Fundamentals of Modern Optics
Winter Term 2013/2014

Prof. Thomas Pertsch
Abbe School of Photonics
Friedrich-Schiller-Universität Jena

This script is based on the lecture series “Theoretische Optik” by Prof. Falk Lederer at the FSU Jena and was adapted to English by Prof. Stefan Skupin for the international education program of the Abbe School of Photonics.

Table of content
0. Introduction ........................................................................................................... 4
1. (Ray optics - geometrical optics, covered by lecture Introduction to Optical Modeling) ................................................................. 16
1.1 Introduction ........................................................................................................ 16
1.2 Postulates ......................................................................................................... 16
1.3 Simple rules for propagation of light ................................................................ 17
1.4 Simple optical components .......................................................................... 17
1.5 Ray tracing in inhomogeneous media (graded-index - GRIN optics) ............... 21
1.5.1 Ray equation .............................................................................................. 21
1.5.2 The eikonal equation ................................................................................ 23
1.6 Matrix optics .................................................................................................... 24
1.6.1 The ray-transfer-matrix ............................................................................. 24
1.6.2 Matrices of optical elements ...................................................................... 24
1.6.3 Cascaded elements .................................................................................... 25
2. Optical fields in dispersive and isotropic media .............................................. 26
2.1 Maxwell’s equations ....................................................................................... 26
2.1.1 Adaption to optics .................................................................................... 26
2.1.2 Temporal dependence of the fields ............................................................ 29
2.1.3 Maxwell’s equations in Fourier domain .................................................... 30
2.1.4 From Maxwell’s equations to the wave equation ..................................... 30
2.1.5 Decoupling of the vectorial wave equation ............................................. 31
2.2 Optical properties of matter ......................................................................... 32
2.2.1 Basics ....................................................................................................... 33
2.2.2 Dielectric polarization and susceptibility .................................................. 36
2.2.3 Conductive current and conductivity ....................................................... 37
2.2.4 The generalized complex dielectric function .......................................... 39
2.2.5 Material models in time domain ............................................................... 43
2.3 The Poynting vector and energy balance ...................................................... 44
2.3.1 Time averaged Poynting vector .............................................................. 44
2.3.2 Time averaged energy balance ............................................................... 46
2.4 Waves in homogeneous isotropic media ...................................................... 49
2.4.1 Transversal waves .................................................................................... 49
2.4.2 Longitudinal waves .................................................................................. 50
2.4.3 Plane wave solutions in different frequency regimes .............................. 51
2.4.4 Time averaged Poynting vector of plane waves ....................................... 56
2.5 Beams and pulses - analogy of diffraction and dispersion ......................... 56
2.5.1 Diffraction of monochromatic beams in homogeneous isotropic media .... 58
2.5.2 Propagation of Gaussian beams ............................................................... 70
2.5.3 Gaussian optics ....................................................................................... 77
2.5.4 Gaussian modes in a resonator ............................................................... 81
2.5.5 Pulse propagation .................................................................................... 86
2.6 (The Kramers-Kronig relation, covered by lecture Structure of Matter) ......... 99
3. Diffraction theory ............................................................................................. 103
3.1 Interaction with plane masks ........................................................................ 103
3.2 Propagation using different approximations .............................................. 104
3.2.1 The general case - small aperture ......................................................... 104
3.2.2 Fresnel approximation (paraxial approximation) .................................... 104
3.2.3 Paraxial Fraunhofer approximation (far field approximation) ............... 105
3.2.4 Non-paraxial Fraunhofer approximation .............................................. 107
3.3 Fraunhofer diffraction at plane masks (paraxial) ......................................... 107
3.3.1 Fraunhofer diffraction pattern .............................................................. 107
3.4 Remarks on Fresnel diffraction .................................................................. 112
4. Fourier optics - optical filtering ..................................................................... 113
4.1 Imaging of arbitrary optical field with thin lens ......................................... 113
4.1.1 Transfer function of a thin lens .............................................................. 113
4.1.2 Optical imaging ..................................................................................... 114
4.2 Optical filtering and image processing ....................................................... 116
4.2.1 The 4f-setup .......................................................................................... 116
4.2.2 Examples of aperture functions .............................................................. 119
4.2.3 Optical resolution .................................................................................. 119
5. The polarization of electromagnetic waves .................................................. 122
5.1 Introduction .................................................................................................. 122
5.2 Polarization of normal modes in isotropic media ....................................... 122
5.3 Polarization states ....................................................................................... 123
6. Principles of optics in crystals ..................................................................... 125
6.1 Susceptibility and dielectric tensor ............................................................ 125
6.2 The optical classification of crystals ......................................................... 127
6.3 The index ellipsoid ...................................................................................... 128
6.4 Normal modes in anisotropic media .......................................................... 129
6.4.1 Normal modes propagating in principal directions ................................ 130
6.4.2 Normal modes for arbitrary propagation direction ................................ 131
6.4.3 Normal surfaces of normal modes ....................................................... 135
6.4.4 Special case: uniaxial crystals ............................................................... 137
7. Optical fields in isotropic, dispersive and piecewise homogeneous media ...... 140
7.1 Basics .......................................................................................................... 140
7.1.1 Definition of the problem ....................................................................... 140
7.1.2 Decoupling of the vectorial wave equation .......................................... 141
7.1.3 Interfaces and symmetries ................................................................. 142
7.1.4 Transition conditions ............................................................................ 142
7.2 Fields in a layer system → matrix method ................................................ 143
7.2.1 Fields in one homogeneous layer ....................................................... 143
7.2.2 The fields in a system of layers ............................................................. 145
7.3 Reflection – transmission problem for layer systems .................................. 147
7.3.1 General layer systems ......................................................................... 147
7.3.2 Single interface .................................................................................... 153
7.3.3 Periodic multi-layer systems - Bragg-mirrors - 1D photonic crystals .... 160
7.3.4 Fabry-Perot-resonators ....................................................................... 167
7.4 Guided waves in layer systems ................................................................. 173
7.4.1 Field structure of guided waves ........................................................... 173
0. Introduction

- 'optique' (Greek) ➔ lore of light ➔ 'what is light'?
- Is light a wave or a particle (photon)?

