the values in the problem give



$$\gamma = 10^8 \text{ s}^{-1} \rightarrow \Gamma = \frac{1}{\gamma} = 10^{-8} \text{ s}$$

The total averaged power radiated by the damped oscillator is

$$\langle P_{rad} \rangle = \frac{\mu_o \omega^4 q^2 x_o^2}{12\pi c} e^{-\gamma t} = 1.82 \cdot 10^{-12} \cdot e^{-10^8 t}$$

■

S3.3 A box filled with a collection of identical atoms is illuminated by a plane EM wave oscillating at frequency ω , $\vec{E}^{(+)} = \vec{E}_o^{(+)} e^{-i\omega t}$. The interaction of the atoms with the wave can be described by the conventional Lorentz model without any need to use the local field correction, although we do use a damping constant γ_{tot} that includes the different loss mechanisms. The (averaged) power density transfer between the external field and the medium is described by the formula

$$\langle P_{transfer} \rangle = \langle \vec{E} \cdot d\vec{P}/dt \rangle$$

where $\vec{P} = -Nq\vec{x}$ is the medium polarization density and N the atom density. The interaction conditions are assumed to be homogeneous, isotropic, non-dispersive, and linear.

a) Write an expression for $\langle P_{transfer} \rangle$, expressed in terms of the material susceptibility, $\chi^{(1)}$, and the incident intensity.

b) Using the Lorentz model and assuming that the only resonant frequency relevant to describe the interaction is ω_o , write an equation for the polarization density. From this equation, obtain an expression for the material susceptibility $\chi^{(1)}$. Particularize this expression for the case of nearly resonant interaction, assuming that the external frequency is very close to ω_o , $\omega \approx \omega_o$.

c) Substitute the expression for the susceptibility obtained in b) to obtain the power transfer between the external field and the medium as a function of the medium parameters and the incident intensity.

Solution

a) The polarization density for a homogeneous, isotropic, non-dispersive material in the linear regime is given by

$$\vec{P} = \epsilon_o \chi^{(1)} \vec{E}$$

The averaged value of the power transfer is given by the averaged value of the product of the electric field and the temporal derivative of the polarization density. As we have seen in this chapter, the average value of the product of two real fields can be obtained from its complex representation:



$$\langle P_{transfer} \rangle = \frac{1}{2} \text{Re} \left[E^{(+)} \cdot \frac{dP^{(-)}}{dt} \right]$$

The complex representation of the polarization density is obtained by considering the complex character of the susceptibility:

$$P^{(+)} = \epsilon_o \left(\chi_{real}^{(1)} + i\chi_{im}^{(1)} \right) E_o^{(+)} e^{-i\omega t} \rightarrow P^{(-)} = \epsilon_o \left(\chi_{real}^{(1)} - i\chi_{im}^{(1)} \right) E_o^{(-)} e^{i\omega t}$$

The temporal derivative is

$$\frac{dP^{(-)}}{dt} = i\omega \epsilon_o \left(\chi_{real}^{(1)} - i\chi_{im}^{(1)} \right) E_o^{(-)} e^{i\omega t}$$

The averaged power transfer:

$$\langle P_{transfer} \rangle = \frac{1}{2} \text{Re} \left[E^{(+)} \cdot \frac{dP^{(-)}}{dt} \right] = \frac{1}{2} \text{Re} \left[\omega \epsilon_o |E_o|^2 (\chi_{im}^{(1)} + i\chi_{real}^{(1)}) \right] = \frac{\omega \epsilon_o \chi_{im}^{(1)} |E_o|^2}{2}$$

From this expression, we see that the imaginary part of the susceptibility controls the energy transfer between the incident light and the collection of atoms. Under conditions of negligible imaginary susceptibility, the energy is not absorbed by the atoms but is instead generally scattered, and the transmitted part suffers a phase lag determined by the real part of the susceptibility.

Considering that we have a plane wave illuminating the medium, its intensity is

$$I_{inc} = \frac{1}{2} \epsilon_o c |E_o|^2$$

and the power transfer average can be rewritten as

$$\langle P_{transfer} \rangle = \frac{\omega \chi_{im}^{(1)}}{c} I_{inc}$$

b) To find an expression for the susceptibility in the quasi-resonant case, we will start with Equation [3.41], which was derived in Section 3.4.4:

$$\chi^{(1)}(\omega) = \frac{\omega_p^2}{(\omega_o^2 - \omega^2) - i\gamma_{tot}\omega} \quad \text{where, } \omega_p \text{ is the plasma frequency} \quad \omega_p^2 = \frac{Nq^2}{\epsilon_o m}$$

In a quasi-resonant interaction, $\omega_o \approx \omega$, and we can approximate the term in the denominator $(\omega_o^2 - \omega^2) = (\omega_o + \omega)(\omega_o - \omega) \approx 2\omega_o(\omega_o - \omega)$:



$$\chi^{(1)}(\omega) \approx \frac{\omega_p^2}{2\omega_o(\omega_o - \omega) - i\gamma_{tot}\omega_o} = \omega_p^2 \frac{2\omega_o(\omega_o - \omega) - i\gamma_{tot}\omega_o}{4\omega_o^2(\omega_o - \omega)^2 + \gamma_{tot}^2\omega_o^2}$$

Rewriting this expression as

$$\chi^{(1)}(\omega) = \frac{\omega_p^2}{\gamma_{tot}} \cdot \frac{2\omega_o(\omega_o - \omega)/\gamma_{tot} + i\omega_o}{(2\omega_o(\omega_o - \omega)/\gamma_{tot})^2 + \omega_o^2}$$

we can see that the imaginary part has a Lorentzian profile centred in the resonant frequency:

$$\chi_{im}^{(1)} = \frac{\omega_p^2}{\omega_o\gamma_{tot}} \cdot \frac{1}{1 + (2(\omega_o - \omega)/\gamma_{tot})^2}$$

c) Finally, we can substitute the imaginary susceptibility relation into the expression of the power transfer:

$$\langle P_{transfer} \rangle = \frac{\omega\omega_p^2}{\omega_o\gamma_{tot}c} \cdot \frac{I_{inc}}{1 + (2(\omega_o - \omega)/\gamma_{tot})^2}$$

■

S3.4 A plane EM wave is incident on an oscillating dipole forcing it to oscillate. Determine the averaged value of the magnetic force $\vec{F} = -q(\dot{\vec{x}} \wedge \vec{B}_{ext})$ exerted by the fields on the dipole, and calculate the radiation pressure at resonance using the expression

$$P_{rad} = \frac{F}{\sigma_{ext}} \quad (\sigma_{ext} \text{ is the extinction cross section of the forced dipole oscillator and } q = 1.6 \cdot 10^{-19} \text{ C}).$$

Solution

Since the force is a bilinear product of vectors, the average value can be calculated in terms of their complex representation:

$$\langle \vec{F} \rangle = \frac{1}{2} \text{Re} \left[-q \dot{\vec{x}}^{(+)} \wedge \vec{B}_{ext}^{(-)} \right]$$

The movement of the electron is opposite to the direction of the electric field, so the direction of the electron velocity is $\dot{\vec{x}} = \dot{x}(-\hat{e})$:

$$-q \dot{\vec{x}}^{(+)} \wedge \vec{B}_{ext}^{(-)} = q \dot{x}^{(+)} B_{ext}^{(-)} \left[\hat{e} \wedge \hat{b} \right] = q \dot{x}^{(+)} B_{ext}^{(-)} \hat{u}$$

where we use the transverse character of the plane wave. As we can see, the force is exerted in the direction of the wave propagation.



The electron's velocity can be calculated by considering Lorentz's model and assuming that the electric force is much larger than the magnetic force:

$$\frac{d^2 \vec{x}}{dt^2} + \gamma \frac{d\vec{x}}{dt} + \omega_o^2 \vec{x} = -\frac{q}{m} \vec{E}_{ext}$$

(electron charge: $-q$).

$$\text{Writing } \vec{E}_{ext} = \text{Re}[\hat{e} E_o^{(+)} e^{-i\omega t}] \text{ and } \vec{x} = \text{Re}[(-\hat{e}) x_o^{(+)} e^{-i\omega t}]$$

we obtain an equation for the modulus of the displacement:

$$\frac{d^2 x^{(+)}}{dt^2} + \gamma \frac{dx^{(+)}}{dt} + \omega_o^2 x^{(+)} = \frac{q}{m} E_{ext}^{(+)}$$

With solution

$$x^{(+)} = \frac{q E_o^{(+)} / m}{(\omega_o^2 - \omega^2) - i\gamma\omega} e^{-i\omega t}$$

the temporal derivative gives

$$\dot{x}^{(+)} = -i \frac{q E_o^{(+)} \omega}{m} \frac{(\omega_o^2 - \omega^2) + i\gamma\omega}{(\omega_o^2 - \omega^2)^2 + \gamma^2 \omega^2} e^{-i\omega t} = \frac{q E_o^{(+)} \omega}{m [(\omega_o^2 - \omega^2)^2 + \gamma^2 \omega^2]} [\gamma\omega - i(\omega_o^2 - \omega^2)] e^{-i\omega t}$$

$$\text{and } B_{ext}^{(-)} = \frac{E_o^{(-)}}{c} e^{i\omega t}$$

The magnetic force is then

$$\langle \vec{F} \rangle = \frac{1}{2} \text{Re} [q \dot{x}^{(+)} B_{ext}^{(-)}] \hat{u} = \frac{q^2 \gamma \omega^2 |E_{ext}|^2}{2mc [(\omega_o^2 - \omega^2)^2 + \gamma^2 \omega^2]} \hat{u}$$

which, in terms of the incident intensity $I_{inc} = \epsilon_o c |E_{ext}|^2 / 2$, reads

$$\langle F \rangle = \frac{q^2 \gamma \omega^2}{m \epsilon_o c^2 [(\omega_o^2 - \omega^2)^2 + \gamma^2 \omega^2]} \cdot I_{inc} = \frac{\sigma_{ext} I_{inc}}{c}$$

where we use Equation [3.28], defining the extinction coefficient.

Finally, the radiation pressure can be calculated:

$$P_{rad} = \frac{\langle F \rangle}{\sigma_{ext}} = \frac{I_{inc}}{c}$$

This expression is equivalent to the one presented in Chapter 1.

■

**S3.5** Derive the equation for energy conservation in macroscopic media**Solution**

Starting with the macroscopic Maxwell equations

$$\vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad \text{and} \quad \vec{\nabla} \wedge \vec{H} = \vec{j}_{cond} + \frac{\partial \vec{D}}{\partial t}$$

make the scalar product of the first equation by $\vec{H}$ and of the second by $\vec{E}$:

$$\begin{aligned} \vec{H} \cdot (\vec{\nabla} \wedge \vec{E}) &= -\vec{H} \cdot \frac{\partial \vec{B}}{\partial t} \\ \vec{E} \cdot (\vec{\nabla} \wedge \vec{H}) &= \vec{E} \cdot \vec{j}_{cond} + \vec{E} \cdot \frac{\partial \vec{D}}{\partial t} \end{aligned}$$

Using the identity $\vec{\nabla} \cdot (\vec{E} \wedge \vec{H}) = \vec{H} \cdot (\vec{\nabla} \wedge \vec{E}) - \vec{E} \cdot (\vec{\nabla} \wedge \vec{H})$, we obtain the relation

$$\vec{\nabla} \cdot (\vec{E} \wedge \vec{H}) = -\vec{H} \cdot \frac{\partial \vec{B}}{\partial t} - \vec{E} \cdot \frac{\partial \vec{D}}{\partial t} - \vec{E} \cdot \vec{j}_{cond}$$

which is equivalent to Equation [1.15A] in the presence of matter, where we define the following.

Poynting vector: $\vec{S} = \vec{E} \wedge \vec{H}$

Rate of change in the electric field energy density: $\frac{dw_E}{dt} = \vec{E} \cdot \frac{\partial \vec{D}}{\partial t}$

Rate of change in the magnetic field energy density: $\frac{dw_M}{dt} = \vec{H} \cdot \frac{\partial \vec{B}}{\partial t}$

This general relation can be particularized in different ways, depending on how the material equations are considered.

For instance, inserting the general relation $\vec{B} = \mu_o \vec{H} + \mu_o \vec{M}$ and $\vec{D} = \epsilon_o \vec{E} + \vec{P}$, we can rewrite the energy conservation equation (valid for isotropic non-dispersive media) as

$$\vec{\nabla} \cdot \vec{S} = -\frac{1}{2} \frac{\partial (\mu_o H^2 + \epsilon_o E^2)}{\partial t} - \mu_o \vec{H} \cdot \frac{\partial \vec{M}}{\partial t} - \vec{E} \cdot \frac{\partial \vec{P}}{\partial t} - \vec{E} \cdot \vec{j}_{cond}$$

where the first term on the right-hand side reduces to the temporal change rate of the electromagnetic energy density in vacuum, with $M = 0$ and $P = 0$. The second and third terms correspond to the work done to magnetize and polarize the medium.

The same equation can be cast in different form:



$$\begin{aligned}\vec{\nabla} \cdot \vec{S} &= -\vec{H} \cdot \frac{\partial \vec{B}}{\partial t} - \vec{E} \cdot \frac{\partial \vec{D}}{\partial t} - \vec{E} \cdot \vec{j}_{cond} = -\frac{\partial(\vec{B} \cdot \vec{H})}{\partial t} - \frac{\partial(\vec{E} \cdot \vec{D})}{\partial t} + \vec{B} \frac{\partial \vec{H}}{\partial t} + \vec{D} \frac{\partial \vec{E}}{\partial t} - \vec{E} \cdot \vec{j}_{cond} = \\ &= -\frac{1}{2} \frac{\partial(\vec{B} \cdot \vec{H} + \vec{E} \cdot \vec{D})}{\partial t} + \left(\vec{B} \frac{\partial \vec{H}}{\partial t} - \frac{1}{2} \frac{\partial(\vec{B} \cdot \vec{H})}{\partial t} \right) + \left(\vec{D} \frac{\partial \vec{E}}{\partial t} - \frac{1}{2} \frac{\partial(\vec{E} \cdot \vec{D})}{\partial t} \right) - \vec{E} \cdot \vec{j}_{cond}\end{aligned}$$

Defining the energy density of the electromagnetic field in the medium as

$$w_{em} = \frac{1}{2} (\vec{B} \cdot \vec{H} + \vec{D} \cdot \vec{E})$$

and using again the material relations $\vec{B} = \mu_o \vec{H} + \mu_o \vec{M}$ and $\vec{D} = \epsilon_o \vec{E} + \vec{P}$, the above equation can be written as

$$\vec{\nabla} \cdot \vec{S} + \frac{\partial w_{em}}{\partial t} = -\frac{1}{2} \left(\vec{H} \frac{\partial \vec{M}}{\partial t} - \vec{M} \frac{\partial \vec{H}}{\partial t} \right) - \frac{1}{2} \left(\vec{E} \frac{\partial \vec{P}}{\partial t} - \vec{P} \frac{\partial \vec{E}}{\partial t} \right) - \vec{E} \cdot \vec{j}_{cond}$$

The first and second terms to the right represent the work done per unit volume and unit time to polarize and magnetize a medium in the nonlinear regime of interaction. These terms are null in the case of linear interaction, leading to the equation

$$\vec{\nabla} \cdot \vec{S} = -\frac{\partial w_{em}}{\partial t} - \vec{E} \cdot \vec{j}_{cond}$$

which is equivalent to Equation [1.16A] in vacuum. ■

S3.6 The complex dielectric constant of the semiconductor cadmium telluride (CdTe) at 500 nm is given by $\epsilon_r = 8.92 + i2.29$. Calculate at this wavelength:

a) phase velocity of light propagating in CdTe

b) absorption coefficient

c) reflectivity at normal incidence

Solution

a) To obtain the phase velocity, we have to calculate the index of refraction of the material from the values of the complex dielectric constant:

$$n_{real} = \sqrt{\frac{\epsilon_{real} + \sqrt{\epsilon_{real}^2 + \epsilon_{im}^2}}{2}} = 3.01 \quad n_{im} = \sqrt{\frac{-\epsilon_{real} + \sqrt{\epsilon_{real}^2 + \epsilon_{im}^2}}{2}} = 0.38$$

Phase velocity:

$$v = \frac{c}{n_{real}} = 0.996 \cdot 10^8 \text{ m/s}$$



b) The absorption coefficient depends on the value of the imaginary part of the index of refraction:

$$\alpha = \frac{2n_{im}\omega}{c} = \frac{4\pi n_{im}}{\lambda} = 9.5 \cdot 10^6 \text{ m}^{-1}$$

c) The reflectivity at normal incidence is given by the expression

$$R = \left| \frac{n-1}{n+1} \right|^2 = \frac{(n_{real}-1) + in_{im}}{(n_{real}+1) + in_{im}} \cdot \frac{(n_{real}-1) - in_{im}}{(n_{real}+1) - in_{im}} = \frac{(n_{real}-1)^2 + n_{im}^2}{(n_{real}+1)^2 + n_{im}^2}$$

Directly substituting the values of the index of refraction give

$$R = 0.2579.$$

The reflectivity of CdTe at a 500 nm wavelength and normal incidence is $R = 25.79\%$. ■

S3.7 The real part of the index of refraction of a given material is $n_{real} = 1.53$ at the wavelength of 590 nm, and it absorbs or scatters 98% of the incident light at a depth of 2 μm . What is the complex dielectric constant of the material at this wavelength?

Solution

The light propagating in the absorptive material follows the Beer–Lambert law:

$$I(z) = I_o e^{-\alpha z}$$

At a distance of 2 μm , it absorbs 98% of the incident radiation, so 2% will be transmitted and $I(L) = 0.02I_o$:

$$0.02I_o = I_o e^{-\alpha L} \rightarrow \alpha = -\frac{\ln 0.02}{L} = 1956011.5 \text{ m}^{-1}$$

The imaginary index of refraction is directly related to the absorption coefficient:

$$\alpha = \frac{2n_{im}\omega}{c} = \frac{4\pi n_{im}}{\lambda} \rightarrow n_{im} = \frac{\lambda\alpha}{4\pi} = 9.18 \cdot 10^{-2}$$

The dielectric constant is calculated directly from the index of refraction:

$$\epsilon_{real} = n_{real}^2 - n_{im}^2 = 2.332$$

$$\epsilon_{im} = 2n_{real}n_{im} = 0.28$$

■

S3.8 A plate with parallel sides is made of optical glass with index of refraction $n = n_{real} + in_{im}$. The thickness of the plate is much larger than the coherence length of the light illuminating it, so interference effects are negligible and we can just add the intensities at each interface.



a) Calculate the total reflectivity and transmission by the plate in the presence of absorption, expressed in terms of the reflectivity at a single layer R .

b) If, additionally, the medium is strongly absorbing, such that multiple reflections are negligible ($\alpha L \gg 1$), determine the transmittance and reflectance of the plate in this situation.

Solution

The reflectivity at the interface will depend on the index of refraction at each side and the incidence angle. For instance, the reflectivity at normal incidence on a surface separating air and in a medium of index n is given by the expression used in problem S3.6 c). We assume that we know the reflectivity, R , of the interface in the problem conditions.

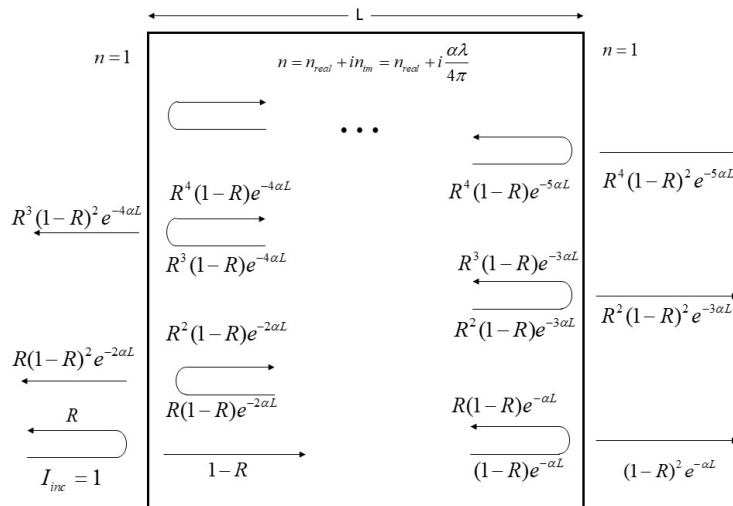
In Chapter 1, we considered the interaction of a coherent wave in a layered medium, where the superposition of reflected and transmitted fields was used together with the boundary conditions to derive the reflection and transmission coefficients.

In this problem, the illumination is incoherent, so we have to add the intensities at each interface.

If the incident intensity is $I_{\text{inc}} = 1$, the reflected intensity from the first interface is $I_{R1} = R$, and the transmitted intensity at this interface is $I_{T1} = (1-R)$. Since the medium is absorptive, the intensity across one length of the crystal satisfies the Beer–Lambert law $I(L) = I(0)e^{-\alpha L}$.

a) To calculate the total reflectivity and transmission, we need to consider the multiple reflections of the beam inside the plate.

The figure shows the first reflections and transmissions together with the value of the intensity at each step, considering that the incident intensity is 1.





The total reflectivity is obtained by summing all the terms leaving the plate to the left. By inspection of the few terms shown in the figure, this sum can be expressed as

$$\mathfrak{R} = I_R / I_{inc} = R + R(1-R)^2 e^{-2\alpha L} \sum_{m=0}^{\infty} R^{2m} e^{-2m\alpha L}$$

By summing this geometric series, we obtain

$$\mathfrak{R} = R + \frac{R(1-R)^2 e^{-2\alpha L}}{1 - R^2 e^{-2\alpha L}}$$

In the same way, the transmission is

$$T = I_T / I_{inc} = (1-R)^2 e^{-\alpha L} \sum_{m=0}^{\infty} R^{2m} e^{-2m\alpha L} = \frac{(1-R)^2 e^{-\alpha L}}{1 - R^2 e^{-2\alpha L}}$$

In the limit of zero absorption, one can check that the condition $\mathfrak{R} + T = 1$ holds.

In the general case, these expressions allow us to obtain the absorbed intensity through the relation $I_{abs} = I_{inc} - I_R - I_T$

b) In this case, the amount of light reflected at the second layer $I_{R2} = RI_2 = R(1-R)e^{-\alpha L}$ is assumed to be already negligible under our approximation.

The intensity transmitted by the plate (remember that we are considering only one reflection) is

$$I_T = I_2(1-R) = (1-R)^2 e^{-\alpha L}$$

which corresponds to the first term in the series obtained in a).

The sum of the reflected and transmitted radiation is

$$I_R + I_T = R + (1-R)^2 e^{-\alpha L} = R + e^{-\alpha L} - 2Re^{-\alpha L} + R^2 e^{-\alpha L} = 1 - (1-R - (1-R)e^{-\alpha L}) - (1-R)Re^{-\alpha L}$$

The absorbed intensity is equal to $I_{abs} = (1-R)(1 - e^{-\alpha L}) = 1 - R - (1-R)Re^{-\alpha L}$

$$I_R + I_T + I_{abs} = 1 - (1-R)e^{-\alpha L} \approx 1 \quad \text{for } \alpha L \gg 1$$

■

S3.9 The power density transferred from an external EM plane wave oscillating at frequency ω and a collection of identical atoms was described in problem S3.3. A more refined model, the rate equation model, is used to describe the interaction of light with atoms. In the so-called two-level approximation, we assume that the frequency of the external field, ω , is very close to a resonant transition, ω_b , in the atoms of the medium and that the population changes to other levels can be neglected. The rate of change of the populations of the two levels involved in the transition are given by



$$\frac{dN_2}{dt} = -\frac{dN_1}{dt} = W_{12}N_1 - W_{21}N_2$$

Here W_{12} and W_{21} are defined as the absorption and stimulated transition rates respectively. In a non-degenerate system $W_{12}=W_{21}$. This equation describes the increase of population of the upper level by the process of absorption and its depletion by the process of stimulated emission. Show that within this model the stimulated transition rate in this system can be expressed in terms of the extinction cross-section of the individual atoms of the medium.

Solution

In a semiclassical approach the atoms are described by an atomic Hamiltonian with stationary eigenstates, characterized by their energy $E = \hbar\omega$, and characterized by a set of quantum numbers. The state j for the collection of atoms is populated with a certain density number, N_j .

The energy density of the medium (in the two-level approximation) can be expressed as $u = N_1E_1 + N_2E_2$. Under the influence of the external EM wave the state of the atom can change through transitions between the energy levels and leading to a temporal variation of the populations.

Absorption of photons at a rate W_{12} increases the medium energy density per unit time by an amount $W_{12}\hbar\omega$ while stimulated emission decreases the power density by an amount $W_{21}\hbar\omega$. Hence, the net power transfer between the two states induced by the external field can be written as

$$\langle P_{trans} \rangle = \frac{du}{dt} = W_{12}\hbar\omega N_1 - W_{21}\hbar\omega N_2 = W_{21}\hbar\omega (N_1 - N_2)$$

here we used the fact that $\frac{dN}{dt} = \frac{1}{\hbar\omega} \frac{du}{dt}$

With this result a direct comparison can be established between this result and that obtained in problem S3.4.

$$\langle P_{transfer} \rangle = \frac{\omega\omega_p^2}{\omega_o\gamma_{tot}c} \cdot \frac{I_{inc}}{1 + (2(\omega_o - \omega)/\gamma_{tot})^2} = W_{12}\hbar\omega(N_1 - N_2)$$

From equation [3.17], we can rewrite the plasma frequency in terms of the spontaneous emission decay rate:

$$\gamma = \frac{\omega_o^2 q^2}{6\pi m \epsilon_o c^3} = \frac{\omega_o^2 \omega_p^2}{6\pi c^3 N} \rightarrow \omega_p^2 = \frac{\gamma 6\pi c^3 N}{\omega_o^2}$$

For the two-level approximation the possible transitions are related to the population difference, so we can assume that the classical variable N , corresponding to the density of dipoles can be substituted by $N=(N_1-N_2)$.



With this change

$$\langle P_{transfer} \rangle = \frac{\omega 6\pi c^2 \gamma}{\omega_o^3 \gamma_{tot}} \cdot \frac{I_{inc} (N_1 - N_2)}{1 + (2(\omega_o - \omega)/\gamma_{tot})^2} = W_{12} \hbar \omega (N_1 - N_2)$$

And we obtain an expression for the transition probability:

$$W_{12} = \frac{6\pi c^2 \gamma}{\hbar \omega_o^3 \gamma_{tot}} \cdot \frac{I_{inc}}{1 + (2(\omega_o - \omega)/\gamma_{tot})^2}$$

To relate this expression with the extinction cross-section we need to consider its frequency dependence for quasi-resonant conditions. Under this approximation, equation [3.28] is transformed, giving

$$\sigma_{ext} \approx \frac{q^2 \gamma_{tot} \omega_o^2}{\epsilon_o c m} \cdot \frac{1}{(2\omega_o(\omega_o - \omega))^2 + \gamma_{tot}^2 \omega_o^2} = \frac{q^2}{\epsilon_o c m \gamma_{tot}} \cdot \frac{1}{1 + (2(\omega_o - \omega)/\gamma_{tot})^2}$$

Using again the expression for the spontaneous emission decay rate,

$$\sigma_{ext} \approx \frac{\gamma 6\pi c^2}{\omega_o^2 \gamma_{tot}} \cdot \frac{1}{1 + (2(\omega_o - \omega)/\gamma_{tot})^2}$$

This result can be replaced in the expression for W_{12} to give

$$W_{12} = \sigma_{ext} \cdot \frac{I_{inc}}{\hbar \omega_o} = \sigma_{ext} \cdot \phi$$

Here $\phi = I_{inc}/\hbar \omega_o$ corresponds to the incident photon flux. These results show that the stimulated emission/absorption probability is proportional to the scattering cross-section of the single atom and to the incident photon flux.

For the case of degenerate energy states with degeneracy factors g_1 and g_2 :

$$W_{12} = \frac{g_1}{g_2} W_{21} = \sigma_{ext} \cdot \phi$$

The rate of change of the atomic population can be expressed, alternatively, in terms of the Einstein coefficients. The Einstein coefficient of stimulated emission, B_{21} , defined as

$$\frac{dN_1}{dt} = -\frac{dN_2}{dt} = B_{21} \cdot u \cdot N_2$$

Can be obtained in terms of the extinction coefficient, using the relation between the intensity and the energy density

$$\frac{dN_1}{dt} = W_{21} N_2 = \sigma_{ext} \phi N_2 = \sigma_{ext} \frac{I}{\hbar \omega} N_2 = \sigma_{ext} \frac{cu}{\hbar \omega} N_2 = B_{21} u N_2$$

■



3.7 Proposed problems

P3.1 A single linear dipole, located at the origin of our system of reference, with constant amplitude, p_0 , rotates in the xOy-plane in the anti-clockwise sense around the z-axis with a constant angular velocity, ω . Using the general expression of the radiated fields of a generic dipole, $\vec{p} = p_0 \hat{d}$, in the electric dipole (long wavelength) and radiation zone approximation, we have

$$\vec{e}_{rad} = \frac{\mu_0 \ddot{\vec{p}}}{4\pi R} \left[\hat{u} \wedge (\hat{u} \wedge \hat{d}) \right]$$

$$\vec{b}_{rad} = \frac{\mu_0 \ddot{\vec{p}}}{4\pi R c} \left[\hat{d} \wedge \hat{u} \right]$$

Write the expression for the total radiated electric and magnetic fields and the total intensity along the Z-axis.

P3.2 To properly describe the dispersive properties of noble metals in the visible, the dielectric function has to account for the contributions of both the free electrons and the bound electrons in the metal. The expression for the dielectric constant includes a Drude-type contribution and two Lorentz-type resonant dispersion terms. A model describing the dispersive behaviour of gold in the 200–1200 nm range is given by

$$\varepsilon(\tilde{\omega}) = 1 - \frac{6.965^2}{\tilde{\omega}^2 + i0.0573\tilde{\omega}} - \frac{4.25^2}{\tilde{\omega}^2 - 3.75^2 + i1.42\tilde{\omega}} - \frac{5.5^2}{\tilde{\omega}^2 - 5^2 + i1.42\tilde{\omega}}$$

Here $\tilde{\omega} = 1/\lambda$, where λ is in microns.

Plot the different contributions separately and describe the observed behaviour with respect to the situation where all terms are included. Describe the effect of the Lorentz-type resonances in the global dispersion relation, indicating the values of the two resonant wavelengths of the Lorentz-type oscillators

P3.3 Find the ratio of the total power radiated by a dipole oscillator when emitting at the wavelengths $\lambda_1 = 400$ nm and $\lambda_2 = 700$ nm in the Rayleigh scattering regime.

P3.4 Obtain the wave equation for the magnetic intensity, H , from the macroscopic Maxwell equations for a medium that is homogeneous, isotropic, non-dispersive, and in the linear regime.

P3.5 Consider a dielectric medium with no conductivity. Give an expression for the real part of the index of refraction and absorption coefficient in terms of the real and imaginary parts of the electric susceptibility for:

- a) weakly absorbing medium satisfying the relation $\chi_{im} \ll 1 + \chi_{real}$
- b) strongly absorbing media, for which $\chi_{im} \gg 1 + \chi_{real}$

P3.6 Enter into the webpage <http://refractiveindex.info/> and check the index of refraction for the following materials: BK7, Methane (CH₄), and gold (Au). Comment on the differences found.



P3.7 A smooth and flat interface between a metal with relative permittivity ϵ_m and a lossless dielectric with relative permittivity ϵ_d , can support localized surface modes called surface plasmons. The dispersion relation for incident TM polarization is given by

$$\beta = k_o \sqrt{\frac{\epsilon_m \epsilon_d}{\epsilon_m + \epsilon_d}}$$

Here β is the wavevector component along the surface. The surface plasmon frequency corresponds to a resonance of this dispersion relation, and corresponds to a pole in the expression for β . Find the value of the surface plasmon frequency for the situation of an interface between a lossless metal with plasma frequency ω_p and air.

P3.8 a) Obtain an expression for the transmission (for light at normal incidence) through a slab as the one described in the solved problem S3.8, for a strongly absorbing medium (i.e. $\alpha L \gg 1$), where the multiple reflections on the slab can be neglected.

b) Use the formula derived in a) to obtain the transmission of a 10 μm thick slab of silicon at the wavelength of 633 nm ($n_{\text{Si}}(633) = 3.8823 + i 0.019589$).

P3.9 a) The absorption of an optical medium can be quantified in terms of the optical density (OD) defined as

$$OD = -\log_{10} \left(\frac{I(l)}{I(0)} \right)$$

Show that this expression is equivalent to $OD = 0.434\alpha L$.

b) Calculate the transmission and the OD through a set of two identical plates of GaAs wafers 20 μm thick for incident radiation at the wavelength of 1064 nm and at normal incidence. ($n_{\text{GaAs}} = 3.4622 + i 0.014163$)

P3.10 Repeat problem P3.7 considering now a lossy metal with $\Gamma = 0.1\omega_p$. (Use the expressions for conducting media derived in section 3.4.4)



→ 4



Crystal Optics

Introduction

Chapter 3 established the material relations for different types of materials. The solution of the wave equation for the case of isotropic crystals showed that plane wave solutions correspond to normal modes of propagation in these media. The effect of the isotropic medium is reflected in the form of the dispersion relation in the material medium, $k(\omega) = n(\omega)k_o$, where the index of refraction is independent of the propagation direction:

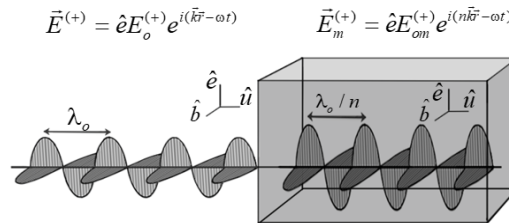


Figure 4.1
Propagation modes
inside the crystal

These solutions exist for any propagation direction and polarization state, and they correspond to propagation modes in isotropic media such as glass, gases, liquids, and crystals with cubic symmetry.

In the general case of crystalline solids, the medium is **anisotropic**, and we observe that the optical properties depend on the propagation direction inside the crystal and the polarization state.

In anisotropic media, the polarization density, P , that is generated in the medium for a given incident electric field is not parallel to the electric field and, consequently, the displacement vector, D , is also not parallel to E . This is due to the fact that the crystalline structure of the lattice impedes the free movement of electrons in every direction in space. The angle between the D and E vectors is called the **birefringence angle**.



This fact has some immediate consequences, which are derived directly from Maxwell's equations.

1) Assuming there are no free charges:

$$\vec{\nabla} \cdot \vec{D} = 0$$

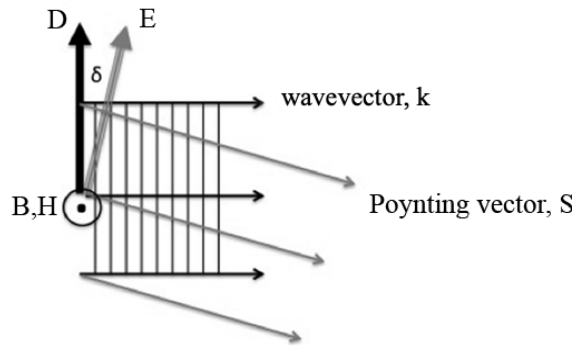
$$\vec{\nabla} \cdot \vec{B} = 0$$

When expressed in the reciprocal space, these relations have a clear geometrical significance: the displacement vector and magnetic field are always perpendicular to the wave vector propagation direction.

2) The magnetic intensity (H) is parallel to the magnetic field (B) in non-magnetic materials. From the definition of the Poynting vector, we see that the energy flows in the direction perpendicular to the plane containing E and H . Since E is not parallel to D , we extract the conclusion that in anisotropic media **the direction of energy flow is not parallel to the direction of wavefront propagation.**

Figure 2 summarizes the relative orientation and propagation directions of the different magnitudes.

Figure 4.2
Relative orientation
and propagation di-
rection of some rele-
vant magnitudes



4.1 Normal modes of propagation in birefringent media

To study the propagation of optical radiation in crystals, we begin by deriving the wave equation for anisotropic, linear, homogeneous, monochromatic, nonconductive ($\sigma = 0$), and non-magnetic ($\mu_r = 1$) media. Under these conditions, the wave equation reads

$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \mu_o \frac{\partial^2 \vec{D}}{\partial t^2} = 0 \quad [4.1]$$

The relation between D and E for anisotropic and monochromatic radiation is

$$D_i(\omega) = \epsilon_o E_i(\omega) + \epsilon_o \chi_{ij}^{(1)}(\omega) E_j(\omega) = \epsilon_o (\delta_{ij} + \chi_{ij}^{(1)}(\omega)) E_j(\omega) = \epsilon_o \epsilon_{ij}(\omega) E_j(\omega)$$

We will omit the frequency argument in the following, assuming that we are always considering monochromatic beams. The dielectric tensor relating the displacement vector and electric field is

$$\begin{array}{cc} \text{General reference system} & \text{Crystallographic Reference System} \\ \varepsilon_{ij} = \delta_{ij} + \chi_{ij}^{(1)} = \begin{pmatrix} \varepsilon'_{xx} & \varepsilon'_{xy} & \varepsilon'_{xz} \\ \varepsilon'_{yx} & \varepsilon'_{yy} & \varepsilon'_{yz} \\ \varepsilon'_{zx} & \varepsilon'_{zy} & \varepsilon'_{zz} \end{pmatrix} & \rightarrow \varepsilon_{ij} = \begin{pmatrix} \varepsilon_{xx} & 0 & 0 \\ 0 & \varepsilon_{yy} & 0 \\ 0 & 0 & \varepsilon_{zz} \end{pmatrix} \end{array}$$

In a general reference system, we should retain all nine components. By choosing the crystallographic system of reference, the matrix becomes diagonal, where the z -axis is called the **optical axis**. In this system:

$$D_i = \varepsilon_o \delta_{ij} \varepsilon_{ij} E_j = \varepsilon_o \varepsilon_{ii} E_i \quad [4.2]$$

Assuming plane wave solutions for the electric field and displacement vector:

$$\vec{E} = \hat{e} E_o^{(+)} e^{i(\vec{k}\vec{r} - \omega t)} \quad \text{and} \quad \vec{D} = \hat{d} D_o^{(+)} e^{i(\vec{k}\vec{r} - \omega t)}$$

and using [4.2], the wave Equation [4.1] is written as

$$\begin{aligned} i\vec{k} \wedge (i\vec{k} \wedge \vec{E}) - \mu_o \omega^2 \vec{D} &= 0 \\ \vec{k} \cdot (\vec{k} \cdot \vec{E}) - k^2 \vec{E} + \mu_o \varepsilon_o \omega^2 \vec{E} &= 0 \rightarrow \boxed{k_i (k_j E_j) - (k^2 - k_o^2 \varepsilon_{ii}) E_i = 0} \end{aligned} \quad [4.3]$$

This equation, equivalent to the wave equation, is the **Fresnel equation**, which can also be expressed in matrix form as $T_{ij} E_j = 0$:

$$\begin{pmatrix} k_o^2 \varepsilon_x - k^2 + k_x^2 & k_x k_y & k_x k_z \\ k_y k_x & k_o^2 \varepsilon_y - k^2 + k_y^2 & k_y k_z \\ k_z k_x & k_z k_y & k_o^2 \varepsilon_z - k^2 + k_z^2 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad [4.4]$$

The system has a non-trivial solution if $\det(T) = 0$. The determinant is given by the expression

$$\begin{aligned} k^2 (\varepsilon_x k_x^2 + \varepsilon_y k_y^2 + \varepsilon_z k_z^2) - k_o^2 [k_x^2 (\varepsilon_x \varepsilon_y + \varepsilon_x \varepsilon_z) + k_y^2 (\varepsilon_y \varepsilon_x + \varepsilon_y \varepsilon_z) + k_z^2 (\varepsilon_z \varepsilon_x + \varepsilon_z \varepsilon_y)] + \\ + k_o^4 \varepsilon_x \varepsilon_y \varepsilon_z = 0 \end{aligned} \quad [4A.5]$$

The solutions to this expression correspond to the values of the wave vector for the different normal modes of propagation in the medium.



Basic solutions

According to the symmetry of their crystalline structure, crystals can be classified in the following groups.

A. Isotropic crystals

Crystalline class: Cubic

These crystals are characterized by a single dielectric constant value, independently of the propagation direction $\epsilon_x = \epsilon_y = \epsilon_z = \epsilon$.

B. Uniaxial crystals

Crystalline classes: Hexagonal, Tetragonal, Trigonal

In uniaxial crystals, we distinguish two different dielectric constant values

- Ordinary: $\epsilon_{or} = \epsilon_x = \epsilon_y$
- Extraordinary: $\epsilon_e = \epsilon_z$

The extraordinary value corresponds to the value along the optical axis.

C. Biaxial crystals

Crystalline class: Monoclinic, Triclinic and Orthorhombic

In biaxial crystals, all dielectric constant terms are different: $\epsilon_x \neq \epsilon_y \neq \epsilon_z$.