  **D.J. Lovell, Optical Anecdotes**

- Light is the origin and requirement for life ➔ photosynthesis
- 90% of information we get is visual

A) Origin of light
- atomic system ➔ determines properties of light (e.g. statistics, frequency, line width)
- optical system ➔ other properties of light (e.g. intensity, duration, ...)
- invention of laser in 1958 ➔ very important development

  **Schawlow and Townes, Phys. Rev. (1958).**

- laser ➔ artificial light source with new and unmatched properties (e.g. coherent, directed, focused, monochromatic)
- applications of laser: fiber-communication, DVD, surgery, microscopy, material processing, ...
Fiber laser: Limpert, Tünnermann, IAP Jena, ~10kW CW (world record)

B) Propagation of light through matter

- light-matter interaction

<table>
<thead>
<tr>
<th>dispersion</th>
<th>diffraction</th>
<th>absorption</th>
<th>scattering</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequency</td>
<td>spatial</td>
<td>center of</td>
<td>wavelength</td>
</tr>
<tr>
<td>spectrum</td>
<td>frequency</td>
<td>frequency</td>
<td>spectrum</td>
</tr>
</tbody>
</table>

- matter is the medium of propagation \( \Rightarrow \) the properties of the medium (natural or artificial) determine the propagation of light
- light is the means to study the matter (spectroscopy) \( \Rightarrow \) measurement methods (interferometer)

- design media with desired properties: glasses, polymers, semiconductors, compounded media (effective media, photonic crystals, meta-materials)

C) Light can modify matter

- light induces physical, chemical and biological processes
- used for lithography, material processing, or modification of biological objects (bio-photonics)

Hole “drilled” with a fs laser at Institute of Applied Physics, FSU Jena.

Two-dimensional photonic crystal membrane.
D) Optics in our daily life

John reached over and shut off the alarm clock. He turned on the lights and got up. Downstairs, he began to make his morning coffee and turned on the television to check the weather forecast. Checking the time on the kitchen clock, he poured his coffee and went to the solarium to sit and read the newspaper.

Upstairs, the kids were getting ready for school. Julie was listening to a favorite song while getting dressed. Stevie felt sick, so his mother, Sarah, checked his temperature. Julie would go to school, but Stevie would stay home.

John drove to work in his new car, a high-tech showcase. He drove across a bridge, noticing the emergency telephones along the side of the freeway. He encountered traffic signals, highway signs, and a police officer scanning for speeders.

Awaiting John in his office were several telephone messages and a fax. He turned on his computer, checked some reference data on a CD-ROM, and printed it to look at later. After copying some last-minute handouts, he went to the conference room to make a presentation.

Meanwhile, Julie was walking to school. As she passed the neighbors’ house, a security light came on. On the next block she passed a construction site for a new apartment building, then a block of medical offices. A few blocks away was the factory where her uncle worked.

At school, Julie’s first class was biology. The students looked at microbes in water samples they had collected on a nature walk the previous day. On the wall, they had also done some birdwatching and taken still and video pictures of the plants and wildlife. The teacher put on her glasses to read Julie’s lab report.

At lunchtime, John left his office to do some grocery shopping. At the checkout counter he paid with a credit card. Among his purchases were a bag of apples, a bottle of wine, and a carton of milk. Each was labeled with a bar code.

At home, Stevie was watching a movie on the large-screen television. With her sick son occupied, Sarah connected her laptop computer to the office network. Modern technology let her do her work, despite having to stay home with the child—and at least John was still doing the shopping.

E) Optics in telecommunications

- transmitting data (Terabit/s in one fiber) over transatlantic distances

1000 m telecommunication fiber is installed every second.
F) Optics in medicine, life sciences

Figure 1. Rotation of an intracellular object inside *E. coli* dense plant cell using the rotating line tweezers. The rod-shaped structure was trapped using 25 mW power and rotated at a speed of 4 Hz. The direction of rotation is shown by arrow (a). Clockwise rotation by angles of 45° (b), 145° (c), and 235° (d). All the images were recorded with the same magnification.

G) Optical sensors and light sources
- new light sources to reduce energy consumption
- new projection techniques

Deutscher Zukunftspreis 2008 - IOF Jena + OSRAM
H) Micro- and nano-optics

- ultra small camera

Insect inspired camera system develop at Fraunhofer Institute IOF Jena

I) Relativistic optics

Figure 3. Two relativistic lasers: (a) Helios circa 1980 at LLNL was the first relativistic laser with \( z_0 \approx 1 \) at a millihertz repetition rate. [Courtesy of LLNL] (b) The \( \lambda^2 \) laser at the University of Michigan has an \( z_0 \approx 1 \) at a kilohertz repetition rate.

Figure 4. Relativistic rectification in plasma: (a) a high-intensity pulse before it enters the plasma and (b) the \( y \times \tilde{H} \) that pushes the first plasma electrons. The electrons drag the heavy ions behind them like a horse pulling a cart. The electrostatic field that is generated is almost as large as the transverse laser field.

Field is created within a plasma, the medium cannot be broken down further and can support these extremely high fields. This effect, predicted by Tajima and Dawson in 1980,\(^9\) had to await the development of relativistic intensity lasers to be demonstrated.\(^{14}\) Electron energies as high as 200 MeV (Ref. 15) have been demonstrated by use of a 50- TW laser focused on a gas jet. Because the electrons drag the ions, a beam of ions is also observed concomitant with the electron beam. The generation of protons with energy as high as 56 MeV has been reported by LLNL.\(^{15}\)
K) What is light?
- electromagnetic wave (c = 3*10^8 m/s)
- amplitude and phase $\rightarrow$ complex description
- polarization, coherence

<table>
<thead>
<tr>
<th>Region</th>
<th>Wavelength (nanometers)</th>
<th>Wavelength (centimeters)</th>
<th>Frequency (Hz)</th>
<th>Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radio</td>
<td>&gt; 10^9</td>
<td>&gt; 10</td>
<td>&lt; 3 x 10^9</td>
<td>&lt; 10^-5</td>
</tr>
<tr>
<td>Microwave</td>
<td>10^9 - 10^6</td>
<td>10 - 0.01</td>
<td>3 x 10^9 - 3 x 10^12</td>
<td>10^-3 - 0.01</td>
</tr>
<tr>
<td>Infrared</td>
<td>10^6 - 700</td>
<td>0.01 - 7 x 10^5</td>
<td>3 x 10^12 - 4.3 x 10^14</td>
<td>0.01 - 2</td>
</tr>
<tr>
<td>Visible</td>
<td>700 - 400</td>
<td>7 x 10^-9 - 4 x 10^-6</td>
<td>4.3 x 10^14 - 7.5 x 10^14</td>
<td>2 - 3</td>
</tr>
<tr>
<td>Ultraviolet</td>
<td>400 - 1</td>
<td>4 x 10^-5 - 10^-7</td>
<td>7.5 x 10^14 - 3 x 10^17</td>
<td>3 - 10^3</td>
</tr>
<tr>
<td>X-Rays</td>
<td>1 - 0.01</td>
<td>10^-7 - 10^-9</td>
<td>3 x 10^-7 - 3 x 10^-9</td>
<td>10^-3 - 10^9</td>
</tr>
<tr>
<td>Gamma Rays</td>
<td>&lt; 0.01</td>
<td>&lt; 10^-9</td>
<td>&gt; 3 x 10^-9</td>
<td>&gt; 10^9</td>
</tr>
</tbody>
</table>

L) Schematic of optics
- geometrical optics
  - $\lambda << $ size of objects $\rightarrow$ daily experiences
  - optical instruments, optical imaging
  - intensity, direction, coherence, phase, polarization, photons

- wave optics
  - $\lambda \approx $ size of objects $\rightarrow$ interference, diffraction, dispersion, coherence
  - laser, holography, resolution, pulse propagation
  - intensity, direction, coherence, phase, polarization, photons

- electromagnetic optics
  - reflection, transmission, guided waves, resonators
  - laser, integrated optics, photonic crystals, Bragg mirrors...
  - intensity, direction, coherence, phase, polarization, photons

- quantum optics
  - small number of photons, fluctuations, light-matter interaction
  - intensity, direction, coherence, phase, polarization, photons

- in this lecture
  - electromagnetic optics and wave optics
  - no quantum optics $\rightarrow$ advanced lecture
M) Literature

- Fundamental
  3. Hecht, 'Optik', Oldenbourg, 2001
  7. Sommerfeld, 'Optik'

- Advanced
  4. Karthe, Müller, 'Integrierte Optik', Teubner
  5. Diels, Rudolph, 'Ultrashort Laser Pulse Phenomena', Academic
  7. Snyder, Love, 'Optical Waveguide Theory', Chapman&Hall

1. (Ray optics - geometrical optics, covered by lecture Introduction to Optical Modeling)

The topic of "Ray optics – geometrical optics" is not covered in the course "Fundamentals of modern optics". This topic will be covered rather by the course "Introduction to optical modeling". The following part of the script which is devoted to this topic is just included in the script for consistency.

1.1 Introduction
- Ray optics or geometrical optics is the simplest theory for doing optics.
- In this theory, propagation of light in various optical media can be described by simple geometrical rules.
- Ray optics is based on a very rough approximation ($\lambda \to 0$, no wave phenomena), but we can explain almost all daily life experiences involving light (shadows, mirrors, etc.).
- In particular, we can describe optical imaging with ray optics approach.
- In isotropic media, the direction of rays corresponds to the direction of energy flow.

What is covered in this chapter?
- It gives fundamental postulates of the theory.
- It derives simple rules for propagation of light (rays).
- It introduces simple optical components.
- It introduces light propagation in inhomogeneous media (graded-index (GRIN) optics).
- It introduces paraxial matrix optics.

1.2 Postulates
A) Light propagates as rays. Those rays are emitted by light-sources and are observable by optical detectors.

B) The optical medium is characterized by a function $n(r)$, the so-called refractive index ($n(r) \geq 1$ - meta-materials $n(r) < 0$)

$$n = \frac{c}{c_n}$$

where $c_n$ is the speed of light in the medium.