We will concentrate on analysing the normal modes of propagation of isotropic and anisotropic crystals, but we will not consider the more complex case of biaxial crystals in this course.

4.1.1 Normal mode solutions for isotropic crystals

For isotropic crystals, all coefficients are the same, and expression [4.5] reduces to

$$\begin{aligned}
 k^2 \epsilon k^2 - k_o^2 2\epsilon^2 k^2 + k_o^4 \epsilon^3 &= 0 \\
 k^4 - 2\epsilon k_o^2 k^2 + k_o^4 \epsilon^2 &= 0
 \end{aligned}$$

which can be rewritten as

$$(k^2 - \epsilon k_o^2)^2 = 0$$

The solutions for the wave vector are then

$$k = \pm \sqrt{\epsilon} k_o = \pm n k_o \quad [4.6]$$

In the case of isotropic media, we know that the normal mode of propagation corresponds to a uniform wave vector with a homogeneous index of refraction, n , which depends on frequency but is constant in all directions.

4.1.2 Normal modes of propagation for uniaxial crystals

The determinant of the Fresnel equation for the case of uniaxial crystals is expressed in terms of the ordinary and extraordinary dielectric constants:

$$k^2(\epsilon_{or}k_{\perp}^2 + \epsilon_e k_z^2) - k_O^2(k_{\perp}^2 \epsilon_{or}^2 + k^2 \epsilon_{or} \epsilon_e + k_z^2 \epsilon_{or} \epsilon_e) + k_O^4 \epsilon_{or}^2 \epsilon_e = 0$$

Regrouping some terms, we have

$$k^2(\epsilon_{or}k_{\perp}^2 + \epsilon_e k_z^2) - k_O^2 \epsilon_{or}(k_{\perp}^2 \epsilon_{or} + k_z^2 \epsilon_e) - k_O^2 k^2 \epsilon_{or} \epsilon_e + k_O^4 \epsilon_{or}^2 \epsilon_e = 0$$

The expression can be written as

$$(\epsilon_{or}k_{\perp}^2 + \epsilon_e k_z^2)(k^2 - \epsilon_{or}k_O^2) - k_O^2 \epsilon_{or} \epsilon_e (k^2 - \epsilon_{or}k_O^2) = 0$$

and finally expressed as

$$(k^2 - \epsilon_{or}k_O^2)(\epsilon_{or}k_{\perp}^2 + \epsilon_e k_z^2 - \epsilon_{or} \epsilon_e k_O^2) = 0 \quad [4.7A]$$

where the transverse wave vector is given by $k_{\perp}^2 = k_x^2 + k_y^2$.

Expression [4.7A] has two different solutions corresponding to the vanishing of each of the parentheses.

A. Ordinary mode of propagation, k_{or}

$$k_{or}^2 - \epsilon_{or}k_O^2 = 0 \quad \rightarrow \quad k_{or} = \pm \sqrt{\epsilon_{or}} k_O \quad [4.7B]$$

This solution, called the **ordinary mode of propagation**, corresponds to an isotropic-like case with constant index of refraction, n_{or} , given by

$$\text{Ordinary Index: } n_{or} = \sqrt{\epsilon_{or}} \quad [4.7C]$$

B. Extraordinary mode of propagation, k_e

The extraordinary mode of propagation, k_e , is found from the relation

$$\epsilon_{or}k_{e\perp}^2 + \epsilon_e k_{ez}^2 - k_O^2 \epsilon_{or} \epsilon_e = 0$$

which is rewritten in the form

$$\frac{k_{e\perp}^2}{\epsilon_e} + \frac{k_{ez}^2}{\epsilon_{or}} = k_O^2 \quad [4.8A]$$

In order to find the expression for the extraordinary index of refraction, we assume that the wave vector propagates in a generic direction that is expressed in spherical coordinates:

$$\vec{k}_e = k_e (\cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k})$$

where we choose the z -axis as directed along the optical axis.

The modulus of the extraordinary wave vector is $k_e = n_e k_O$, so:



$$k_{e\perp}^2 = k_{ex}^2 + k_{ey}^2 = n_e^2 k_O^2 (\cos^2 \phi \sin^2 \theta + \sin^2 \phi \sin^2 \theta) = n_e^2 k_O^2 \sin^2 \theta$$

$$k_{ez}^2 = n_e^2 k_O^2 \cos^2 \theta$$

Equation [4.8A] is expressed as

$$\frac{n_e^2 \sin^2 \theta}{\epsilon_e} + \frac{n_e^2 \cos^2 \theta}{\epsilon_{or}} = 1$$

thus giving the final form of the extraordinary index of refraction. Opposite to the ordinary mode, the extraordinary index of refraction depends on the angle of propagation with respect to the optical axis:

$$n_e(\theta) = \frac{1}{\sqrt{\frac{\sin^2 \theta}{\epsilon_e} + \frac{\cos^2 \theta}{\epsilon_{or}}}} \quad [4.8B]$$

This expression takes particular values at two specific propagation angles:

$$\text{At } \theta = 0^\circ \rightarrow n_e(0) = \sqrt{\epsilon_{or}} = n_{or}$$

$$\text{At } \theta = 90^\circ \rightarrow n_e(90) = \sqrt{\epsilon_e} = n_e$$

Any uniaxial crystal is characterized by these two different indices, which are defined by the ordinary and extraordinary dielectric constants. From these values, we classify uniaxial crystals as

- Positive Uniaxial Crystal: Crystal with $n_e > n_{or}$
- Negative Uniaxial Crystal: Crystal with $n_e < n_{or}$

And define the medium birefringence: $|n_e - n_{or}|$

For ordinary modes, the value of the index of refraction for different propagation directions can be represented as a sphere, and for extraordinary modes as a revolution ellipsoid around the optical axis.

Figure 3 shows the plot of the refractive indices for both types of crystals. For any generic propagation direction, which is defined by the azimuthal angle with respect to the optical axis, θ , the intersections with the circumference and the ellipse provide the values of the ordinary and extraordinary refractive indices at that particular propagation direction.

The plots shown in Figure 3 correspond to a particular value of the wavelength. As we have seen in previous chapters, material dispersion is at the origin of the wavelength dependence of the refractive indices, so the most generic expression for the ordinary and extraordinary refractive indices is the following.

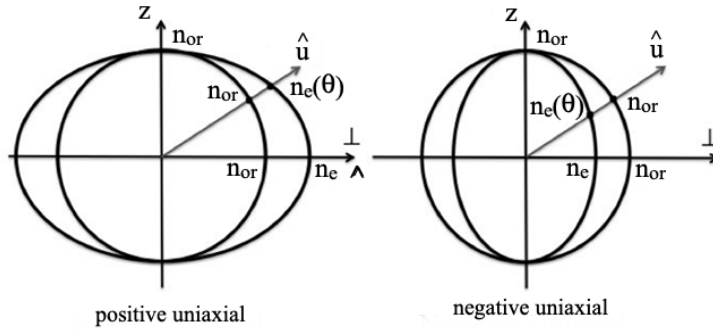


Figure 4.3
Ordinary and extraordinary refractive indices for positive and negative uniaxial crystals

$$\text{Ordinary Mode: } n_{or}(\omega) = \sqrt{\epsilon_{or}(\omega)} \quad \rightarrow \quad k_{or}(\omega) = n_{or}(\omega)k_o$$

$$\text{Extraordinary Mode: } n_e(\omega, \theta) = \frac{1}{\sqrt{\frac{\sin^2 \theta}{\epsilon_e(\omega)} + \frac{\cos^2 \theta}{\epsilon_{or}(\omega)}}} \quad \rightarrow \quad k_e(\omega, \theta) = n_e(\omega, \theta)k_o$$

In the normal dispersion regime, the index of refraction increases with frequency, so different ellipses can be represented for different wavelengths. For any specific crystal, the values of the refractive indices, $n_{or}(\lambda)$ and $n_e(\lambda)$, can be found in standard optical handbooks such, as Palik, E.D. (1998). *Handbook of Optical Constants of Solids*, Academic Press.

The expression for the normal modes in uniaxial crystals can be expressed in terms of the indices of refraction as follows.

$$\text{Ordinary Mode: } \vec{E}_{or}^{(+)} = \hat{e}_{or} E_{0or}^{(+)} \exp[i(\vec{k}_{or} \vec{r} - \omega t)] \quad [4.9]$$

$$\text{Extraordinary Mode: } \vec{E}_e^{(+)} = \hat{e}_e E_{0e}^{(+)} \exp[i(\vec{k}_e \vec{r} - \omega t)]$$

4.2 Polarization state of the normal modes

Once we have obtained the expressions for the dispersion relation corresponding to the ordinary and extraordinary modes, we can proceed now to determine the polarization directions of each one. To do this, we start with the Fresnel equation [4.4] to find the values of the corresponding eigenvectors.

4.2.1 Isotropic case

In the case of isotropic crystals, we have $\epsilon_x = \epsilon_y = \epsilon_z = \epsilon$ and $k^2 = \epsilon k_o^2$

Substituting in [4.4] gives

$$\begin{pmatrix} k_x^2 & k_x k_y & k_x k_z \\ k_y k_x & k_y^2 & k_y k_z \\ k_z k_x & k_z k_y & k_z^2 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$



We obtain from this the relation:

$$k_x E_x + k_y E_y + k_z E_z = 0 \quad \rightarrow \quad \boxed{\vec{k} \cdot \vec{E} = 0} \quad [4.10]$$

This relation retrieves the well-known result for the isotropic case: the transverse character of the electric field with respect to the direction of wavefront propagation. Any state of polarization can be defined for this solution, since there are no restrictions imposed by Fresnel's equation in this case.

4.2.2 Uniaxial crystals

1. Ordinary Mode

This solution is obtained by considering an arbitrary propagation direction inside the crystal (defined in terms of spherical coordinates, with the z -axis corresponding to the optical axis). The ordinary wave vector is written as

$$\begin{aligned} \vec{k}_{or} &= n_{or} k_o \hat{u} = \sqrt{\epsilon_{or}} k_o (\cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k}) = \\ &= k_{or} (\cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k}) \end{aligned}$$

Fresnel's equation in this case is

$$\begin{pmatrix} k_o^2 \epsilon_{or} - k_{or}^2 + k_{orx}^2 & k_{orx} k_{ory} & k_{orx} k_{orz} \\ k_{ory} k_{orx} & k_o^2 \epsilon_{or} - k_{or}^2 + k_{ory}^2 & k_{ory} k_{orz} \\ k_{orz} k_{orx} & k_{orz} k_{ory} & k_o^2 \epsilon_e - k_{or}^2 + k_{orz}^2 \end{pmatrix} \begin{pmatrix} E_{orx} \\ E_{ory} \\ E_{orz} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

which results in the following three equations:

$$\cos^2 \phi \sin^2 \theta E_{orx} + \sin^2 \theta \cos \phi \sin \phi E_{ory} + \cos \phi \sin \theta \cos \theta E_{orz} = 0 \quad [1]$$

$$\sin^2 \theta \cos \phi \sin \phi E_{orx} + \sin^2 \phi \sin^2 \theta E_{ory} + \sin \phi \sin \theta \cos \theta E_{orz} = 0 \quad [2]$$

$$k_{or}^2 \cos \phi \sin \theta \cos \theta E_{orx} + k_{or}^2 \sin \phi \sin \theta \cos \theta E_{ory} + [k_o^2 (n_e^2 - n_{or}^2) + k_{or}^2 \cos^2 \theta] E_{orz} = 0 \quad [3]$$

From [1], we obtain

$$E_{orx} \cos \phi + E_{ory} \sin \phi + \frac{\cos \theta}{\sin \theta} E_{orz} = 0 \quad [4]$$

and [3] can be rewritten as

$$E_{orx} \cos \phi + E_{ory} \sin \phi + \frac{k_o^2 (n_e^2 - n_{or}^2)}{k_{or}^2 \sin \theta \cos \theta} E_{orz} + \frac{\cos \theta}{\sin \theta} E_{orz} = 0$$

Using [4], this equation reduces to:

$$k_o^2 (n_e^2 - n_{or}^2) E_{orz} = 0$$

Since, in uniaxial crystals, n_e is different from n_{or} , it must be true that $E_{orz} = 0$.

Then, from [4], we obtain

$$\frac{E_{orx}}{E_{ory}} = \frac{-\sin \phi}{\cos \phi} \quad [4.11A]$$

The polarization unitary vector for ordinary modes is then

$$\hat{e}_{or} = -\sin \phi \hat{i} + \cos \phi \hat{j} \quad [4.11B]$$

Being perpendicular to both, $\hat{u}$ and $\hat{k}$, from this result, we can conclude that **the ordinary mode is always polarized in the direction perpendicular to the plane containing the optical axis and the wave vector propagation direction** (plane of incidence).

2. Extraordinary mode

For this mode, we have the relations

$$\vec{k}_e(\theta) = k_o n_e(\theta) \hat{u} = k_e(\theta) (\cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k})$$

with

$$k_o^2 = k_e^2(\theta) \left[\frac{\sin^2 \theta}{\epsilon_e} + \frac{\cos^2 \theta}{\epsilon_{or}} \right] \quad [4.12A]$$

The Fresnel equation for the extraordinary modes reads as

$$\begin{pmatrix} k_o^2 \epsilon_{or} - k_e^2 + k_{ex} & k_{ex} k_{ey} & k_{ex} k_{ez} \\ k_{ey} k_{ex} & k_o^2 \epsilon_{or} - k_e^2 + k_{ey}^2 & k_{ey} k_{ez} \\ k_{ez} k_{ex} & k_{ez} k_{ey} & k_o^2 \epsilon_e - k_e^2 + k_{ez}^2 \end{pmatrix} \begin{pmatrix} E_{ex} \\ E_{ey} \\ E_{ez} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

In this expression, we substitute the expression [4.12A] to obtain the following three equations.

$$\left[k_e^2 \frac{\epsilon_{or} \sin^2 \theta}{\epsilon_e} + k_e^2 \cos^2 \theta - k_e^2 + k_e^2 \cos^2 \phi \sin^2 \theta \right] E_{ex} + k_e^2 \sin^2 \theta \cos \phi \sin \phi E_{ey} + k_e^2 \cos \phi \sin \theta \cos \theta E_{ez} = 0 \quad [5]$$

$$k_e^2 \sin^2 \theta \cos \phi \sin \phi E_{ex} + \left[k_e^2 \frac{\epsilon_{or} \sin^2 \theta}{\epsilon_e} + k_e^2 \cos^2 \theta - k_e^2 + k_e^2 \sin^2 \phi \sin^2 \theta \right] E_{ey} + k_e^2 \sin \phi \sin \theta \cos \theta E_{ez} = 0 \quad [6]$$

$$k_e^2 \cos \phi \cos \theta \sin \theta E_{ex} + k_e^2 \cos \theta \sin \theta \sin \phi E_{ey} + \left[k_e^2 \sin^2 \theta + k_e^2 \frac{\epsilon_e \cos^2 \theta}{\epsilon_{or}} - k_e^2 + k_e^2 \cos^2 \theta \right] E_{ez} = 0 \quad [7]$$

which can be rewritten, respectively, as

$$\left[\frac{\epsilon_{or} \sin^2 \theta}{\epsilon_e} - \sin^2 \theta \sin^2 \phi \right] E_{ex} + \sin^2 \theta \cos \phi \sin \phi E_{ey} + \sin \theta \cos \phi \cos \theta E_{ez} = 0$$



$$\sin^2 \theta \cos \phi \sin \phi E_{ex} + \left[\frac{\epsilon_{or} \sin^2 \theta}{\epsilon_e} - \sin^2 \theta \cos^2 \phi \right] E_{ey} + \sin \phi \sin \theta \cos \theta E_{ez} = 0$$

$$\cos \phi \cos \theta \sin \theta E_{ex} + \cos \theta \sin \theta \sin \phi E_{ey} + \frac{\epsilon_e \cos^2 \theta}{\epsilon_{or}} E_{ez} = 0$$

and further simplified to obtain

$$\left[\frac{\epsilon_{or}}{\epsilon_e} - \sin^2 \phi \right] E_{ex} + \cos \phi \sin \phi E_{ey} + \frac{\cos \phi \cos \theta}{\sin \theta} E_{ez} = 0 \quad [8]$$

$$\cos \phi \sin \phi E_{ex} + \left[\frac{\epsilon_{or}}{\epsilon_e} - \cos^2 \phi \right] E_{ey} + \frac{\sin \phi \cos \theta}{\sin \theta} E_{ez} = 0 \quad [9]$$

$$\cos \phi E_{ex} + \sin \phi E_{ey} + \frac{\epsilon_e \cos \theta}{\epsilon_{or} \sin \theta} E_{ez} = 0 \quad [10]$$

By performing the operation [8]–[10]·cosφ, we obtain

$$\left[\frac{\epsilon_{or}}{\epsilon_e} - 1 \right] E_{ex} + \frac{\cos \phi \cos \theta}{\sin \theta} \left[1 - \frac{\epsilon_e}{\epsilon_{or}} \right] E_{ez} = 0 \rightarrow \boxed{E_{ex} = -\frac{\epsilon_e \cos \phi \cos \theta}{\epsilon_{or} \sin \theta} E_{ez}}$$

And from the operation [9]–[10]·sinφ, we have

$$\left[\frac{\epsilon_{or}}{\epsilon_e} - 1 \right] E_{ey} + \frac{\sin \phi \cos \theta}{\sin \theta} \left[1 - \frac{\epsilon_e}{\epsilon_{or}} \right] E_{ez} = 0 \rightarrow \boxed{E_{ey} = -\frac{\epsilon_e \sin \phi \cos \theta}{\epsilon_{or} \sin \theta} E_{ez}}$$

From these relations, we can readily find an expression for the extraordinary eigenvector:

$$\vec{E}_e = \left(\frac{\cos \theta \cos \phi}{\epsilon_{or}} \hat{i} + \frac{\cos \theta \sin \phi}{\epsilon_{or}} \hat{j} - \frac{\sin \theta}{\epsilon_e} \hat{k} \right)$$

$$\hat{e}_e = \frac{\vec{E}_e}{|\vec{E}_e|} = \frac{\left(\frac{\cos \theta \cos \phi}{\epsilon_{or}} \hat{i} + \frac{\cos \theta \sin \phi}{\epsilon_{or}} \hat{j} - \frac{\sin \theta}{\epsilon_e} \hat{k} \right)}{\sqrt{\frac{\cos^2 \theta}{\epsilon_{or}^2} + \frac{\sin^2 \theta}{\epsilon_e^2}}}$$

This vector is perpendicular to the direction of polarization of the ordinary mode. Thus, it is contained in the incidence plane. One can check that **the extraordinary mode is polarized in the direction tangent to the ellipse of extraordinary refractive index.**

4.3 Ray propagation in uniaxial crystals

A light ray propagates in a uniaxial crystal with wave vector directed along a given direction (expressed in spherical coordinates, where the z -axis corresponds to the optical axis of the crystal):

$$\hat{u} = (\cos \phi \sin \theta \hat{i} + \sin \phi \sin \theta \hat{j} + \cos \theta \hat{k})$$

We can define a unitary vector that is perpendicular to this direction and contained in the incidence plane (the plane containing $\hat{u}$ and $\hat{k}$):

$$\hat{u}_\perp = (\cos \phi \cos \theta \hat{i} + \sin \phi \cos \theta \hat{j} - \sin \theta \hat{k})$$

From our previous results, we can summarize in the following way the polarization states of the two normal modes propagating in the medium.

Ordinary Mode: polarized in the plane perpendicular to the incidence plane.

$$\hat{e}_{or} = (-\sin \phi \hat{i} + \cos \phi \hat{j})$$

Extraordinary Mode: polarization direction contained in the incidence plane and tangent to the ellipse of extraordinary refractive index at the point where the unitary vector $\hat{u}$ intersects the ellipse. The extraordinary unitary vector forms an angle δ with respect to the perpendicular vector $\hat{u}_\perp$.

$$\hat{e}_e = \frac{\left(\frac{\cos \theta \cos \phi}{\epsilon_{or}} \hat{i} + \frac{\cos \theta \sin \phi}{\epsilon_{or}} \hat{j} - \frac{\sin \theta}{\epsilon_e} \hat{k} \right)}{\sqrt{\frac{\cos^2 \theta}{\epsilon_{or}^2} + \frac{\sin^2 \theta}{\epsilon_e^2}}}$$

The birefringence angle can thus be calculated from the relation

$$\hat{e}_e \cdot \hat{u}_\perp = \cos \delta \quad \rightarrow \quad \cos \delta = \frac{\frac{\cos^2 \theta}{\epsilon_{or}} + \frac{\sin^2 \theta}{\epsilon_e}}{\sqrt{\frac{\cos^2 \theta}{\epsilon_{or}^2} + \frac{\sin^2 \theta}{\epsilon_e^2}}} = \frac{\epsilon_e \cos^2 \theta + \epsilon_{or} \sin^2 \theta}{\sqrt{\epsilon_e^2 \cos^2 \theta + \epsilon_{or}^2 \sin^2 \theta}}$$

Ray energy propagation

Once we know the polarization direction of the modes, we can calculate the propagation direction of the energy wavefronts, which are associated with the Poynting vector. For plane monochromatic waves, we find the relation

$$\vec{S}_{or} = |\vec{S}_{or}| \hat{S}_{or} = \vec{E}_{or} \wedge \vec{H}_{or} = |\vec{E}_{or}| |\vec{H}_{or}| (\hat{e}_{or} \wedge \hat{h}_{or}) = |\vec{E}_{or}| |\vec{H}_{or}| (\hat{e}_{or} \wedge (\hat{u} \wedge \hat{e}_{or})) = |\vec{E}_{or}| |\vec{H}_{or}| \hat{u}$$

where $\hat{h}_{or}$ corresponds to the direction of the magnetic intensity vector. Therefore, the energy wavefronts (perpendicular to $\hat{S}$) in this case propagate parallel to the phase wavefronts (perpendicular to $\hat{u}$).



For the extraordinary mode:

$$\vec{S}_e = |\vec{E}_e| |\vec{H}_e| (\hat{e}_e \wedge \hat{h}_e) = |\vec{E}_e| |\vec{H}_e| (\hat{e}_e \wedge \frac{(\hat{u} \wedge \hat{e}_e)}{\cos \delta}) = |\vec{E}_e| |\vec{H}_e| \frac{(\hat{u} - \sin \delta \hat{e}_e)}{\cos \delta} = |\vec{E}_e| |\vec{H}_e| \hat{S}_e$$

with

$$\hat{S}_e = \frac{\hat{u} - \sin \delta \hat{e}_e}{\cos \delta}$$

where we define the unitary vector in the direction of the Poynting vector for the extraordinary mode. By direct scalar product, we can check that this propagating vector forms an angle δ with the wavefront propagation direction:

$$\hat{S}_e \cdot \hat{u} = \frac{\hat{u} \cdot \hat{u} - \sin \delta \hat{e}_e \cdot \hat{u}}{\cos \delta} = \frac{1 - \sin^2 \delta}{\cos \delta} = \cos \delta$$

Figure 4 summarizes all these aspects for positive and negative uniaxial crystals.

Figure 4a
Polarization direction of the ordinary mode

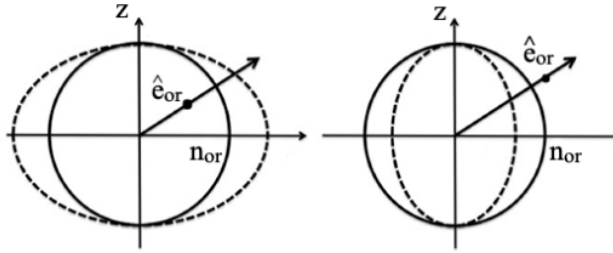
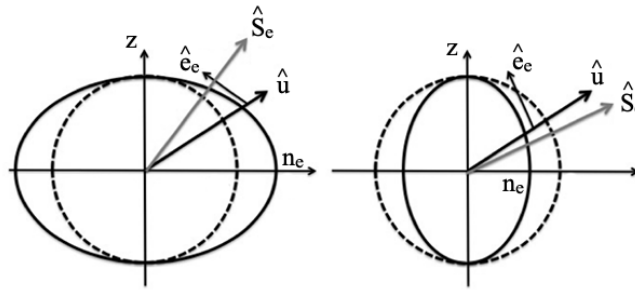


Figure 4b
Polarization direction and Poynting vector of the extraordinary mode



For extraordinary waves, we find the surprising result that the energy wavefronts do not propagate parallel to the phase wavefronts, but instead form an angle between them equal to the birefringence angle. This leads to a spatial separation of the ordinary and extraordinary modes as they propagate inside the uniaxial crystal. This effect is called **spatial walk-off**.

One may wonder about the mechanism that generates the two normal solutions when a beam from an isotropic medium is incident on the uniaxial medium. We provide here only a partial answer to this question, since a highly detailed treatment of this question is beyond the scope of our work here.

At the interface, Snell's law should apply for each mode separately. Assuming that the index of refraction of the incident medium is 1, we have

$$\sin \theta_i = n_{or} \sin \theta_t$$

$$\sin \theta_i = n_e(\theta_i) \sin \theta_t$$

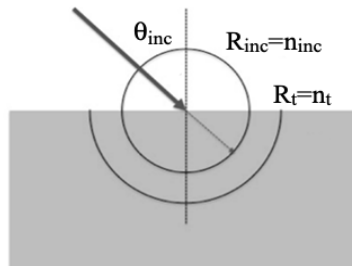
The ordinary and extraordinary modes will be generated only if the corresponding polarization direction is present in the incident beam (within the incidence plane for extraordinary mode and perpendicular to the incidence plane for ordinary mode).

Graphical solution

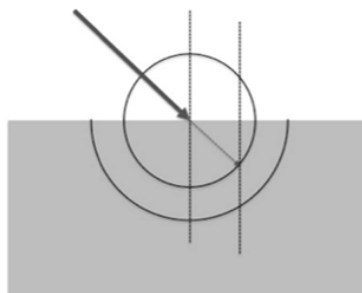
A graphical solution allows determining the propagation direction of the ordinary and extraordinary modes generated at the surface of separation between an isotropic and an anisotropic medium.

The general method for the case of isotropic to isotropic propagation is summarized in the following.

1. Centred at the incident point on the surface, draw two circles with radii proportional to the corresponding refractive indices. Draw the incident beam with the proper angle of incidence, and extend it until it intersects the circle of index n_{inc} .

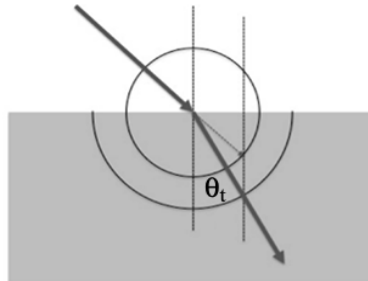


2. Parallel to the normal at the incidence point, draw a line that passes through the intersection point.





3) The transmitted beam can be drawn from the incident point at the surface to the intersection of the line drawn in step 2), with the circle having a radius proportional to n_T .



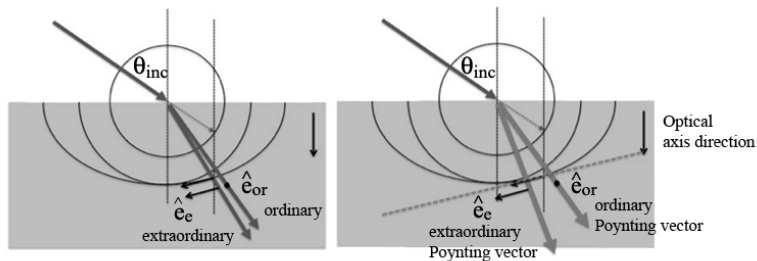
This same general procedure can be used to trace the rays generated at the surface between isotropic and anisotropic media. In this case, a circle and an ellipse should be drawn in the anisotropic medium. The resulting wave vector and energy propagation directions will depend on how the surface is cut with respect to the optical axis of the crystal. The following examples show the general tendency.

The trajectory of the ordinary and extraordinary wave vectors for a positive uniaxial crystal is summarized in the next figures. The polarization state of each mode is also indicated (remember that the extraordinary polarization is tangent to the ellipse at the intersection point).

The propagation of the energy (in the direction of the Poynting vector) is perpendicular to the tangent to the electric field polarization direction and does not coincide with the wave vector propagation direction for the extraordinary mode, as seen in the figure.

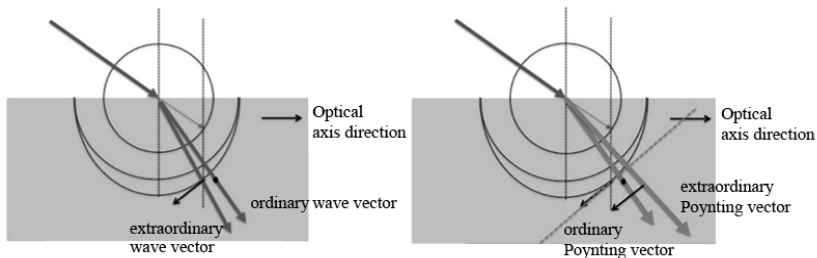
Example 1

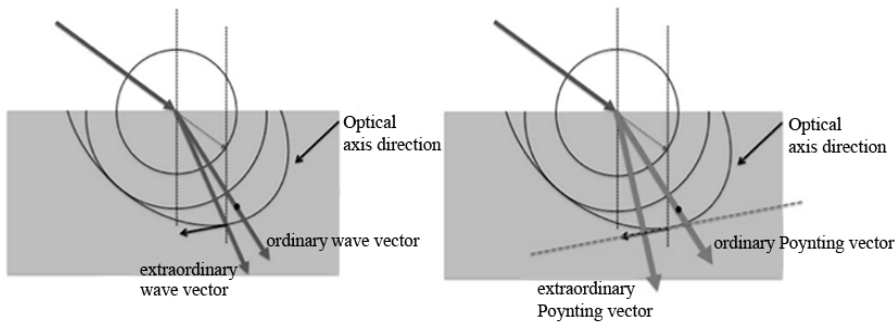
Positive uniaxial with optical axis perpendicular to the incident surface



Example 2

Positive uniaxial with optical axis parallel to the incident surface





Example 3
Positive uniaxial
with arbitrary direc-
tion of the optical
axis

4.4 Applications of wave propagation in uniaxial crystals

4.4.1 Retarders

The properties of uniaxial crystals that we have discussed constitute the basis for the operation of retarders. **Retarders** are devices used to change the polarization state of an incident polarized light beam.

A retarder is made from uniaxial crystals cut with their optical axis parallel to the retarder surface, as shown in Figure 5.

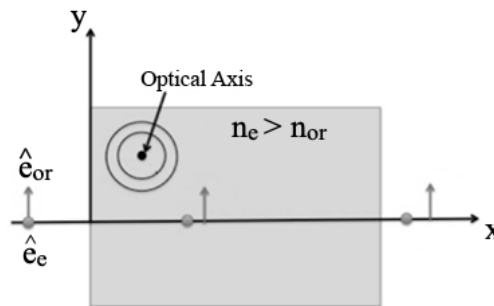


Figure 4.5
A retarder made from a
positive uniaxial crystal in-
troduces a relative phase
between the ordinary and
extraordinary components
of the incident beam.

In the configuration shown, a beam propagating along the OX axis and incident normal to the surface generates two modes propagating inside the crystal.

- Ordinary mode: polarized along the direction of the OY axis.
- Extraordinary mode: polarized along the direction of the OZ axis.

At the same time, the following conditions must hold.

- The ordinary and extraordinary modes will propagate collinearly in the OX direction.
- The index of refraction of the extraordinary mode takes its maximum value, $n_e(90^\circ) = n_e$, while the ordinary mode index is n_{or} .



Assuming that the incident wave has an arbitrary polarization state, its representation in the plane $x = 0$ is

$$\vec{E}_{in}^{(+)} = [E_{or} e^{i\phi_{or}} \hat{j} + E_e e^{i\phi_e} \hat{k}] e^{-i\omega t}$$

This beam will be converted inside the medium into two mutually orthogonal modes, each component propagating with different phase velocity: $v_{or} = c/n_{or}$ and $v_e = c/n_e$

Due to this difference in phase velocity, the modes will get out of phase (retarded in phase one with respect to the other) as they propagate inside the crystal. If the crystal length is L , the field at the output is

$$\vec{E}_{out}^{(+)} = [E_{or} e^{i\phi_{or}} e^{ik_{or}L} \hat{j} + E_e e^{i\phi_e} e^{ik_eL} \hat{k}] e^{-i\omega t}$$

which can be rewritten as

$$\vec{E}_{out}^{(+)} = [E_{or} \hat{j} + E_e e^{i\delta} e^{i\delta_R} \hat{k}] e^{i(k_{or}L - \omega t + \phi_{or})}$$

where $\delta = \phi_e - \phi_{or}$ is the initial relative phase difference, and $\delta_R = (k_e - k_{or})L$ is a relative phase introduced by the retarder.

We can make a physical interpretation of this result if we consider the initial state of polarization with respect to the ordinary (OY) axis. As explained in Chapter 1, and with respect to this axis, the polarization state is given by

$$\hat{e}_{in} = (E_{or} \hat{j} + e^{i\delta} E_e \hat{k}) = E_o (\cos \theta \hat{j} + e^{i\delta} \sin \theta \hat{k})$$

where $E_o^2 = E_{or}^2 + E_e^2$

The transformation of the initial Jones vector by the retarder can be written in matrix form as

$$\begin{pmatrix} e_{or}(L) \\ e_e(L) \end{pmatrix} = e^{ik_{or}L} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta_R} \end{pmatrix} \begin{pmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{pmatrix} \rightarrow \boxed{\vec{e}_{out}(L) = R(\delta_R, L) \hat{e}_{in}}$$

where we define the retarder Jones matrix $R(\delta_R, L)$, which depends on the value of the relative phase introduced by the retarder:

$$\delta_R = (k_e - k_{or})L = \frac{(n_e - n_{or})2\pi L}{\lambda} \quad [4.13]$$

From these results, we can conclude that, for a given crystal, the length of the retarder can be selected to obtain the desired phase difference between the ordinary and extraordinary components. Consequently, to select the polarization state of the output wave, we use

$$\delta_{out} = \delta + \delta_R$$

It can also be seen that the effect of the retarder is wavelength-dependent.

In particular, we find the following two specific retarders that are widely used in any optical system with polarization control.

A. Half-wavelength plate

For a given wavelength, the length of a half-wavelength plate is selected to introduce a relative delay $\delta_R = \pi$. Directly substituting this into in Equation [4.13] gives

$$\pi = \frac{|n_e - n_{or}| 2\pi L_{\lambda/2}}{\lambda} \rightarrow L_{\lambda/2} = \frac{\lambda}{2|n_e - n_{or}|}$$

The effect of the retarder on an incident linearly polarized and circular polarized light is obtained directly from knowing the introduced retarder phase.

Linear polarization is converted to linear polarized light that is rotated with respect to the incident one, changing its polarization angle (measured with respect to the ordinary axis) from $+\theta$ to $-\theta$. Note that linear polarization rotation can also be obtained with single linear polarizers (Malus's law). However, the energy of the incident wave is not changed when using a half-wavelength plate.

In general, half-wavelength plates are used whenever we want to rotate the polarization axis of the light beam; and single linear polarizers are used instead when we want to modulate the intensity of the light beam.

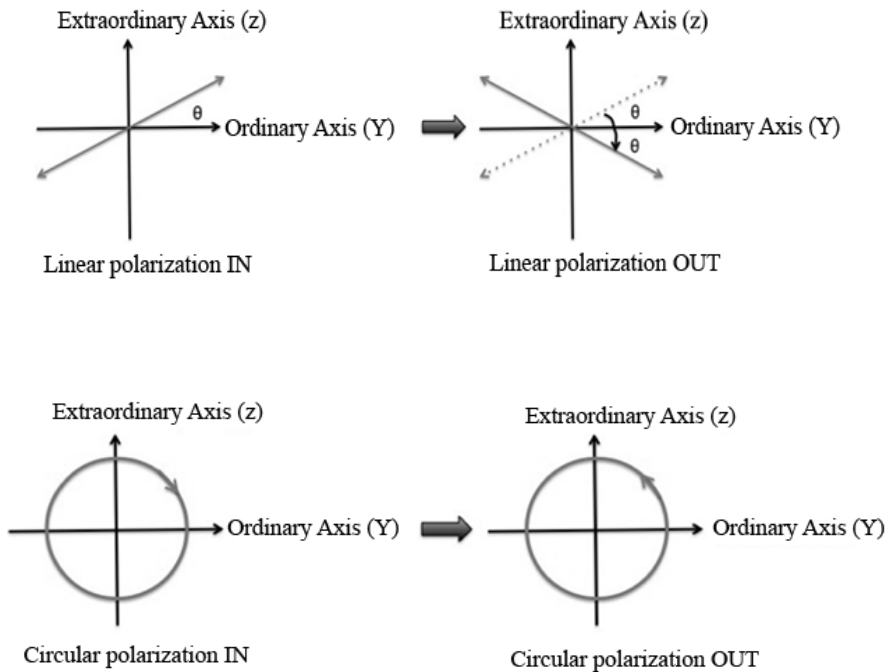


Figure 4.6
Effect of a half-wave-length plate on a light beam with linear (top) and circular (bottom) polarization

When circular polarized light is incident on the half-wavelength plate, its rotation direction is reversed, changing from right circular to left circular polarization and vice versa.



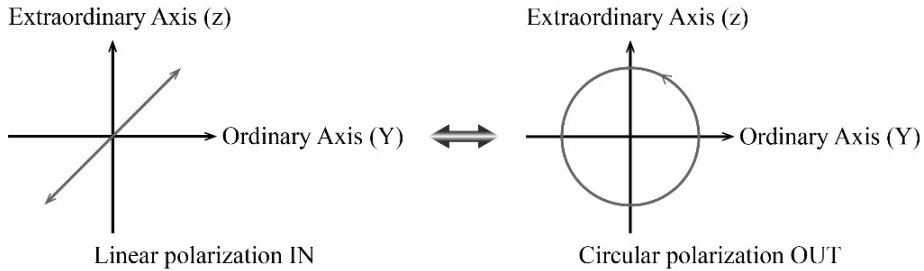
B. Quarter-wavelength plate

A quarter-wavelength plate introduces a relative phase difference of $\delta_R = \pi/2$, corresponding to a retarder thickness of

$$\frac{\pi}{2} = \frac{|n_e - n_{or}| 2\pi L_{\lambda/4}}{\lambda} \rightarrow L_{\lambda/4} = \frac{\lambda}{4|n_e - n_{or}|}$$

Quarter-wavelength plates are often used to convert linear polarization into circular polarization and vice versa (as will be studied in detail in the collection of problems).

Figure 4.7
Effect of a quarter-wavelength plate on a light beam with linear polarization



4.4.2 Conservation of linear momentum using uniaxial crystals

Another important application of uniaxial crystals is when we use these crystals to achieve momentum conservation during interactions among multiple frequency waves. This momentum conservation allows for efficient energy transfer between the waves through a nonlinear interaction process, and it is widely used in photonics technology.

The conservation of momentum, also called phase matching, is expressed as a relation between the wave vectors of the frequencies involved. Next, we present an example of momentum conservation relations. In this example, we consider collinear propagation of the waves (more complicated situations may arise, and the interested reader will find extended discussion on these topics in the recommended bibliography section).

Example

Effective interaction between a wave of frequency ω and a wave of frequency 2ω requires that the following condition be met:

$$2k(\omega) = k(2\omega)$$

This relationship translates into a condition between refractive indices:

$$2n(\omega)\omega/c = n(2\omega)2\omega/c \text{ leading to } n(\omega) = n(2\omega)$$

In general, these conditions cannot be met due to material dispersion. In the normal dispersion region, the index of refraction grows monotonically with frequency; so no possibility exists for momentum conservation in normal isotropic materials. However, these

conditions can be met in anisotropic media for particular propagation directions and polarization states.

To see how this is possible, consider example 1. If we plot the refractive indices at frequencies ω and 2ω using a positive uniaxial crystal, we can see that there is a particular direction of propagation where the extraordinary index at ω is equal to the ordinary index at 2ω .