C) optical path length ~ delay
   i) homogeneous media
   \[ \int n(r) \, ds \]
   ii) inhomogeneous media
   \[ \int \frac{1}{n(r)} \, ds \]
D) Fermat’s principle

\[ \frac{\delta}{\delta n(r)} \int ds = 0 \]

Rays of light choose the optical path with the shortest delay.

1.3 Simple rules for propagation of light

A) Homogeneous media
   - \( n = \text{const.} \rightarrow \text{minimum delay} = \text{minimum distance} \)
   - Rays of light propagate on straight lines.

B) Reflection by a mirror (metal, dielectric coating)
   - The reflected ray lies in the plane of incidence.
   - The angle of reflection equals the angle of incidence.

C) Reflection and refraction by an interface
   - Incident ray → reflected ray plus refracted ray
   - The reflected ray obeys b).
   - The refracted ray lies in the plane of incidence.

   - The angle of refraction \( \theta_2 \) depends on the angle of incidence \( \theta_1 \) and is given by Snell’s law:
     \[ n_1 \sin \theta_1 = n_2 \sin \theta_2 \]
   - no information about amplitude ratio.

1.4 Simple optical components

A) Mirror
   i) Planar mirror
      - Rays originating from \( P_1 \) are reflected and seem to originate from \( P_2 \).

   ii) Parabolic mirror
      - Parallel rays converge in the focal point (focal length \( f \)).
      - Applications: Telescope, collimator

   iii) Elliptic mirror
      - Rays originating from focal point \( P_1 \) converge in the second focal point \( P_2 \)

   iv) Spherical mirror
      - Neither imaging like elliptical mirror nor focusing like parabolic mirror
      - parallel rays cross the optical axis at different points
      - connecting line of intersections of rays → caustic
parallel, paraxial rays converge to the focal point \( f = (-R) / 2 \)

- for paraxial rays the spherical mirror acts as a focusing as well as an imaging optical element. paraxial rays emitted in point \( P_1 \) are reflected and converge in point \( P_2 \)

\[
\frac{1}{z_1} + \frac{1}{z_2} \approx \frac{2}{(-R)}
\]

(paraxial imaging: imaging formula and magnification)

\[
m = -\frac{z_2}{z_1}
\]

(Proof given in exercises)

B) Planar interface

Snell’s law: \( n_1 \sin \theta_1 = n_2 \sin \theta_2 \)

for paraxial rays: \( n_1 \theta_1 = n_2 \theta_2 \)

- external reflection \( (n_1 < n_2) \): ray refracted away from the interface
- internal reflection \( (n_1 > n_2) \): ray refracted towards the interface
- total internal reflection (TIR) for:

\[
\theta_2 = \frac{\pi}{2} \rightarrow \sin \theta_1 = \sin \theta_{\text{TIR}} = \frac{n_2}{n_1}
\]

C) Spherical interface (paraxial)

paraxial imaging

\[
m = -\frac{z_2}{z_1}
\]

(Proof: exercise)

\[
\theta_2 \approx \frac{n_1}{n_2} \theta_1 - n_2 - n_1 \frac{y}{R} (*)
\]

D) Spherical thin lens (paraxial)
two spherical interfaces \((R_1, R_2, \Delta)\) apply (*) two times and assume \(y=\text{const} (\Delta \text{ small})\)

\[
0_z \approx 0_i - \frac{y}{f}
\]

with focal length: \(\frac{1}{f} = (n - 1) \left( \frac{1}{R_1} - \frac{1}{R_2} \right)\)

\[
\frac{1}{z_1} + \frac{1}{z_2} \approx \frac{1}{f}
\]

( imaging formula) \(m = \frac{z_2}{z_1}\) (magnification)

(compare to spherical mirror)

### 1.5 Ray tracing in inhomogeneous media (graded-index - GRIN optics)

- \(n(r)\) - continuous function, fabricated by, e.g., doping
- curved trajectories \(\rightarrow\) graded-index layer can act as, e.g., a lens

#### 1.5.1 Ray equation

Starting point: we minimize the optical path or the delay (Fermat)

\[
\delta \int_A^n n(r) ds = 0
\]

computation:

\[
L = \int_A^n [n(r)] ds
\]

variation of the path: \(r(s) + \delta r(s)\)

\[
\delta L = \int_A^n \delta n ds + \int_A^n n \delta ds
\]

\[
\delta n = \nabla n \cdot \delta r
\]

\[
\delta ds = \sqrt{\left(\frac{dr}{ds} + d\delta r\right)^2} - \sqrt{\left(\frac{dr}{ds}\right)^2}
\]

\[
\approx ds \sqrt{1 + 2 \frac{dr}{ds} \cdot \frac{d\delta r}{ds} + \left(\frac{d\delta r}{ds}\right)^2} - ds
\]

\[
\approx ds \left( 1 + \frac{dr}{ds} \cdot \frac{d\delta r}{ds} - ds \right)
\]

\[
= ds \frac{dr}{ds} \frac{d\delta r}{ds}
\]

\[
\delta L = \int_A^n \left( \nabla n \cdot \delta r + n \frac{dr}{ds} \frac{d\delta r}{ds} \right) ds
\]

integration by parts and \(A, B\) fix

\[
\delta L = \int_A^n \left( \nabla n \delta r + n \frac{dr}{ds} \frac{d\delta r}{ds} \right) ds
\]

\[
\delta L = 0 \quad \text{for arbitrary variation}
\]

\[
\text{Possible solutions:}
\]

A) \(x(z), y(z)\) and \(ds = dz \sqrt{1 + (dx/dz)^2 + (dy/dz)^2}\)

- solve for \(x(z), y(z)\)
- paraxial rays \(\rightarrow\) \((ds \approx dz)\)
\[ \frac{d}{dz} n(x,y,z) \frac{dx}{dz} \approx \frac{dn}{dx} \]
\[ \frac{d}{dz} n(x,y,z) \frac{dy}{dz} \approx \frac{dn}{dy} \]

B) homogeneous media
- straight lines

C) graded-index layer \( n(y) \) - paraxial, SELFOC

paraxial \rightarrow \frac{dy}{dz} \ll 1 \text{ and } dz \approx ds

\[ n(y) = n_0^2 \left( 1 - \alpha^2 y^2 \right) \Rightarrow n(y) \approx n_0 \left( 1 - \frac{1}{2} \alpha^2 y^2 \right) \text{ for } \alpha \ll 1 \]

\[ \frac{d}{ds} n(y) \frac{dy}{ds} \approx \frac{d}{dz} n(y) \frac{dy}{dz} \approx n(y) \frac{d^2 y}{dz^2} = \frac{1}{n(y)} \frac{dn(y)}{dy} \]

for \( n(y) - n_0 \ll 1 \):

\[ \frac{d^2 y}{dz^2} = -\alpha^2 y \rightarrow \]

\[ y(z) = y_0 \cos \alpha z + \frac{\theta_0}{\alpha} \sin \alpha z \]

\[ \theta(z) = \frac{dy}{dz} = -y_0 \alpha \sin \alpha z + \theta_0 \cos \alpha z \]

1.5.2 The eikonal equation
- bridge between geometrical optics and wave
- eikonal \( S(r) = \text{constant} \rightarrow \) planes perpendicular to rays
- from \( S(r) \) we can determine direction of rays \( \sim \) \( \text{grad} \ S(r) \) (like potential)

\[ \left( \text{grad} S(r) \right)^2 = n(r)^2 \]

Remark: it is possible to derive Fermat’s principle from eikonal equation
- geometrical optics: Fermat’s or eikonal equation

\[ S(r_0) - S(r_1) = \int_{r_0}^{r_1} \text{grad} S(r) \cdot ds = \int_{r_0}^{r_1} n(r) ds \]

eikonal \rightarrow \text{optical path length} \sim \text{phase of the wave}

1.6 Matrix optics
- technique for paraxial ray tracing through optical systems
- propagation in a single plane only
- rays are characterized by the distance to the optical axis \((y)\) and their inclination \((\theta)\) \rightarrow two algebraic equation \(2 \times 2\) matrix

Advantage: we can trace a ray through an optical system of many elements by multiplication of matrices.

1.6.1 The ray-transfer-matrix

\[ y_2 = Ay_1 + B \theta_1 \]
\[ \theta_2 = Cy_1 + D \theta_1 \]

A=0: same \( \theta_1 \rightarrow \) same \( y_2 \rightarrow \) focusing
D=0: same \( y_1 \rightarrow \) same \( \theta_2 \rightarrow \) collimation

1.6.2 Matrices of optical elements
A) free space

\[ M = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \]

B) refraction on planar interface
2. Optical fields in dispersive and isotropic media

2.1 Maxwell’s equations

Our general starting point is the set of Maxwell’s equations. They are the basis of the electromagnetic approach to optics developed in this lecture.

2.1.1 Adaption to optics

The notation of Maxwell’s equations is different for different disciplines of science and engineering which rely on these equations to describe electromagnet phenomena at different frequency ranges. Even though Maxwell’s equations are valid for all frequencies, the physics of light matter interaction is different for different frequencies. Since light matter interaction must be included in the Maxwell’s equations to solve them consistently, different ways have been established how to write down Maxwell’s equations for different frequency ranges. Here we follow a notation which was established for a convenient notation at frequencies close to visible light.

Maxwell’s equations (macroscopic)

In a rigorous way the electromagnetic theory is developed starting from the properties of electromagnetic fields in vacuum. In vacuum one could write down Maxwell’s equations in there so-called pure microscopic form, which includes the interaction with any kind of matter based on the consideration of point charges. Obviously this is inadequate for the description of light in condensed matter, since the number of point charges which would need to be taken into account to describe a macroscopic object, would exceed all imaginable computational resources.

To solve this problem one uses an averaging procedure, which summarizes to influence of many point charges on the electromagnetic field in a homogeneously distributed response of the solid state on the excitation by the light. In turn, also the electromagnetic fields are averaged over some adequate volume. For optics this procedure is justified, since any kind of available experimental detector could not resolve the very fine spatial details of the fields in between the point charges, e.g. ions or electrons, which are lost by this averaging.