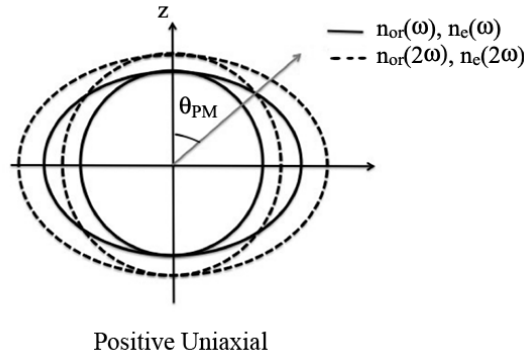


Figure 4.8
Phase-matching angle for SHG in a positive uniaxial crystal

In this particular direction, θ_{PM} , we obtain

$$n_e(\omega, \theta_{PM}) = n_{or}(2\omega).$$

From the expression of the refractive index, we can get the relation for the phase matching angle:

$$\frac{1}{n_e^2(\omega, \theta_{PM})} = \frac{1}{n_{or}^2(2\omega)} \rightarrow \frac{\sin^2(\theta_{PM})}{n_e^2(\omega)} + \frac{\cos^2(\theta_{PM})}{n_{or}^2(\omega)} = \frac{1}{n_{or}^2(2\omega)}.$$

After writing $\sin^2(\theta) = 1 - \cos^2(\theta)$, the phase-matching angle condition is obtained:

$$\cos^2(\theta_{PM}) = \frac{\frac{1}{n_{or}^2(2\omega)} - \frac{1}{n_e^2(\omega)}}{\frac{1}{n_{or}^2(\omega)} - \frac{1}{n_e^2(\omega)}}.$$

The required condition will be fulfilled by two beams with frequencies ω and 2ω and with the proper polarization propagating in this particular direction.

Additional examples will be proposed in the collection of problems.



4.5 Solved problems

S4.1 A quartz crystal is used to construct a retarder. The index of refraction as a function of wavelength is given by the expression

$$n_{or}^2(\lambda) = 1.28604141 + \frac{1.07044083 \cdot \lambda^2}{\lambda^2 - 1.00585997 \cdot 10^{-2}} + \frac{1.10202242 \cdot \lambda^2}{\lambda^2 - 100} \quad (\lambda \text{ in microns})$$

$$n_e^2(\lambda) = 1.28851804 + \frac{1.09509924 \cdot \lambda^2}{\lambda^2 - 1.02101864 \cdot 10^{-2}} + \frac{1.15662475 \cdot \lambda^2}{\lambda^2 - 100} \quad (\lambda \text{ in microns})$$

a) Represent in a single plot the ordinary index of refraction, the extraordinary index of refraction, and the extraordinary index for a wave propagating at an angle $\theta = 30^\circ$ inside the crystal in the wavelength range 400–1100 nm.

b) Calculate the crystal thickness needed to construct a half-wavelength plate for a wavelength of 600 nm.

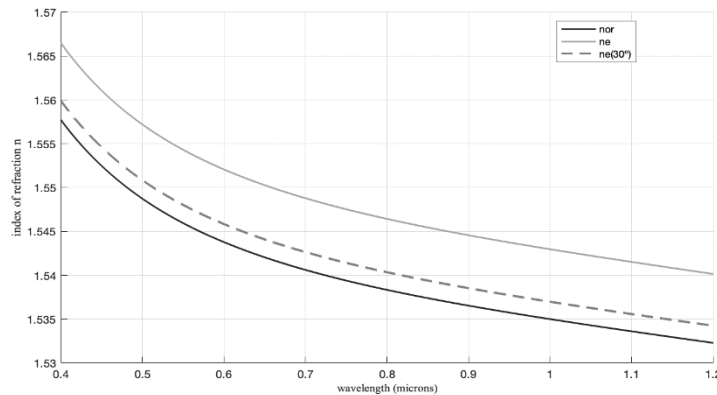
c) What is the phase difference introduced by this retarder in a plane wave of wavelength 500 nm incident normal to the surface? Obtain the polarization state of the transmitted beam if it is incident with linear polarization forming 40° with the ordinary axis of the retarder.

Solution

a) The ordinary and extraordinary indices of refraction of quartz are represented directly from the Sellmeier equation as a function of wavelength. The extraordinary index, n_e , corresponds to the refractive index value of the extraordinary mode when it propagates in a plane perpendicular to the optical axis ($\theta = 90^\circ$).

To represent the index of refraction of the extraordinary mode propagating at an angle $\theta = 30^\circ$, we use the expression

$$n_e(\theta, \lambda) = \frac{1}{\sqrt{(\sin \theta / n_e(\lambda))^2 + (\cos \theta / n_{or}(\lambda))^2}} \rightarrow n_e(30, \lambda) = \frac{1}{\sqrt{(\sin 30 / n_e(\lambda))^2 + (\cos 30 / n_{or}(\lambda))^2}}$$



The plot shows n_{or} , n_e and $n_e(30)$. Because $n_e > n_{or}$, we know that quartz is a positive uniaxial crystal.

b) From the Sellmeier equation, and at 600 nm, we obtain $n_e(600) = 1.5529$ and $n_{or}(600) = 1.5438$

The phase difference introduced by the retarder is

$$\delta_R(\lambda) = (k_e(\lambda) - k_{or}(\lambda))L = (n_e(\lambda) - n_{or}(\lambda))k_o L = \frac{(n_e(\lambda) - n_{or}(\lambda))2\pi L}{\lambda}$$

The crystal thickness is calculated directly from this expression. Using the fact that $\delta_R = \pi$ for a half-wavelength plate, we have

$$L = \frac{\lambda}{2(n_e(\lambda) - n_{or}(\lambda))} = \frac{600 \cdot 10^{-9}}{2(1.5529 - 1.5438)} = 3.2967 \cdot 10^{-5} \text{ m} = 32.97 \text{ } \mu\text{m}$$

c) Since the retarder phase is strongly dependent on wavelength, a half-wave plate designed at a specific wavelength will behave in a different way for another wavelength.

If we consider two different wavelengths, the relative phases introduced by the retarder at these two wavelengths are

$$\delta_R(\lambda_1) = \frac{2\pi L \Delta n_1}{\lambda_1} \text{ and } \delta_R(\lambda_2) = \frac{2\pi L \Delta n_2}{\lambda_2} \text{ where } \Delta n_i = n_e(\lambda_i) - n_{or}(\lambda_i)$$

and their ratios are

$$\frac{\delta_R(\lambda_1)}{\delta_R(\lambda_2)} = \frac{\Delta n_1 \cdot \lambda_2}{\Delta n_2 \cdot \lambda_1}$$

Using the fact that $\delta_R(600) = \pi$, the phase retardation at 500 nm is calculated from this expression and the Sellmeier equation:

$$\delta_R(500) = \delta_R(600) \frac{\Delta n_{500} 600}{\Delta n_{600} 500} = 1.21\pi$$

■

S4.2 An optical wave propagating along the OX axis is incident on a half-wavelength plate whose ordinary axis is oriented in the OY direction while the extraordinary axis moves in the OZ direction. Find the polarization of the outgoing wave when the incident wave polarization is:

- Linearly polarized along the ordinary axis
- Linearly polarized along the extraordinary axis
- Linearly polarized forming an angle θ with the ordinary axis
- Right circular polarization
- Left circular polarization

**Solution**

The expression for a plane wave propagating in the direction normal to the waveplate (the OX axis) and with components E_{or} and E_e along the ordinary and extraordinary directions, respectively, is

$$\vec{E}_{in}^{(+)}(x) = (E_{or}^{(+)} \hat{j} + E_e^{(+)} \hat{k}) e^{i(k_o x - \omega t)} = (E_{or} \hat{j} + e^{i\delta_{in}} E_e \hat{k}) e^{i(k_o x - \omega t + \phi_x)}$$

Where $\delta_{in} = \phi_y - \phi_x$ is the initial phase difference. We define our incident Jones vector in terms of these directions, considering our horizontal axis to be along the ordinary direction of the waveplate and the vertical axis along the extraordinary direction.

The action of the waveplate is given by the general retarder matrix with $\delta_R = (k_e - k_{or})L = \pi$:

$$R_{\lambda/2} = e^{ik_{or}L} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi} \end{pmatrix} = e^{ik_{or}L} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The output polarization state is given by the action of this matrix on the incident Jones vector in each of the requested situations.

a) In this case, $\hat{e}_{in} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and the incident field amplitude is E_{or} .

The input wave at $x = 0$ is $E_{in}^{(+)}(0) = E_{or} \hat{j} e^{i(\phi_x - \omega t)}$.

The output Jones vector: $\vec{e}_{out} = R_{\lambda/2} \cdot \hat{e}_{in} = e^{ik_{or}L} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e^{ik_{or}L} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

As expected, when the polarization of the incident wave lies along the ordinary axis, there is no relative phase retardation between the ordinary and extraordinary components. Furthermore, the output wave maintains its linear polarization state with an added phase, $k_{or}L$, due to the medium of thickness L .

The output beam is $E_{out}^{(+)}(L) = E_{or} \hat{j} e^{i(\phi_x + k_{or}L - \omega t)}$.

b) In this case, the Jones vector of the incident wave is $\hat{e}_{in} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, with field amplitude E_e .

The output Jones vector: $\vec{e}_{out} = e^{ik_{or}L} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = e^{ik_{or}L} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = e^{i(k_{or}L + \pi)} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Using the relation $\delta_R = (k_e - k_{or})L = \pi \rightarrow k_{or}L + \pi = k_e L$,

we obtain

$$\vec{e}_{out} = e^{ik_e L} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ and } E_{out}^{(+)}(L) = E_e \hat{k} e^{i(\phi_x + k_e L - \omega t)}.$$

Similarly to case a), the polarization state is not changed and an additional phase, $k_e L$, is added to the field.

c) In this case, the electric field incident is not parallel to the ordinary or extraordinary axis, so we have a mixture of ordinary and extraordinary components inside the retarder. If the incident electric field has magnitude E_o , the ordinary and extraordinary components correspond to the projections along the corresponding axes, giving

$$E_{or} = E_o \cos \theta \text{ and } E_e = E_o \sin \theta \text{ with a Jones vector } \hat{e}_{in} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

$$\text{and } \vec{E}_{in}^{(+)}(0) = E_o (\cos \theta \hat{j} + \sin \theta \hat{k}) e^{i(\phi_x - \omega t)}$$

$$\text{Action of the retarder: } \vec{e}_{out} = e^{ik_{or} L} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} = e^{ik_{or} L} \begin{pmatrix} \cos \theta \\ -\sin \theta \end{pmatrix} = e^{ik_{or} L} \begin{pmatrix} \cos(-\theta) \\ \sin(-\theta) \end{pmatrix}$$

Written in this way, we see that the output wave is linearly polarized, forming an angle of $-\theta$ with respect to the ordinary axis. The action of the retarder is to produce a total rotation of the polarization state by an angle 2θ . This rotation is produced without a loss of intensity in the incident wave.

$$\text{The output field is } E_{out}^{(+)}(L) = E_o (\cos \theta \hat{j} - \sin \theta \hat{k}) e^{i(\phi_x + k_{or} L - \omega t)}$$

$$d) \text{ In this case, } \hat{e}_{in} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \text{ and } E_{or} = E_e = E.$$

$$\text{The action of the retarder is } \vec{e}_{out} = e^{ik_{or} L} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = e^{ik_{or} L} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

and the output polarization is circular, changing from right circular to left circular.

e) The last case is just the opposite of case d), and the retarder changes the incident left circular polarization to right circular polarization.

$$\vec{e}_{out} = e^{ik_{or} L} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = e^{ik_{or} L} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

■

S4.3 A circular polarization selector is constructed using a combination of a quarter-wave retarder followed by a linear polarizer oriented at 45° with respect to the retarder axes. Find the polarization matrix for this configuration, and study the transmitted polarization for incident left and right circular polarization.

**Solution**

A circular polarization selector is a device transmitting light of a given circular polarization rotation direction and blocking light with circular polarization in the opposite rotation direction. Such a device is constructed by combining two polarizers in series: a quarter-wavelength plate and a linear polarizer oriented at 45° .

As explained in 4.4.1 the quarter-wavelength retarder acts on an incident circular polarized wave by transforming it into a linear polarized wave oriented at $+45^\circ$ or -45° , depending on the rotation direction. Thus, the linear polarizer placed at 45° should block the radiation for one of the two situations. The orientation of the linear polarizer will determine whether we have a right circular or left circular polarization selector.

To check that this configuration works, we introduce the corresponding matrices for these elements.

$$\text{Quarter-wavelength plate: } R_\delta = e^{i\phi_{or}} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix} \xrightarrow{\lambda/4} R_{\pi/2} = e^{i\phi_{or}} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

$$\text{Linear polarizer: } P_L(\theta) = \begin{pmatrix} \cos^2 \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \sin^2 \theta \end{pmatrix} \rightarrow P_L(45^\circ) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

The combined action is obtained by direct matrix multiplication:

$$\vec{e}_{out} = P_L(45^\circ) \cdot R_{\pi/2} \cdot \hat{e}_in = S_R \cdot \hat{e}_in,$$

$$\text{where } S_R = \frac{e^{i\phi_{or}}}{2} \begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix}.$$

We can check immediately that, when acting on incident circular polarized states, we obtain

$$S_R \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \frac{e^{i\phi_{or}}}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$S_R \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and S_R corresponds to a right circular polarization selector.

To obtain left circular polarization, we have to change the orientation of the linear polarizer:

$$S_L = P_L(-45^\circ) R_{\pi/2} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} e^{i\phi_{or}} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} = \frac{e^{i\phi_{or}}}{2} \begin{pmatrix} 1 & -i \\ -1 & i \end{pmatrix}$$

giving, in this case, no transmission if the incident radiation is right circular and linearly polarized light at -45° if it has left circular polarization:

$$S_L \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$S_L \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} = \frac{e^{i\phi_{or}}}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

■

S4.4 Consider a plane wave with arbitrary polarization and modulus $|E_o| = 1\text{V/m}$, defined by its Jones vector:

$$\hat{e}_{in} = \begin{pmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{pmatrix}$$

Stokes's parameters are an alternative way to express the polarization state of a light wave. These parameters can be calculated experimentally by measuring the transmitted intensity of the light wave through different elements. Calculate the transmitted intensity of the incident wave through the following elements.

- Linear polarizer at 0° : $I(0)$
- Linear polarizer at 90° : $I(90)$
- Linear polarizer at 45° : $I(45)$
- Linear polarizer at -45° : $I(-45)$
- Left circular polarization selector: $I(L)$
- Right circular polarization selector: $I(R)$

Obtain a general expression for the Stokes's parameters defined in terms of the incident intensity, I_o , as:

$$S_1 = \frac{I(0) - I(90)}{I_o}$$

$$S_2 = \frac{I(45) - I(-45)}{I_o}$$

$$S_3 = \frac{I(L) - I(R)}{I_o}$$

Solution

The incident intensity of the plane wave is $I_o = \epsilon_o c |E|^2 / 2 = \epsilon_o c / 2$.

Transmission through the different elements.



a) Linear polarizer at 0° :

$$\vec{e}_{out} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{pmatrix} = \cos \theta \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\text{The transmitted intensity: } I(0) = \frac{1}{2} \varepsilon_o c (\vec{e}_{out} \cdot \vec{e}_{out}^*) = \frac{1}{2} \varepsilon_o c \cos^2 \theta$$

b) Linear polarizer at 90° :

$$\vec{e}_{out} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{pmatrix} = e^{i\delta} \sin \theta \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\text{and } I(90) = \frac{1}{2} \varepsilon_o c (\vec{e}_{out} \cdot \vec{e}_{out}^*) = \frac{1}{2} \varepsilon_o c \sin^2 \theta$$

c) Linear polarizer at 45° :

$$\vec{e}_{out} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \cos \theta + e^{i\delta} \sin \theta \\ \cos \theta + e^{i\delta} \sin \theta \end{pmatrix} = \frac{\cos \theta + e^{i\delta} \sin \theta}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right]$$

The transmitted intensity in this case is

$$I(45) = \frac{1}{2} \varepsilon_o c \left(\frac{\cos \theta + e^{i\delta} \sin \theta}{\sqrt{2}} \right) \left(\frac{\cos \theta + e^{-i\delta} \sin \theta}{\sqrt{2}} \right) = \frac{1}{2} \varepsilon_o c \left(\frac{1 + 2 \cos \theta \sin \theta \cos \delta}{2} \right)$$

d) Linear polarizer at -45° :

$$\vec{e}_{out} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{pmatrix} = \frac{\cos \theta - e^{i\delta} \sin \theta}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right]$$

Repeating the steps in c), we obtain

$$I(-45) = \frac{1}{2} \varepsilon_o c \left(\frac{1 - 2 \cos \theta \sin \theta \cos \delta}{2} \right)$$

e) To consider the effect of the left circular selector, we will use the results obtained in the previous problem, S3, to get

$$\vec{e}_{out} = S_L \begin{pmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{pmatrix} = \frac{e^{i\phi_{or}} (\cos \theta - i e^{i\delta} \sin \theta)}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right]$$

The transmitted intensity is obtained after few algebraic steps:

$$I(L) = \frac{1}{2} \varepsilon_o c \left(\frac{\cos \theta - i e^{i\delta} \sin \theta}{\sqrt{2}} \right) \left(\frac{\cos \theta + i e^{-i\delta} \sin \theta}{\sqrt{2}} \right) = \frac{1}{2} \varepsilon_o c \left(\frac{1 + 2 \cos \theta \sin \theta \sin \delta}{2} \right)$$

f) The effect of the right circular selector is obtained in the same way:

$$\vec{e}_{out} = S_R \cdot \hat{e}_{in} = \frac{e^{i\phi_{or}}}{2} \begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix} \begin{pmatrix} \cos \theta \\ e^{i\delta} \sin \theta \end{pmatrix} = \frac{e^{i\phi_{or}} (\cos \theta + i e^{i\delta} \sin \theta)}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right]$$

and

$$I(L) = \frac{1}{2} \varepsilon_o c \left(\frac{1 - 2 \cos \theta \sin \theta \sin \delta}{2} \right)$$

Directly applying the definitions in terms of the Jones vector components, the expressions for Stoke's parameters are

$$S_1 = (I(0) - I(90)) / I_o = \cos^2 \theta - \sin^2 \theta = \cos(2\theta)$$

$$S_2 = (I(45) - I(-45)) / I_o = 2 \cos \theta \sin \theta \cos \delta = \sin(2\theta) \cos \delta$$

$$S_3 = (I(L) - I(R)) / I_o = 2 \cos \theta \sin \theta \sin \delta = \sin(2\theta) \sin \delta$$

From this result, one can confirm that the relation $S_1^2 + S_2^2 + S_3^2 = 1$ holds. Thus, the points indicating the polarization state of a wave can be placed over a unitary sphere (Poincaré sphere).

■

S4.5 The polarization state of a beam is defined in terms of what we call the laboratory frame of reference, defined by two orthogonal axes fixed to the optical table and named horizontal (H) and vertical (V) axes. The angle θ of the associated Jones vector is defined with respect to the H axis. A retarder can be inserted into the beam path, with its ordinary axis forming an arbitrary angle with respect to the horizontal axis. Describe the effect of the retarder and find the associated matrix in this more general case.

Solution

The effect of the retarder has been previously defined in terms of the ordinary and extraordinary components incident on the retarder and, hence, the matrix R_δ (introduced in this chapter) is expressed in what we call the retarder frame of reference, defined by two orthogonal axes in the directions of the ordinary and extraordinary axes of the retarder.

Any incident polarization Jones vector $\hat{e}_{in} = \begin{pmatrix} e_{or} \\ e_e \end{pmatrix}$ will be transformed by the retarder according to the rule $\vec{e}_{out} = R_{\delta_R} \hat{e}_{in}$.



Up to now, we have considered the simplified case where the laboratory and retarder frames of reference were coincident. In this problem, we explore a more general situation that is quite often found in real applications, where the retarder is rotated by an angle ψ with respect to the laboratory frame of reference.

Our initial polarization vector in the laboratory frame of reference can be written as

$$\hat{e}_{in}^{lab} = \begin{pmatrix} e_H \\ e_V \end{pmatrix}$$

The vector in the retarder frame of reference is related to this one through a rotation of the coordinate system by an angle ψ and, hence, we can write

$$\hat{e}_{in}^{ret} = W(\psi) \hat{e}_{in}^{lab} \rightarrow \begin{pmatrix} e_{or} \\ e_e \end{pmatrix} = \begin{pmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{pmatrix} \begin{pmatrix} e_H \\ e_V \end{pmatrix}$$

where $W(\psi)$ is the rotation matrix.

The effect of the retarder can be applied in the retarder frame of reference, giving

$$\bar{\hat{e}}_{out}^{ret} = R_{\delta_R} \hat{e}_{in}^{ret} = R_{\delta_R} W(\psi) \hat{e}_{in}^{lab}$$

Finally, we convert our polarization state back to the laboratory frame of reference by applying a rotation of $-\psi$:

$$\bar{\hat{e}}_{out}^{lab} = W(-\psi) \bar{\hat{e}}_{out}^{ret} = W(-\psi) R_{\delta_R} W(\psi) \hat{e}_{in}^{lab} = R_{\delta_R}^\psi \hat{e}_{in}^{lab}$$

We can see from this expression that the output Jones vector in the laboratory frame of reference can be obtained from the incident Jones vector in this frame of reference through a unitary transformation of the retarder matrix:

$$R_{\delta_R}^\psi = W(-\psi) R_{\delta_R} W(\psi) = \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix} e^{ik_{or}L} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta_R} \end{pmatrix} \begin{pmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{pmatrix}$$

$$R_{\delta_R}^\psi = e^{ik_{or}L} \begin{pmatrix} \cos^2 \psi + e^{i\delta_R} \sin^2 \psi & \cos \psi \sin \psi (1 - e^{i\delta_R}) \\ \cos \psi \sin \psi (1 - e^{i\delta_R}) & \sin^2 \psi + e^{i\delta_R} \cos^2 \psi \end{pmatrix}$$

■

S4.6 Consider a plane wave that is linearly polarized along the vertical axis in the laboratory frame of reference. We use a half-wave plate to rotate the polarization state of the wave. The retarder is rotated with its ordinary axis forming an angle $\psi = 30^\circ$ with respect to the horizontal axis. Obtain the polarization Jones vector after the retarder.

Solution

The incident Jones vector in the laboratory frame of reference is $\hat{e}_{in}^{lab} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

If the retarder is not rotated, this state should correspond to incidence along the extraordinary axis. Consequently, no polarization change will be observed at the output, as discussed in S2 b).

Since the retarder has been rotated, its effect on the incident beam polarization state is obtained by direct application of the result in problem S5:

$$\vec{e}_{out}^{lab} = R_{\pi}^{30} \vec{e}_{in}^{lab} = e^{ik_{or}L} \begin{pmatrix} \cos^2 30 - \sin^2 30 & 2 \cos 30 \sin 30 \\ 2 \cos 30 \sin 30 & \sin^2 30 - \cos^2 30 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = e^{ik_{or}L} \begin{pmatrix} 2 \cos 30 \\ \sin^2 30 - \cos^2 30 \end{pmatrix}$$

Rewriting this expression:

$$\vec{e}_{out}^{lab} = e^{ik_{or}L} \begin{pmatrix} \sin 60 \\ -\cos 60 \end{pmatrix} = e^{ik_{or}L} \begin{pmatrix} \cos 30 \\ -\sin 30 \end{pmatrix} = e^{ik_{or}L} \begin{pmatrix} \cos(-30) \\ \sin(-30) \end{pmatrix}$$

Apart from an additional phase term, which will not influence the intensity, we can identify in this result that the transmitted wave will be linearly polarized, forming an angle of -30° with respect to the horizontal axis. Since the angle with respect to this axis was 90° at the entrance, the retarder has rotated the incident polarization by 120° .

This corresponds to twice the angle forming the incident polarization state with respect to the ordinary axis of the retarder (60°), and it is in accordance with what is described in Figure 6 (top). ■

S4.7 A plane wave propagates along the z -axis with its polarization state characterized by a Jones vector, defined by the parameters $\theta = 30^\circ$ and $\delta = \pi/6$ rad. To transform it into a linearly polarized wave along the horizontal direction (x -axis), we use two retarders placed in series. In the first retarder, we can select any value of the retarder's phase δ_R by controlling a voltage applied to the device. Its ordinary axis is aligned along the horizontal direction and the voltage applied to this retarder is adjusted to obtain a linearly polarized wave at the output. The second retarder is a half-wave plate, which is used to rotate the linearly polarized wave after the first retarder to obtain the desired horizontal linear polarization state at the output.

a) In order to attain this goal, what phase should be introduced into the first retarder and what orientation should the ordinary axis of the half-wavelength plate have with respect to the horizontal axis?

b) To obtain a circularly polarized light at the output, we replace the half-wave plate with a different retarder (keeping the first retarder, as chosen in a). In terms of the working principle of retarders, justify which retarder type is needed in this case, as well as the orientation with respect to the horizontal direction.

Solution

a) The incident Jones vector $\vec{e}_{in}^{lab} = \begin{pmatrix} \cos 30 \\ e^{i\pi/6} \sin 30 \end{pmatrix}$ corresponds to elliptical polarization.



Action of the first retarder: this retarder is oriented with its ordinary axis along the horizontal direction, and its matrix is given by

$$R_1 = e^{i\phi_{or}} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta_R} \end{pmatrix}$$

(we will take $\phi_{or} = 0$ without losing generality).

The transmitted beam after the first retarder is

$$\hat{e}_{out1} = R_1 \cdot \hat{e}_{in}^{lab} = \begin{pmatrix} \cos 30 \\ e^{i\delta_{out}} \sin 30 \end{pmatrix} \text{ with } \delta_{out} = \frac{\pi}{6} + \delta_R$$

To obtain linear polarization we have two possible solutions:

$$\delta_{out} = 0, 2\pi, \dots \rightarrow \delta_R = \delta_{out} - \pi/6 = 11\pi/6$$

$$\delta_{out} = \pi, 3\pi, \dots \rightarrow \delta_R = \delta_{out} - \pi/6 = 5\pi/6$$

For $\delta_R = 11\pi/6$, the transmitted beam will be linearly polarized, with $\theta = 30^\circ$.

For $\delta_R = 5\pi/6$, the transmitted beam will be linearly polarized, with $\theta = -30^\circ$.

The orientation of the second retarder can be obtained without calculations by considering the effect of a half-wave plate on linearly polarized light: the transmitted beam will have linear polarization rotated by an angle 2α with respect to the incident polarization, where α is the angle of the incident wave with respect to the retarder axis.

Following this rule to obtain linear polarization in the horizontal direction, we should place the retarder axis at

$$\psi = 15^\circ \text{ for } \delta_R = 11\pi/6$$

$$\psi = -15^\circ \text{ for } \delta_R = 5\pi/6$$

This result can be obtained from the general result derived in S5.

The effect of a rotated retarder can be expressed in terms of the matrix R_π^ψ . For the half-

$$\text{wave plate, this simplifies to } R_\pi^\psi = \begin{pmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{pmatrix}$$

For $\delta_R = 11\pi/6$:

$$\bar{e}_{out}^{lab} = R_\pi^\psi \cdot \hat{e}_{in}^{lab} \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{pmatrix} \begin{pmatrix} \cos 30 \\ \sin 30 \end{pmatrix}$$

$$\cos 2\psi \cos 30 + \sin 2\psi \sin 30 = 1 \rightarrow 2\psi = 30 \rightarrow \psi = 15^\circ$$

$$\sin 2\psi \cos 30 - \cos 2\psi \sin 30 = 0 \rightarrow \tan 2\psi = \tan 30 \rightarrow \psi = 15^\circ$$

For $\delta_R = 5\pi/6$:

$$\vec{e}_{out}^{lab} = R_\pi^w \cdot \hat{e}_{in}^{lab} \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{pmatrix} \begin{pmatrix} \cos(-30) \\ \sin(-30) \end{pmatrix}$$

$$\cos 2\psi \cos 30 - \sin 2\psi \sin 30 = 1 \rightarrow 2\psi = -30 \rightarrow \psi = -15^\circ$$

$$\sin 2\psi \cos 30 + \cos 2\psi \sin 30 = 0 \rightarrow \tan 2\psi = -\tan 30 \rightarrow \psi = -15^\circ$$

b) For circular polarization $\delta_{tot} = \mp\pi/2$ and $\theta = 45^\circ$, this is achieved if the retarder is a quarter-wave plate and the incident linearly polarized light is oriented at 45° with respect to the ordinary axis.

For $\delta_R = 11\pi/6$, the incident beam is polarized at 30° , so the retarder should be placed with the ordinary axis at an angle $\psi = -15^\circ$ with respect to the horizontal.

For $\delta_R = 5\pi/6$, the incident beam is polarized at -30° and the retarder should be placed with the ordinary axis at an angle $\psi = -75^\circ$ with respect to the horizontal.

This result could be checked directly using the matrix approach.

■

S4.8 A negative uniaxial crystal, KDP (potassium dideuterium phosphate), is used to generate a second harmonic (SH) signal at the wavelength $\lambda_{SH} = 532$ nm from a fundamental wave at $\lambda_f = 1064$ nm. We want to calculate the phase-matching angle in the following two different configurations.

- Type I phase-matching: the momentum conservation condition to be satisfied is $2k(1064) = k(532)$. Find the only possibility for the polarization of both beams, and find the phase-matching angle.
- Type II collinear phase-matching: in this configuration at the fundamental wavelength with different polarization, two beams mix collinearly to generate radiation at the SH wavelength. The efficiency of the process is governed by the momentum conservation relation $k_{or}(1064) + k_e(\theta, 1064) = k(532)$. Find the polarization of the SH beam and calculate the phase-matching angle.

KDP refractive indices: $n_{or}(532) = 1.5123$, $n_e(532) = 1.4705$, $n_{or}(1064) = 1.4938$ and $n_e(1064) = 1.4599$.

Solution

In the process of second harmonic generation (SHG), we want to transfer energy from the fundamental beam to the second harmonic. An effective energy transfer between the waves requires simultaneous energy and momentum conservation.



When light is studied in terms of photons, its energy is $E = \hbar\omega$ and its momentum is $\vec{p} = \hbar\vec{k}$. The energy conservation requires that two photons of the fundamental frequency should be destroyed in order to create a photon at the second harmonic frequency. This can be written as

$$\omega_{SH} = 2\omega_{fund} \rightarrow \lambda_{SH} = \lambda_{fund} / 2$$

The momentum conservation is a relation between the wave vectors of the interacting waves. In a collinear interaction, all waves propagate in the same direction and we can write a scalar relation. This momentum conservation condition can be satisfied in different ways, the simplest being the Type I configuration.

Type I configuration

For an SHG process, the configuration is given by the relation

$$2k_{fund} = k_{SH}$$

This relation can be expressed in terms of the index of refraction at each wavelength:

$$2n_{fund} \frac{2\pi}{\lambda_{fund}} = n_{SH} \frac{2\pi}{\lambda_{SH}} = n_{SH} \frac{4\pi}{\lambda_{fund}} \rightarrow n_{fund} = n_{SH}$$

where the energy conservation relation has been used.

From this expression, we can see that momentum conservation in this process requires that the index of refraction at the two wavelengths must be equal.

This condition is forbidden in isotropic materials, due to material dispersion. However, it can be obtained in anisotropic materials under particular conditions of the polarization state and of the propagation direction for the waves interacting inside the crystal.

The general idea is that, to achieve this condition, the polarization of the fundamental and SH must be different. Furthermore, the variation of the extraordinary index of refraction with angle allows, in some situations, finding a point satisfying the required momentum conservation condition (phase-matching).

Because the index of refraction increases with frequency in a normal dispersion regime, it will always happen that

$$n_{or}(fund) < n_{or}(SH) \text{ and } n_e(fund, \theta) < n_e(SH, \theta).$$

The figure shows this plot for a generic negative uniaxial crystal in two different situations.

- If the medium birefringence is too small (left plot), the curves of the fundamental and SH do not intersect at any angle. Under these conditions, the crystal is not useful for achieving phase-matching.

- If the medium birefringence is large enough (right plot), the curve of the ordinary SH and of the extraordinary fundamental intersect at a particular propagation direction. At this point, the momentum conservation condition is satisfied.

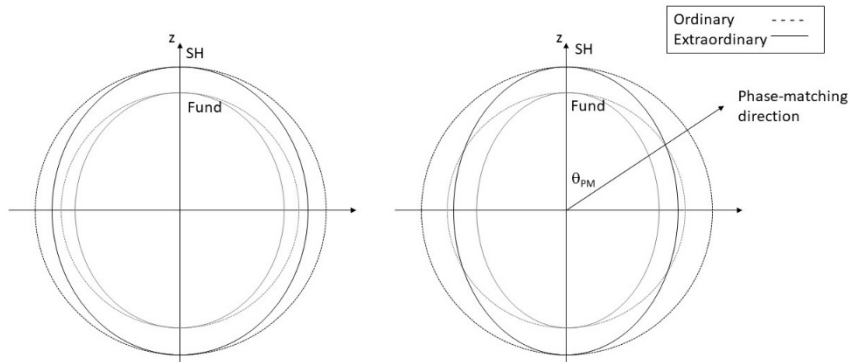


Figure 4.9
Type I phase-matching

From the figure, we can see that phase-matching for negative uniaxial crystals will be possible if $n_{or}(\text{fund}) > n_e(\text{SH})$.

From the values of the indices of refraction of KDP at 1064 nm and at 532 nm: $n_{or}(1064) = 1.4938 > n_e(532) = 1.4705$. In type I configuration, this condition holds, and phase matching is possible in KDP for the propagation angle satisfying the condition:

$$n_{or}(1064) = n_e(532, \theta_{PM})$$

Substituting the expression for the extraordinary refractive index as a function of the angle:

$$\frac{1}{n_{or}^2(1064)} = \frac{1}{n_e^2(532, \theta_{PM})} = \frac{\sin^2 \theta_{PM}}{n_e^2(532)} + \frac{\cos^2 \theta_{PM}}{n_{or}^2(532)} = \frac{1 - \cos^2 \theta_{PM}}{n_e^2(532)} + \frac{\cos^2 \theta_{PM}}{n_{or}^2(532)}$$

$$\cos^2 \theta_{PM} = \frac{1/n_{or}^2(1064) - 1/n_e^2(532)}{1/n_{or}^2(532) - 1/n_e^2(532)} = 0.5677 \rightarrow \theta_{PM}^{Type I} = 41.1^\circ$$

Type II Configuration

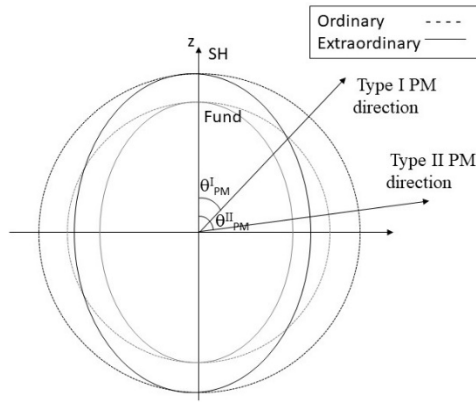
For type II collinear interaction, we have a mixture of the ordinary and extraordinary fundamental modes propagating collinearly. From the momentum conservation condition, we can find the relation between the refractive indices:

$$k_{or}(1064) + k_e(1064, \theta) = k(532)$$

$$\frac{n_{or}(1064)2\pi}{1064} + \frac{n_e(1064, \theta)2\pi}{1064} = \frac{n(532)2\pi}{532} \rightarrow n(532) = \frac{n_{or}(1064) + n_e(1064, \theta)}{2}$$



The index of refraction of the SH must be the averaged value of its ordinary and extraordinary value. For this condition to happen, the birefringence of the crystal should be large enough. This situation is plotted in the figure.



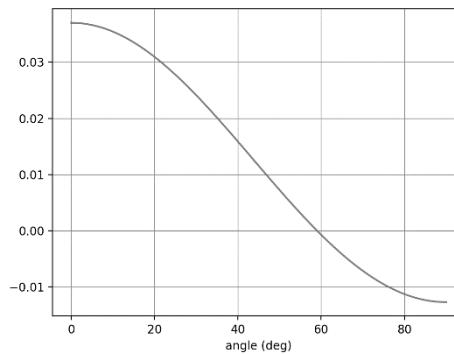
From the figure, we see that the SH index of refraction must be extraordinary. At the phase-matching angle, its value lies between the ordinary and extraordinary indices of the fundamental, thus satisfying the averaged-index condition that is just derived.

$$\text{The type II PM condition is } n_e(532, \theta_{PM}^{II}) = \frac{n_{or}(1064) + n_e(1064, \theta_{PM}^{II})}{2}$$

From the figure, we can see that the Type II interaction angle will always be larger than the Type I phase-matching angle.

The angle of phase-matching can also be calculated numerically. The next plot represents the function $f(\theta) = 2n_e(532, \theta) - n_{or}(1064) - n_e(1064, \theta)$ as a function of the angle θ . The value of the phase-matching angle can be obtained from the plot, since

$$f(\theta_{PM}) = 0, \text{ at a phase matching close to } 60^\circ.$$



4.6 Proposed problems

P4.1 An optical wave propagates inside quartz in the direction defined by the angles $\theta = 20^\circ$ and $\phi = 30^\circ$. The vacuum wavelength is 800 nm, and its field amplitude is $E = 100$ V/m. Write the expression for the ordinary and extraordinary normal modes propagating in this medium. Find the direction of the Poynting vector for each of these modes.

P4.2 An optical wave of wavelength 589 nm is incident normal to the surface of a calcite retarder ($n_{or} = 1.6583$; $n_e = 1.4864$).

- a) What is the minimum crystal thickness needed to obtain a relative phase delay between the ordinary and extraordinary components of π ; $\pi/4$ and $\pi/2$?
- b) For an input wave linearly polarized at $\theta = 45^\circ$ with respect to the ordinary axis, calculate the polarization state at the output of the retarder for the three cases considered in a).

P4.3 A quarter-wavelength plate is oriented in such a way that its ordinary axis lies on the OY axis and its extraordinary axis along the OZ axis. An incident wave propagates along the OX axis. Find the polarization of the outgoing wave when the incident polarization state is:

- a) Linearly polarized along the ordinary axis
- b) Linearly polarized along the extraordinary axis
- c) Linearly polarized forming an angle θ with the ordinary axis
- d) Right circular polarization
- e) Left circular polarization

P4.4 A plane wave with linear polarization is incident on a circular polarizer selector (see solved problem S3). Study the effect of this polarizer as a function of the angle of the incident linearly polarized wave.

P4.5 We construct a circular polarization selector as explained in solved problem S3. At a given instant of time, the linear polarizer starts rotating with constant angular velocity ω . Determine the transmitted polarization state for incident right circular polarization.

P4.6 A wave plate made from a positive uniaxial crystal has its optical axis parallel to the plate surface and perpendicular to the incidence plane. It introduces a relative phase change $\delta_R = (n_e - n_{or})k_o L$ for a beam incident normal to the surface. We change the incidence angle from zero to a value of θ_{in} .

- a) Describe the trajectories of the ordinary and extraordinary beam in the retarder.
- b) Find an expression for the relative phase change in this case.

P4.7 Spectral filters with a very narrow spectral bandwidth can be obtained with the use of birefringent crystals. A Šolc filter is made using an integer number of retarder pairs inserted between crossed polarizers in such a way that the sequence of orientations of the



ordinary axis of each retarder pair with respect to the initial horizontal polarizer direction is given by $[\theta, -\theta]$. The total sequence of elements is $[0^\circ, \theta, -\theta, \dots, \theta, -\theta, 90^\circ]$.

- Find the matrix corresponding to a given pair of retarders.
- Find the matrix for a Šolc filter with a single retarder pair.
- If the retarders are half-wave plates, determine the transmitted intensity for incident radiation that is linearly polarized along the horizontal axis.

P4.8 A quartz crystal is used to generate radiation at 400 nm (second harmonic) from radiation at 800 nm (fundamental) through the process of second harmonic generation. The refractive indices of the crystal at these wavelengths are $n_{or}(400) = 1.5577$, $n_e(400) = 1.5673$, $n_{or}(800) = 1.5383$ and $n_e(800) = 1.5472$. Describe and justify the possibility of obtaining the required phase-matching using this crystal.

P4.9 A positive uniaxial crystal is used to obtain phase-matching in the process of converting radiation from a beam of wavelength 1064 nm to its second harmonic at 532 nm using a type-II collinear phase-matching configuration. In this configuration, the incident radiation is a mixture of extraordinary and ordinary polarized waves at 1064 nm. The phase matching condition is

$$k_{or}(1064) + k_e(1064) = k(532)$$

- Discuss which polarization state of the second harmonic radiation is the only one possible that satisfies the phase-matching condition.
- Determine the phase-matching angle in this configuration. The indices of refraction are $n_{or}(1064) = 1.535$, $n_e(1064) = 1.544$, $n_{or}(532) = 1.538$ and $n_e(532) = 1.549$.

P4.10 A DKDP crystal (negative uniaxial) generates UV radiation at $\lambda_3 = 354.67$ nm through the process of sum-frequency generation using the superposition of radiation at $\lambda_1 = 1064$ nm and $\lambda_2 = 532$ nm. To obtain an efficient conversion, the following conditions must be fulfilled.

Energy conservation: $\omega_1 + \omega_2 = \omega_3$

Momentum conservation: $k_{or}(1064) + k_{or}(532) = k_e(354.67, \theta_{PM})$

Find the phase-matching angle, assuming that the interaction is collinear, so that all beams propagate in the same direction.

$\lambda(\text{nm})$	n_{or}	n_e
1064	1.4928	1.4555
532	1.5085	1.4690
354.67	1.5263	1.4841



→ 5



Light pulse propagation

Introduction

In the previous chapters, we dealt exclusively with monochromatic radiation. Chapter 1 covered the basic solutions to the wave equation; Chapter 2 developed paraxial beam solutions for transverse spatial localization of light; and Chapter 3 faced the problem of some temporal dependence appearing due to the damped emission of single dipole emitters. When we consider the superposition of damped oscillators from many sources (typically around 10^{18} atoms/cm³), the random phases result in a continuous emission of light with added fluctuations due to the statistical character of the emission process. This is the typical emission of light from natural sources such as the Sun. Thermal sources emit over a broad frequency spectrum (for instance, Sun emits from IR to UV), although the light is not emitted in the form of a light pulse but instead in a continuous way.

This emission is completely opposite to that of modern laser sources emitting radiation pulses with durations in the picosecond, nanosecond, or femtosecond range at a given repetition frequency rate. Such pulses have very great power and are extensively used in a large number of photonics applications, such as telecommunications, material processing, and medical surgery, among others.

In this chapter, we will learn the basic description of light pulses and how they propagate in dispersive materials. In the same way that light beams suffer diffraction during propagation, we will see that the effect of dispersion is to introduce a temporal broadening of the pulses.

In the first place, let us consider how we can get a pulse from the superposition of monochromatic waves of different frequencies. A monochromatic wave represents a temporal infinite solution of the wave equation. If we add two monochromatic waves with slightly different frequencies, ω and $\omega + \Delta\omega$, we obtain a modulated signal showing the characteristic modulation at a frequency equal to $\Delta\omega/2$. By adding more and



more frequencies with a well-defined phase relation between them, we obtain a modulating envelope that becomes narrower and narrower as the frequency bandwidth increases. The field confinement disappears, however, when the phase relation between the waves is randomized. This phase dependence between the different frequencies is fundamental for discerning a light pulse (for instance, emitted from ultrashort lasers) from a broadband continuous wave emitting source like the Sun. The mechanism responsible for locking the phases represents a major advance in photonics over recent decades.

5.1 Mathematical description of light pulses

The electric field corresponding to a light pulse is expressed mathematically as a temporal-dependent function. As we have learned, light propagates while oscillating at very high frequencies of the order of 10^{15} Hz, which allows expressing our pulse in what is called a quasi-monochromatic approximation, specifically as a carrier wave at a large optical frequency that is modulated in time by a temporal envelope:

$$\vec{E}(z, t) = \vec{E}_o^{(+)}(z, t) e^{-i\omega_0 t}$$

where $E_o(z, t)$ represents the complex amplitude, also known as the **temporal pulse envelope**, and ω_0 is the carrier frequency.

The validity of this approximation can be justified by considering the period of a single cycle oscillation: $T = 2\pi/\omega \approx 2 \cdot 10^{-15}$ s. This expression will be valid for pulse envelope durations larger than T . (We will not consider transverse dependence of the light pulses in this chapter, since the more detailed spatio-temporal dependence is highly complicated.) We assume in the following that our beam can be considered spatially homogeneous in the transverse plane.

For the moment, we will consider the properties of the light pulse at the plane $z = 0$ and concentrate on the basic properties of its temporal dependence. Later in this chapter, we will consider the effects of pulse propagation. The expression for the light pulse can be written in terms of its real amplitude and temporal phase, which are, respectively, $|E_o^{(+)}(t)|$ and $\varphi(t)$ in the complex representation:

$$E^{(+)}(t) = E_o^{(+)}(t) e^{-i\omega_0 t} = |E_o^{(+)}(t)| e^{i\varphi(t)} e^{-i\omega_0 t} \quad [5.1A]$$

The real electric field is given by the real part of the analytic signal, as usual:

$$E(t) = \text{Re}[E^{(+)}(t)] \quad [5.1B]$$

We will not consider polarization effects in this chapter and thus will omit the vector character of the fields in the expressions.

The temporal intensity of the pulse is proportional to the modulus squared of the real amplitude:

$$I(t) \propto |E_o^{(+)}|^2$$



An alternative way to write the light pulse is to express the scalar wave function in terms of intensity:

$$U^{(+)}(t) = \sqrt{I(t)} e^{i\varphi(t)} e^{-i\omega_0 t} \quad [5.1C]$$

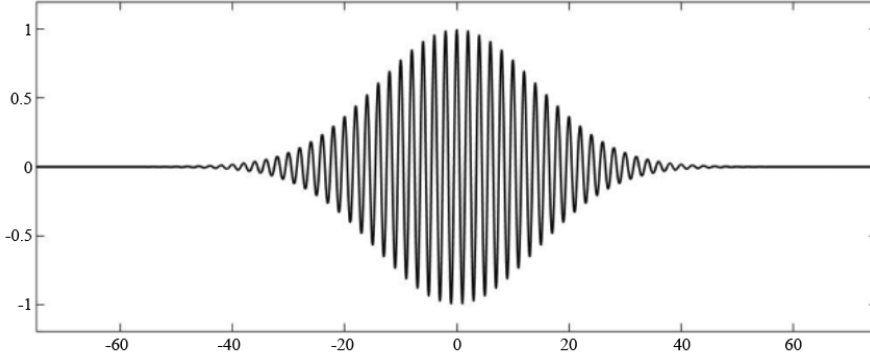


Figure 5.1
The electric field of a light pulse showing the fast oscillations of the carrier frequency modulated by the temporal envelope

When working with pulses, it is usually advisable to work in the frequency space, which is related to the temporal space by means of a Fourier transform. We define the frequency representation of the temporal field as

$$\tilde{E}(\omega) = \int_{-\infty}^{\infty} E(t) e^{i\omega t} dt \quad \Leftrightarrow \quad E(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}(\omega) e^{-i\omega t} d\omega \quad [5.2]$$

These expressions mathematically represent the physical fact that a pulse is obtained as a coherent superposition of different frequencies. The electric field appearing in these relations corresponds to the real electric field and, hence, its Fourier transform, $\tilde{E}(\omega)$, will be complex. The reality of the temporal field imposes a relation between the negative and positive components of the Fourier transform:

$$\tilde{E}(-\omega) = \tilde{E}(\omega)^*$$

In order to work with complex representations of the temporal fields, we define the positive part of the spectral field:

$$\tilde{E}^{(+)}(\omega) = \begin{cases} \tilde{E}(\omega) & \text{if } \omega \geq 0 \\ 0 & \text{if } \omega < 0 \end{cases}$$

The Fourier transform (FT) of this function is complex and corresponds to the analytic representation of the real electric field. This means that the functions $E^{(+)}(t)$ and $\tilde{E}^{(+)}(\omega)$ are a complex FT pair:

$$\tilde{E}^{(+)}(\omega) = \int_{-\infty}^{\infty} E^{(+)}(t) e^{i\omega t} dt \quad \Leftrightarrow \quad E^{(+)}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}^{(+)}(\omega) e^{-i\omega t} d\omega \quad [5.3]$$



If, for the expression $E^{(+)}(t)$, we factor out an oscillatory term proportional to the carrier frequency, we obtain

$$E^{(+)}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E^{(+)}(\omega) e^{-i(\omega-\omega_0)t} e^{-i\omega_0 t} d\omega = \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} E^{(+)}(\omega_0 + \Omega) e^{-i\Omega t} d\Omega \right] e^{-i\omega_0 t} \quad [5.4A]$$

where we define the translated frequency $\Omega = \omega - \omega_0$.

Comparing Equations [5.1A] and [5.4A], we recognize that the term in brackets is the temporal complex amplitude. The Fourier transform of the complex amplitude is hence the translated version of the electric field Fourier transform:

$$E_o^{(+)}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}^{(+)}(\Omega) e^{-i\Omega t} d\Omega \quad \Leftrightarrow \quad \tilde{E}^{(+)}(\Omega) = \int_{-\infty}^{\infty} E_o^{(+)}(t) e^{i\Omega t} dt \quad [5.4B]$$

where the term $\tilde{E}(\Omega)$ is written in this simplified way, although it refers to the complex amplitude corresponding to the real frequency $\omega = \Omega + \omega_0$.

We will mostly use this notation to refer to the evolution of the pulse envelope, obviating the fact that the total analytic function is obtained through multiplying the envelope by the carrier frequency exponential (similar to what we did in the derivation of the paraxial wave solutions). With these considerations in mind, we will usually write the light pulse in either the temporal or frequency domain.

Light pulse in the **temporal representation**:

$$E^{(+)}(t) = |E_o^{(+)}(t)| e^{i\varphi(t)} e^{-i\omega_0 t} \Rightarrow I(t) \propto |E_o^{(+)}(t)|^2 \quad [5.5A]$$

where

- $|E_o^{(+)}(t)|$ is the temporal pulse envelope;
- $\varphi(t)$ is the temporal phase; and
- $I(t)$ is the pulse intensity.

Light pulse in the **spectral representation**:

$$\tilde{E}^{(+)}(\omega_0 + \Omega) = |\tilde{E}_o^{(+)}(\Omega)| e^{i\psi(\Omega)} \Rightarrow S(\Omega) \propto |\tilde{E}_o^{(+)}(\Omega)|^2 \quad [5.5B]$$

where

- $|\tilde{E}_o^{(+)}(\Omega)|$ is the spectral pulse envelope;
- $\psi(\Omega)$ is the spectral phase; and
- $S(\Omega)$ is the pulse spectrum.



5.2 Pulse duration and bandwidth

As we saw in Chapter 2 when dealing with Gaussian beams, there is no single expression for defining the beam radius. In the same way, we can use different expressions to define the pulse duration or the spectrum bandwidth. If the temporal pulses have a Gaussian envelope profile, we can define the $1/e$, $1/e^2$ or FWHM pulse durations, as we did with optical beam radius. For pulse profiles other than Gaussian but with a smooth single-pulse profile, the FWHM convention can usually be used. For pulses of more complex shapes or multi-peaked profiles, we define the general pulse duration as the standard deviation:

$$\Delta t = \sqrt{\langle t^2 \rangle - \langle t \rangle^2}$$

with

$$\langle f \rangle = \frac{\int_{-\infty}^{\infty} f I(t) dt}{\int_{-\infty}^{\infty} I(t) dt}$$

An equivalent expression is defined to obtain the width of the spectrum that defines the range of frequencies present in the pulse.

Time-Bandwidth product (TBP)

The time-bandwidth product of the pulse depends on the temporal and spectral phase of the pulse. In general, it holds:

$$TBP = \Delta\tau \cdot \Delta\Omega \geq K \quad [5.6]$$

where K is a constant that depends on the particular pulse profile. This is a general relation linking the FT pairs.

The minimum value of TBP corresponds to the situation with zero or constant temporal and spectral phases when all the frequencies forming the pulse are added in phase. The pulse in this case is called **Fourier transform-limited**, and it represents the minimum pulse duration that can be attained for a given frequency spectrum.

Fourier-transform-limited pulse: $\Delta t = K/\Delta\Omega$

A temporal-dependent spectral phase means that the different frequency components of the pulse overlap with some relative phase difference between them. As we showed previously, this leads to a broadening of the pulse.

Pulse with non-constant temporal or spectral phase: $\Delta t > K/\Delta\Omega$.

Example 1: TBP of Fourier-transform-limited Gaussian pulse

A Gaussian pulse with pulse duration ($1/e$) in amplitude τ_0 is written as $E_o^{(+)}(t) = E_o e^{-t^2/\tau_0^2}$



Pulse duration: $\Delta t = \tau_o$

To obtain the bandwidth we calculate the spectral amplitude:

$$\tilde{E}_0^{(+)}(\Omega) = \int_{-\infty}^{\infty} E_0^{(+)}(t) e^{i\Omega t} dt = E_o \int_{-\infty}^{\infty} e^{-t^2/\tau_o^2} e^{i\Omega t} dt$$

The integral can be solved using

$$\int_{-\infty}^{\infty} e^{-At^2} e^{-2Bt} dt = \frac{\sqrt{\pi}}{\sqrt{A}} e^{B^2/A} \quad (\text{Re}(A) > 0) \quad [5.7]$$

We obtain

$$\tilde{E}_0^{(+)}(\Omega) = E_o \sqrt{\pi} \tau_o e^{-\Omega^2 \tau_o^2 / 4} = E_o \sqrt{\pi} \tau_o e^{-\Omega^2 / (2/\tau_o)^2}$$

The bandwidth is then

$$\Delta\Omega = 2/\tau_o$$

and

$$\text{TBP} = 2$$

Example 2: TBP of a Gaussian pulse with quadratic temporal phase dependence

The pulse in this case is expressed as

$$E_0^{(+)}(t) = E_o e^{-\frac{t^2}{\tau_o^2}(1+i\alpha)}$$

The spectral amplitude is obtained by direct Fourier transform:

$$\tilde{E}_0^{(+)}(\Omega) = \int_{-\infty}^{\infty} E_0^{(+)}(t) e^{i\Omega t} dt = \frac{E_o^{(+)} \sqrt{\pi} \tau_o}{\sqrt{1+i\alpha}} e^{-\frac{\Omega^2 \tau_o^2}{4(1+i\alpha)}}$$

By separating the amplitude and phase in the exponential function, we obtain the expression

$$\tilde{E}_0^{(+)}(\Omega) = \frac{E_o^{(+)} \sqrt{\pi} \tau_o}{\sqrt{1+i\alpha}} e^{-\frac{\Omega^2 \tau_o^2}{4(1+\alpha^2)}} e^{i \frac{\Omega^2 \tau_o^2 \alpha}{4(1+\alpha^2)}}$$

The frequency bandwidth is obtained directly from the gaussian amplitude:

$$\tilde{E}_o^{(+)}(\Omega) \propto e^{-\frac{\Omega^2}{\Delta\Omega^2}} \quad \text{with} \quad \Delta\Omega = \frac{2\sqrt{1+\alpha^2}}{\tau_o}$$

The TBP becomes:

$$\text{TBP} = \Delta t \cdot \Delta\Omega = 2\sqrt{1+\alpha^2}$$



From this result, we can infer that, for a fixed frequency content (fixed frequency bandwidth), a pulse with quadratic temporal dependence in the temporal phase will result in a longer pulse, with its duration being $\sqrt{1 + \alpha^2}$ times longer than a Fourier-transform-limited pulse with the same bandwidth.

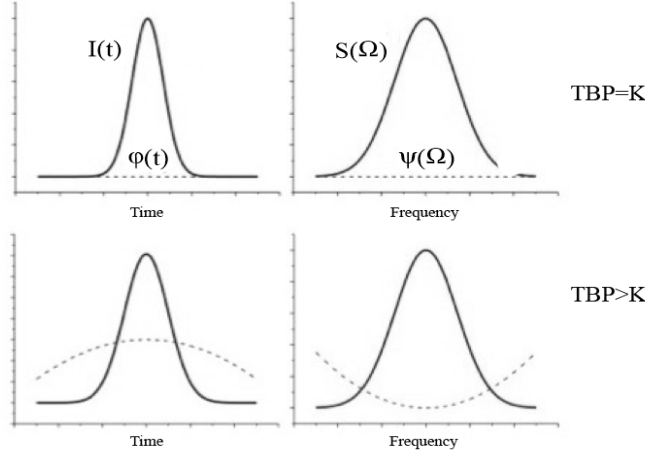


Figure 5.2
Two Gaussian pulses with identical spectrum and different phases. (Top) Fourier-transform-limited pulse. (Bottom) pulse with quadratic phase dependence leading to longer pulses in the time domain for a given spectrum.

5.3 Instantaneous frequency and chirp

In the preceding section, we saw that an optical pulse can be expressed as a temporally modulated plane wave with carrier frequency (ω_0). This frequency is usually associated with the maximum at the pulse spectrum. In this section, we will explore more closely the role played by the temporal phase term in the temporal distribution of the electric field.

Fast variations of the temporal phase can modify the instantaneous frequency across the pulse, thus leading to frequency modulation effects.

The light pulse can be written in terms of its amplitude and total phase:

$$E^{(+)}(t) = |E_o^{(+)}(t)| e^{i\phi(t)} e^{-i\omega_0 t} = |E_o^{(+)}(t)| e^{-i(\omega_0 t - \phi(t))} = |E_o^{(+)}(t)| e^{-i\Phi(t)}$$

The **instantaneous frequency** of the pulse is defined through the relation

$$\omega_{inst}(t) = \frac{d\Phi(t)}{dt} = \omega_0 - \frac{d\phi(t)}{dt} \quad [5.8]$$

This definition allows us to see that when the temporal phase of the amplitude, $\phi(t)$, is constant or zero, the instantaneous frequency is constant across the pulse and corresponds to the carrier frequency. This is the case for the Fourier-transform-limited pulses shown in Figure 1. Any temporal dependence of the temporal phase will lead to a change in the instantaneous frequency.

Writing out the temporal phase in general form, we have the expansion



$$\varphi(t) = \varphi_0 + \varphi_1 t + \frac{1}{2} \varphi_2 t^2 + \frac{1}{3!} \varphi_3 t^3 + \dots$$

Using [5.8], the instantaneous frequency becomes

$$\omega_{inst}(t) = \omega_0 - \varphi_1 - \varphi_2 t - \frac{1}{2} \varphi_3 t^2 - \dots$$

When the instantaneous frequency depends on time, the pulse becomes frequency modulated and we say it is **chirped**.

In the case where $\varphi_2 \gg \varphi_3, \varphi_4, \dots$, the instantaneous frequency varies linearly with time and the pulse is called a **linearly chirped pulse**. In this case, we define the **chirp coefficient** C through the relation

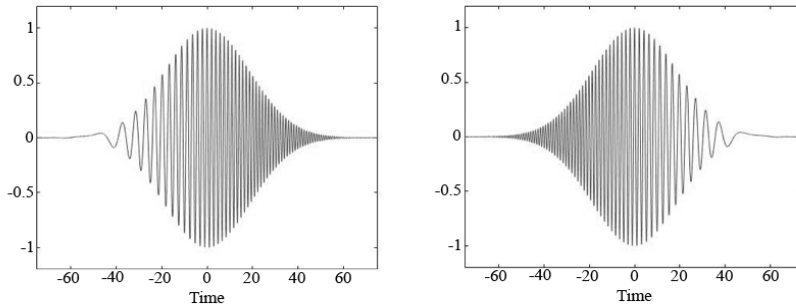
$$\omega_{inst}(t) = \omega_0' + Ct \quad [5.9]$$

where $\omega_0' = \omega_0 - \varphi_1$ and $C = -\varphi_2$. The dimensions of C are $[T]^{-2}$, and they are usually expressed in fs^{-2} because the effect of the chirp on the pulses generally becomes more significant for pulses with pulse durations of the order of femtoseconds.

Pulses with $C > 0$ are called **up-chirped pulses**, because the instantaneous frequency increases with time. The pulses with $C < 0$ are called **down-chirped pulses**, since the instantaneous frequency decreases with time.

Figure 3 shows a representation of up-chirped and down-chirped pulses. The presence of chirp in a pulse indicates unequivocally that the pulse is not Fourier-transform-limited.

Figure 5.3
Pulses with linear
up-chirp (left) and down-
chirp (right)



For the case of Gaussian pulses, the mathematical representation of a linearly chirped pulse corresponds to the expression in Example 2 (for the particular case $\varphi_0 = \varphi_1 = 0$):

$$E^{(+)}(t) = E_0 e^{-\frac{t^2}{\tau_0^2}(1+i\alpha)} e^{-i\omega_0 t}$$

We can identify that this corresponds to an up-chirped pulse with chirp coefficient:

$$C = \frac{2\alpha}{\tau_0^2}$$



The variable α (dimensionless) is called the **chirp parameter** of the Gaussian pulse. Next, we will see that light pulses acquire chirp as they propagate in dispersive media.

5.4 Optical pulse propagation in dispersive media

Dispersive media are characterized by a frequency-dependent index of refraction $n(\omega)$. This frequency dependence affects the behaviour of an optical pulse propagating in such a medium. As we have seen, any pulse can be decomposed into a series of frequency components that, when added together coherently, constitute the pulse profile. A change in refractive index implies a change in phase velocity and, hence, we can infer that the different spectral components will change their relative phases upon propagation. The final consequence of this dephasing is the appearance of pulse chirp and the temporal broadening of the pulse. In order to describe this phenomenon, we consider a light pulse with initial envelope $E^{(+)}(0, t)$, and our goal will be to trace the evolution of the pulse along the propagation distance $E^{(+)}(z, t)$.

The starting point for studying pulse propagation is the dispersion relation of the medium:

$$k(\omega) = \frac{n(\omega)\omega}{c} \quad [5.10A]$$

Under the assumption of quasi-monochromatic representation, the typical bandwidth of the pulse envelope should be much smaller than the carrier-envelope frequency, $\Delta\Omega \ll \omega_0$. Consequently, the dispersion relation can be expanded around the carrier frequency:

$$k(\omega) = k(\omega_0) + \left[\frac{\partial k}{\partial \omega} \right]_{\omega_0} (\omega - \omega_0) + \frac{1}{2} \left[\frac{\partial^2 k}{\partial \omega^2} \right]_{\omega_0} (\omega - \omega_0)^2 + \dots \quad [5.10B]$$

Expressed alternatively in terms of Ω , we have

$$k(\Omega) = k(0) + \left[\frac{\partial k}{\partial \omega} \right]_{\omega_0} \Omega + \frac{1}{2} \left[\frac{\partial^2 k}{\partial \omega^2} \right]_{\omega_0} \Omega^2 + \dots \quad [5.10C]$$

The expressions that we derive from the expansion up to the second order in frequency constitute the so-called second-order approximation pulse propagation models.

The terms in brackets are essential to the theory and are related to the **group velocity** and **group velocity dispersion coefficient** of the pulse, as described below.

Group velocity (u)

The group velocity of the pulse is defined as

$$u(\omega) = \left[\frac{\partial k}{\partial \omega} \right]^{-1} = \left[\frac{\omega dn(\omega)/d\omega}{c} + \frac{n(\omega)}{c} \right]^{-1}$$

$$u(\omega) = \frac{c}{n(\omega) + \frac{dn(\omega)}{d\omega} \omega}$$



Since the index of refraction is usually expressed in terms of wavelength, it is more convenient to rewrite the group velocity in terms of wavelength. Using the relation

$$\omega = 2\pi c / \lambda \Rightarrow \frac{d\omega}{d\lambda} = -\frac{2\pi c}{\lambda^2} d\lambda = -\frac{\omega}{\lambda} d\lambda \Rightarrow \omega / d\omega = -\lambda / d\lambda$$

(where λ refers to vacuum wavelength), we can write

$$u(\lambda) = \frac{c}{n(\lambda) - \lambda \frac{dn(\lambda)}{d\lambda}} \quad [5.11A]$$

This expression allows us to obtain the group velocity at any wavelength from the expression of its index of refraction.

Group-velocity dispersion coefficient (g)

The group velocity dispersion coefficient can be expressed in terms of the group velocity:

$$g(\omega) = \left[\frac{\partial^2 k}{\partial \omega^2} \right] = \frac{\partial}{\partial \omega} \left[\frac{1}{u(\omega)} \right] = -\frac{1}{u^2(\omega)} \frac{du(\omega)}{d\omega}$$

This parameter is related to the change in the group velocity with frequency. We distinguish the following two different regimes.

Normal group velocity dispersion: in this regime, the group velocity decreases with frequency, so low frequency components travel faster.

Anomalous group velocity dispersion: the group velocity increases with frequency, meaning that high frequency components travel faster.

In order to obtain a plot of this function, it is better repeat what we did with the group velocity and express it as a function of the wavelength:

$$g(\lambda) = \frac{\lambda^3}{2\pi c^2} \left[\frac{d^2 n(\lambda)}{d\lambda^2} \right] \quad [5.11B]$$

The units of this coefficient are s^2/m , but they are most often expressed in fs^2/mm .

In terms of u and g , the dispersion relation is written up to the second order as

$$k(\omega_o + \Omega) = k(\omega_o) + \frac{\Omega}{u_o} + \frac{1}{2} g_o \Omega^2$$

Basic equation for pulse propagation

The equation defining the propagation of a light pulse in a dispersive medium is obtained from the general wave equation derived in Chapter 3, assuming an isotropic, dispersive, homogeneous, and non-conductive medium:



$$\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{E}) + \mu_o \frac{\partial^2 \vec{D}}{\partial t^2} = 0$$

Since the medium is dispersive, the relation between the electric field and the electric displacement for the case of linear interaction is

$$\vec{D}(\vec{r}, t) = \epsilon_o \int_{-\infty}^t \epsilon(\vec{r}, t - t') \vec{E}(t') dt'$$

Inserting this expression into the wave equation and using the expansion of the dispersion relation around the carrier frequency, an equation for the pulse propagation is obtained. The detailed treatment is beyond the scope of this introductory course (interested readers can find all the details in, for instance, the book Diels, J.C., and Rudolph, W. (2006). *Ultrashort Laser Pulse Phenomena* Academic Press). The resulting equation is

$$\left\{ \frac{\partial}{\partial z'} + \frac{1}{u} \frac{\partial}{\partial t'} + \frac{ig}{2} \frac{\partial^2}{\partial t'^2} - \frac{i}{2k_o} \left[\frac{\partial^2}{\partial z'^2} - \frac{1}{u^2} \frac{\partial^2}{\partial t'^2} \right] \right\} E_o^{(+)}(t', z') - D = 0 \quad [5.12]$$

$$D = \frac{1}{2k_o} \sum_{m=3}^{\infty} \frac{(-i)^{m+1}}{m!} \left(\frac{\partial^m k}{\partial \omega^m} \right)_{\omega_o} \frac{\partial^m E_o^{(+)}(t', z')}{\partial t'^2}$$

where the D term includes higher-order dispersion terms. When the duration of the pulses is larger than a few tenths of femtoseconds, this term can be neglected when using the **equation for pulse propagation up to second-order in dispersion**. This equation is expressed in the laboratory frame of reference (z', t'), but it can be simplified if we introduce a system of local coordinates (z, t) and the retarded frame of reference, which is a reference system that moves with the pulse at group velocity u . The change in coordinates to the new system is given by

$$z = z'$$

$$t = t' - z/u.$$

Under conditions when the pulse envelope variations over distances of the order of the wavelength or over times of the order of the optical period are much smaller than the value of the envelope itself (slowly varying envelope approximation, SVEA, in space and time) we obtain the simplified equation:

$$\boxed{\frac{\partial E_o^{(+)}}{\partial z} + \frac{ig}{2} \frac{\partial^2 E_o^{(+)}}{\partial t^2} = 0} \quad [5.13]$$

This is the basic equation we will use to determine the effect of the propagation on a given temporal pulse profile.

The general solution to this equation can be more easily obtained in the frequency space. In terms of its spectral components, the electric pulse envelope is written as in [5.4B], where we now add the dependence on the propagation variable z :



$$E_o^{(+)}(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}^{(+)}(z, \Omega) e^{-i\Omega t} d\Omega \quad [5.14]$$

Substituting this expression into [5.13]:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ \frac{\partial \tilde{E}^{(+)}}{\partial z} - \frac{ig\Omega^2}{2} \tilde{E}^{(+)} \right\} e^{-i\Omega t} d\Omega = 0$$

This relation is true for any arbitrary envelope if it satisfies

$$\frac{\partial \tilde{E}^{(+)}}{\partial z} - \frac{ig\Omega^2}{2} \tilde{E}^{(+)} = 0$$

The solution to this equation is easily found:

$$\boxed{\tilde{E}^{(+)}(z, \Omega) = \tilde{E}^{(+)}(0, \Omega) e^{i\frac{g\Omega^2 z}{2}}} \quad [5.15]$$

Written as

$\tilde{E}^{(+)}(z, \Omega) = \tilde{E}^{(+)}(0, \Omega) H(z, \Omega)$, the function $H(z, \Omega) = e^{ig\Omega^2 z/2}$ is called the **transfer function for pulse propagation in dispersive media**.

The general equation in the time domain is obtained directly through [5.14].

A close inspection of the solution can give us some physical insight into the effects on any pulse propagating in a dispersive medium. Let us consider an initial pulse with arbitrary pulse envelope at $z = 0$; $E_o^{(+)}(0, t)$. By direct Fourier transformation, we obtain its spectral amplitude and phase:

$$\tilde{E}^{(+)}(0, \Omega) = \left| \tilde{E}^{(+)}(0, \Omega) \right| e^{i\psi(0, \Omega)}$$

and its spectrum $S(0, \Omega) = \left| \tilde{E}^{(+)}(0, \Omega) \right|^2$.

By applying the result [5.15] to this pulse, we see immediately the following consequences:

- Due to the effect of propagation we add a phase term to the initial spectral envelope. Since this phase term does not change the absolute value of the complex envelope, **the spectrum of the pulse is not changed upon propagation**. This means that no new frequencies are added to the pulse spectrum.
- The spectral phase at the output will be that of the input plus the one added by the effect of propagation:

$$\psi(z, \Omega) = \psi(0, \Omega) + \frac{g\Omega^2 z}{2}$$



This phase term is proportional to the propagated distance and to the square of the frequency. This quadratic frequency phase will lead to a quadratic temporal phase in the time domain, as seen previously in Example 2 and consequently, the **pulse becomes chirped**.

In the normal group velocity dispersion regime, $g > 0$, and the pulse is up-chirped during propagation; while in the anomalous group velocity dispersion regime, $g < 0$, the pulse will be down-chirped.

- As a consequence of dephasing the different spectral components, **the pulse broadens in the temporal domain**.
- If $g = 0$ the pulse propagates without distortion in the medium.

The presence of group velocity dispersion (g) gives rise to a dephasing of the different frequencies composing the pulse. As a consequence of this dephasing, the pulse gets chirped and broadened during propagation.

Example 3: *Propagation of an initially Fourier-transform-limited Gaussian pulse in a dispersive medium*

Consider the initial pulse

$$E_o^{(+)}(0, t) = E_o e^{-t^2/\tau_o^2}$$

Its spectral amplitude at $z = 0$ was calculated in Example 1:

$$\tilde{E}^{(+)}(0, \Omega) = \int_{-\infty}^{\infty} E_o e^{-t^2/\tau_o^2} e^{i\Omega t} dt = E_o \tau_o \sqrt{\pi} e^{-\frac{\Omega^2 \tau_o^2}{4}}$$

The propagated spectral envelope at distance z is given by [5.15]:

$$\tilde{E}^{(+)}(z, \Omega) = \tilde{E}^{(+)}(0, \Omega) e^{i \frac{g \Omega^2 z}{2}} = E_o \tau_o \sqrt{\pi} e^{-\Omega^2 \left[\frac{\tau_o^2}{4} - i \frac{gz}{2} \right]}$$

The propagated solution in the temporal domain can be found directly:

$$E_o^{(+)}(z, t) = \frac{1}{2\pi} E_o \tau_o \sqrt{\pi} \int_{-\infty}^{\infty} e^{-\Omega^2 \left[\frac{\tau_o^2}{4} - i \frac{gz}{2} \right]} e^{-i\Omega t} d\Omega$$

The integral can be solved using [5.7]:

$$E_o^{(+)}(z, t) = \frac{E_o}{2\pi} \sqrt{\pi} \tau_o \frac{\sqrt{\pi}}{\sqrt{\frac{\tau_o^2}{4} - i \frac{gz}{2}}} e^{-\frac{t^2}{4 \left[\frac{\tau_o^2}{4} - i \frac{gz}{2} \right]}}$$

Rearranging terms, we obtain



$$E_o^{(+)}(z, t) = \frac{E_o}{\sqrt{1 - i \frac{z}{L_D}}} e^{-\frac{t^2}{\tau_o^2(1 - i \frac{z}{L_D})}}$$

where we define the group velocity dispersion length, L_D : $L_D = \frac{\tau_o^2}{2g}$

This expression can be rewritten as

$$E_o^{(+)}(z, t) = E_o \frac{\left(1 + \frac{z^2}{L_D^2}\right)^{1/4} e^{i \frac{\text{atg}(z/L_D)}{2}}}{\sqrt{1 + \frac{z^2}{L_D^2}}} e^{-\frac{t^2}{\tau_o^2(1 + \frac{z^2}{L_D^2})}} e^{-i \frac{t^2 z}{\tau_o^2 L_D(1 + \frac{z^2}{L_D^2})}}$$

by defining the following parameters.

Pulse duration: $\tau(z) = \tau_o \sqrt{1 + \frac{z^2}{L_D^2}}$

Chirp Parameter: $\alpha(z) = \frac{L_D}{z(1 + \frac{L_D^2}{z^2})}$

Phase-shift: $\xi(z) = \text{atg}\left(\frac{z}{L_D}\right) / 2$

The Gaussian pulse after distance z can be written as

$$E_o^{(+)}(z, t) = E_o \sqrt{\frac{\tau_o}{\tau(z)}} e^{-\frac{t^2}{\tau^2(z)}} e^{-i \left(\frac{\alpha(z)t^2}{\tau_o^2} - \xi(z) \right)}$$

This expression is quite similar to that obtained for the Gaussian optical beam in Chapter 2.

The solution of the propagation Equation [5.13] can be obtained in the temporal domain by writing the inverse Fourier transform of Equation [5.15]:

$$E_o^{(+)}(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}^{(+)}(z, \Omega) e^{-i\Omega t} d\Omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}^{(+)}(0, \Omega) e^{i \frac{g\Omega^2 z}{2}} e^{-i\Omega t} d\Omega$$

The spectral envelope at $z = 0$ can be written in terms of the pulse envelope profile at $z = 0$ to obtain



$$E_0^{(+)}(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} E_0^{(+)}(0, t') e^{i\Omega t'} dt' \right\} e^{i\frac{g\Omega^2 z}{2}} e^{-i\Omega t} d\Omega$$

Interchanging the integration order, we rewrite this equation in the form

$$E_0^{(+)}(z, t) = \int_{-\infty}^{\infty} E_0^{(+)}(0, t') \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\frac{g\Omega^2 z}{2}} e^{-i\Omega(t-t')} d\Omega \right\} dt'$$

We define the function

$$h(t, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\frac{g\Omega^2 z}{2}} e^{-i\Omega t} d\Omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(z, \Omega) e^{-i\Omega t} d\Omega$$

which thus corresponds to the **impulse response function or temporal propagator of the light pulses in dispersive media**. As is evident from this expression, the impulse response function and the transfer function are a Fourier transform pair.

$$\boxed{E_o^{(+)}(z, t) = \int_{-\infty}^{\infty} h(t - t', z) E_0^{(+)}(0, t') dt'} \quad [5.16]$$

By performing the integral, the impulse response function can be expressed alternatively as

$$h(z, t) = \frac{1}{2\pi} \frac{\sqrt{2\pi}}{\sqrt{-igz}} e^{-i\frac{t^2}{2gz}} = \frac{e^{-it^2/(2gz)}}{\sqrt{-i2\pi gz}}$$

From these relations, we can observe the close similitude between the general solution to the parabolic wave equation for beams (outlined at the end of Chapter 2) and the solution of the pulse propagation equation in the second-order approximation of the dispersion, which is derived here. The mathematical formulation of these two problems is in fact very similar, and in both cases the physical origin of the broadening of the solutions is a result of dephasing either the different plane waves (for beams) of the light beam or the monochromatic beams (pulses) forming the light pulse.



5.5 Solved problems

S5.1 The index of refraction of fused silica is given by the expression

$$n^2(\lambda) = 1 + \frac{0.6961663 \cdot \lambda^2}{\lambda^2 - 0.0684043^2} + \frac{0.4079426 \cdot \lambda^2}{\lambda^2 - 0.1162414^2} + \frac{0.8974794 \cdot \lambda^2}{\lambda^2 - 9.896161^2}$$

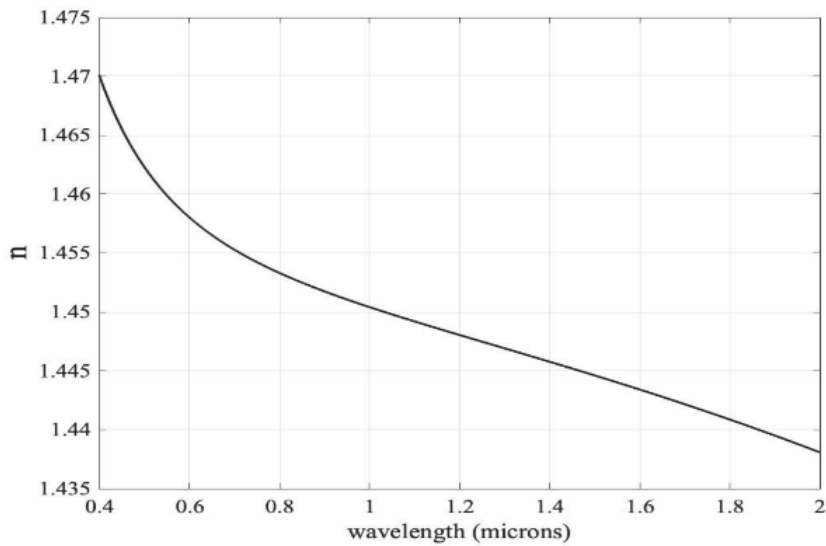
Obtain a plot of the following magnitudes in the range 400 nm to 2000 nm.

- Phase velocity
- Group velocity (u)
- Group velocity dispersion (g)

Solution

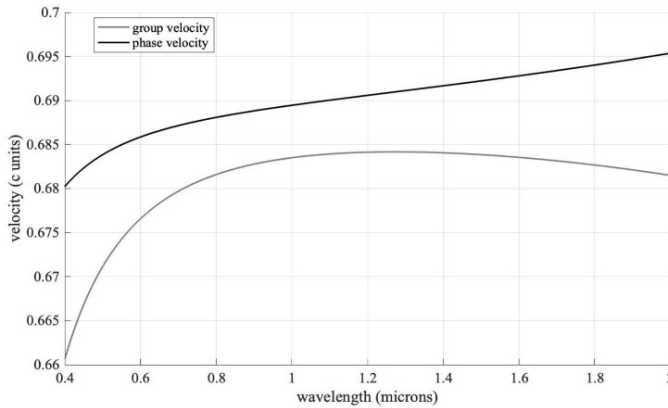
From the Sellmeier equation, we can extract relevant information concerning the optical dispersive properties of a given crystal.

The index of refraction corresponds to the real part of n :

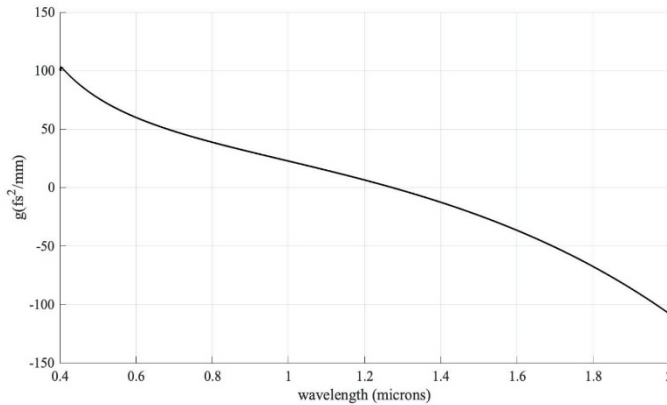


The phase velocity is defined as $v(\lambda) = c/n(\lambda)$, while the group velocity can be calculated from the expression $u(\lambda) = \frac{c}{n(\lambda) - \lambda \frac{dn(\lambda)}{d\lambda}}$:

$$u(\lambda) = \frac{c}{n(\lambda) - \lambda \frac{dn(\lambda)}{d\lambda}}$$



The group velocity dispersion coefficient $g(\lambda) = \frac{\lambda^3}{2\pi c^2} \left[\frac{d^2 n(\lambda)}{d\lambda^2} \right]$:



■

S5.2 A down-chirped optical pulse with Gaussian temporal envelope and carrier central wavelength $\lambda_o = 800 \text{ nm}$ has a time-bandwidth product (TBP) value of $\tau \Delta\Omega = 3.5$, where τ is the $1/e$ pulse duration in amplitude of the field envelope $E_o(t)$, and $\Delta\Omega$ is the corresponding $1/e$ frequency bandwidth of its Fourier transform $E_o(\Omega)$. The pulse frequency bandwidth is $\Delta\Omega = 3 \cdot 10^{13} \text{ rad/s}$.

- Calculate the pulse duration, the chirp coefficient, the chirp parameter and the instantaneous frequency of the pulse at this instant of time.
- Determine the frequency bandwidth of the spectrum at FWHM, and express it as a wavelength bandwidth, $\Delta\lambda$.
- Explain briefly how this pulse can be compressed, and give the value of the minimum pulse duration that we could obtain after compression.

**Solution**

a) The pulse duration can be calculated directly from the TBP, since we know the bandwidth of the pulse:

$$\tau = \frac{TBP}{\Delta\Omega} = \frac{3.5}{3 \cdot 10^{13}} = 1.167 \cdot 10^{-13} \text{ s} = 116.7 \text{ fs}$$

For Gaussian pulses, the TBP equals 2 if the pulse duration and bandwidth are defined at 1/e in amplitude. Since the value of the TBP is 3.5, we know that the pulse is chirped; and since it is down-chirped, the value of the chirp parameter is negative.