These averaged electromagnetic equations have been rigorously derived in a number of fundamental text books on electro-dynamic theory. Here we will not redo this derivation. We will rather start directly from the averaged Maxwell’s equations equation.

\[
\begin{align*}
\text{rot} \mathbf{E}(r,t) &= -\frac{\partial \mathbf{B}(r,t)}{\partial t} \\
\text{div} \mathbf{D}(r,t) &= \rho_{\text{ext}}(r,t) \\
\text{rot} \mathbf{H}(r,t) &= \mathbf{j}_{\text{ext}}(r,t) + \frac{\partial \mathbf{D}(r,t)}{\partial t} \\
\text{div} \mathbf{B}(r,t) &= 0
\end{align*}
\]

- electric field \( \mathbf{E}(r,t) \) [V/m]
- magnetic flux density \( B(r,t) \) [Vs/m²] or [tesla]
- dielectric flux density \( D(r,t) \) [As/m²]
- magnetic field \( H(r,t) \) [A/m]
- external charge density \( \rho_{\text{ext}}(r,t) \) [As/m³]
- macroscopic current density \( j_{\text{makr}}(r,t) \) [A/m²]

**Auxiliary fields**

The "cost" of the introduction of macroscopic Maxwell’s equations is the occurrence of two additional fields, the dielectric flux density \( D(r,t) \) and the magnetic field \( H(r,t) \). These two fields are related to the electric field \( E(r,t) \) and magnetic flux density \( B(r,t) \) by two other new fields.

\[
\begin{align*}
D(r,t) &= \varepsilon_0 E(r,t) + P(r,t) \\
H(r,t) &= \frac{1}{\mu_0} [B(r,t) - M(r,t)]
\end{align*}
\]

- dielectric polarization \( P(r,t) \) [As/m²],
- magnetic polarization \( M(r,t) \) [Vs/m²] (magnetization)
- electric constant (vacuum permittivity)
  \( \varepsilon_0 = \frac{1}{\mu_0 c^2} \approx 8.854 \times 10^{-12} \text{ As/Vm} \)
- magnetic constant (vacuum permeability)
  \( \mu_0 = 4\pi \times 10^{-7} \text{ Vs/Am} \)

**Light matter interaction**

In order to solve this set of equations, i.e. Maxwell’s equations and auxiliary field equations one needs to connect the dielectric flux density \( D(r,t) \) and the magnetic field \( H(r,t) \) to the electric field \( E(r,t) \) and the magnetic flux density \( B(r,t) \). This is achieved by modeling the material properties by introducing the material equations.

- The effect of the medium gives rise to polarization \( P(r,t) = f[E] \) and magnetization \( M(r,t) = f[B] \).
- In optics, we generally deal with non-magnetizable media → \( M(r,t) = 0 \)

Furthermore we need to introduce sources of the fields into our model. This is achieved by the so-called source terms which are inhomogeneities and hence they define unique solutions of the equations.

- free charge density \( \rho_{\text{ext}}(r,t) \) [As/m³]
- macroscopic current density \( j_{\text{makr}}(r,t) = j_{\text{cond}}(r,t) + j_{\text{conv}}(r,t) \) [A/m²]
- conductive current density \( j_{\text{cond}}(r,t) = f[E] \)
- convective current density \( j_{\text{conv}}(r,t) = \rho_{\text{ext}}(r,t) V(r,t) \)
- In optics, we generally have no free charges which change at speeds comparable to the frequency of light:
  \( \rho_{\text{ext}}(r,t) = 0 \implies j_{\text{conv}}(r,t) = 0 \)

- With the above simplifications, we can formulate Maxwell’s equations in the context of optics:

\[
\begin{align*}
\nabla \times E(r,t) &= -\mu_0 \frac{\partial H(r,t)}{\partial t} & \varepsilon_0 \nabla \times E(r,t) &= -\nabla P(r,t) \\
\nabla \times H(r,t) &= j(r,t) + \frac{\partial P(r,t)}{\partial t} + \frac{\varepsilon_0}{\varepsilon_0} \frac{\partial E(r,t)}{\partial t} & \nabla \cdot H(r,t) &= 0
\end{align*}
\]

- In optics, the medium (or more precisely the mathematical material model) determines the dependence of the polarization on the electric field \( P(E) \) and the dependence of the (conductive) current density on the electric field \( j(E) \).

- Once we have specified these relations, we can solve Maxwell’s equations consistently.

**Example:**

- In vacuum, both polarization and current density are zero, and we can solve Maxwell’s equations directly (most simple material model).

**Remark:**

- We can define a bound charge density
  \( \rho_b(r,t) = -\nabla \cdot P(r,t) \)
- and a bound current density
  \( j_b(r,t) = \frac{\partial P(r,t)}{\partial t} \)

- This essentially means that we can describe the same physics in two different ways (see generalized complex dielectric function below).

**Complex field formalism:**
Maxwell’s equations are also valid for complex fields and are easier to solve.

This fact can be exploited to simplify calculations, because it is easier to deal with complex exponential functions \([\exp(ix)]\) than with trigonometric functions \([\cos(x)\text{ and }\sin(x)]\).

In this lecture

- convention in this lecture

real physical field:

complex mathematical representation:

They are related by

\[
E_c(r,t) = \frac{1}{2} \left[ E(r,t) + E^*(r,t) \right] = \text{Re}[E(r,t)]
\]

Remark: This relation can be defined differently in different textbooks.

This means in general: For calculations we use the complex fields \([E(r,t)]\) and for physical results we go back to real fields by simply omitting the imaginary part. This works because Maxwell’s equations are linear and no multiplications of fields occur.

Therefore, be careful when multiplications of fields are required \(\rightarrow\) go back to real quantities before! (relevant for, e.g., calculation of Poynting vector, see Chapter below).

### 2.1.2 Temporal dependence of the fields

When it comes to time dependence of the electromagnetic field, we can distinguish two different types of light:

#### A) Monochromatic light \(\leftrightarrow\) stationary fields

- Harmonic dependence on temporal coordinate
- \(\sim \exp(-i\omega t)\) \(\rightarrow\) phase is fixed \(\rightarrow\) coherent, infinite wave train e.g.:
  \[
  E(r,t) = E(r) \exp(-i\omega t)
  \]

- Monochromatic light approximates very well the typical output of a continuous wave (CW) laser. Once we know the frequency we have to compute the spatial dependence of the (stationary) fields only.