In Example 2 of Section 5.2, we derived the expression for the TBP of a chirped pulse. From this relation, we can obtain the chirp parameter:

$$TBP = 2\sqrt{1+\alpha^2} \rightarrow |\alpha| = \sqrt{\left(\frac{TBP}{2}\right)^2 - 1} = 1.436$$

For our down-chirped pulse, the chirp parameter is $\alpha = -1.436$.

Writing the expression of our Gaussian beam as $E_o^{(+)}(t) = E_o e^{-\frac{t^2(1+i\alpha)}{\tau^2}}$, the chirp coefficient is obtained as

$$C = \frac{2\alpha}{\tau^2} = -2.11 \cdot 10^{-4} \text{ fs}^{-2}$$

The instantaneous frequency is

$$\omega_{inst}(t) = \omega_o + Ct = \frac{2\pi c}{\lambda} + Ct = (2.35 - 2.11 \cdot 10^{-4} t) \text{ fs}^{-1}, \text{ with } t \text{ given in fs.}$$

b) The spectrum is expressed as

$$S(\Omega) = \left| \tilde{E}_o^{(+)}(\Omega) \right|^2 = E_o^2 e^{-\frac{2\Omega^2}{\Delta\Omega^2}} = S_o e^{-\frac{2\Omega^2}{\Delta\Omega^2}}$$

To obtain the FWHM frequency bandwidth, we look for the value of the spectrum whose value is $S(\Omega_{1/2}) = S_o/2$:

$$\frac{1}{2} = e^{-\frac{2\Omega_{1/2}^2}{\Delta\Omega^2}} \rightarrow \ln 2 = \frac{2\Omega_{1/2}^2}{\Delta\Omega^2} \rightarrow \Omega_{1/2} = \sqrt{\frac{\ln 2}{2}} \cdot \Delta\Omega$$

$$\text{The FWHM is } \Delta\Omega_{FWHM} = 2\Omega_{1/2} = \sqrt{2 \ln 2} \cdot \Delta\Omega = 3.53 \cdot 10^{13} \text{ rad / s ,}$$

which can be expressed as a wavelength bandwidth using the relation between Ω and λ :



$$\Omega = \frac{2\pi c}{\lambda} \rightarrow \Delta\Omega = -\frac{2\pi c}{\lambda^2} \Delta\lambda$$

$$|\Delta\lambda| = \frac{\lambda^2 \Delta\Omega}{2\pi c} = 11.98 \cdot 10^{-9} m = 12 \text{ nm}$$

c) The fact that the TBP of the pulse is greater than 2 indicates that this pulse is not Fourier-transform-limited. The pulse could be compressed if we let it propagate in a medium that compensates the chirp present in the pulse. Since the pulse is down-chirped, we need a medium that introduces a positive chirp (positive GVD coefficient, g). The medium length necessary to compensate the initial chirp will be explained in one of the following problems.

The minimum pulse duration will be achieved when the TBP equals 2, and then

$$\tau_{\min} = \frac{2}{\Delta\Omega} = 66.67 \text{ fs}$$

■

S5.3 A Gaussian pulse with pulse duration τ and chirp parameter α propagates in a medium of group velocity dispersion coefficient g . Determine the expression for the pulse amplitude as a function of the distance z propagated in the medium; and, from this expression, obtain the pulse duration as a function of z .

Solution

The expression for the Gaussian pulse envelope at $z = 0$ is

$$E_o^{(+)}(0, t) = E_o e^{-\frac{t^2}{\tau^2}(1+i\alpha)}$$

To calculate the propagated field, we will use the transfer function formalism. We start by calculating the spectral amplitude of the pulse:

$$\tilde{E}_o^{(+)}(0, \Omega) = \int_{-\infty}^{\infty} E_o e^{-\frac{t^2}{\tau^2}(1+i\alpha)} e^{i\Omega t} dt = \frac{E_o \tau \sqrt{\pi}}{\sqrt{1+i\alpha}} e^{-\frac{\Omega^2 \tau^2}{4(1+i\alpha)}}$$

where we used the integral $\int_{-\infty}^{\infty} e^{-Ax^2 - 2Bx} dx = \frac{\sqrt{\pi}}{\sqrt{A}} e^{B^2/A} \quad (\text{Re}(A) > 0).$

The propagation of the spectral amplitude is directly obtained by multiplying it by the transfer function:

$$\tilde{E}_o^{(+)}(z, \Omega) = \tilde{E}_o^{(+)}(0, \Omega) e^{i\frac{g\Omega^2 z}{2}} = \frac{E_o \tau \sqrt{\pi}}{(1+i\alpha)^{1/2}} e^{-\frac{\Omega^2 \tau^2}{4(1+i\alpha)}} e^{i\frac{g\Omega^2 z}{2}}$$

The solution in real space is obtained by performing the Fourier transform:



$$E_o^{(+)}(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}_o^{(+)}(z, \Omega) e^{-i\Omega t} d\Omega$$

This Fourier transform can be evaluated using the same integral as before. After a few algebraic steps, we reach the final result, which can be expressed in the form

$$E_o^{(+)}(z, t) = E_o \sqrt{\frac{\tau}{\tau(z)}} e^{-\frac{t^2}{\tau^2(z)}} e^{-i \left[\frac{t^2 \alpha(z)}{\tau^2} - \frac{1}{2} \arctan \left(\frac{z/L_D}{1 + \alpha z/L_D} \right) \right]}$$

where we introduce the pulse duration and pulse chirp parameter in terms of the group velocity dispersion length $L_D = \tau^2/2g$:

$$\tau(z) = \tau \sqrt{1 + \frac{2\alpha z}{L_D} + \frac{z^2(1 + \alpha^2)}{L_D^2}} = \tau \sqrt{1 + \frac{2 \cdot \text{sign}(\alpha g) |\alpha| z}{|L_D|} + \frac{z^2(1 + \alpha^2)}{L_D^2}}$$

$$\alpha(z) = \frac{\frac{z}{L_D}(1 + \alpha^2) + \alpha}{1 + \frac{2\alpha z}{L_D} + \frac{z^2(1 + \alpha^2)}{L_D^2}}$$

■

S5.4 Starting with the expression of the pulse duration of an initially chirped Gaussian pulse propagating in a dispersive medium with the group velocity dispersion (GVD) coefficient g that was obtained in problem S5.3, complete the following.

- Find the conditions leading to an initial compression of the pulse.
- Calculate the TBP at the initial position.
- Find the distance where the pulse duration is minimum, as well as the TBP at this point.

Solution

a) The representation of the Gaussian pulse at $z = 0$ is given in terms of the initial pulse duration, τ_{in} , and initial chirp parameter, α :

$$E_o^{(+)}(0, t) = E_o e^{-\frac{t^2(1+i\alpha)}{\tau_o^2}}$$

The initial chirp coefficient is $C_o = \frac{2\alpha}{\tau_{in}^2}$.

When propagating in the dispersive medium, the pulse duration, chirp parameter and chirp coefficient will change according to the results obtained in problem S5.3. The



physical mechanism leading to this process is the dephasing of the different frequency components of the pulse during propagation, due to their different propagation velocity. This leads to a general broadening of the pulse during propagation. However, in some situations, the pulse can be shortened when entering the crystal, thus reaching a minimum pulse duration at a particular distance inside the crystal. Beyond this point, the pulse begins increasing its duration again.

To analyse the conditions leading to this behaviour, we start with the expression of the pulse duration derived in the previous problem:

$$\tau(z) = \tau_{in} \sqrt{1 + \frac{2\alpha z}{L_D} + \frac{z^2(1+\alpha^2)}{L_D^2}} = \tau_{in} \sqrt{1 + \frac{2 \cdot \text{sign}(\alpha g) |\alpha| z}{|L_D|} + \frac{z^2(1+\alpha^2)}{L_D^2}}$$

The carrier wavelength of our pulse will dictate if we are in the normal GVD ($g > 0$) regime or in the anomalous GVD ($g < 0$) regime. In addition, the initial chirp of our pulse, α , may be positive (up-chirped pulse) or negative (down-chirped pulse).

An inspection of the equation for pulse duration shows that the first and last terms are always positive, thus leading to pulse lengthening. However, the second term may be positive or negative, depending on the sign of the g and α coefficients. Stated more clearly, the expression has been written in terms of the sign of the product of these two magnitudes, thus leading to the following two different scenarios.

- $\text{Sign}(g \cdot \alpha) > 0$

This condition is found when an initially up-chirped pulse propagates in a normal GVD regime or when an initially down-chirped pulse propagates in an anomalous GVD regime.

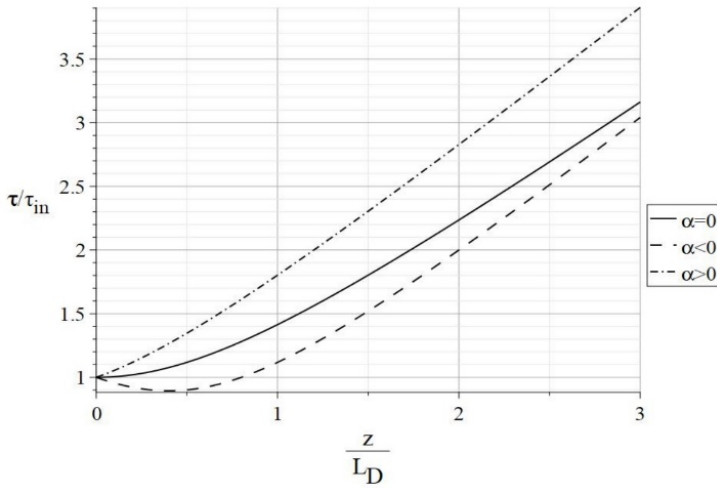
In this situation, the pulse increases its duration from the beginning. The pulse duration increases faster than in the case where no initial chirp is present in the pulse (Fourier-transform-limited pulse).

- $\text{Sign}(g \cdot \alpha) < 0$

When this condition is met, the second term becomes negative and the pulse begins decreasing its duration when it enters the crystal. The reason for this behaviour is that the frequency components of the pulse, initially out of phase, begin coming into phase due to the effect of the material dispersion.

To compress an initially up-chirped pulse, we need to propagate in the anomalous GVD regime, while the opposite happens for an initially down-chirped pulse.

The figure shows the evolution of the pulse duration in a normal dispersive medium ($g > 0$) for zero, positive, and negative initial chirp.



b) To obtain the initial TBP:

$$TBP_{in} = \tau_{in} \cdot \Delta\Omega$$

We have to calculate the pulse bandwidth:

$$\tilde{E}_o^{(+)}(0, \Omega) = \int_{-\infty}^{\infty} E_o e^{-\frac{t^2(1+i\alpha)}{\tau_{in}^2}} e^{i\Omega t} dt = \frac{E_o \tau_{in} \sqrt{\pi}}{\sqrt{1+i\alpha}} e^{-\frac{\Omega^2 \tau_{in}^2}{4(1+i\alpha)}} = \frac{E_o \tau_{in} \sqrt{\pi}}{(1+\alpha^2)^{1/4}} e^{-\frac{\Omega^2 \tau_{in}^2}{4(1+\alpha^2)}} e^{i \left(\frac{\Omega^2 \tau_{in}^2 \alpha}{4(1+\alpha^2)} \frac{\arctan(\alpha)}{2} \right)}$$

where we use $\sqrt{1+i\alpha} = (1+\alpha^2)^{1/4} e^{i \frac{\arctan(\alpha)}{2}}$

From this expression, we obtain the frequency bandwidth: $\Delta\Omega = \frac{2\sqrt{1+\alpha^2}}{\tau_{in}}$

and $TBP_{in} = 2\sqrt{1+\alpha^2}$.

c) To calculate the distance of maximum compression (minimum pulse duration), we assume that we are under the proper initial conditions, i.e., $\text{sign}(g \cdot \alpha) < 0$, and the pulse duration equation is

$$\tau(z) = \tau_{in} \sqrt{1 - \frac{2|\alpha|z}{|L_D|} + \frac{z^2(1+\alpha^2)}{L_D^2}}$$

The minimum is found at the position

$$\frac{d\tau(z)}{dz} = 0 \rightarrow \frac{d(\tau^2(z)/\tau_{in}^2)}{dz} = 0$$



$$-\frac{2|\alpha|}{|L_D|} + \frac{2z_{\min}(1+\alpha^2)}{L_D^2} = 0 \rightarrow z_{\min} = \frac{|\alpha| \cdot |L_D|}{1+\alpha^2} = \frac{|\alpha| \cdot \tau_{in}^2}{2|g|(1+\alpha^2)}$$

The minimum pulse duration corresponds to the pulse duration at this position inside the crystal:

$$\tau(z_{\min}) = \tau_{in} \sqrt{1 - \frac{2|\alpha|z_{\min}}{|L_D|} + \frac{z_{\min}^2(1+\alpha^2)}{L_D^2}} = \frac{\tau_{in}}{\sqrt{1+\alpha^2}}$$

The compression factor is

$$\eta = \frac{\tau_{\min}}{\tau_{in}} = \frac{1}{\sqrt{1+\alpha^2}}$$

The TBP at this point is $TBP_{\min} = \tau_{\min} \cdot \Delta\Omega = \frac{\tau_{in}}{\sqrt{1+\alpha^2}} \cdot \frac{2\sqrt{1+\alpha^2}}{\tau_{in}} = 2$

This indicates that the pulse at this position is Fourier-transform-limited, and it has the minimum pulse duration possible for the given bandwidth.

Further compression should be possible, for instance, by increasing the pulse bandwidth. This is possible using nonlinear interactions in the process called nonlinear phase modulation.

■

S5.5 The amplitude of a Fourier transformed limited Gaussian pulse with duration τ_o ($1/e$ in amplitude) is given by the expression $E^{(+)}(t) = E_o e^{-t^2/\tau_o^2}$.

The intensity autocorrelation of the pulse is defined as $I_{AC}(\tau) = \int_{-\infty}^{\infty} I(t)I(t-\tau)dt$.

Show that, for the given pulse, this function has a Gaussian profile and obtain the value of its full width at half-maximum (FWHM) duration.

Solution

The pulse intensity is proportional to the square of the electric field amplitude:

$$I(t) = I_o e^{-\frac{2t^2}{\tau_o^2}}$$

where I_o is the peak intensity.

Directly substituting the intensity in the autocorrelation integral, we have



$$I_{AC}(\tau) = \int_{-\infty}^{\infty} I_o e^{-\frac{2t^2}{\tau_o^2}} \cdot I_o e^{-\frac{2(t-\tau)^2}{\tau_o^2}} dt = I_o^2 e^{-\frac{2\tau^2}{\tau_o^2}} \int_{-\infty}^{\infty} e^{-\frac{2t^2}{\tau_o^2}} e^{-\frac{2t^2}{\tau_o^2}} e^{\frac{4t\tau}{\tau_o^2}} dt = I_o^2 e^{-\frac{2\tau^2}{\tau_o^2}} \int_{-\infty}^{\infty} e^{-\frac{4t^2}{\tau_o^2}} e^{\frac{4t\tau}{\tau_o^2}} dt$$

Solving the integral:

$$I_{AC}(\tau) = I_o^2 e^{-\frac{2\tau^2}{\tau_o^2}} \frac{\tau_o \sqrt{\pi}}{2} e^{\frac{4\tau^2/\tau_o^4}{4/\tau_o^2}} = \frac{I_o^2 \tau_o \sqrt{\pi}}{2} e^{-\frac{\tau^2}{\tau_o^2}} = I_{ACo} e^{-\frac{\tau^2}{\tau_o^2}}$$

The autocorrelation function also shows a Gaussian profile, which is $\sqrt{2}$ times broader than the intensity profile.

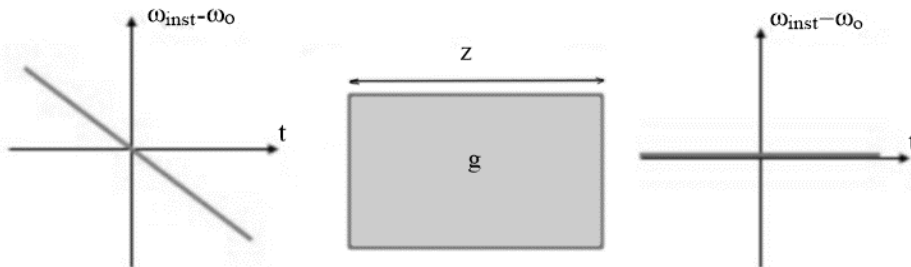
To obtain the FWHM value, we look for the value of τ where the autocorrelation takes the value $I_{ACo}/2$:

$$\frac{I_{ACo}}{2} = I_{ACo} e^{-\frac{\tau_{1/2}^2}{\tau_o^2}} \rightarrow \ln 2 = \frac{\tau_{1/2}^2}{\tau_o^2} \rightarrow \tau_{1/2} = \tau_o \sqrt{\ln 2}$$

and $\tau_{AC}^{FWHM} = 2\tau_{1/2} = 2\tau_o \sqrt{\ln 2}$.

■

S5.6 A Gaussian pulse propagates in a dispersive material with GVD coefficient g at the carrier frequency of the pulse. The plot shows the instantaneous frequency at the entrance and at the output of the crystal.



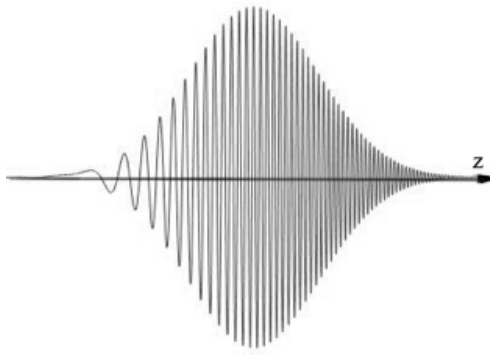
- Explain what the spatial appearance will be for the pulse entering the crystal.
- What is the sign of the GVD coefficient in this material?
- The frequency bandwidth of the spectrum (1/e in amplitude) is $\Delta\Omega = 7 \cdot 10^{13}$ rad/s, and the pulse duration at the entrance of the crystal is 50 fs. What is the pulse duration (1/e in amplitude) at the output of the crystal and what is the value of the initial chirp parameter, α ?



Solution

a) As we can see from the plots, the instantaneous frequency at the entrance of the crystal decreases with time. This indicates that our initial pulse is initially down-chirped. A down-chirped pulse can be represented either as a function of time or as a function of space. Since the time is measured with respect to the location of the maximum, negative times correspond to the part of the beam occurring previous to the maximum and, hence, it is spatially located at the front of the beam. The temporal and spatial representations are then inverted, and one should always pay attention to what is represented on the horizontal axis when using these kinds of plots.

A down-chirped pulse represented in space:



b) As we can see from the plot, the incident pulse is down-chirped. However, at the output, the instantaneous frequency vanishes, indicating that it has no chirp. Since the initial chirp has been compensated by the material, it should introduce a positive chirp. This corresponds to a normal GVD and, hence, to a $g > 0$.

c) The output pulse is Fourier-transform-limited, and this means that its TBP has a well-defined value. If we take the $(1/e)$ in amplitude criterion for the measurement of the pulse duration and bandwidth, we have

$$TBP_{out} = \tau_o \cdot \Delta\Omega = 2$$

$$\tau_o = \frac{2}{\Delta\Omega} = 28.57 \cdot 10^{-15} \text{ s} = 28.57 \text{ fs}$$

To obtain the chirp parameter, we will use the fact that the TBP at the entrance corresponds to a down-chirped pulse, and it should be

$$TBP_{in} = \tau_{in} \cdot \Delta\Omega = 2\sqrt{1 + \alpha_{in}^2}$$

Since the pulse is down-chirped the chirp parameter is negative, we therefore have

$$\alpha_{in} = -\sqrt{\frac{\tau_{in}^2 \cdot \Delta\Omega^2}{4} - 1} = -1.436 \quad \blacksquare$$



S5.7 A square light pulse of duration τ_o and carrier wavelength ω_o is represented by the function $E^{(+)}(t) = E_o^{(+)} e^{-i\omega_o t} = E_o \text{rect}(t/\tau_o) e^{-i\omega_o t}$,

$$\text{where the function } \text{rect}(x/a) = \begin{cases} 1 & \text{for } |x| \leq a/2 \\ 0 & \text{for } |x| > a/2 \end{cases}$$

Calculate the spectrum of the pulse $S(\Omega)$ and its time bandwidth product, TBP.

Solution

The spectral amplitude of the pulse is obtained directly from the temporal field envelope:

$$\begin{aligned} \tilde{E}^{(+)}(\Omega) &= \int_{-\infty}^{\infty} E_o^{(+)}(t) e^{i\Omega t} dt = \int_{-\tau_o/2}^{\tau_o/2} E_o e^{i\Omega t} dt \\ \tilde{E}^{(+)}(\Omega) &= \frac{E_o}{i\Omega} \left[e^{i\frac{\Omega\tau_o}{2}} - e^{-i\frac{\Omega\tau_o}{2}} \right] = \frac{2E_o \sin\left(\frac{\tau_o\Omega}{2}\right)}{\Omega} \\ \tilde{E}^{(+)}(\Omega) &= \tau_o E_o \frac{\sin(\tau_o\Omega/2)}{\tau_o\Omega/2} = \tau_o E_o \text{sinc}\left(\frac{\tau_o\Omega}{2}\right) \end{aligned}$$

The spectrum is

$$S(\Omega) = |\tilde{E}(\Omega)|^2 = \tau_o^2 E_o^2 \text{sinc}^2(\tau_o\Omega/2)$$

The zeros of the *sinc* function are found at $\frac{\tau_o\Omega_m}{2} = \pm m\pi \rightarrow \Omega_m = \pm \frac{2m\pi}{\tau_o}$

By defining the bandwidth as the full width of the central peak, which is defined as the distance between the zeros at π and $-\pi$, we get

$$\Delta\Omega = \Omega_1 - \Omega_{-1} = \frac{4\pi}{\tau_o}$$

With this definition, we get $TBP = \tau_o \cdot \Delta\Omega = 4\pi$

Another possible definition for the bandwidth definition is to use the FWHM.

Since the function $\text{sinc}^2(x) = 1/2$ at $x = 0.443\pi$, we get $\Delta x_{FWHM} = 0.886\pi$

$$\text{In our case, } \Delta\Omega_{FWHM} = \frac{2 \cdot 0.886 \cdot \pi}{\tau_o} = \frac{1.772 \cdot \pi}{\tau_o}$$

and $TBP = \tau_o \cdot \Delta\Omega_{FWHM} = 1.77\pi$ ■



S5.8 A femtosecond laser system emits a train of pulses at a repetition rate of $f_{rep} = 10$ KHz. The averaged power emitted by this laser is measured with a slow detector, giving $P_{av} = 500$ mW. Each individual pulse may be considered to have a Gaussian beam profile with a pulse duration ($1/e$ in amplitude) of $\tau_o = 50$ fs.

a) Calculate the energy per pulse.

b) The spatial profile of the beam is top-hat, with a constant spatial intensity distribution over a radius $w_o = 1$ mm. Obtain the expression for the instantaneous power $P(t)$ and the peak power per pulse.

c) Calculate the peak intensity of the pulses emitted by this laser.

Solution

a) To calculate the energy per pulse, we consider that the averaged power measures the equivalent energy emitted by a CW source in 1s. Since this energy should be distributed among the number of pulses given by the repetition rate, the energy per pulse is given by the relation

$$E_{pulse} = \frac{P_{av}}{f_{rep}} = 50 \text{ } \mu\text{J} / \text{pulse}$$

b) Since the beam has a Gaussian temporal dependence and a constant spatial distribution, we can write an expression for the intensity (energy distribution per unit area and time) as a function of position and time:

$$I(r, t) = \begin{cases} I_o e^{-\frac{2r^2}{w_o^2}} & \text{for } r \leq w_o \\ 0 & \text{for } r > w_o \end{cases}$$

The instantaneous power is obtained directly from the intensity distribution:

$$P(t) = \int_0^\infty I(r, t) 2\pi r dr = \int_0^{w_o} I_o e^{-\frac{2r^2}{w_o^2}} 2\pi r dr = I_o \pi w_o^2 e^{-\frac{2r^2}{w_o^2}}$$

The peak power corresponds to its maximum value: $P_o = \pi w_o^2 I_o$.

c) The temporal integration of the instantaneous power function gives the total energy per pulse:

$$E = \int_{-\infty}^\infty P(t) dt = \int_{-\infty}^\infty \pi w_o^2 I_o e^{-\frac{2t^2}{\tau_o^2}} dt = \pi w_o^2 I_o \frac{\tau_o \sqrt{\pi}}{\sqrt{2}}$$

where we use the Gaussian integral $\int_{-\infty}^\infty e^{-\alpha t^2} dt = \sqrt{\frac{\pi}{\alpha}}$.



The peak intensity can be obtained from this expression:

$$I_o = \frac{E\sqrt{2}}{\pi^{3/2}w_o^2\tau_o} = 2.54 \cdot 10^{14} \text{ W / m}^2 = 25.4 \text{ GW / cm}^2$$

Using this value, the peak power is $P_o = \pi w_o^2 I_o = 0.798 \text{ GW}$.

■



5.6 Proposed problems

P5.1 A laser source emits Fourier-transform-limited Gaussian pulses of 50 fs pulse duration ($1/e$ in amplitude) at the wavelength $\lambda = 800$ nm. These pulses propagate in an SBN crystal that is 2 cm long. The GVD coefficient of this crystal at 800 nm is $g = 439$ fs²/mm.

- a) Obtain the spectral bandwidths $\Delta\Omega$ and $\Delta\lambda$ of the pulse entering the crystal.
- b) Calculate the pulse duration and the chirp coefficient at the output of the crystal.

P5.2 A Gaussian pulse of initial pulse duration $\tau_o = 100$ fs, and a wavelength of 800 nm propagates in fused silica.

- a) If the pulse is initially Fourier-transform-limited, find the pulse duration and the chirp coefficient, C , after a distance of $z_1 = 5$ mm, $z_2 = 05$ cm, and $z_3 = 5$ m.
- b) Obtain the pulse spectrum.
- c) Calculate the TBP at the distances specified in a).

P5.3 An unchirped pulse of central wavelength 1550 nm and duration 10 ps is transmitted through a fused-silica optical fibre.

- a) Determine the group velocity dispersion coefficient and group velocity dispersion length (use the Sellmeier equation in problem S5.1).
- b) Determine the pulse duration and chirp parameter at a distance $L = 100$ km.
- c) If the pulse is to be compressed back to the original width of 10 ps using another fibre made of some material exhibiting normal dispersion with a group velocity dispersion coefficient $g = 127.45$ fs²/mm, determine the length of the fibre we should use.

P5.4 A Gaussian pulse propagates in a dispersive medium with initial chirp α . The sign of the product of the GVD coefficient and initial chirp is negative ($\text{sign}(g \cdot \alpha) < 0$), indicating that the pulse will be initially compressed. If we know the distance at which the pulse duration will be minimum and the value of the minimum pulse duration, complete the following.

- a) Determine the value of the chirp coefficient of the pulse at the entrance of the crystal.
- b) Calculate the pulse duration and the spectrum bandwidth at the entrance of the crystal.

P5.5 A Gaussian pulse of pulse duration at $1/e$ in amplitude τ_o and with zero initial phase propagates in a medium with GVD coefficient g . Obtain the expression of the temporal pulse amplitude and intensity at a distance z by using the impulse response function formalism, as defined in Equation [5.16].



P5.6 The electric field of a Fourier-transform-limited Gaussian pulse is described by the expression $E_o^{(+)}(t) = E_o e^{-\frac{t^2}{\tau_o^2}}$.

a) Express the electric field, $E_o^{(+)}(t)$, the intensity, $I(t)$, and the spectral amplitude, $\tilde{E}_o^{(+)}(\Omega)$, in terms of the full-width-half-maximum pulse duration in the intensity parameter τ .

b) Calculate the TBP for this pulse, expressed in terms of the full-width-half-maximum (FWHM) pulse duration and the corresponding FWHM spectrum bandwidth.

P5.7 Calculate the pulse duration (expressed as the FWHM in intensity) and the frequency spectrum bandwidth (at FWHM), which are associated with pulses that have an angular spectrum bandwidth (FWHM) of $\Delta\lambda = 5nm$, $\Delta\lambda = 30nm$ and $\Delta\lambda = 100nm$.

P5.8 A femtosecond laser system emits a train of pulses at a repetition rate of $f_{rep} = 76$ MHz. The averaged power emitted by this laser is measured with a slow detector, giving $P_{av} = 1.6$ W. Each individual pulse may be considered to have a Gaussian beam profile with a pulse duration ($1/e$ in amplitude) of $\tau_o = 110$ fs.

a) Calculate the energy per pulse.

b) The spatial profile of the beam is Gaussian, with a radius ($1/e$ in amplitude) of $w_o = 1$ mm. We consider that this beam radius remains constant over the pulse duration (this assumption implies that the spatial extent of our pulse along the propagation axis is much shorter than the Rayleigh range associated with the Gaussian beam's transverse dimensions). Obtain the expression for the instantaneous power $P(t)$ and the peak power per pulse.

c) Calculate the peak intensity of the pulses emitted by this laser.



→6



Basic diffraction phenomena

Introduction

The more basic solutions to the wave equation, i.e., plane waves and spherical waves are mathematical representations that are rarely found in real optical beams. This is a consequence of the multi-emission character of optical radiation from real light sources. Thus, the optical wavefronts are modified by the propagation in nonhomogeneous media, the reflection from non-perfectly flat and reflecting surfaces, and many other phenomena. Propagation of these optical wavefronts in free space modifies them further in a non-trivial way. In order to control the optical beams in any applied device, it is fundamental to understand how the beam propagates across the optical system. In some situations, taking a geometrical optics approach can be an effective tool for performing this task. However, in most situations, a more rigorous treatment is needed. The more complex wavefronts can be obtained as the coherent superposition of plane monochromatic waves, and the study of these phenomena forms the basis of Fourier optics. In this chapter, we will consider in further detail the effects of the superposition of monochromatic plane waves (and also wavelets, as we will see), thus leading to beam solutions localized transversally to the propagation direction. The basic physical effect leading to changes in the optical wavefronts during propagation is called **diffraction**.

Diffraction involves studying the evolution of the spatial optical wavefronts in generic optical systems, thus imposing limitations on the final resolution of these systems. In this chapter, we will introduce the basic concepts of the scalar theory of diffraction in the Fresnel and Fraunhofer approximations.

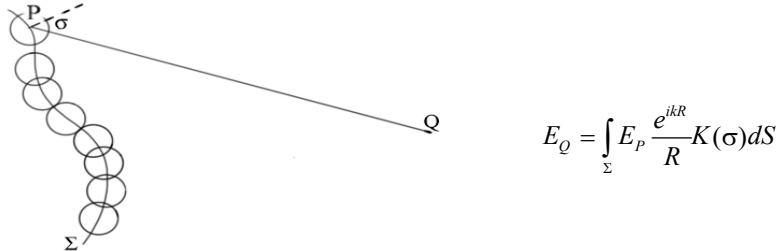
The process of diffraction can be studied from two different perspectives, for which we have the following two distinct formulations.

1. Kirchhoff–Rayleigh–Sommerfeld integral equations for the propagation of optical beams

The basic postulate is **Fresnel's superposition principle**: each point of an optical wavefront at a particular instant of time acts as a secondary source of a spherical wave-

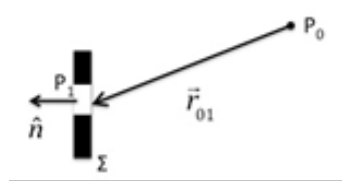


front of the same frequency (wavelet). The amplitude of the optical field at any point located after the wavefront will be the result of the superposition of all previous wavelets when considering their amplitude and relative phases.



The term $K(\sigma)$ is called the obliquity factor, and it takes into account the propagation at a given angle.

The general theory based on this principle is given by the scalar theory of diffraction, developed by Kirchhoff, Rayleigh and Sommerfeld. The main result of this theory gives the optical disturbance at a given point in space, P_0 , which is caused by the optical distribution at a previous plane $E(P_1) = U(P_1)\exp(-i\omega t)$:



$$U(P_0) = \frac{1}{i\lambda_m} \iint_{\Sigma} U(P_1) \frac{\exp(ikr_{01})}{r_{01}} \cos(\hat{n} \cdot \hat{r}_{01}) dS \quad [6.1]$$

where λ_m is the wavelength in the medium, and the different directions are explained in the figure. This expression is the mathematical formulation of Fresnel's theorem.

2. Angular spectrum representation

An alternative way to study diffraction effects is to represent the angular spectrum using the plane wave expansion of the optical fields and the concept of spatial frequency.

This approach is very similar to the one adopted in the treatment of optical pulse propagation. Hence, it is more suitable to this course. The following sections of this chapter will elaborate on the formulation of these concepts. Our intention here is to introduce the basic concepts of diffraction theory and Fourier optics (a detailed discussion on this subject can be found in the book Goodman, J.W. (1998). *Introduction to Fourier Optics*. Mc Graw Hill).



6.1 Angular spectrum representation of optical beams

As we have seen in Chapter 2, the basic equation governing the propagation of a monochromatic optical beam in isotropic media is the Helmholtz equation:

$$\vec{\nabla} \wedge \vec{\nabla} \wedge \vec{E} - k^2 \vec{E} = 0 \quad \xrightarrow{\vec{\nabla} \cdot \vec{E} = 0} \quad \nabla^2 \vec{E} + k^2 \vec{E} = 0 \quad [6.2]$$

Let us consider the propagation of a monochromatic optical beam along a given direction, considered to be the OZ axis. Assuming that the light disturbance is generated by sources placed at the spatial region $z < 0$, we want to determine the evolution of the optical wavefront in the half-space $z > 0$. In particular, we will determine the relationship between the beam profiles at different z planes.

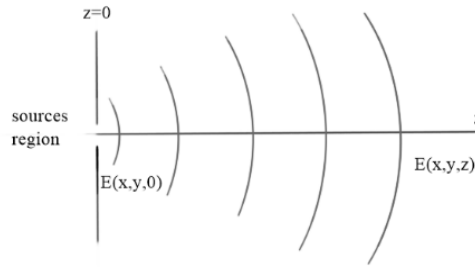


Figure 6.1
Field propagation in
the $z > 0$ region

The angular spectrum representation is based on a decomposition of the light disturbance at a given z plane in terms of plane waves. In order to find this relation, let us assume that our optical beam allows for a 2D Fourier decomposition in the transverse plane:

$$\begin{aligned} \vec{E}(\vec{r}, t) &= \hat{e} E(x, y, z) e^{-i\omega t} \\ E(x, y, z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{E}(f_x, f_y, z) e^{i2\pi(f_x x + f_y y)} df_x df_y \end{aligned} \quad [6.3]$$

where f_x and f_y are the spatial frequencies associated with the wave field.

(!) In order to simplify the notation, we will omit in this chapter the (+) superscript in the fields indicating the complex representation; neither will we consider polarization effects. In the rest of this chapter, we will refer only to the spatial dependence of the amplitude $E(x, y, z)$, assuming that the total field can be written directly from it as $\vec{E}(\vec{r}, t) = \hat{e} E(x, y, z) e^{-i\omega t}$.

Directly substituting the expression [6.3] into the Helmholtz Equation [6.2] gives us

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\frac{\partial^2 \tilde{E}(f_x, f_y, z)}{\partial z^2} - (2\pi)^2 (f_x^2 + f_y^2) \tilde{E}(f_x, f_y, z) + k^2 \tilde{E}(f_x, f_y, z) \right] e^{i2\pi(f_x x + f_y y)} df_x df_y = 0$$

This relation holds for any arbitrary light field if the term in the bracket vanishes. We write this relation in the form



$$\frac{\partial^2 \tilde{E}(f_x, f_y, z)}{\partial z^2} + w^2 \tilde{E}(f_x, f_y, z) = 0 \quad [6.4A]$$

with $w^2 = k^2 - (2\pi)^2(f_x^2 + f_y^2)$.

The value of w can be real or complex, depending on the values of its spatial frequencies:

$$w = \begin{cases} \sqrt{k^2 - (2\pi)^2(f_x^2 + f_y^2)} & \rightarrow (f_x^2 + f_y^2) \leq \left(\frac{k}{2\pi}\right)^2 \\ i\sqrt{(2\pi)^2(f_x^2 + f_y^2) - k^2} & \rightarrow (f_x^2 + f_y^2) > \left(\frac{k}{2\pi}\right)^2 \end{cases} \quad [6.4B]$$

The general solution to Equation [6.4A] is

$$\tilde{E}(f_x, f_y, z) = a(f_x, f_y)e^{iwz} + b(f_x, f_y)e^{-iwz}$$

In order to represent physical solutions to the light field, we need to consider the geometry of our problem. If we seek solutions that are valid in the whole semi-infinite $z > 0$ plane (propagating to the right), then $b(f_x, f_y) = 0$ (since, for $z > 0$, the exponential term would increase exponentially when w becomes imaginary). For valid solutions propagating to the left in the semi-infinite plane $z > 0$ (in this case, we should consider the sources placed at $z > 0$), we need $a(f_x, f_y) = 0$ instead. The whole solution would correspond to radiation confined between two planes in which we have a superposition of light propagating in opposite directions.

We will consider the solutions propagating in the semi-infinite plane $z > 0$ and, hence, $b(f_x, f_y) = 0$. After writing w in slightly different form, we have

$$w = 2\pi m \rightarrow m = \begin{cases} \sqrt{\frac{1}{\lambda_m^2} - (f_x^2 + f_y^2)} & \rightarrow (f_x^2 + f_y^2) \leq \frac{1}{\lambda_m^2} \\ i\sqrt{(f_x^2 + f_y^2) - \frac{1}{\lambda_m^2}} & \rightarrow (f_x^2 + f_y^2) > \frac{1}{\lambda_m^2} \end{cases} \quad [6.4C]$$

We obtain a solution for the propagated field as:

$$E(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{E}(f_x, f_y, z) e^{i2\pi(f_x x + f_y y)} df_x df_y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(f_x, f_y) e^{i2\pi m z} e^{i2\pi(f_x x + f_y y)} df_x df_y \quad [6.5]$$

The expression can be inverted to obtain, for $z = 0$,

$$a(f_x, f_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x, y, 0) e^{-i2\pi(f_x x + f_y y)} dx dy = \tilde{E}(f_x, f_y, 0) \quad [6.6]$$

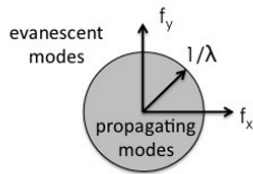


From this expression, we can see that $a(f_x, f_y)$ represents the 2D Fourier transform of the light field at $z = 0$ and, consequently, it carries the information of the spatial frequency components that were initially present in our beam (the initial spatial angular spectrum of the light field).

The effect of the propagation is thus included in the exponential factor $\exp(i2\pi mz)$. This factor decreases exponentially for high enough frequencies when $(f_x^2 + f_y^2) > (1/\lambda^2)$. These frequencies are called **evanescent modes**, and they will not contribute to the field at $z > 0$, thereby representing a loss in resolution of the original light field during free space propagation.

We can then understand the free-space propagation of optical light fields in the following way.

- Any light field distribution at a given plane $z = 0$ can be decomposed into a set of spatial frequency components (its initial spatial angular spectrum), $a(f_x, f_y)$.
- Each propagating spatial frequency component evolves according to Equation [6.4A], acquiring a phase $\exp(i2\pi mz)$, which increases with propagation distance. This phase is different for every spatial frequency, so the different components will go out of phase during propagation, thus leading to a transverse broadening of the initial profile, which we call diffraction.
- The higher frequencies are exponentially reduced and do not contribute to the formation of the new wavefront, thus leading to a loss of resolution. The evanescent frequencies are those satisfying the relation



$$f_x^2 + f_y^2 > \frac{1}{\lambda_m^2}$$

The consequence of this limit is that we cannot use free-space propagation to propagate the details of any image smaller than the wavelength over distances larger than the wavelength itself. Free space propagation behaves as a low-pass spatial frequency filter.

◆ *Comment*

The imaging techniques at distances shorter than a few wavelengths allow retrieving images at resolutions of a few nanometres, which are far beyond the diffraction limit, but these techniques allow capturing the information from the evanescent modes when capturing the image. This forms the basis of near-field optics, a very active research field in photonics.



The propagation of the light field can also be formulated in terms of an impulse response function (or field propagator) if we insert relation [6.6] into [6.5]:

$$E(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x', y', 0) e^{-i2\pi(f_x x' + f_y y')} dx' dy' \right\} e^{i2\pi(f_x x + f_y y + mz)} df_x df_y$$

Inverting the order of integration:

$$E(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x', y', 0) \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i2\pi m z} e^{i2\pi(f_x(x-x') + f_y(y-y'))} df_x df_y \right\} dx' dy' \quad [6.7]$$

$$E(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x-x', y-y', z) E(x', y', 0) dx' dy'$$

The **impulse response function** for free-space propagation is defined as

$$h(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i2\pi m z} e^{i2\pi(f_x x + f_y y)} df_x df_y \quad [6.8]$$

6.1.1 The angular representation as a superposition of plane waves

We have formulated the problem of diffraction in terms of spatial frequencies, whose physical meaning we will consider in a later section. The same problem can be viewed as a superposition of a collection of plane waves.

Consider a plane wave with unit amplitude propagating in an arbitrary direction, and its wave vector is indicated by a unitary vector expressed in terms of the direction cosines:

$$\hat{u} = (\cos \alpha \hat{i} + \cos \beta \hat{j} + \cos \gamma \hat{k})$$

where α , β and γ are the angles between $\hat{u}$ and the OX, OY and OZ axes, respectively.

Since this vector is unitary, $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$.

The plane wave can be written as

$$U = e^{i\vec{k}\vec{r}} = e^{ik(x\cos\alpha + y\cos\beta + z\cos\gamma)} = e^{i2\pi\left(\frac{\cos\alpha}{\lambda_m}x + \frac{\cos\beta}{\lambda_m}y\right)} e^{i\frac{2\pi}{\lambda_m}\sqrt{1-\cos^2\alpha-\cos^2\beta}z} \quad [6.9]$$

Identifying the spatial frequencies f_x and f_y with

$$f_x = \cos\alpha/\lambda_m \text{ and } f_y = \cos\beta/\lambda_m \quad [6.10]$$

the last term in the expression can be rewritten as

$$\frac{2\pi}{\lambda_m} \sqrt{1-\cos^2\alpha-\cos^2\beta} z = \frac{2\pi}{\lambda_m} \sqrt{1-(\lambda_m f_x)^2 - (\lambda_m f_y)^2} z = 2\pi \sqrt{\frac{1}{\lambda_m^2} - f_x^2 - f_y^2} z = 2\pi m z$$



With these relations, we can interpret Equation [6.5] in terms of plane waves. The electric field in plane z can be seen as composed of a superposition of plane waves with amplitudes $a(f_x, f_y)$. Each spatial frequency relates to a particular propagation angle through expression [6.10].

According to this, the propagated field at z , given by Equation [6.5], can be interpreted as a superposition of plane waves forming angles α and β with respect to the OX and OY axes, respectively. The propagated frequencies are those for which m is real valued, the amplitude of each plane wave being proportional to the Fourier transform of the field at $z = 0$, $a(f_x, f_y)$. Expression [6.5] can be written in terms of these propagation angles:

$$E(x, y, z) = \iint \frac{1}{\lambda_m^2} a(\cos \alpha / \lambda_m, \cos \beta / \lambda_m) e^{i2\pi mz} e^{ik(\cos \alpha x + \cos \beta y)} d \cos \alpha d \cos \beta \quad [6.11]$$

This expression corresponds to the **angular spectrum representation** of the light field.

6.2 The concept of spatial frequency

In the same way that we can think of an optical pulse as being composed of a superposition of different temporal frequencies, any light distribution in a given plane can be understood as a superposition of different spatial frequencies. A spatial frequency f_x can be associated with a pattern with amplitude $\cos(2\pi f_x x)$ and, hence, with spatial periodicity $\Lambda_x = 1/f_x$. Thus, higher spatial frequencies are associated with smaller periods.

The first question we must consider is how to relate this concept with the plane wave propagation at definite angles according to Equation [6.10]. We can view this correspondence in the following two ways.

1. From the first point of view, we can see how different transverse spatial frequencies can be obtained by the superposition of plane waves of a given wavelength λ_m .

Consider two plane waves propagating in the xz -plane at angles $+\alpha$ and $\pi-\alpha$ with respect to the OX axis. The total field is written as

$$E_{TOT}^{(+)} = E_o e^{i(k \cos \alpha x + \sin \alpha z)} + E_o e^{i(-k \cos \alpha x + \sin \alpha z)} = 2E_o \cos(2\pi \frac{\cos \alpha}{\lambda_m} x) e^{ik \sin \alpha z}$$

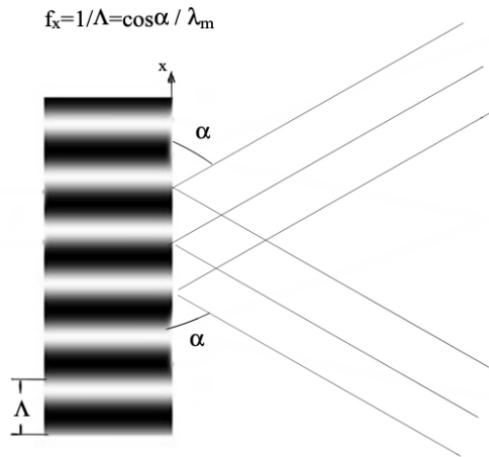
The superposition of these two waves generates a periodic pattern of frequency $f_x = \cos \alpha / \lambda_m$, which corresponds to the previously defined frequency. The smallest period that can be obtained with these waves corresponds to the case $\alpha = 0$, where the two counter-propagating waves generate a pattern with period $f_{x \max} = 1 / \Lambda_{\min} = 1 / \lambda_m$.



In this way, each spatial frequency can be related to a particular propagation angle of the plane wave.

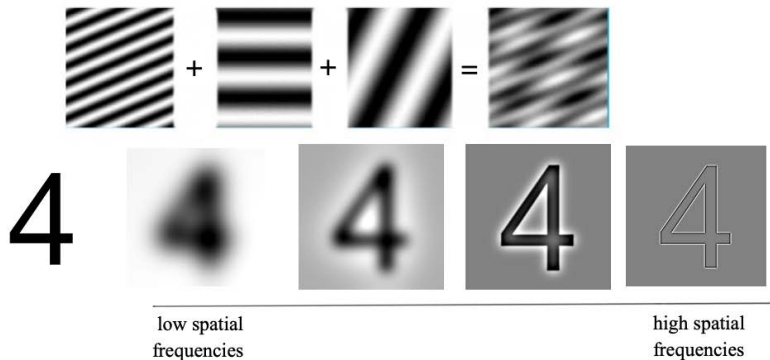
2. From the second point of view, we want to stress the fact that a given periodic pattern can generate the emission of a couple of plane waves at well-defined propagating angles according to relation [6.10]. If we consider a slide with transmittance function $T = \cos(2\pi f_x x)$ and illuminated by a plane wave of wavelength λ_m , the transmitted field will consist of a superposition of two waves emitted at angles α and $\pi - \alpha$, as we will see in the following sections.

Figure 6.2
Relation between spatial frequency and propagation angle of the represented plane wave



Any object with arbitrary transmittance can be obtained as a superposition of periodic cosine patterns with different frequencies. In accordance with this idea, the resulting transmitted field can be understood as having been formed by a superposition of plane waves. The details in the figure are related to the large frequencies, as schematically represented in the following photograph showing the original picture and some filtered images corresponding to low and high frequencies.

Figure 6.3
Frequency content of an image





6.3 The connection between the angular spectrum representation and the scalar Rayleigh–Sommerfeld diffraction theory

In order to establish the connection between the angular spectrum representation (superposition of plane waves) and the Rayleigh–Sommerfeld theory (superposition of spherical wavelets), we start with the angular spectrum representation of an outgoing spherical wave, which we present here without any demonstration:

$$\frac{e^{ikr}}{r} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{i}{m} e^{i2\pi mz} e^{i2\pi(f_x x + f_y y)} df_x df_y \quad [6.12]$$

Partial derivation with respect to z gives us

$$\frac{\partial}{\partial z} \left(\frac{e^{ikr}}{r} \right) = -2\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i2\pi mz} e^{i2\pi(f_x x + f_y y)} df_x df_y = -2\pi h(x, y, z)$$

The impulse response function can be written as

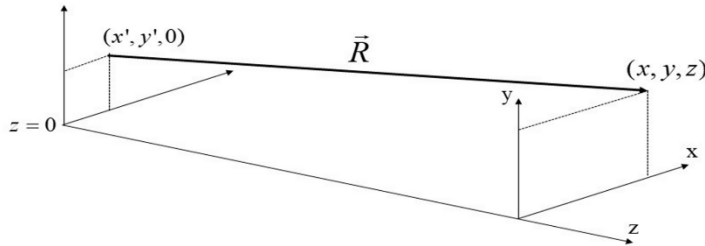
$$h(x, y, z) = -\frac{1}{2\pi} \frac{\partial}{\partial z} \left(\frac{e^{ikr}}{r} \right) = -\frac{1}{2\pi} \frac{\partial}{\partial r} \left[\frac{e^{ikr}}{r} \right] \frac{\partial r}{\partial z}$$

with $r = \sqrt{x^2 + y^2 + z^2}$.

Performing the derivatives:

$$h(x, y, z) = -\frac{1}{2\pi} \left[ik - \frac{1}{r} \right] \frac{z}{r} \frac{e^{ikr}}{r} \quad [6.13]$$

In order to use this expression in [6.7] to obtain $E(x, y, z)$, we have to use the translated function $h(x-x', y-y', z)$, where the coordinates (x', y') correspond to points in the $z = 0$ plane, and (x, y) to points in the observation plane.



with $R = \sqrt{(x-x')^2 + (y-y')^2 + z^2}$.

We obtain the expression for the impulse response function:



$$h(x-x', y-y', z) = -\frac{1}{2\pi} \left[ik - \frac{1}{R} \right] \frac{z}{R} \frac{e^{ikR}}{R}$$

Using this expression, we obtain the general relation:

$$E(x, y, z) = \frac{-1}{2\pi} \int_{-\infty}^{\infty} \int E(x', y', 0) \left[ik - \frac{1}{R} \right] \frac{z}{R} \frac{e^{ikR}}{R} dx' dy'$$

The factor in brackets simplifies when the observation distance is $R \gg \lambda$, $\left[ik - \frac{1}{R} \right] \approx ik$, leading to the **Rayleigh–Sommerfeld diffraction formula**:

$$E(x, y, z) = \frac{1}{i\lambda} \int_{-\infty}^{\infty} \int E(x', y', 0) \frac{z}{R} \frac{e^{ikR}}{R} dx' dy' \quad [6.14]$$

This expression corresponds to the mathematical formulation of Fresnel's superposition principle. The propagated field at z can be obtained from the field distribution at $z = 0$. This formulation is valid in the scalar approximation and can be simplified under some assumptions that lead to diffraction theory in the Fresnel and Fraunhofer approximations. This formulation will be developed further in the next section.

6.4 Diffraction theory in the Fresnel and Fraunhofer approximation

In this section, we will consider the diffraction from a finite aperture (or slide), of characteristic size Σ and illuminated by monochromatic light. The transmission from the aperture, placed at $z = 0$, is given by

$$E(x', y', 0) = E_{inc}(x', y', 0) \cdot t(x', y') = E_{inc}(x', y', 0) |t(x', y')| e^{i\phi(x', y')} \quad [6.15]$$

where $t(x', y')$ is the transmittance function of the screen (or slide), and $E_{inc}(x', y', 0)$ is the field incident on the slide, which can represent variations in either amplitude or phase. For opaque screens, we assume that the transmittance function takes values 1 whenever light is transmitted, and 0 when the screen stops it.

We consider diffraction effects in the paraxial region, where the following conditions hold.

- The observation plane is placed at distances $z > \Sigma$.
- The observation region of interest in the plane z is close to the origin ($x = 0; y = 0$).

Under these conditions, R can be expanded as

$$\begin{aligned} R &= z \sqrt{1 + \frac{(x-x')^2 + (y-y')^2}{z^2}} \approx \\ &\approx z \left(1 + \frac{1}{2} \frac{(x-x')^2 + (y-y')^2}{z^2} - \frac{1}{8} \left[\frac{(x-x')^2 + (y-y')^2}{z^2} \right]^2 + \dots \right) \end{aligned}$$



Under Fresnel's approximation, we retain terms up to the first order for distances $R \approx z$ and up to the second order for phase terms. The impulse response function can be simplified to

$$h(x-x', y-y', z) = \frac{1}{i\lambda} \frac{z}{R} \frac{e^{ikR}}{R} \approx \frac{1}{i\lambda} \frac{e^{ikz}}{z} e^{ik \frac{(x-x')^2 + (y-y')^2}{2z}}$$

With this approximation, we obtain from [6.14]:

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} \int_{-\infty}^{\infty} \int E(x', y', 0) e^{i \frac{k}{2z} [(x-x')^2 + (y-y')^2]} dx' dy'$$

Rewritten after expanding the square brackets to give the Fresnel **diffraction formula**, we have

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} e^{i \frac{k}{2z} (x^2 + y^2)} \int_{-\infty}^{\infty} \int E(x', y', 0) e^{i \frac{k}{2z} (x'^2 + y'^2)} e^{-i \frac{k}{z} (xx' + yy')} dx' dy' \quad [6.16]$$

This is the basic equation for obtaining the propagated wavefront in plane z of a given field distribution at $z = 0$ in the Fresnel approximation, which is extensively used in physical optics. An example of field propagation is shown in Figure 4 where a uniform plane wave illuminates a screen with a slit of width w . At distances close to the aperture, the field profile changes strongly during propagation. This is called the **near field** region. At larger distances, the diffracted pulse acquires a constant profile shape and its dimension increases linearly during propagation. This region is called the **far field** or **Fraunhofer region**.

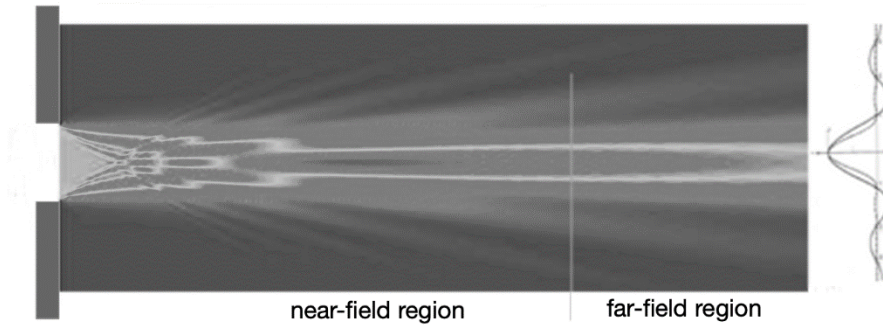


Figure 6.4
Diffraction from a
rectangular aperture

The diffraction pattern in the far field (or Fraunhofer) region is obtained from the Fresnel diffraction formula when an additional condition is imposed.

$$\text{Fraunhofer region: } z \gg \frac{k(x'^2 + y'^2)_{\max}}{2},$$



where $(x'^2 + y'^2)_{\max}$ refers to the maximum extension of the initial field distribution in the $z = 0$ plane.

In this region, the exponential factor $e^{ik\frac{(x'^2 + y'^2)}{2z}}$ in [6.16] can be approximated by 1 and simplifies to **Fraunhofer's diffraction formula**:

$$\boxed{\begin{aligned} E(x, y, z) &= \frac{e^{ikz} e^{i\frac{k(x^2 + y^2)}{2z}}}{i\lambda z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x', y', 0) e^{-i\frac{2\pi}{\lambda z}(xx' + yy')} dx' dy' \\ E(x, y, z) &= \frac{e^{i(kz + \frac{k(x^2 + y^2)}{2z})}}{i\lambda z} \tilde{E}(f_x, f_y, 0) \Big|_{f_x = \frac{x}{\lambda z}; f_y = \frac{y}{\lambda z}} \end{aligned}} \quad [6.17]$$

The diffraction pattern in the Fraunhofer region is proportional to the Fourier transform distribution at the aperture ($z = 0$) plane, where the spatial frequencies correspond to the transformation $f_x = x/\lambda z$ and $f_y = y/\lambda z$. This transformation means that we should calculate the Fourier transform and then substitute the corresponding spatial frequencies f_x and f_y with these transformations in order to obtain the field as a function of (x, y, z) .

Example 1 Diffraction by a rectangular aperture of dimensions $L_x \times L_y$

Consider a rectangular aperture of dimensions L_x and L_y illuminated by a uniform plane wave. The transmittance at the $z = 0$ plane can be written as

$$E(x', y', 0) = \text{rect}\left(\frac{x'}{L_x}\right) \cdot \text{rect}\left(\frac{y'}{L_y}\right)$$

$$\text{with } \text{rect}\left(\frac{x}{L_x}\right) = \begin{cases} 1 & |x| \leq \frac{L_x}{2} \\ 0 & |x| > \frac{L_x}{2} \end{cases}.$$

To obtain the Fraunhofer diffraction pattern [6.17], we calculate the Fourier transform:

$$\tilde{E}(f_x, f_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{rect}\left(\frac{x'}{L_x}\right) \text{rect}\left(\frac{y'}{L_y}\right) e^{-i2\pi(f_x x' + f_y y')} dx' dy' = \int_{-L_x/2}^{L_x/2} e^{-i2\pi f_x x'} dx' \int_{-L_y/2}^{L_y/2} e^{-i2\pi f_y y'} dy'$$

Evaluating the integrals:

$$\tilde{E}(f_x, f_y) = \frac{(e^{-i\pi f_x L_x} - e^{i\pi f_x L_x})}{-i2\pi f_x} \cdot \frac{(e^{-i\pi f_y L_y} - e^{i\pi f_y L_y})}{-i2\pi f_y} = \frac{\sin(\pi f_x L_x)}{\pi f_x} \cdot \frac{\sin(\pi f_y L_y)}{\pi f_y}$$



$$\tilde{E}(f_x, f_y) = L_x L_y \text{sinc}(\pi f_x x) \text{sinc}(\pi f_y y) \quad \text{with} \quad \text{sinc}(a) = \sin(a) / a.$$

The Fraunhofer diffraction pattern:

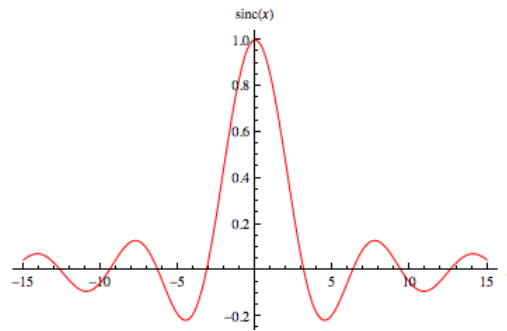
$$E(x, y, z) = \frac{e^{ikz(1 + \frac{x^2 + y^2}{2z^2})}}{i\lambda z} L_x L_y \text{sinc}\left(\frac{\pi L_x x}{\lambda z}\right) \text{sinc}\left(\frac{\pi L_y y}{\lambda z}\right)$$

and the intensity pattern

$$I(x, y, z) = I_o \frac{L_x^2 L_y^2}{\lambda^2 z^2} \text{sinc}^2\left(\frac{\pi L_x x}{\lambda z}\right) \text{sinc}^2\left(\frac{\pi L_y y}{\lambda z}\right)$$

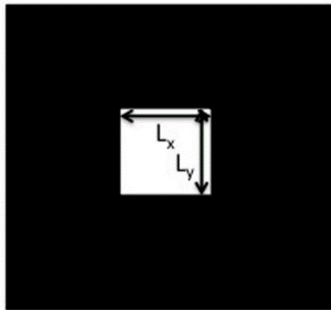
where $I_o = (1/2)\epsilon_0 c n$ for the plane waves illuminating the slide.

The sinc(a) function has zero values at the positions $a = \pm m\pi$ with $m = 1, 2, 3, \dots$, and its maximum value occurs at $a = 0$

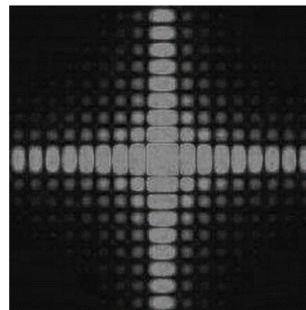


The intensity pattern shows zeros at the positions

$$x_{\text{zeros}} = m \frac{\lambda z}{L_x} \quad y_{\text{zeros}} = m \frac{\lambda z}{L_y}$$



object



Fraunhofer diffraction pattern

Figure 6.5
Diffraction pattern of a
rectangular aperture

**Example 2** *Sinusoidal amplitude grating*

We consider in this example a rectangular aperture with a sinusoidal amplitude transmittance illuminated by a plane wave. The field at $z = 0$ plane is

$$E(x', y', 0) = \left[\frac{1}{2} + \frac{m}{2} \cos(2\pi f_o x') \right] \text{rect} \frac{x'}{L_x} \text{rect} \frac{y'}{L_y}$$

which can be written as the product of two functions:

$$E(x', y', 0) = T(x') \cdot P(x', y')$$

where the function P defines the global size of the object and is called the pupil function. The Fourier transform of this object is given by the convolution of the Fourier transforms:

$$\tilde{E}(f_x, f_y) = [\tilde{T}(f_x) \otimes \tilde{P}(f_x, f_y)]$$

where

$$\begin{aligned} \tilde{T}(f_x) &= \frac{1}{2} \delta(f_x) + \frac{m}{4} \delta(f_x - f_o) + \frac{m}{4} \delta(f_x + f_o) \\ \tilde{P}(f_x, f_y) &= L_x L_y \text{sinc}(\pi f_x L_x) \text{sinc}(\pi f_y L_y) \end{aligned}$$

Performing the convolution, we obtain the Fourier transform for this case:

$$\tilde{E}(f_x, f_y, 0) = \frac{L_x L_y}{2} \text{sinc}(\pi f_y L_y) \left[\text{sinc}(\pi f_x L_x) + \frac{m}{2} \text{sinc}(\pi(f_x - f_o)L_x) + \frac{m}{2} \text{sinc}(\pi(f_x + f_o)L_x) \right]$$

We also obtain the Fraunhofer diffraction pattern:

$$\begin{aligned} E(x, y, z) &= \frac{e^{ikz(1 + \frac{x^2 + y^2}{2z^2})}}{i\lambda z} \cdot \frac{L_x L_y}{2} \text{sinc}\left(\frac{\pi L_y y}{\lambda z}\right) \cdot \\ &\quad \cdot \left\{ \text{sinc}\left(\frac{\pi L_x x}{\lambda z}\right) + \frac{m}{2} \text{sinc}\left[\frac{\pi L_x}{\lambda z}(x - f_o \lambda z)\right] + \frac{m}{2} \text{sinc}\left[\frac{\pi L_x}{\lambda z}(x + f_o \lambda z)\right] \right\} \end{aligned}$$

If the original function is defined without the pupil function (with infinite extension), the Fraunhofer diffraction pattern shows three delta functions at $x = 0$ and $x = \pm f_o \lambda z$. This result seems to be unphysical, due to the singular behaviour of a delta function. However, one can see that when the pupil function is considered, the diffraction pattern of the pupil function appears at the position of the delta functions.

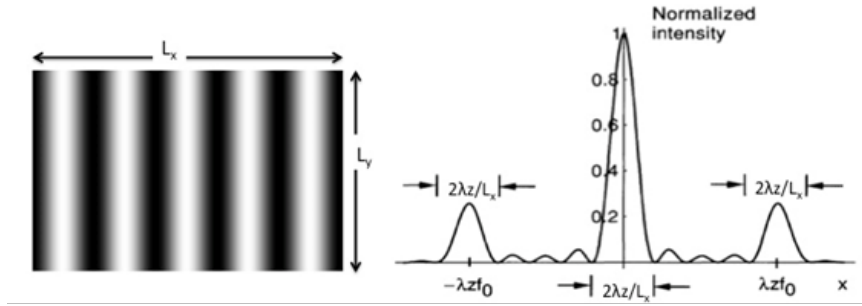


Figure 6.6
Diffraction pattern of a sinusoidal amplitude grating

6.5 Fourier transforms by lenses: Introduction to image processing

Fresnel's diffraction formula is applied whenever we want to find the propagated field at a given distance if the field at $z = 0$ is known, including propagation through different optical elements such as lenses or apertures. A very important result is obtained when we study the propagation of the transmitted field by a slide of transmittance $T(x', y')$ that is illuminated by coherent monochromatic light.

If we consider our lens placed at a distance d' from the object ($z = 0$) plane, the field distribution at the rear focal plane of the lens is given by the following (for which a detailed derivation is presented in Appendix F at the end of the book).

$$E(x, y, d' + f) = A e^{i \frac{k}{2f} (1 - \frac{d'}{f}) (x^2 + y^2)} \tilde{E}(f_x, f_y) \Big|_{f_x = x/\lambda f; f_y = y/\lambda f}$$

From this expression, we see that **the Fourier transform of an object is placed at its rear focal length**. Additionally, if the object is placed at the front focal length, the phase factors cancel and the field at the focal plane is

$$E(x', y', f + f) = A \tilde{E}(f_x, f_y, 0) \Big|_{f_x = x/\lambda f; f_y = y/\lambda f} \quad [6.18]$$

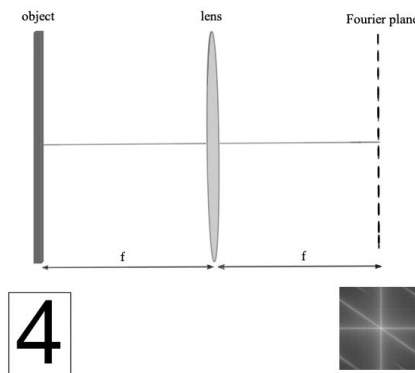


Figure 6.7
Fourier transform by a lens

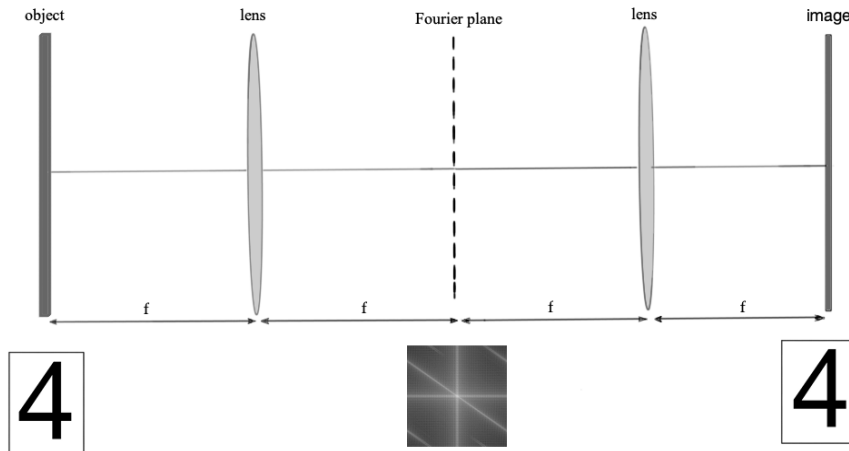


This property of the lenses can be used in many different ways to manipulate or to extract information from a given image by introducing changes in its optical spectrum.

6.5.1 Basic examples of signal processing: The $4f$ correlator

The $4f$ correlator constitutes the most basic signal processing system and consists of a set of two lenses of identical focal length, f , separated by a distance of $2f$. The input plane is at a distance f to the left of the first lens and the output plane is at a distance f to the right of the second lens. We can imagine the operation of this system in the following way. The first lens performs the optical Fourier transform of the object placed at the input plane. The Fourier plane is located at distance f from the first lens. The second lens performs the optical Fourier transform of the field located at the Fourier plane, which will be located at the image plane.

Figure 6.8
 $4f$ correlator

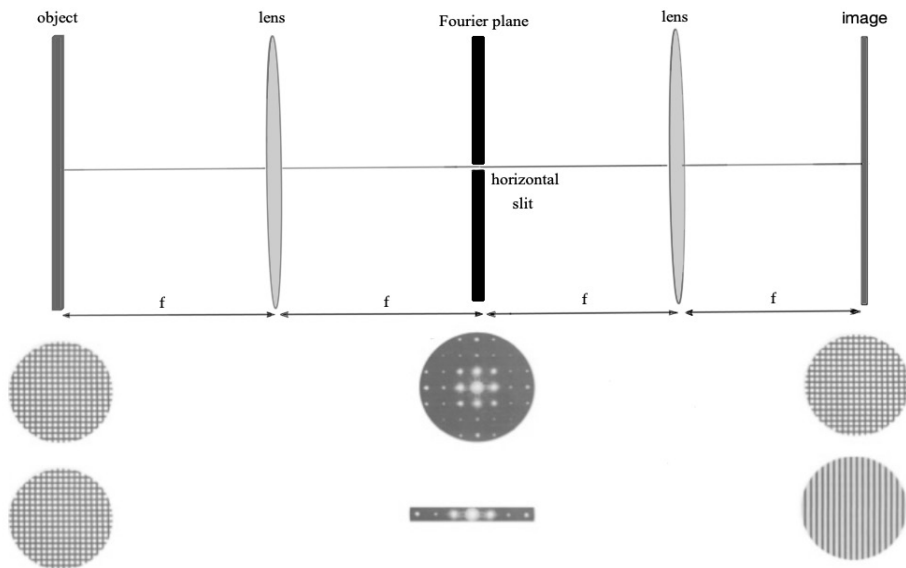


Acting on the spatial spectrum located in the Fourier plane, we can modify the aspect of the image coming out of the system.

Below are some basic examples of signal processing.

1. Spatial filtering

If we eliminate part of the frequency spectrum of the object by placing some masks or obstacles in the Fourier plane, the reconstructed image can be manipulated as shown in the image.



Spatial filtering is a widely used technique for obtaining optical beams with high spatial quality. A laser beam propagated across optical systems acquires high frequency modulations due to the presence of dust or by means of diffraction through small obstacles, thus acquiring a granular aspect known as speckle. For many applications, a uniform spatial beam is desirable, for which the beam can be cleaned by placing a pinhole with the proper diameter at the centre of the Fourier plane of the $4f$ correlator. The pinhole blocks the high frequencies, and the reconstructed beam looks much more homogeneous, as seen in the figure.

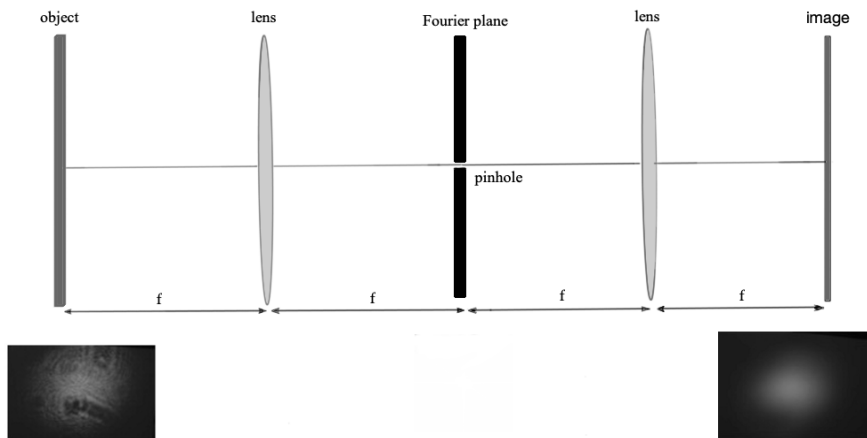


Figure 6.9
Spatial filtering



2. Improvement of image characteristics: Phase-contrast microscopy

Many highly transparent objects have a transmittance function for which the amplitude is almost constant, thus introducing mainly phase variations into the incident field. When these objects are observed through a microscope, the contrast is very low, and they can hardly be seen because the light incident on them is transmitted mainly without variation and very little amount of light is diffracted due to the phase variations. Zernike solved this problem, which once impeded the microscopic observation of microorganisms.

Let us consider a highly transparent object whose transmittance function can be expressed as $t(x,y) = e^{i\phi(x,y)} \approx 1 + i\phi(x,y)$.

The intensity transmitted by the object is hence almost constant if the phase variations are small: $I \approx 1 + \phi^2(x,y) + \dots \approx 1$, and the resulting image shows very poor contrast.

Zernike's technique transforms phase variations into intensity variations by exploiting the concept of image processing. At the Fourier plane, the signal of these kinds of objects has a strong signal at the zero frequency and weak components at other frequencies, given by

$$\tilde{E}(f_x, f_y) = \delta(f_x, f_y) + FT(i\phi)$$

where FT means Fourier transform. By placing a phase retarder with $\delta = \pm\pi/2$ at the zero-frequency position, the resulting field at the Fourier plane is

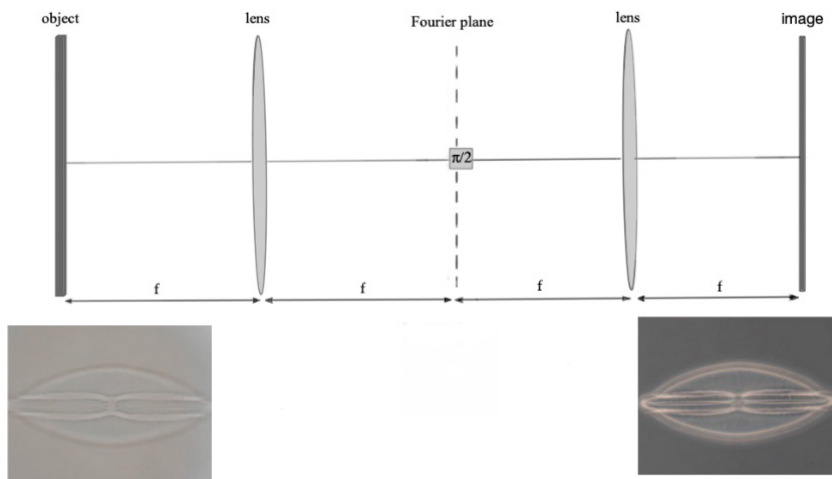
$$\tilde{E}_T(f_x, f_y) = i\delta(f_x, f_y) + FT(i\phi)$$

Which, after propagation to the image plane, results in

$$E_{out}(x, y) = i + i\phi(x, y) \Rightarrow I_{out}(x, y) = 1 + 2\phi(x, y) + \phi^2(x, y)$$

The linear term appearing in the image intensity considerably increases the final image contrast.

Figure 6.10
Phase-contrast
microscopy





3. Frequency domain synthesis

By placing transmittance masks in the Fourier plane, it is possible to perform a whole series of operations on the object of the form:

$$I(x, y) = K \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x', y') \cdot m(x - x', y - y') dx' dy' \right|^2 \quad [6.19]$$

where $I(x, y)$ is the intensity distribution at the image plane, K is a complex constant, $E(x', y')$ is the field distribution in the object plane, and the function $m(x, y)$ takes into account the effect of the mask.

The working principle of this process, called **frequency domain synthesis**, is most easily seen in the Fourier plane.

The field distribution of the object at the Fourier plane is given by

$$K_o \tilde{E}(x / \lambda_m f, y / \lambda_m f)$$

At this plane, we also place our mask, whose transmittance is proportional to the Fourier transform of the mask function $m(x, y)$:

$$K_m \tilde{m}(x / \lambda_m f, y / \lambda_m f)$$

At the output of the Fourier plane, the field distribution is the product of these two fields:

$$\tilde{E}_{\text{Fourier plane}}(x, y) = K_o K_m \tilde{E}(x / \lambda_m f, y / \lambda_m f) \tilde{m}(x / \lambda_m f, y / \lambda_m f)$$

The effect of the second lens on this field distribution results in the intensity given by [6.19] in the image plane.

With this basic idea, different operations can be performed on the object by using the adequate mask. These masks were first implemented by direct recording on photographic plates (Vander Lugt filters), and they can actually be performed by inserting spatial light modulators in the Fourier plane or even directly by using computer-generated masks and digital processing.

One particular example of these techniques is object recognition. If the mask is the Fourier transform of a given object, then Equation [6.19] corresponds to the cross-correlation function. If the incident object coincides with the one used to generate the mask, we obtain a peak in the image plane.

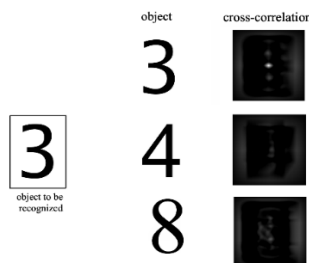


Figure 6.11
Number recognition by
cross-correlation
operation.



6.6 Resolution limit in optical systems

The preceding sections have so far not considered the limited size of any aperture in optical systems, and this course will not treat it in detail. The effect of the finite aperture is introduced through the use of pupil functions.

The transmittance function of the object is thus multiplied by the pupil function, and the corresponding diffraction pattern in the Fraunhofer approximation will be given by the convolution of the Fourier transforms of the object and the pupil function:

$$E_o(x, y) = E(x, y)P(x, y)$$

$$\text{Fraunhofer diffraction pattern: } \tilde{E}_o(x, y) = K\tilde{E}(f_x, f_y) \otimes \tilde{P}(f_x, f_y) \big|_{f_x = x/\lambda_m z; f_y = y/\lambda_m z}$$

Consider the simplest example of an unlimited plane wave with a Fourier transform given by a delta function, i.e., a single point in the Fourier plane. When a circular aperture of radius R blocks the plane wave, the diffraction pattern changes from a single point to show the diffraction pattern of a circular aperture consisting (see Appendix E) of a Bessel function with a central disk diameter of $1.22 \lambda_m f/R$, which is usually called the *Airy function*.

In Example 2 of this chapter, and also explained previously in Section 4, the transmittance function with sinusoidal profile shows a Fraunhofer diffraction pattern consisting of three delta functions placed at $x = 0$ and $x = \pm \lambda_m z f_o$. When the pupil function is added, the result is that the FT of the pupil function becomes convoluted with that of the object and, hence, each delta function is substituted by the FT of the pupil function.

This is a general case in optical systems, where the pupils and stops of the different elements result in diffraction patterns in the propagated fields, which blur the images and thus result in a loss of resolution. Let us try to evaluate the minimum separation between two point sources at a distance z from an imaging system of diameter D so that these sources can be resolved at the image plane. The two-point resolution criterion established by Lord Rayleigh has been used as a quality factor for optical systems with incoherent sources, particularly in astronomical applications. The Rayleigh criterion asserts that these objects will be resolved by the diffraction-limited system when the centre of the Airy disk generated by one source falls on the first zero of the Airy disk generated by the second. The minimum resolvable separation of the geometrical images is therefore

$$s = \frac{0,61\lambda_m z}{R} = \frac{1,22\lambda_m z}{D}$$

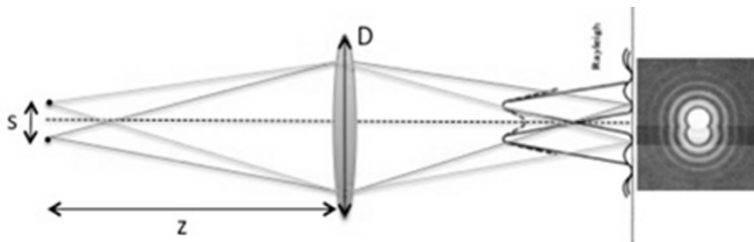


Figure 6.12
Rayleigh criterion's minimum distinguishable separation between two point sources imaged by a lens of diameter D



6.7 Diffraction from an array of periodic elements

In order to define the basic characteristics of the Fraunhofer diffraction pattern in arrays of periodic elements, consider the following functions and basic operations.

The comb function corresponds to a set of delta functions separated at regular intervals:

$$\text{comb}(x) = \sum_{n=-\infty}^{\infty} \delta(x - n)$$

To represent a set of delta functions separated by distance a , we use the scaled comb function:

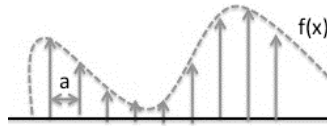
$$\text{comb}(x/a) = \sum_{n=-\infty}^{\infty} \delta\left(\frac{x}{a} - n\right) = \sum_{n=-\infty}^{\infty} \delta\left(\frac{x - na}{a}\right)$$

The Fourier transform of a comb function is also a comb function:

$$FT[\text{comb}(x/a)] = a \cdot \text{comb}(f_x a)$$

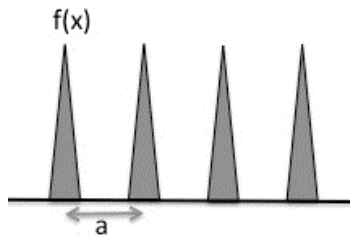
1. **Sampling** is an operation that provides the value of a given function at a regular spacing interval within the period a (value of the function at the positions $x = na$):

$$S(f(x), a) = \frac{1}{a} f(x) \text{comb}(x/a) = \frac{1}{a} f(x) \sum_n \delta\left(\frac{x - na}{a}\right) = \sum_n f(na)$$



2. **Repeat** gives a periodic repetition of a particular function (which we will call the unit cell function) at the positions $x = na$:

$$C(f(x), a) = \frac{1}{a} (f(x) \otimes \text{comb}(x/a)) = \sum_n f(x - na)$$





Any 2D periodic array of unit cell elements can be written in the following way:

$$A(x, y) = \left[g(x, y) \otimes \frac{1}{ab} \text{comb}\left(\frac{x}{a}, \frac{y}{b}\right) \right] \cdot P(x, y)$$

where $g(x, y)$ is the unit cell function, a, b are the periods in the x and y directions, and $P(x, y)$ is the pupil function, which corresponds to the total shape and size of the array.