#### B) Polychromatic light \(\leftrightarrow\) non-stationary fields

- Finite wave train
- With the help of Fourier transformation we can decompose the fields into infinite wave trains and use all the results from case A) (see next section)

### 2.1.3 Maxwell’s equations in Fourier domain

We want to plug the Fourier decompositions of our fields into Maxwell’s equations in order to get a more simple description. For this purpose, we need to know how a time derivative transforms into Fourier space. Here we used integration by parts:

\[
\frac{1}{2\pi^2} \int_{-\infty}^{\infty} dt \left[ \frac{\partial}{\partial t} E(r,t) \right] \exp(i\omega t) = -\omega \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \left[ \frac{\partial}{\partial t} E(r,t) \right] \exp(i\omega t) = -\omega E(r,\omega)
\]

\(\Rightarrow\) rule: 

\[
\frac{\partial}{\partial t} \rightarrow -i\omega
\]

Now we can write Maxwell’s equations in Fourier domain:

\[
\begin{align*}
\text{rot} \bar{E}(r,\omega) &= i\omega \varepsilon_0 \bar{H}(r,\omega) \\
\varepsilon_0 \text{div} \bar{E}(r,\omega) &= -\text{div} \bar{P}(r,\omega)
\end{align*}
\]

\[
\begin{align*}
\text{rot} \bar{H}(r,\omega) &= j(r,\omega) - i\omega \bar{P}(r,\omega) - i\omega \varepsilon_0 \bar{E}(r,\omega) \\
\text{div} H(r,\omega) &= 0
\end{align*}
\]

### 2.1.4 From Maxwell’s equations to the wave equation

Maxwell's equations provide the basis to derive all possible mathematical solutions. However we are interested in radiation fields which can be described more easily by an adapted equation, which is the wave equation. From Maxwell’s equations it is straightforward to derive the wave equation by using the two curl equations.

#### A) Time domain derivation

We start from applying the curl a second time on \(\text{rot} E(r,t) = \ldots\) and substitute \(\text{rot} \bar{H}\)

\[
\begin{align*}
\text{rot} \text{rot} E(r,t) &= -\mu_0 \text{rot} \frac{\partial}{\partial t} \bar{H}(r,t) \\
&= -\mu_0 \frac{\partial}{\partial t} \left[ j(r,t) + \frac{\partial P(r,t)}{\partial t} + \varepsilon_0 \frac{\partial E(r,t)}{\partial t} \right]
\end{align*}
\]

And find the wave equation for the electric field

\[
\begin{align*}
\text{rot} \text{rot} E(r,t) + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} E(r,t) &= -\mu_0 \frac{\partial j(r,t)}{\partial t} - \mu_0 \frac{\partial^2 P(r,t)}{\partial t^2}
\end{align*}
\]

The blue terms require knowledge of the material model. Additionally, we have to make sure that all other Maxwell’s equations are fulfilled, in particular:

\[
\text{div} \left[ \varepsilon_0 E(r,t) + P(r,t) \right] = 0
\]
Once we have solved the wave equation, we know the electric field. From that we can easily compute the magnetic field:

\[
\frac{\mathbf{H}(\mathbf{r}, t)}{\partial t} = -\frac{1}{\mu_0} \text{rot} \mathbf{E}(\mathbf{r}, t)
\]

Remarks:
- analog procedure possible for \( \mathbf{H} \), i.e., we can derive a wave equation for the magnetic field
- generally, the wave equation for \( \mathbf{E} \) is more convenient, because we have \( \mathbf{P}(\mathbf{E}) \) given from the material model
- however, for inhomogeneous media \( \mathbf{H} \) can be the better choice for the numerical solution of the partial differential equation since it form a hermitian operator
- analog procedure possible for \( \mathbf{H} \rightarrow \mathbf{E} \)
- generally, wave equation for \( \mathbf{E} \) is more convenient, because \( \mathbf{P}(\mathbf{E}) \) given
- for inhomogeneous media \( \mathbf{H} \) can be better choice

**B) Frequency domain derivation**

We can do the same procedure in the Fourier domain and find

\[
\text{rot rot} \ \mathbf{E}(\mathbf{r}, \omega) - \frac{\omega^2}{c^2} \mathbf{E}(\mathbf{r}, \omega) = -\omega \mu_0 \mathbf{j}(\mathbf{r}, \omega) + \mu_0 \omega^2 \mathbf{P}(\mathbf{r}, \omega)
\]

and

\[
\text{div} [\varepsilon_0 \mathbf{E}(\mathbf{r}, \omega) + \mathbf{P}(\mathbf{r}, \omega)] = 0
\]

- magnetic field:

\[
\mathbf{H}(\mathbf{r}, \omega) = -\frac{i}{\omega \mu_0} \text{rot} \ \mathbf{E}(\mathbf{r}, \omega)
\]

- transferring the results from the Fourier domain to the time domain
  - for stationary fields: take solution and multiply by \( e^{-i \omega t} \).
  - for non-stationary fields and linear media \( \Rightarrow \) inverse Fourier transformation

\[
\mathbf{E}(\mathbf{r}, t) = \frac{1}{2\pi} \int \mathbf{E}(\mathbf{r}, \omega) \exp(-i \omega t) d\omega
\]

**2.1.5 Decoupling of the vectorial wave equation**

- For arbitrary isotropic media generally all 3 field components are coupled.
- For problems with translational invariance in at least one direction, as e.g. for homogeneous infinite media, layers or interfaces, there can be decoupling of the components. Let’s assume invariance in the \( y \)-direction and propagation only in the \( x \)-\( z \)-plane. Then all spatial derivatives along the \( y \)-direction disappear and the operators in the wave equation simplify.