The Fraunhofer diffraction pattern of such an object will be proportional to its Fourier transform:

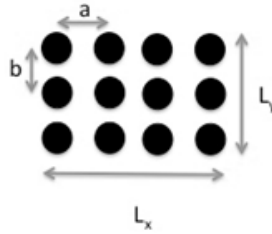
$$E(x, y, z) \propto \left[\tilde{g}(f_x, f_y) \cdot \text{comb}(af_x, bf_y) \right] \otimes \tilde{P}(f_x, f_y) \Big|_{f_x = x/\lambda_m z; f_y = y/\lambda_m z}$$

The diffraction pattern consists of a 2D periodic array of elements with the following characteristics.

- The periodicity of the array is $\lambda_m z/a$ and $\lambda_m z/b$ in the x and y directions, respectively.
- The repeated element is the Fourier transform of the pupil function.
- The intensity of each element is proportional to the value of the Fourier transform of the unit cell function at this particular point.

Example: The pattern shown in the figure is comprised of a set of circular apertures of radius r_0 forming a rectangular pattern of spatial periods a and b in the x and y directions, respectively. The pupil function is a rectangle of sizes L_x and L_y . This periodic array can be represented as

$$E(x, y, 0) = \left[\text{circ}(r/r_0) \otimes \frac{1}{ab} \text{comb}\left(\frac{x}{a}, \frac{y}{b}\right) \right] \text{rect}\left(\frac{x}{L_x}\right) \text{rect}\left(\frac{y}{L_y}\right)$$



The Fraunhofer diffraction pattern:

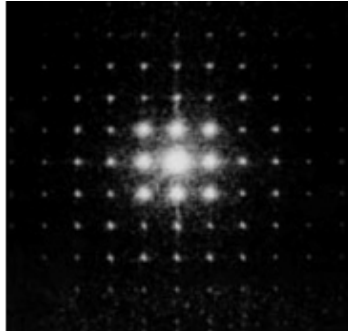
$$E(x, y, z) = \frac{e^{ikz} e^{i\frac{k}{2z}(x^2 + y^2)}}{i\lambda_m z} \left[r_0^2 \frac{J_1(2\pi r r_0 / \lambda_m z)}{r r_0 / \lambda_m z} \cdot \text{comb}(af_x, bf_y) \right] \otimes L_x L_y \text{sinc}(\pi f_x L_x) \text{sinc}(\pi f_y L_y) \Big|_{f_x = x/\lambda_m z; f_y = y/\lambda_m z}$$

which can be written as

$$E(x, y, z) = K \frac{J_1(2\pi r r_0 / \lambda_m z)}{r r_0 / \lambda_m z} \cdot \sum_n \text{sinc} \left(\pi \left(\frac{x}{\lambda_m z} - \frac{n}{a} \right) L_x \right) \sum_m \text{sinc} \left(\pi \left(\frac{y}{\lambda_m z} - \frac{m}{b} \right) L_y \right) =$$

$$K \frac{J_1(2\pi r r_0 / \lambda_m z)}{r r_0 / \lambda_m z} \sum_n \sin c \left(\frac{\pi L_x}{\lambda_m z} \left(x - \frac{\lambda_m z n}{a} \right) \right) \sum_m \sin c \left(\frac{\pi L_y}{\lambda_m z} \left(y - \frac{\lambda_m z m}{b} \right) \right)$$

where K includes all constants. The diffraction pattern is shown in the figure below.



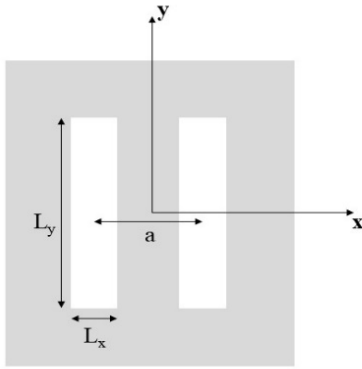


6.8 Solved problems

S6.1 Calculate the Fraunhofer diffraction pattern of a double-slit screen illuminated by a monochromatic plane wave with amplitude $E_o = 1 \text{ V/m}$.

Solution

The two-slit system schematized in the figure is illuminated by a plane wave with uniform intensity.



The field distribution at the object plane is given by the product of the illuminating wave and the transmittance function. In our chosen coordinate system, it is expressed as

$$E_{in}(x', y', 0) = 1 \cdot t(x', y') = \left[\text{rect}\left(\frac{x' - a/2}{L_x}\right) + \text{rect}\left(\frac{x' + a/2}{L_x}\right) \right] \cdot \text{rect}\left(\frac{y'}{L_y}\right)$$

To calculate the Fraunhofer diffraction pattern, we have to calculate the Fourier transform of this distribution:

$$\tilde{E}(f_x, f_y, 0) = \left[L_x \text{sinc}(\pi L_x f_x) e^{-i2\pi a f_x / 2} + L_x \text{sinc}(\pi L_x f_x) e^{i2\pi a f_x / 2} \right] L_y \text{sinc}(\pi L_y f_y)$$

where we use the shifting property of the Fourier transform (see Appendix D).

This expression can be rewritten as

$$\tilde{E}(f_x, f_y, 0) = L_x L_y \text{sinc}(\pi L_x f_x) \text{sinc}(\pi L_y f_y) 2 \cos(\pi a f_x) \text{ used}$$

The Fraunhofer diffraction pattern at a distance z is

$$E(x, y, z) = \frac{e^{ik\left(z + \frac{x^2 + y^2}{2z}\right)}}{i\lambda z} \left[\tilde{E}(f_x, f_y, 0) \right]_{\substack{f_x = x/\lambda z \\ f_y = y/\lambda z}}$$



$$E(x, y, z) = \frac{e^{ik\left(z + \frac{x^2 + y^2}{2z}\right)}}{i\lambda z} 2L_x L_y \operatorname{sinc}\left(\frac{\pi L_x x}{\lambda z}\right) \operatorname{sinc}\left(\frac{\pi L_y y}{\lambda z}\right) \cos\left(\frac{\pi a x}{\lambda z}\right)$$

The intensity of the diffraction pattern:

$$I(x, y, z) = \left(\frac{1}{2} \varepsilon_o c\right) \frac{4L_x^2 L_y^2}{\lambda^2 z^2} \operatorname{sinc}^2\left(\frac{\pi L_x x}{\lambda z}\right) \operatorname{sinc}^2\left(\frac{\pi L_y y}{\lambda z}\right) \cos^2\left(\frac{\pi a x}{\lambda z}\right)$$

The intensity distribution along the x -axis is

$$I(x, 0, z) = \left(\frac{1}{2} \varepsilon_o c\right) \frac{4L_x^2 L_y^2}{\lambda^2 z^2} \operatorname{sinc}^2\left(\frac{\pi L_x x}{\lambda z}\right) \cos^2\left(\frac{\pi a x}{\lambda z}\right)$$

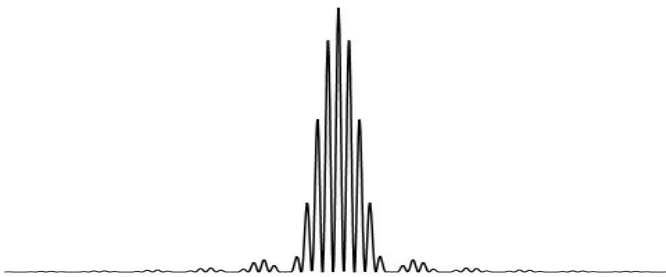
The zeros of the *sinc* function appear at the positions $x_{\text{zero}} = \pm m \frac{\lambda z}{L_x}$ $m = 0, 1, 2, \dots$ and

the separation between minima is $\Delta x_{\text{zero}} = \frac{\lambda z}{L_x}$.

While the zeros of the *cos* function appear at $x_{\text{zero}} = \pm \frac{(2m+1)\lambda z}{2a}$, and the separation

between minima is $\Delta x_{\text{zero}} = \frac{\lambda z}{a}$.

Since $a > L_x$ the intensity distribution shows the oscillatory $\cos^2\left(\frac{\pi a x}{\lambda z}\right)$ function modulated by the $\operatorname{sinc}^2\left(\frac{\pi L_x x}{\lambda z}\right)$ function.



Intensity distribution
of the double slit

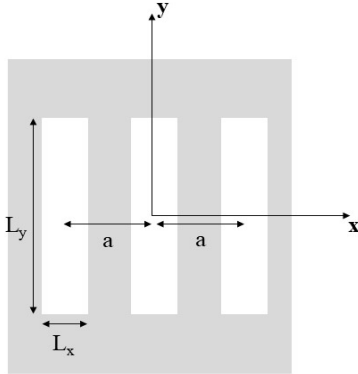
The periodic $\cos^2(\pi a x / \lambda z)$ function corresponds to the interference pattern of the double slit and would be the intensity pattern in the limiting case of infinitely narrow slits. The finite size of the slit produces the modulation pattern, which is responsible for the strong decrease in the intensity pattern at axial distances from the centre $|x| > \lambda z / L_x$.

■



S6.2 Calculate the Fraunhofer diffraction pattern of a three-slit screen of width L_x and height $L_y \gg L_x$, separated by a distance D between their centres and illuminated by a monochromatic plane wave with amplitude $E_o = 1$ V/m.

Solution



The field distribution at the plane $z = 0$ is given by the transmittance function, since it is illuminated by a uniform plane wave of unit amplitude:

$$E(x', y', 0) = \left[\text{rect}\left(\frac{x'}{L_x}\right) + \text{rect}\left(\frac{x'-a}{L_x}\right) + \text{rect}\left(\frac{x'+a}{L_x}\right) \right] \text{rect}\left(\frac{y'}{L_y}\right)$$

The Fourier transform is

$$\begin{aligned} \tilde{E}(f_x, f_y, 0) &= L_x L_y \text{sinc}(\pi f_x L_x) \text{sinc}(\pi f_y L_y) \left[1 + e^{-i2\pi a f_x} + e^{i2\pi a f_x} \right] = \\ &= L_x L_y \text{sinc}(\pi f_x L_x) \text{sinc}(\pi f_y L_y) (1 + 2 \cos(2\pi a f_x)) \end{aligned}$$

The function $1 + 2 \cos(x)$ can be rewritten in the form

$$\begin{aligned} 1 + 2 \cos(x) &= 1 + 2 \left(2 \cos^2(x/2) - 1 \right) = 4 \cos^2(x/2) - 1 = 3 - 4 \sin^2(x/2) \\ &= \frac{3 \sin(x/2) - 4 \sin^3(x/2)}{\sin(x/2)} = \frac{\sin(3x/2)}{\sin(x/2)} \end{aligned}$$

giving

$$\tilde{E}(f_x, f_y, 0) = L_x L_y \text{sinc}(\pi f_x L_x) \text{sinc}(\pi f_y L_y) \frac{\sin(3\pi a f_x)}{\sin(\pi a f_x)}$$

The Fraunhofer diffraction pattern:



$$E(x, y, z) = \frac{e^{ik\left(z + \frac{x^2 + y^2}{2z}\right)}}{i\lambda z} \left[\tilde{E}(f_x, f_y, 0) \right]_{\substack{f_x = x/\lambda z \\ f_y = y/\lambda z}}$$

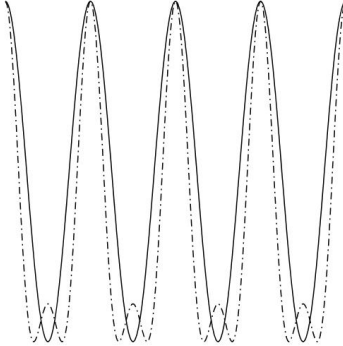
Substituting the Fourier transform:

$$E(x, y, z) = \frac{e^{ik\left(z + \frac{x^2 + y^2}{2z}\right)}}{i\lambda z} L_x L_y \text{sinc}\left(\frac{\pi L_x x}{\lambda z}\right) \text{sinc}\left(\frac{\pi L_y y}{\lambda z}\right) \frac{\sin(3\pi ax/\lambda z)}{\sin(\pi ax/\lambda z)}$$

We obtain the intensity distribution of the Fraunhofer distribution pattern:

$$I(x, y, z) = \left(\frac{1}{2} \epsilon_o c\right) \frac{L_x^2 L_y^2}{\lambda^2 z^2} \text{sinc}^2\left(\frac{\pi L_x x}{\lambda z}\right) \text{sinc}^2\left(\frac{\pi L_y y}{\lambda z}\right) \left[\frac{\sin(3\pi ax/\lambda z)}{\sin(\pi ax/\lambda z)} \right]^2$$

When comparing this result with the one obtained in the previous problem, we can see that the modulation function is the same. This is due to the fact that the dimensions and shape of the single slits are identical. The Fourier transform of the slit (unit cell) gives the modulation pattern, while the fast oscillations of the intensity are generated by the periodic distribution of slits in the array, i.e., the number of slits and separation between them. If the separation is the same, the maxima positions will be the same, as we can observe by comparing the normalized functions $\cos^2(\pi ax/\lambda z)$ (continuous line) and $(\sin(3\pi ax/\lambda z) / 3 \sin(\pi ax/\lambda z))^2$ (dot-dashed line).

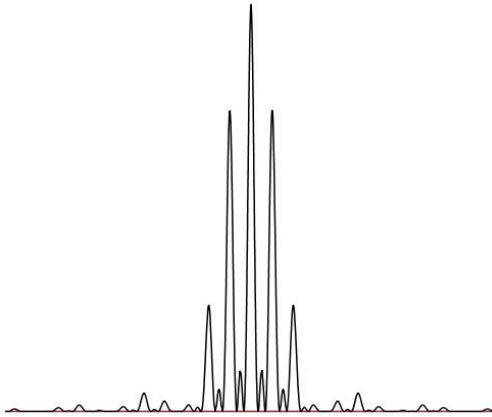


Increasing the number of slits reduces the width of the maxima. As the number of slits increases, the peaks become narrower and narrower.

The intensity of the Fraunhofer diffraction pattern on-axis

$$I(x, 0, z) = \left(\frac{1}{2} \epsilon_o c\right) \frac{L_x^2 L_y^2}{\lambda^2 z^2} \text{sinc}^2\left(\frac{\pi L_x x}{\lambda z}\right) \left[\frac{\sin(3\pi ax/\lambda z)}{\sin(\pi ax/\lambda z)} \right]^2$$

is shown in the figure below.

Diffraction pattern
of 3 slits

■

S6.3 Calculate the Fraunhofer diffraction pattern of a rectangular screen of sides L_x L_y , illuminated by a monochromatic, unit amplitude plane wave.

Solution

Contrary to the previous problems, where light is transmitted by the rectangular apertures, in this problem we have an opaque screen of rectangular shape blocking the light path.

The transmittance function is now written as $t(x', y', 0) = 1 - \text{rect}(x/L_x) \text{rect}(y/L_y)$.

Since the illuminating wave is uniform with unit amplitude, the field distribution is equal to this transmittance function. To calculate the Fraunhofer diffraction pattern, let us calculate the Fourier transform:

$$\tilde{E}(f_x, f_y, 0) = \delta(f_x, f_y) - L_x L_y \text{sinc}(\pi L_x f_x) \text{sinc}(\pi L_y f_y)$$

And the Fraunhofer diffraction pattern:

$$E(x, y, z) = \frac{e^{ik\left(z + \frac{x^2 + y^2}{2z}\right)}}{i\lambda z} \left(\delta\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right) + e^{i\pi} L_x L_y \text{sinc}\left(\frac{\pi L_x x}{\lambda z}\right) \text{sinc}\left(\frac{\pi L_y y}{\lambda z}\right) \right)$$

This diffraction pattern is very similar to the diffraction pattern of a rectangular aperture, which corresponds to the complementary aperture, except for the presence of a strong spot on axis and an extra phase factor.

This is a generic result known as Babinet's principle, in which the diffraction pattern of an opaque body is identical to that from a hole of the same size and shape, except for the overall forward beam intensity.

■



S6.4 White light is incident on a diffraction grating with spatial frequency $f = 800$ lines/cm.

a) At which angle will the red light appear if it is of wavelength 620 nm in the second diffraction order?

b) If the diffraction grating is rectangular with dimensions L_x and $L_y < L_x$, qualitatively explain the aspect of its diffraction pattern.

Solution

a) The diffraction pattern of a diffraction grating corresponds to a particular case of diffraction by periodic objects.

In this simple case, we are considering neither the slit shape nor the finite aperture size, and the grating is modelled simply as a comb function of the proper spacing. We also consider infinitely long apertures along the vertical direction, so we will not consider the dependence on y in this problem.

The spacing between lines is $a = 1/f = 0.00125$ cm = 12.5 μ m.

Assuming uniform illumination, the field distribution out of the grating is

$$E(x', 0) = \frac{1}{a} \text{comb}\left(\frac{x'}{a}\right) = \frac{1}{a} \sum_{n=-\infty}^{\infty} \delta\left(\frac{x'}{a} - n\right) = \frac{1}{a} \sum_{n=-\infty}^{\infty} \delta\left(\frac{x' - na}{a}\right) = \sum_{n=-\infty}^{\infty} \delta(x' - na).$$

The Fourier transform gives

$$\tilde{E}(f_x, 0) = \text{comb}(af_x).$$

The Fraunhofer diffraction pattern:

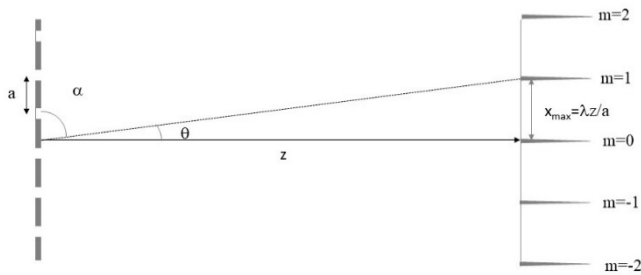
$$E(x, z) = \frac{e^{ik\left(z + \frac{x^2}{2z}\right)}}{i\lambda z} \text{comb}\left(\frac{ax}{\lambda z}\right) = \frac{e^{ik\left(z + \frac{x^2}{2z}\right)}}{i\lambda z} \sum_{m=-\infty}^{\infty} \delta\left(\frac{a}{\lambda z}\left(x - \frac{m\lambda z}{a}\right)\right) = \frac{e^{ik\left(z + \frac{x^2}{2z}\right)}}{ia} \sum_{m=-\infty}^{\infty} \delta\left(x - \frac{m\lambda z}{a}\right)$$

The diffraction pattern will appear along the x -axis as a set of maxima located at the positions $x_{\max} = \pm m \frac{\lambda z}{a}$, where $m = 0, 1, 2, \dots$ is the diffraction order.

The angle of propagation of these maxima can be calculated, assuming that the observation plane $z \gg a$:

$$\sin \theta_m \approx \frac{x_m}{z} = m \frac{\lambda}{a} = m\lambda f$$

In terms of the angle with respect to the x -axis, α : $\cos \alpha_1 = \lambda f \rightarrow f = \cos \alpha_1 / \lambda$, in accordance with the comments in Section 6.2.



Diffraction from an infinite diffraction grating

In our case, the wavelength 620 nm in the second diffraction order appears at the angle

$$\sin \theta_2 = 2\lambda f = 0.0992 \rightarrow \theta_2 = 5.69^\circ$$

b) In a more realistic case, the diffraction grating has a finite aperture. The effect of including the aperture to the diffraction pattern is that the delta functions are substituted by the diffraction pattern of the aperture. The diffraction pattern would consist of a periodic distribution of intensity patterns in the form

$$I \propto \text{sinc}^2\left(\frac{\pi L_x x}{\lambda z}\right) \text{sinc}^2\left(\frac{\pi L_y y}{\lambda z}\right)$$

Additionally, the particular shape of each individual slit would also introduce a modulation of the overall intensity pattern proportional to the diffraction pattern of the slit shape function.

■

S6.5 A plane wave of light with unit amplitude and wavelength $\lambda = 446 \text{ nm}$ is incident normally on a slit of width $L_x = 1 \text{ mm}$ and height $L_y = 3 \text{ mm}$. A lens is mounted at a distance f behind the slit (f being the focal length of the lens), and a screen is located in its back focal plane at a distance $f = 100 \text{ cm}$ after the lens. The size of the lens is large enough that we can neglect the effects of the finite aperture imposed by the lens diameter.

- Find the diffraction pattern at the focal plane of the lens.
- At what distance x (in cm) from the centre of the diffraction pattern of the slit does the fourth minimum along the x -axis appear?
- Changing the wavelength, we observe that the third minimum of the diffraction pattern now appears at the position found in b). What is the new wavelength?

Solution

a) The system is mounted in an f-f configuration. The Fresnel propagation of the wavefronts in this configuration shows that, at the focal plane of the lens, we have the



Fraunhofer diffraction pattern of the object with an added phase term, which cancels for $d = d' = f$ (see the Appendix F for details).

Thus, using this configuration gives us a way to obtain the Fraunhofer diffraction pattern at a fixed plane: the back focal plane of the lens.

The diffraction pattern of the slit is written as

$$E(x, y, f) = \frac{1}{i\lambda f} \left[\tilde{E}(f_x, f_y, 0) \right]_{\substack{f_x = x/\lambda f \\ f_y = y/\lambda f}} = \frac{L_x L_y}{i\lambda f} \operatorname{sinc} \left(\frac{\pi L_x x}{\lambda f} \right) \operatorname{sinc} \left(\frac{\pi L_y y}{\lambda f} \right)$$

What we can see in the screen is the intensity pattern:

$$I(x, y, f) = \left(\frac{1}{2} \epsilon_0 c \right) \frac{L_x^2 L_y^2}{\lambda^2 f^2} \operatorname{sinc}^2 \left(\frac{\pi L_x x}{\lambda f} \right) \operatorname{sinc}^2 \left(\frac{\pi L_y y}{\lambda f} \right)$$

b) The minima of the intensity pattern along the x -axis occurs at the zeros of the sinc-squared function:

$$\frac{\pi L_x x_m}{\lambda f} = \pm m\pi \quad \text{for } m = 1, 2, \dots$$

The fourth minimum appears at $x_4 = \frac{4\lambda f}{L_x} = 1.78 \cdot 10^{-3} \text{ m} = 1.78 \text{ mm}$.

c) By changing the wavelength illuminating the slide, the position of the minima will change according to the given formula:

$$x'_m = \pm m \frac{\lambda' f}{L_x}$$

The conditions of the problem require that $x_4 = x'_3$:

$$\frac{4\lambda f}{L_x} = \frac{3\lambda' f}{L_x} \rightarrow \lambda' = \frac{4\lambda}{3} = 594.67 \text{ nm}$$

■

S6.6 Write the expressions for the transfer function obtained using the paraxial wave equation for beams and the angular spectrum representation. Discuss the similarities and differences between these two transfer functions, and show the limit at which both expressions are coincident.

Solution

The transfer function, $H(f_x, f_y, z)$, relates the frequency components of the field at different positions in linear systems:



$$\tilde{E}(f_x, f_y, z) = \tilde{E}(f_x, f_y, 0) \cdot H(f_x, f_y, z)$$

The transfer function associated with the paraxial wave equation for beams:

$$\frac{\partial E_o^{(+)}}{\partial z} - \frac{i}{2k} \nabla_{\perp}^2 E_o^{(+)} = 0$$

This was derived in Chapter 2 (Eq. [2.15]):

$$\tilde{E}_o^{(+)}(f_x, f_y, z) = \tilde{E}_o^{(+)}(f_x, f_y, 0) \cdot e^{-i \frac{(k_x^2 + k_y^2)z}{2k}} = \tilde{E}_o^{(+)}(f_x, f_y, 0) \cdot e^{-i \lambda \pi (f_x^2 + f_y^2)z}$$

The total field propagated at distance z in the paraxial approximation is given by

$$\tilde{E}^{(+)}(f_x, f_y, z) = e^{ikz} \tilde{E}_o^{(+)}(f_x, f_y, z) = \tilde{E}_o^{(+)}(f_x, f_y, 0) \cdot e^{i(kz - \lambda \pi (f_x^2 + f_y^2)z)}$$

When we consider the angular spectrum representation, starting with the Helmholtz equation allows us to directly obtain the transfer function using Equation [6.5]:

$$\tilde{E}^{(+)}(f_x, f_y, z) = \tilde{E}^{(+)}(f_x, f_y, 0) e^{i2\pi m z} = \tilde{E}^{(+)}(f_x, f_y, 0) e^{i2\pi z \sqrt{\frac{1}{\lambda^2} - (f_x^2 + f_y^2)}}$$

The transfer function in the paraxial wave approximation is always a pure phase term, while the angular spectrum representation also contains evanescent terms for large enough frequencies that satisfy the relation $(f_x^2 + f_y^2) > \frac{1}{\lambda^2}$

These evanescent fields propagate along the object surface and, hence, do not satisfy the paraxiality criterion of propagation close to the propagation axis. In the paraxial region, the spatial frequencies involved should be small, leading to a pure phase factor in the transfer function.

Moreover, the condition satisfied by the spatial frequencies related to propagation angles in the paraxial relation is $(f_x^2 + f_y^2) \ll \frac{1}{\lambda^2}$, and we can write

$$\sqrt{\frac{1}{\lambda^2} - (f_x^2 + f_y^2)} = \frac{1}{\lambda} \sqrt{1 - \lambda^2 (f_x^2 + f_y^2)} \approx \frac{1}{\lambda} \left(1 - \frac{\lambda^2 (f_x^2 + f_y^2)}{2} \right)$$

Under this approximation, we obtain

$$\tilde{E}^{(+)}(f_x, f_y, z) = \tilde{E}^{(+)}(f_x, f_y, 0) \cdot e^{i(kz - \lambda \pi (f_x^2 + f_y^2)z)}$$

which coincides with the expression obtained under the paraxial wave approximation. A particular conclusion of this limit is that the paraxial approximation cannot be used to study the propagation of field distributions containing very large frequency components of the order or close to $1/\lambda$.

■



S6.7 In a spatial-filter using a $4f$ correlator setup, we want to eliminate the high frequencies of our optical beam corresponding to values higher than $1/(25\lambda)$. The lenses used in our device are microscope objectives of X20 with focal length $f = 8 \text{ mm}$. Find the position in the Fourier plane where each spatial frequency will be located and the radius of the aperture we have to use in order to filter our beam.

Solution

In a $4f$ correlator setup, the Fraunhofer diffraction pattern is placed at the back focal plane of the transforming lens, which in this case is the microscope objective.

Assuming that our initial field distribution is $E^{(+)}(x', y', 0)$, and neglecting the effects of the finite apertures of the objectives at the Fourier plane, we will obtain a field distribution given by the expression

$$\tilde{E}^{(+)}(x, y, f) = A \cdot \left[\tilde{E}(f_x, f_y, 0) \right]_{\substack{f_x = x/\lambda f \\ f_y = y/\lambda f}}$$

The scaling factor allows us to relate each point in the focal plane with a particular spatial frequency:

$$x = f_x \lambda f \text{ and } y = f_y \lambda f$$

$$\text{At a particular distance: } r = \sqrt{x^2 + y^2} = \lambda f \sqrt{f_x^2 + f_y^2}$$

The spatial frequencies $f_x = f_y = 1/(25\lambda)$ will then be placed in the Fourier plane at

$$r = \frac{\lambda f \sqrt{2}}{25\lambda} = \frac{\sqrt{2}}{25} f = 0.45 \text{ mm} = 450 \text{ }\mu\text{m}$$

To eliminate the high frequencies, as desired, a pinhole of 450 microns should be placed at the focal plane of the lens. The propagated field after the pinhole will be modified, since the large frequencies will be missing. In the case of having an initial field distribution consisting of a Gaussian beam with strong high-frequency modulations, this technique allows us to *clean* the beam by eliminating the high frequencies and obtaining a much smoother Gaussian beam at the output.

■

S6.8 Two lenses of diameters D_1 and D_2 and equal focal length, f , are used to image two points placed at a very large distance. Assuming that the light coming from each of the objects can be approximated by a plane wave when reaching the lens, determine what is the minimum angular separation that is measurable in each case.

Solution

A lens with infinite diameter should focus a perfect plane wave onto a single point in the Fourier plane if all aberrations can be perfectly corrected. Even in the case of lenses



with perfectly corrected aberrations, the finite diameter of the lens would impose a limiting condition on the resolution of distant objects, due to the diffraction pattern induced by the pupil function of the lens.

A plane wave illuminating a perfectly corrected lens of diameter D will give a diffraction pattern in the focal plane, corresponding to the Airy function (see Appendix E):

$$I(r, f) \propto \left[\frac{J_1(\pi D r / \lambda f)}{D r / 2 \lambda z} \right]^2$$

The diameter of the central spot would be $\phi = \frac{1.22 \lambda f}{D/2}$

Applying the Rayleigh criterion, the minimum resolvable separation between the images, s , would correspond to half this diameter. Thus, the maximum of the diffraction pattern of one of the objects coincides with the first zero of the diffraction pattern of the other:

$$s = \frac{\phi}{2} = \frac{1.22 \lambda f}{D}$$

Given this distance, the minimum resolvable angular separation between the distant objects is

$$\theta_{\min} \approx \frac{s}{f} = \frac{1.22 \lambda}{D}$$

Using this relation, the minimum resolvable angular separation using lenses of different diameters will be

$$\frac{\theta_{\min 1}}{\theta_{\min 2}} = \frac{D_2}{D_1} \rightarrow \theta_{\min 1} = \frac{D_2}{D_1} \theta_{\min 2}$$

■



6.9 Proposed problems

P6.1 A one-dimensional object is illuminated by radiation of wavelength λ_0 . The object has periodic elements of spatial periods in the x direction: $\Delta x = \lambda_0/5$. Determine the amplitude of the Fourier components of the field amplitude for this particular frequency at $z = \lambda_0$.

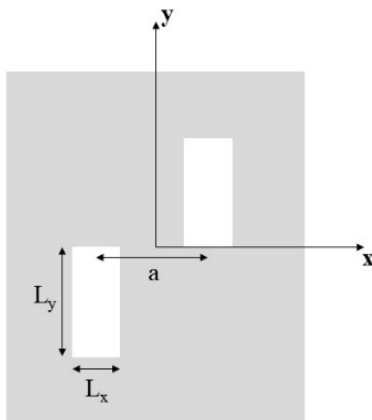
P6.2 A two-dimensional object, illuminated with monochromatic radiation of wavelength λ_0 , has a periodicity giving maximum spatial frequencies of $f_x = 0.5/\lambda_0$ in the x direction. Find the range of frequency components in the y direction for which the corresponding Fourier components $\tilde{E}(0.5/\lambda_0, f_y, z)$, correspond to propagating modes.

P6.3 Discuss the validity region of Fresnel's approximation, which is the region along the transverse plane for which the contribution of the third-order and higher-order terms in the binomial expansion of R produce phase changes much less than 1 rad when inserted in the Rayleigh–Sommerfeld formula (Equation [6.14]).

P6.4 A slit of height $L_y = 0.3$ mm and width $L_x = 0.1$ mm is illuminated by radiation of wavelength $\lambda_0 = 462$ nm. The observation plane is placed at a distance of $D = 2$ m. Determine if the resulting diffraction pattern satisfies the condition for being in the Fraunhofer diffraction region (far field). Calculate the Fraunhofer diffraction pattern and determine the angular width of the central intensity maximum in the x direction.

P6.5 An opaque screen has two slits of 0.10 mm width and 0.5 mm height, separated by a distance (between centres) of 0.25 mm and illuminated by a monochromatic plane wave of wavelength 532 nm. The observation plane is placed at a distance of $D = 5$ m. Determine the Fraunhofer diffraction pattern and calculate how many bright fringes will appear inside the central diffraction peak.

P6.6 Find the Fraunhofer diffraction pattern of the two-slit distribution (shown in the figure) when it is illuminated by a plane wave of unit constant amplitude.





P6.7 The diameter of an observatory telescope is 5 m.

a) Determine the angular resolution limit at the wavelengths of 400 nm, 500 nm and 600 nm.

b) What is the minimum distance between two objects placed at the Moon's surface that can be resolved by this telescope at the given wavelengths?

P6.8 White light is incident on a diffraction grating with 1000 lines/cm. At what angle will the red light appear if it has a wavelength of 650 nm in the first diffraction order?

P6.9 The light emitted by a Na lamp has two spectral components at the wavelengths 589.5923 nm and 589.9953 nm. If we are using a diffraction grating with 10000 lines/cm.

a) What will be the separation between these two lines in a screen placed at a distance of 1 m from the grating?

b) Assuming that the aperture is fully and homogeneously illuminated, what is the minimum aperture needed to resolve these two lines at this distance? (Consider that we are in the Fraunhofer diffraction regime.)

P6.10 A set of circular holes of radius R is placed along a horizontal line, with their centres separated by a fixed distance D . The whole linear chain of holes has a total length L and is illuminated by a uniform monochromatic plane wave. Qualitatively explain the general aspect of the corresponding Fraunhofer diffraction pattern.

P6.11 Calculate the Fraunhofer diffraction pattern of an N -slit screen of width L_x and height $L_y \gg L_x$, separated by a distance D between their centres and illuminated by a monochromatic plane wave with amplitude $E_o = 1$ V/m.

P6.12 A diffraction grating generally consists of some periodic distribution of the transmittance function with a given period. Consider the following three different types of gratings assumed to be infinitely extended in the x and y directions and with a 1D periodic distribution along the x -axis.

- Diffraction grating with transmittance: $t(x) = \frac{1}{2} \left[1 + m \cdot \cos \left(\frac{2\pi x}{a} \right) \right]$.
- Diffraction grating with negligible width (delta-type) apertures, separated by a distance a .
- Diffraction grating with rectangular apertures of width L_x , separated by a distance a .

If we consider that each of these apertures is illuminated by a uniform plane wave of constant intensity, calculate their corresponding Fraunhofer diffraction patterns.

Qualitatively discuss how the diffraction pattern of each of these structures would be modified when considering a circular aperture of radius $R \gg a$ that limits the effective size of our gratings.







Appendix A

The Laser

The theory of the laser—which is the acronym for “light amplification by stimulated emission of radiation”—was developed by Basov and Prokhorov in Russia and Townes and Schawlow in USA. The first laser was constructed by Maimann in 1960 using a ruby crystal. The laser was a natural extension of the maser, which was used as a source for infrared radiation. Since the invention of this source, the field of photonics and its applications have grown considerably. The laser is an optical equivalent to an electronic amplifier with feedback consisting of an amplifying medium and a resonator.

Research on thermal radiation at the end of the 19th century constitutes the first step towards understanding atomic systems. Bohr postulates the following. 1) Atoms may exist in steady states characterized by a discrete set of possible energy values, $E_1, E_2, \dots$. In these states, the atoms neither emit nor absorb energy. 2) The radiation emitted or absorbed by the atom corresponds to monochromatic radiation of frequency

$$E_j - E_i = \hbar\omega_{ij}$$

where $\hbar$ is Planck’s constant and E_i, E_j are generic atomic energy levels with $E_j > E_i$. Energy is emitted and absorbed by atoms in discrete amounts, which we call *photons*.

In his paper *On the quantum theory of radiation*, Einstein considered the interaction of atoms with thermal radiation and postulated the basic processes for light–matter interaction. Let us assume a gas formed by a collection of identical atoms, with each atom having a pair of energy levels E_1 and E_2 that satisfy the condition $E_2 - E_1 = \hbar\omega$. Energy conserving transitions can occur between levels 1 and 2, in which photons of frequency ω are emitted or absorbed by the atoms making transitions between levels 1 and 2. Under the presence of light with a given energy density $u(\omega)$, the probabilities of photon absorption and emission are defined as follows.

a) Spontaneous emission: An atom in state 2 has a finite probability to spontaneously decay into level 1, emitting a photon of energy $\hbar\omega$ in an arbitrary direction. The probability per unit time for this process to occur is



$$P_{sp} = A_{21}$$

b) Stimulated absorption: The atom absorbs a photon of frequency ω and makes a transition from level 1 to state 2. The probability per unit time for this radiation induced process is

$$P_{abs} = B_{12}u(\omega)$$

c) Stimulated emission: The presence of radiation at frequency ω induces or stimulates a transition from level 2 to level 1 at a rate proportional to the radiative energy density. The emitted photon is emitted in the same state of the incident photon stimulating the transition. The stimulated emission probability per unit time is

$$P_{st} = B_{21}u(\omega)$$

The Einstein coefficients A , B_{12} and B_{21} depend on the properties of the two atomic states.

The effect of these basic processes on the atomic level populations can be studied under the assumption that the total number of atoms is large enough for the individual absorption and emission processes to produce a smooth temporal dependence of the populations, whose rate of change can be defined through the rate equations:

$$\frac{dN_1}{dt} = -\frac{dN_2}{dt} = N_2 A_{21} + B_{21}u(\omega)N_2 - B_{12}u(\omega)N_1$$

where we are assuming that transitions are taking place only between levels 1 and 2.

For a collection of atoms in equilibrium with thermal radiation the net rate of change should be zero, and we get

$$u(\omega) = \frac{A_{21}/B_{21}}{(B_{12}/B_{21})(N_1/N_2) - 1} = \frac{A_{21}/B_{21}}{(B_{12}/B_{21})e^{\hbar\omega/kT} - 1}$$

where the level populations in the state of thermal equilibrium obey the Boltzmann distribution:

$$N_1/N_2 = e^{\hbar(E_2 - E_1)/kT} = e^{\hbar\omega/kT},$$

where k is Boltzmann's constant.

By taking the limit of this expression when $T \rightarrow \infty$, we obtain that $B_{12} = B_{21}$ (this relation holds for non-degenerate population levels). For the case of degenerate levels with degeneracies g_1 and g_2 , the more general relation $g_1 B_{12} = g_2 B_{21}$ holds.

At the limit $\hbar\omega \ll kT$ ($\omega \rightarrow 0$), the energy density should correspond to the classical Rayleigh-Jeans formula:

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3} kT$$

From this, the ratio between the spontaneous and stimulated emission rates is deduced:



$$\frac{A_{21}}{B_{21}} = \frac{\hbar\omega^3}{\pi^2 c^3}$$

This relation explains the difficulties in constructing a light source based on stimulated emission. The frequency dependence proportional to ω^3 states that, as the frequency increases, it is more difficult to achieve a situation where stimulated emission prevails over spontaneous emission rates. Although this relation has been derived under thermal radiation conditions, it turns out to be a general relation between these coefficients.

Putting all these relations together, we get the expression

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3} \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}$$

which corresponds to Planck's radiation law.

To fabricate a laser, we will not have the isotropic radiation related to thermal radiation but instead some directional radiation that we want to amplify by stimulated emission.

To calculate the amplification rate achieved in the gain medium (crystal, liquid, or gas) of transverse area S , let us consider that it is illuminated by a plane wave. The volume contained in a slab of thickness dz is $dV = S \cdot dz$ and, denoting the number of atoms per unit volume as N , we can write the number of atoms in the volume dV as $dN = N \cdot dV$. From this amount, a fraction $dN_i = N_i dV$, will populate the level $i = 1, 2$ and $N = N_1 + N_2$. According to Einstein's model, the total radiation power of the given volume element of the medium is equal to

$$dP = dP_{21}^{sp} + dP_{21}^{st} - dP_{12}^{abs} = A_{21}\hbar\omega dN_2 + B_{21}\hbar\omega u(\omega)dN_2 - B_{12}\hbar\omega u(\omega)dN_1$$

The spectral intensity $I(\omega)$ is related to the spectral energy density in the following way.

$$dI(\omega) = \frac{dE(\omega)}{Sdt} = \frac{dE(\omega) \cdot dz}{dV \cdot dt} = u(\omega) \cdot c$$

and the total intensity for a source of spectral bandwidth (considered narrow) $\Delta\omega$ is given by $I = I(\omega)\Delta\omega = cu(\omega)\Delta\omega$. The increment of light intensity in the slab dz is

$$dI = \frac{dP}{S} = \frac{dP}{dV} dz = \left[A_{21}N_2 + B(N_2 - N_1) \frac{I}{c\Delta\omega} \right] \hbar\omega dz$$

where we use the relation between absorption and stimulated coefficients. If the intensity is amplified such that the spontaneous emission term may be neglected, we obtain

$$dI = GIdz \rightarrow I(z) = I_0 e^{Gz}$$

where G is the gain coefficient $G = B(N_2 - N_1)\hbar\omega / c\Delta\omega$.

It is clear from this expression that light amplification will appear only in the case of $N_2 > N_1$. This condition, known as *population inversion*, is a necessary condition for achieving laser emission. Under usual thermal radiation conditions, the population of



the lower energy levels is always larger, so the medium is absorbing instead of amplifying the incident radiation.

Thus, the first requirement to obtain a laser is to achieve population inversion. This is usually achieved by illuminating the gain medium with light (optical pumping), by using electric discharges, electric currents, or chemical reactions.

By reinjecting the amplified intensity back into the medium by means of an external resonator, the stimulated emission process is further enhanced until a stable output emission of radiation is obtained. Laser emission requires that the gain in the medium exceeds the losses. Assuming a simple system of cavity length L in which the losses are caused only by the output mirrors of the resonator, with reflectivity values R_1 and R_2 , the laser condition is achieved when the following conditions hold.

Amplitude condition: $R_1 R_2 e^{GL} \geq 1$

Phase condition: $L = n\lambda/2, \quad n = 1, 2, 3, \dots$

In summary, a laser requires a gain medium in which population inversion is achieved through the pumping mechanism. The population inversion produces a large amount of photons through a stimulated emission process that emits them in the same quantum state (direction and phase), and the external resonators produce the feedback, thus allowing the gain to exceed the losses and obtain stable output radiation.

The different types of lasers differ in the gain media used, the cavity configurations implemented, and the pumping schemes.

The basic characteristics of the continuous wave (CW) laser radiation are:

- High monochromaticity: relative bandwidths of the order of $\Delta\nu/\nu \sim 10^9$ are typical for CW lasers and much larger values can be obtained in stabilized systems.
- High directionality: the laser output of many lasers is close to a Gaussian beam.
- High spatial and temporal coherence: possibility to obtain interference phenomena.

Many current photonics applications require the use of high intensity or ultra-short pulse duration sources. Lasers can be made to emit pulses through the mechanism of phase-locking. By these and other techniques such as pulse chirped amplification, lasers emitting pulses as short as a few femtoseconds or powers of TW/cm^2 are routinely used in many laboratories. Interested readers will find specific references that extensively cover the theory of lasers in the proposed bibliography.







Appendix B

Classical radiation from elementary dipoles

Elementary dipole

An elementary electric dipole is formed by two oppositely charged particles separated by a small distance. Given the position of the negatively charged particle with respect to the positively charged particle ($\vec{x}$), the elementary dipole moment is defined as $\vec{p} = -q\vec{x} = p\hat{d}$.



The emission of electromagnetic radiation from an elementary dipole is obtained in classical electrodynamics from the microscopic Maxwell equations:

$$\vec{\nabla} \cdot \vec{e}(\vec{r}, t) = \frac{\rho(\vec{r}, t)}{\epsilon_o} \quad (1) \quad \vec{\nabla} \wedge \vec{e}(\vec{r}, t) = -\frac{\partial \vec{b}(\vec{r}, t)}{\partial t} \quad (3)$$

$$\vec{\nabla} \cdot \vec{b}(\vec{r}, t) = 0 \quad (2) \quad \vec{\nabla} \wedge \vec{b}(\vec{r}, t) = \mu_o \vec{j}(\vec{r}, t) + \mu_o \epsilon_o \frac{\partial \vec{e}(\vec{r}, t)}{\partial t} \quad (4)$$

(!) We will omit the spatial and temporal dependence of the fields in the expressions below.

Potentials

According to equations (1) and (2), the electric and magnetic fields can always be written in terms of the following additional functions, called electromagnetic potentials.

Vector potential $\vec{a}$:

$$\vec{b} = \vec{\nabla} \wedge \vec{a} \quad (5)$$

Scalar potential ϕ :

$$\vec{e} = -\vec{\nabla}\phi - \frac{\partial \vec{a}}{\partial t} \quad (6)$$



Substituting these into Maxwell's equations (3) and (4) gives the equations of motion for the potentials:

$$\begin{aligned}\nabla^2\phi + \frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{a}) &= \frac{-\rho}{\epsilon_o} \\ \nabla^2\vec{a} - \frac{1}{c^2} \frac{\partial^2\vec{a}}{\partial t^2} - \vec{\nabla}[\vec{\nabla} \cdot \vec{a} + \mu_o\epsilon_o \frac{\partial\phi}{\partial t}] &= -\mu_o\vec{j}\end{aligned}$$

The electromagnetic potentials are not unique, since we can use a given pair of potentials to define a new pair of valid potentials through the substitution

$$\begin{aligned}\vec{a}' &= \vec{a} + \vec{\nabla}\psi \\ \phi' &= \phi - \frac{\partial\psi}{\partial t}\end{aligned}$$

Each selection of an (a, ϕ) constitutes a gauge transformation.

The most common transformations are the following.

Coulomb gauge: $\vec{\nabla} \cdot \vec{a} = 0$

Lorenz gauge: $\vec{\nabla} \cdot \vec{a} + \mu_o\epsilon_o \frac{\partial\phi}{\partial t} = 0$

In the Lorenz gauge, the equations of motion for the potentials are written as

$$\begin{aligned}\nabla^2\phi - \mu_o\epsilon_o \frac{\partial^2\phi}{\partial t^2} &= \frac{-\rho}{\epsilon_o} \\ \nabla^2\vec{a} - \mu_o\epsilon_o \frac{\partial^2\vec{a}}{\partial t^2} &= -\mu_o\vec{j}\end{aligned}$$

The general solutions to these equations are expressed in terms of the retarded potentials:

$$\begin{aligned}\phi(\vec{r}, t) &= \frac{1}{4\pi\epsilon_o} \int_V \frac{\rho(\vec{r}', t - R/c)}{R} d^3r', & \vec{a}(\vec{r}, t) &= \frac{\mu_o}{4\pi} \int_V \frac{\vec{j}(\vec{r}', t - R/c)}{R} d^3r', \\ R &= |\vec{r} - \vec{r}'|\end{aligned}$$

Classical radiation problems can be solved with these equations. From a given distribution, the potentials are calculated and the resulting electromagnetic field is obtained from Eqs. (5) and (6).

Hertz vector potential

An alternative way to obtain the radiation by dipoles can be obtained in terms of a new potential: The Hertz potential.



In the equations for the potentials, the source terms are provided by the elementary charges and current densities. These are always related by the continuity equation:

$$\vec{\nabla} \cdot \vec{j} + \frac{\partial \rho}{\partial t} = 0$$

We can express for an electric dipole moment, $\vec{p}$, the corresponding relations

$$\rho = -\vec{\nabla} \cdot \vec{p} \quad \text{and} \quad \vec{j} = \frac{\partial \vec{p}}{\partial t}$$

Defining a new potential (Π) through the relations

$$\vec{a} = \mu_o \frac{\partial \vec{\Pi}}{\partial t} \quad \text{and} \quad \phi = -\frac{\nabla \Pi}{\epsilon_o}$$

we obtain an equation of motion for the new potential:

$$\nabla^2 \vec{\Pi} - \mu_o \epsilon_o \frac{\partial^2 \vec{\Pi}}{\partial t^2} = -\vec{p} \quad \text{whose solution is} \quad \rightarrow \quad \vec{\Pi}(\vec{r}, t) = \frac{1}{4\pi} \int_V \frac{\vec{p}(\vec{r}', t - R/c)}{R} d^3 r'$$

Once the solution for the Hertz vector potential is found, the resulting electric and magnetic field vectors are obtained as

$$\begin{aligned} \vec{e}(\vec{r}, t) &= \frac{1}{\epsilon_o} \vec{\nabla} [\vec{\nabla} \cdot \vec{\Pi}(\vec{r}, t)] - \mu_o \frac{\partial^2 \vec{\Pi}(\vec{r}, t)}{\partial t^2} \\ \vec{b}(\vec{r}, t) &= \mu_o [\vec{\nabla} \wedge \frac{\partial \vec{\Pi}(\vec{r}, t)}{\partial t}] \end{aligned}$$

Example: electromagnetic field of an elemental Hertzian dipole

Consider an elemental linear electric dipole at position r_o oscillating along a fixed direction with a corresponding dipole moment expressed by the relation

$$\vec{p}(\vec{r}, t) = p(t) \delta(\vec{r} - \vec{r}_o) \hat{d}$$

The solution of the Hertz potential associated with this dipole moment is

$$\vec{\Pi}(\vec{r}, t) = \frac{1}{4\pi R} p(t - R/c) \hat{d} \quad R = |\vec{r} - \vec{r}_o|$$

The corresponding electric and magnetic fields are found by direct substitution into the expressions shown previously:

$$\begin{aligned} \vec{e} &= \frac{1}{\epsilon_o} \vec{\nabla} [\vec{\nabla} \cdot \vec{\Pi}] - \mu_o \frac{\partial^2 \vec{\Pi}}{\partial t^2} \\ \vec{e} &= \frac{1}{\epsilon_o} \vec{\nabla} [\vec{\nabla} \cdot (\frac{p \hat{d}}{4\pi R})] - \mu_o \frac{\partial^2}{\partial t^2} (\frac{p \hat{d}}{4\pi R}) \end{aligned}$$

(!) We omit the $(t-R/c)$ dependence on p to simplify the expressions



$$\vec{e} = \frac{1}{\epsilon_o} \vec{\nabla} \left[\frac{(\vec{\nabla} p) \hat{d}}{4\pi R} + \frac{p \hat{d}}{4\pi} \vec{\nabla} \left(\frac{1}{R} \right) \right] - \mu_o \frac{\ddot{p} \hat{d}}{4\pi R}$$

(!) Considering that $\hat{d}$ is constant and $\ddot{p}$ is a second temporal derivative evaluated at retarded time.

To evaluate these expressions, we use the mathematical equivalences

$$\vec{\nabla} p(t - R/c) = \frac{\partial p}{\partial t} \frac{(-\vec{\nabla} R)}{c} = -\frac{\dot{p} \vec{R}}{cR} \quad \text{and} \quad \vec{\nabla} \left(\frac{1}{R} \right) = -\frac{\vec{R}}{R^3}$$

then

$$\begin{aligned} \vec{e} = & \frac{1}{4\pi\epsilon_o} \vec{\nabla} \left[-\frac{\dot{p} \vec{R} \hat{d}}{cR^2} - \frac{p \hat{d} \vec{R}}{R^3} \right] - \mu_o \frac{\ddot{p} \hat{d}}{4\pi R} = \frac{-1}{4\pi\epsilon_o} \left[\frac{(\vec{\nabla} \dot{p}) \hat{d} \vec{R}}{cR^2} + \frac{\dot{p} \vec{\nabla}(\hat{d} \vec{R})}{cR^2} + \frac{\dot{p} \hat{d} \vec{R}}{c} \vec{\nabla} \left(\frac{1}{R^2} \right) \right. \\ & \left. + \frac{(\vec{\nabla} p) \hat{d} \vec{R}}{R^3} + \frac{p \vec{\nabla}(\hat{d} \vec{R})}{R^3} + p \hat{d} \vec{R} \vec{\nabla} \left(\frac{1}{R^3} \right) \right] - \mu_o \frac{\ddot{p} \hat{d}}{4\pi R} \end{aligned}$$

Now, consider that

$$\begin{aligned} \vec{\nabla}(\hat{d} \vec{R}) &= \hat{d} \wedge (\vec{\nabla} \wedge \vec{R}) + \vec{R} \wedge (\vec{\nabla} \wedge \hat{d}) + (\hat{d} \vec{\nabla}) \vec{R} + (\vec{R} \vec{\nabla}) \hat{d} = \\ &= (\hat{d} \vec{\nabla}) \vec{R} = (d_x \frac{\partial}{\partial x} + d_y \frac{\partial}{\partial y} + d_z \frac{\partial}{\partial z}) \vec{R} = \hat{d} \end{aligned}$$

to obtain

$$\begin{aligned} \vec{e} = & \frac{-1}{4\pi\epsilon_o} \left[-\frac{\ddot{p} \vec{R}(\hat{d} \vec{R})}{c^2 R^3} + \frac{\dot{p} \hat{d}}{cR^2} + \frac{\dot{p} \hat{d} \vec{R}}{c} \left(\frac{-2\vec{R}}{R^4} \right) - \frac{\dot{p} \vec{R}(\hat{d} \vec{R})}{cR^4} + \frac{p \hat{d}}{R^3} + p \hat{d} \vec{R} \left(\frac{-3\vec{R}}{R^5} \right) \right] - \mu_o \frac{\ddot{p} \hat{d}}{4\pi R} \\ \vec{e} = & \frac{1}{4\pi\epsilon_o} \left[\left(\frac{\ddot{p}}{c^2 R^3} + \frac{3\dot{p}}{cR^4} + \frac{3p}{R^5} \right) \vec{R}(\hat{d} \vec{R}) - \left(\frac{p}{R^3} + \frac{\dot{p}}{cR^2} \right) \hat{d} \right] - \mu_o \frac{\ddot{p} \hat{d}}{4\pi R} \\ \vec{e} = & \frac{1}{4\pi\epsilon_o} \left[\left(\frac{\ddot{p}}{c^2 R} + \frac{3\dot{p}}{cR^2} + \frac{3p}{R^3} \right) \hat{R}(\hat{d} \hat{R}) - \left(\frac{p}{R^3} + \frac{\dot{p}}{cR^2} + \frac{\ddot{p}}{c^2 R} \right) \hat{d} \right] \end{aligned}$$

The resulting magnetic field is calculated from the expression

$$\begin{aligned} \vec{b}(\vec{r}, t) &= \mu_o \left[\vec{\nabla} \wedge \frac{\partial \vec{\Pi}(\vec{r}, t)}{\partial t} \right] \\ \vec{b} &= \frac{\mu_o}{4\pi} \left(\vec{\nabla} \wedge \frac{\dot{p} \hat{d}}{R} \right) = \frac{\mu_o}{4\pi} \left(\frac{\dot{p}}{R} \vec{\nabla} \wedge \hat{d} + \vec{\nabla} \left(\frac{\dot{p}}{R} \right) \wedge \hat{d} \right) = \frac{\mu_o}{4\pi} \left((\vec{\nabla} \dot{p}) \frac{1}{R} \wedge \hat{d} + \dot{p} \vec{\nabla} \left(\frac{1}{R} \right) \wedge \hat{d} \right) \\ \vec{b} &= \frac{\mu_o}{4\pi} \left(\frac{-\ddot{p}(\vec{R} \wedge \hat{d})}{cR^2} - \frac{\dot{p}(\vec{R} \wedge \hat{d})}{R^3} \right) = \frac{\mu_o}{4\pi} \left(\frac{\ddot{p}}{cR} + \frac{\dot{p}}{R^2} \right) (\hat{d} \wedge \hat{R}) \end{aligned}$$



Summarizing, the electromagnetic field from a Hertzian dipole is

$$\begin{aligned} \vec{e} &= \frac{1}{4\pi\epsilon_0} \left[\left(\frac{\ddot{p}}{c^2 R} + \frac{3\dot{p}}{cR^2} + \frac{3p}{R^3} \right) \hat{R}(\hat{d}\hat{R}) - \left(\frac{p}{R^3} + \frac{\dot{p}}{cR^2} + \frac{\ddot{p}}{c^2 R} \right) \hat{d} \right] \\ \vec{b} &= \frac{\mu_0}{4\pi} \left(\frac{\ddot{p}}{cR} + \frac{\dot{p}}{R^2} \right) (\hat{d} \wedge \hat{R}) \end{aligned} \quad \hat{R} = \frac{\vec{R}}{R}$$

In the radiation zone (for which $l \ll \lambda \ll R$), the only terms making any significant contribution are those that are proportional to $(1/R)$. In this zone, we can write the radiated fields as

$$\begin{aligned} \vec{e}_{rad} &= \frac{\mu_0 \ddot{p}}{4\pi R} [\hat{R}(\hat{d}\hat{R}) - \hat{d}] = \frac{\mu_0 \ddot{p}}{4\pi R} (\hat{R} \wedge (\hat{R} \wedge \hat{d})) \\ \vec{b}_{rad} &= \frac{\mu_0 \ddot{p}}{4\pi c R} (\hat{d} \wedge \hat{R}) \end{aligned}$$

This field can be written in terms of the unitary directions of electric and magnetic fields:

$$\begin{aligned} \vec{e}_{rad} &= \frac{\mu_0 \ddot{p} \sin \theta}{4\pi R} \hat{e} \\ \vec{b}_{rad} &= \frac{\mu_0 \ddot{p} \sin \theta}{4\pi c R} \hat{b} \end{aligned}$$

$$\hat{e} = (\cos \theta \cos \phi \hat{x} + \cos \theta \sin \phi \hat{y} - \sin \theta \hat{z})$$

$$\hat{b} = (-\sin \phi \hat{x} + \cos \phi \hat{y})$$

$$\hat{R} = (\sin \theta \cos \phi \hat{x} + \sin \theta \sin \phi \hat{y} + \cos \theta \hat{z})$$

$$\hat{d} = \hat{z}$$

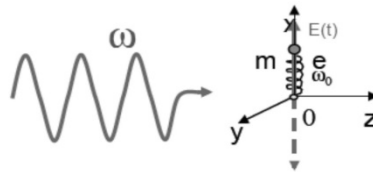




Appendix C

Lorentz model of electron oscillation

In the Lorentz model of electron oscillation, the action of an incident electric field on the atom is studied by considering the driven oscillatory motion of the electronic cloud, which gives rise to a dipole moment and a given polarization vector:



The basic assumptions of this model are the following.

- The electron moves with respect to its equilibrium position in a potential $U(x)$, which can be expanded into the Taylor series:

$$U(x) = U(0) + x \left(\frac{dU}{dx} \right)_{x=0} + \frac{1}{2} x^2 \left(\frac{d^2U}{dx^2} \right)_{x=0} + \frac{1}{3!} x^3 \left(\frac{d^3U}{dx^3} \right)_{x=0} + \dots$$

- The electron motion is damped by the effect of radiation losses, which are phenomenologically included in the model through a friction term:

$$\vec{F}_{fr} = -b \frac{d\vec{x}(t)}{dt}$$

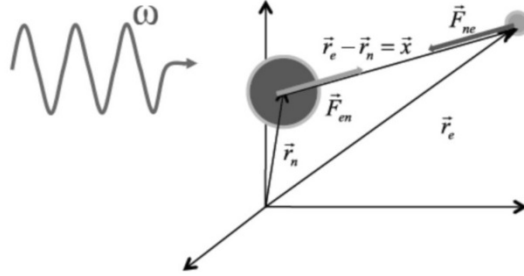
- The electron motion is driven externally by the electric field (a more refined model also includes the magnetic force term, which may be neglected under the electric dipole approximation in the presence of the electric field effect. The inclusion of the magnetic field introduces a nonlinear response):

$$\vec{F}_{ext} = -q \cdot \vec{E}$$

(!) q is the proton charge $q = 1.602 \cdot 10^{-19} \text{ C}$.



To obtain the basic equations governing the electronic motion, we start with the following model for the interaction of a system (electron–proton) with an incident electric field:



Centre of mass: $\vec{R} = \frac{m_e \vec{r}_e + m_n \vec{r}_n}{m_e + m_n}$

Relative electron position with respect to the nucleus (proton): $\vec{x} = \vec{r}_e - \vec{r}_n$

Equations of motion under the applied field:

$$m_n \frac{d^2 \vec{r}_n}{dt^2} = q \vec{E}(\vec{r}_n, t) + \vec{F}_{en}(\vec{x})$$

$$m_e \frac{d^2 \vec{r}_e}{dt^2} = -q \vec{E}(\vec{r}_e, t) + \vec{F}_{ne}(\vec{x})$$

where $\vec{F}_{en}(\vec{x}) = -\vec{F}_{ne}(\vec{x})$ is the Coulomb force between the electron and the proton.

In terms of the centre of mass and relative electron position, these equations can be written as

$$M \frac{d^2 \vec{R}}{dt^2} = -q \left[\vec{E}(\vec{R} + m_n \vec{x}/M, t) - \vec{E}(\vec{R} - m_e \vec{x}/M, t) \right]$$

$$m \frac{d^2 \vec{x}}{dt^2} = -q \left[\frac{m_n}{M} \vec{E}(\vec{R} + m_n \vec{x}/M, t) + \frac{m_e}{M} \vec{E}(\vec{R} - m_e \vec{x}/M, t) \right] + \vec{F}_{ne}(\vec{x})$$

where M is the total mass $M = m_n + m_e$, and m is the reduced mass of the system $m = m_e m_n / M \approx m_e$.

In the optical domain, the variation of the field amplitude over distances much smaller than the wavelength (typical atomic and molecular dimensions) can be considered very small, and we can expand the electric field in a Taylor series expansion around the centre of mass position:



$$M \frac{d^2 \vec{R}}{dt^2} \approx -q \vec{x} \vec{\nabla} \vec{E}(\vec{R}, t) = \vec{\phi} \cdot \vec{\nabla} \vec{E}(\vec{R}, t)$$

$$m \frac{d^2 \vec{x}}{dt^2} \approx -q \vec{E}(\vec{R}, t) + \vec{F}_{ne}(\vec{x}) + \frac{m_n - m_e}{M} \vec{\phi} \cdot \vec{\nabla} \vec{E}(\vec{R}, t)$$

In the electric dipole approximation, the field is constant over the atom dimensions, giving us the term $\vec{\nabla} \vec{E}(\vec{R}, t) \approx 0$. Under this assumption, the centre of mass is not changed by the electric field, and the equation of motion for the electron's position simplifies to

$$m \frac{d^2 \vec{x}}{dt^2} = -q \vec{E}(\vec{R}, t) + \vec{F}_{ne}(\vec{x})$$

When the electron is perturbed by the action of the external field, the effect of the internal force (Coulomb interaction) between the electron and the nucleus can be considered under the assumption that the electron responds to the external field as if it were bound to its nucleus by the harmonic potential $U(x)$, defined above. Thus:

$$\vec{F}_{ne}(\vec{x}) = -\vec{\nabla} U(x) = -x \left(\frac{d^2 U}{dx^2} \right)_{xo} - \frac{1}{2} x^2 \left(\frac{d^3 U}{dx^3} \right)_{xo} + \dots$$

When the external electric field is weak enough, the interaction is linear and only the first term is relevant. For stronger fields, the electron is subjected to nonlinear interactions, and more terms should be included. Since the derivatives are evaluated around the equilibrium position of the undisturbed electron motion, which corresponds to a stable equilibrium point, the second derivative is positive and we write it as

$$\left(\frac{d^2 U}{dx^2} \right)_{xo} = m \omega_o^2$$

As discussed previously, the electron motion is damped by the effect of radiation losses, which are included phenomenologically in this classical model through a friction term. Putting all these together, we arrive at the final expression for the Lorenz model in the electric dipole approximation:

$$m \frac{d^2 \vec{x}}{dt^2} = -q \vec{E}(\vec{R}, t) - m \omega_o^2 \vec{x} - b \frac{d\vec{x}}{dt}$$

This is usually written in the form

$$\frac{d^2 \vec{x}}{dt^2} + \gamma \frac{d\vec{x}}{dt} + \omega_o^2 \vec{x} = \frac{-q \vec{E}(\vec{R}, t)}{m}$$

with $\gamma = b / m$.





Appendix D

Basic properties of Fourier transforms

Given the complex function $E(t)$, it can be expressed as a superposition in terms of exponential complex functions:

$$E(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}(\omega) \exp(-i\omega t) d\omega = \int_{-\infty}^{\infty} \tilde{E}(f) \exp(-i2\pi f t) df$$

The function $E(\omega)$ corresponds to the weight at each frequency and is the Fourier transform of $E(t)$, whose argument is the frequency ω (rad/s) or f (Hz). The relation between $E(t)$ and $E(\omega)$ can be inverted to obtain

$$\begin{aligned} \tilde{E}(\omega) &= FT[E(t)] = \int_{-\infty}^{\infty} E(t) \exp(i\omega t) dt \\ E(t) &= FT^{-1}[\tilde{E}(\omega)] \end{aligned}$$

Basic properties

- **Linearity:** Given two complex functions $G(t)$ and $H(t)$ and two scalars a, b , then

$$FT[aG(t) + bH(t)] = aFT[G(t)] + bFT[H(t)] = a\tilde{G}(\omega) + b\tilde{H}(\omega)$$

- **Similarity:**

$$FT[G(at)] = \frac{1}{|a|} \tilde{G}\left(\frac{\omega}{a}\right)$$

- **Time/frequency shift:**

$$\begin{aligned} FT[G(t - a)] &= \tilde{G}(\omega) \exp(i\omega a) \\ FT^{-1}[\tilde{G}(\omega - b)] &= G(t) \exp(-ibt) \end{aligned}$$



- **Parseval's theorem:**

$$\int_{-\infty}^{\infty} |G(t)|^2 dt = \int_{-\infty}^{\infty} |\tilde{G}(\omega)|^2 d\omega$$

- **Convolution theorem:**

$$FT[G(t) * H(t)] = FT[G(t)]FT[H(t)] = \tilde{G}(\omega)\tilde{H}(\omega)$$

where the symbol * refers to the convolution operation

$$G(t) * H(t) = \int G(t')H(t-t')dt'$$

- **Fourier integral theorem:**

$$FT[FT^{-1}[\tilde{G}(\omega)]] = FT[G(t)] = \tilde{G}(\omega)$$

$$FT^{-1}[FT[G(t)]] = FT^{-1}[\tilde{G}(\omega)] = G(t)$$

Basic Fourier –Transform pairs

G(t)

G(ω)

$$\cos(\omega_o t)$$

$$\frac{1}{2}[\delta(\omega - \omega_o) + \delta(\omega + \omega_o)]$$

$$rect(t) = \begin{cases} 1 \rightarrow |t| \leq 1/2 \\ 0 \rightarrow |t| > 1/2 \end{cases}$$

$$\text{sinc}(\omega/2) = \frac{\sin(\omega/2)}{\omega/2}$$

$$\exp(-t^2)$$

$$\sqrt{\pi} \exp(-\omega^2/4)$$

A useful integral related to Gaussian functions:

$$\int_{-\infty}^{\infty} e^{-Ax^2} e^{-2Bx} dx = \frac{\sqrt{\pi}}{\sqrt{A}} \exp\left(\frac{B^2}{A}\right); \quad \text{Re}(A) > 0$$







Appendix E

Diffraction pattern from a circular aperture

The diffraction pattern from a circular aperture is obtained from the general expression

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} e^{\frac{ikr^2}{2z}} \left[\tilde{E}(f_x, f_y, 0) \right]$$

$f_x = \lambda x / z$
 $f_y = \lambda y / z$

The transmittance function of a circular aperture of radius r_o is

$$circ(r / r_o) = \begin{cases} 1 \rightarrow r \leq r_o \\ 0 \rightarrow r > r_o \end{cases}$$

Thus, the field distribution if the aperture is illuminated by a uniform plane wave is $E(x, y, 0) = Acirc(r/r_o)$.

Let us first calculate the Fourier transform of the circular aperture:

$$\tilde{E}(f_x, f_y, 0) = \iint Acirc(r / r_o) e^{-i2\pi(f_x x + f_y y)} dx dy$$

To solve this integral, it is better to use polar coordinates:

$$\begin{aligned} x &= r \cos \theta & \rightarrow & \quad f_x = f_r \cos \phi \\ y &= r \sin \theta & \rightarrow & \quad f_y = f_r \sin \phi \end{aligned}$$

$$\tilde{E}(f_r, 0) = \iint Acirc(r / r_o) e^{-i2\pi f_r (\cos \theta \cos \phi + \sin \theta \sin \phi)} r dr d\theta$$

$$\tilde{E}(f_r, 0) = \int_0^\infty Acirc(r / r_o) r dr \int_0^{2\pi} e^{-i2\pi f_r (\cos \theta \cos \phi + \sin \theta \sin \phi)} d\theta$$

The second integral is solved in terms of the zero-order Bessel function:

$$\int_0^{2\pi} e^{-i2\pi f_r \cos(\theta - \phi)} d\theta = 2\pi J_0(2\pi f_r)$$



Inserting this solution, we have

$$\tilde{E}(f_r, 0) = \int_0^{r_0} A r 2\pi J_0(2\pi r f_r) dr$$

This is solved by applying the following change in variables:

$$x = 2\pi r f_r \rightarrow r = \frac{x}{2\pi f_r} \rightarrow dr = \frac{dx}{2\pi f_r}$$

$$\tilde{E}(f_r, 0) = \int_0^{2\pi f_r r_0} A \frac{x}{f_r} J_0(x) \frac{dx}{2\pi f_r} = \frac{A}{2\pi f_r^2} \int_0^{x_0=2\pi f_r r_0} x J_0(x) dx = \frac{A r_0}{f_r} J_1(2\pi r_0 f_r)$$

where we have used the following solution in terms of first-order Bessel functions:

$$\int_0^{x_0} x J_0(x) dx = x_0 J_1(x_0)$$

Thus, the Fraunhofer diffraction pattern of the circular aperture is expressed as

$$E(x, y, z) = \frac{e^{ikz}}{i\lambda z} e^{\frac{ikr^2}{2z}} \left[\tilde{E}(f_x, f_y, 0) \right]_{\substack{f_x = \lambda x / z \\ f_y = \lambda y / z}} = \frac{A e^{ikz}}{i\lambda z} e^{\frac{ikr^2}{2z}} \left[\frac{r_0}{f_r} J_1(2\pi r_0 f_r) \right]_{f_r = r / \lambda z}$$

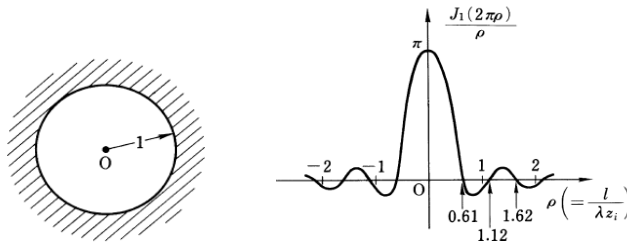
$$E(x, y, z) = \frac{A e^{ikz}}{i\lambda z} e^{\frac{ikr^2}{2z}} r_0^2 \frac{J_1(2\pi r_0 r / \lambda z)}{r r_0 / \lambda z}$$

The diffraction pattern is proportional to the Airy function: $\frac{J_1(2\pi x)}{x}$

This function has the first zeros at $x = \pm 0.61 \rightarrow r = \pm \frac{0.61 \lambda z}{r_0}$

Thus, the diffraction pattern of a circular aperture of radius r_0 is an Airy function, whose diameter between zeros is $\phi = 2r = \frac{1.22 \lambda z}{r_0}$.

Diffraction pattern from a circular aperture









Appendix F

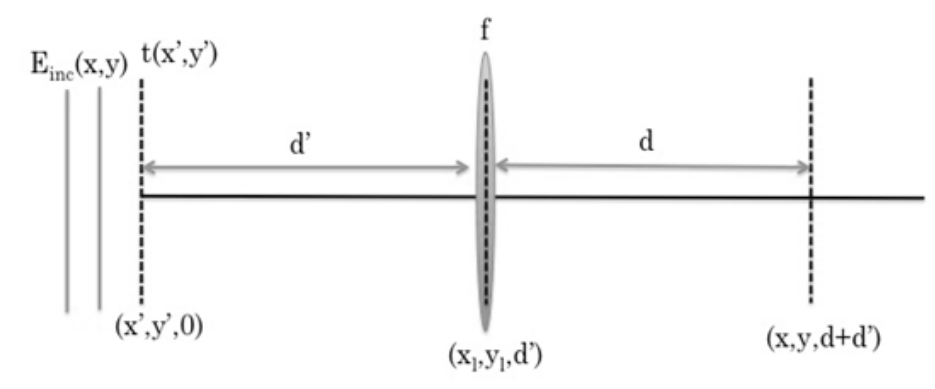
Fourier transforms by lenses

In this appendix, will show that we can obtain an optical implementation of the Fourier transform by using single lenses and applying Fresnel's diffraction formula.

Let us consider the propagation of a general field distribution that is generated by an object with arbitrary transmittance function illuminated by an incident field:

$$E(x', y') = E_{inc}(x', y', 0) \cdot t(x', y')$$

The object is placed at a distance d' to the left of a lens of focal length f , and we want to find out the resulting field at a distance d to the right of the lens.



After propagation of the initial field by a distance d' , the field distribution just before the lens is obtained using Fresnel's diffraction formula:

$$E(x_l, y_l, d') = \frac{e^{ikd'} e^{\frac{k}{2d'}(x_l^2 + y_l^2)}}{i\lambda_m d'} \int_{-\infty}^{\infty} \int E(x', y', 0) e^{i\frac{k}{2d'}(x'^2 + y'^2)} e^{-i\frac{2\pi}{\lambda_m d'}(x'x_l + y'y_l)} dx' dy'$$

where (x_l, y_l) correspond to points in the plane just before the lens. The lens introduces quadratic phase dependence into the optical beam, as discussed in Chapter 2 (this result was derived in the paraxial approximation, which also holds under Fresnel's



conditions). If we assume that there are no intensity variations introduced by the lens (no reflection and no absorption), the transmittance by the lens is

$$t(x_l, y_l) = e^{ikD} e^{-i\frac{k}{2f}(x_l^2 + y_l^2)}, \text{ with } D \text{ being the lens thickness.}$$

The field at the output of the lens is

$$E_{out}(x_l, y_l, d') = E(x_l, y_l, d) \cdot t(x_l, y_l) \cdot P(x_l, y_l)$$

where $P(x_l, y_l)$ is the pupil function of the lens and accounts for the finite size of the aperture. If we assume that the lens is large enough, the effect of this pupil can be ignored, so $P = 1$.

Applying Fresnel's formula again, the field leaving the lens at a distance d will be

$$E(x, y, d + d') = \frac{e^{ikd} e^{i\frac{k}{2d}(x^2 + y^2)}}{i\lambda_m d} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_{out}(x_l, y_l, d') e^{i\frac{k}{2d}(x_l^2 + y_l^2)} e^{-i\frac{2\pi}{\lambda_m d}(x_l x + y_l y)} dx_l dy_l$$

Using the previous expressions, the field can be written in terms of the initial field:

$$E(x, y, d + d') = \frac{e^{ikD} e^{ik(d'+d)}}{-\lambda_m^2 d d'} e^{i\frac{k}{2d}(x^2 + y^2)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x', y', 0) e^{i\frac{k}{2d'}(x'^2 + y'^2)} e^{-i\frac{2\pi}{\lambda_m d'}(x' x_l + y' y_l)} dx' dy' \right\} \\ e^{i\frac{k}{2}\left[\frac{1}{d} + \frac{1}{d'} - \frac{1}{f}\right](x_l^2 + y_l^2)} e^{-i\frac{2\pi}{\lambda_m d}(x_l x + y_l y)} dx_l dy_l$$

We define: $\frac{1}{w} = \frac{1}{d'} + \frac{1}{d} - \frac{1}{f}$ and change the order of integration to obtain

$$E(x, y, d + d') = \frac{e^{ikD} e^{ik(d'+d)}}{-\lambda_m^2 d' d} e^{i\frac{k}{2d}(x^2 + y^2)} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\frac{k}{2w}(x_l^2 + y_l^2)} e^{-i\frac{2\pi}{\lambda_m}\left[\left(\frac{x'}{d'} + \frac{x}{d}\right)x_l + \left(\frac{y'}{d'} + \frac{y}{d}\right)y_l\right]} dx_l dy_l \right\} \\ \cdot E(x', y', 0) e^{i\frac{k}{2d'}(x'^2 + y'^2)}$$

The integral in brackets is of the type $\int_{-\infty}^{\infty} e^{-Ax^2} e^{-2Bx} dx$, and it can be readily evaluated as

we did in previous chapters:

$$\int_{-\infty}^{\infty} e^{i\frac{k}{2w}x_l^2} e^{-i\frac{2\pi}{\lambda_m}\left(\frac{x'}{d'} + \frac{x}{d}\right)x_l} dx_l \cdot \int_{-\infty}^{\infty} e^{i\frac{k}{2w}y_l^2} e^{-i\frac{2\pi}{\lambda_m}\left(\frac{y'}{d'} + \frac{y}{d}\right)y_l} dy_l = i\lambda_m w e^{-i\frac{k w}{2}\left[\left(\frac{x'}{d'} + \frac{x}{d}\right)^2 + \left(\frac{y'}{d'} + \frac{y}{d}\right)^2\right]}$$

Expanding the square brackets and rearranging the terms leads to the expression



$$E(x, y, d + d') = \frac{w}{i\lambda_m d d'} e^{ik(D+d+d')} e^{i\frac{k}{2d}(1-\frac{w}{d})(x^2+y^2)} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x', y', 0) e^{i\frac{k}{2d'}(1-\frac{w}{d'})} e^{-i2\pi(f_x x' + f_y y')} dx' dy' \Big|_{f_x = \frac{xw}{\lambda_m d d'}, f_y = \frac{yw}{\lambda_m d d'}}$$

This is a general result relating the fields at $z = 0$ and $z = d + d'$. When the plane of observation after the lens is the focal plane of the lens (that is, when $d = f$) we obtain the important result

$$d = f \rightarrow w = d'$$

$$E(x, y, d + d') = \frac{w e^{ik(D+d'+d)}}{i\lambda_m d' d} e^{i\frac{k}{2f}(1-\frac{d'}{f})(x^2+y^2)} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x', y', 0) e^{-i2\pi(f_x x' + f_y y')} dx' dy' \Big|_{f_x = \frac{x}{\lambda_m f}, f_y = \frac{y}{\lambda_m f}}$$

The field at the focal plane of a lens corresponds to the Fourier transform of the incident field, which is evaluated at $f_x = x/\lambda_m f$; $f_y = y/\lambda_m f$ and multiplied by a quadratic phase factor.

In the particular case $d = d' = f$, the quadratic factor cancels and we obtain

$$\boxed{E(x, y, f + f) = A \tilde{E}(f_x, f_y, 0) \Big|_{f_x = x/\lambda_m f, f_y = y/\lambda_m f}} \quad A = \frac{w e^{ik(D+d'+d)}}{i\lambda_m d' d}$$





Bibliography

This section includes a list of reference books that can be used for further reading and to gain a deeper understanding of the different subjects within the field of photonics.

General reference books in optics and photonics

These references include classical reference text books in optics and photonics, all of them covering the fundamental aspects of the field.

- Akhmanov, S.A., Nikitin, S.Yu. (1997). Physical Optics. Clarendon Press Oxford.
- Born, M., Wolf, E. (1999). Principles of Optics. Pergamon Press.
- Feynman, R., Leighton, R., Sands, M. (2010). The Feynman Lectures on Physics. New Millenium Edition. Basic Books (Perseus Books LLC).
- Haken, H. (1986). Light (Volume 1): Waves, Photons, Atoms. North-Holland Physics Publishing.
- Hecht, E. (2016). Optics. Pearson.
- Iizuka, K. (2019). Engineering Optics. Springer.
- Ronchi, V. (1970). The Nature of Light. An Historical Survey. Harvard University Press.
- Saleh, B.E.A., Teich, M.C. (2007). Fundamentals of Photonics. John Wiley&Sons.
- Yariv, A. (1989). Quantum Electronics. John Wiley&Sons.

Advanced books on specialized topics

This list includes some advanced books for expanding on the basic concepts introduced in this text. As an introduction to even more concepts not described in the course, it



should be valuable for those readers willing to delve into the different issues covered here.

- Cohen-Tannoudji, C., Dupont-Roc, J., Grynberg, G. (1989). *Photons & Atoms. Introduction to Quantum Electrodynamics*. John Wiley & Sons.
- Diels, J.-C., Rudolph, W. (2006). *Ultrashort Laser Pulse Phenomena*. Elsevier Inc.
- Goodman, J.W. (2005). *Introduction to Fourier Optics*. Roberts & Company Publishers.
- Goodman, J.W. (2015). *Statistical Optics*. John Wiley & Sons.
- Grynberg, G., Aspect, A., Fabre, C. (2010). *Introduction to Quantum Optics. From the semi-classical approach to quantized light*. Cambridge University Press.
- Haken, H. (1986). *Light (Volume 2): Laser Light Dynamics*, North-Holland Physics Publishing.
- Jackson, J.D. (1962) *Classical Electrodynamics*. John Wiley & Sons.
- Loudon, R. (2000). *The Quantum Theory of Light*. Oxford University Press.
- Maier, S.A. (2007). *Plasmonics: Fundamentals and Applications*, Springer.
- Mandel, L., Wolf, E. (1995). *Optical Coherence and Quantum Optics*. Cambridge University Press.
- Novotny, L., Hecht, B. (2007). *Principles of Nano-Optics*. Cambridge University Press.
- Siegman, A.E. (196). *Lasers*. University Science Books.

Handbooks of Optics

Finally, the handbooks provide a huge source of succinctly presented information. They can be used to as an introduction into a particular subject and for general reference.

- Bass, M. (Editor in chief). (2010). *Handbook of Optics. (Volumes 1 to 5)*. McGraw Hill.
- Palik, E.D. (1998). *Handbook of Optical Constants of Solids*, Academic Press.
- Pöpp, J., Tuchin, V.V., Chiou, Heinemann., (Editors). (2012) *Handbook of Biophotonics (Volumes 1 to 3)*. John Wiley & Sons.
- Träger, F. (Editor). (2007). *Springer Handbook of Lasers and Optics*. Springer.