- generally all field components are coupled
- for translational invariance in e.g. the \( y \)-direction and propagation only in the \( x \)-\( z \)-plane \( \Rightarrow \partial / \partial y = 0 \)

\[
\text{rot rot} \ \mathbf{E} = \text{grad} \ \text{div} \ \mathbf{E} - \Delta \mathbf{E} = \begin{pmatrix}
\frac{\partial}{\partial x} \left( \frac{\partial E_x}{\partial x} + \frac{\partial E_z}{\partial z} \right) \\
\frac{\partial}{\partial z} \left( \frac{\partial E_x}{\partial x} + \frac{\partial E_z}{\partial z} \right)
\end{pmatrix}
\]

- decomposition of electric field

\[
\mathbf{E} = \mathbf{E}_\perp + \mathbf{E}_\| \Rightarrow \mathbf{E}_\perp = \begin{pmatrix}
0 \\
E_x
\end{pmatrix}, \quad \mathbf{E}_\| = \begin{pmatrix}
E_x \\
0
\end{pmatrix}
\]

with Nabla operator \( \nabla^{(2)} = \begin{pmatrix}
\partial / \partial x \\
0
\end{pmatrix} \) and Laplace \( \Delta^{(2)} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \)

- gives two decoupled wave equations

\[
\Delta^{(2)} \mathbf{E}_\perp(\mathbf{r}, \omega) + \frac{\omega^2}{c^2} \mathbf{E}_\perp(\mathbf{r}, \omega) = -i \omega \mu_0 \mathbf{j}_\perp(\mathbf{r}, \omega) - \mu_0 \omega^2 \mathbf{P}_\perp(\mathbf{r}, \omega)
\]

\[
\Delta^{(2)} \mathbf{E}_\| (\mathbf{r}, \omega) + \frac{\omega^2}{c^2} \mathbf{E}_\| (\mathbf{r}, \omega) = -i \omega \mu_0 \mathbf{j}_\parallel(\mathbf{r}, \omega) - \mu_0 \omega^2 \mathbf{P}_\parallel(\mathbf{r}, \omega)
\]

- properties
  - propagation of perpendicularly polarized fields \( \mathbf{E}_\perp \) and \( \mathbf{E}_\| \) can be treated separately
  - alternative notations:
    \( \perp \rightarrow s \rightarrow \text{TE (transversal electric)} \)
    \( \parallel \rightarrow p \rightarrow \text{TM (transversal magnetic)} \)

**2.2 Optical properties of matter**

In this chapter we will derive a simple material model for the polarization and the current density. The basic idea is to write down an equation of motion for a single exemplary charged particle and assume that all other particles of the same type behave similarly. More precisely, we will use a driven harmonic oscillator model to describe the motion of bound charges giving rise to a polarization of the medium. For free charges, leading eventually to a current density we will use the same model but without restoring force. In the literature, this simple approach is often called the Drude-Lorentz model (named after Paul Drude and Hendrik Antoon Lorentz).
2.2.1 Basics
We are looking for \( P(E) \) and \( j(E) \). In general, this leads to a many-body problem in solid state theory which is rather complex. However, in many cases phenomenological models are sufficient to describe the necessary phenomena. As already pointed out above, we use the simplest approach, the so-called Drude-Lorentz model for free or bound charge carriers (electrons):

- assume an ensemble of non-coupling, driven, and damped harmonic oscillators
- free charge carriers: metals and excited semiconductors (intraband)
- bound charge carriers: dielectric media and semiconductors (interband)

The Drude-Lorentz model creates a link between cause (electric field) and effect (induced polarization or current). Because the resulting relations \( P(E) \) and \( j(E) \) are linear (no \( E^2 \) etc.), we can use linear response theory.

For the polarization \( P(E) \) (for \( j(E) \) very similar):

- description in both time and frequency domain possible
- \textbf{In time domain:} we introduce the response function \( E(r,t) \rightarrow \text{medium (response function)} \rightarrow P(r,t) \)
  \[
P(t,t') = e_0 \sum_j \int \limits_{-\infty}^{\infty} R_j (r,t-t') E_j (r,t') dt'
\]
  with \( \dot{R} \) being a \( 2^{\text{nd}} \) rank tensor
  \( i = x, y, z \) and summing over \( j = x, y, z \)

- \textbf{In frequency domain:} we introduce the susceptibility \( E(\omega) \rightarrow \text{medium (susceptibility)} \rightarrow \tilde{P}(\omega) \)
  \[
  \tilde{P}(\omega) = e_0 \sum_j \chi_{ij} (\omega) \tilde{E}_j (\omega)
  \]
  - response function and susceptibility are linked via Fourier transform (convolution theorem)
    \[
    R_j (\omega) = \frac{1}{2\pi} \int \limits_{\infty}^{\infty} \chi_{ij} (\omega) \exp(-i\omega t) d\omega
    \]
  - Obviously, things look friendlier in frequency domain. Using the wave equation from before and assuming that there are no currents (\( j = 0 \)) we find

\[
\operatorname{rotrot} \tilde{E}(\omega) - \frac{\omega^2}{c^2} \tilde{E}(\omega) = \mu_0 \varepsilon_0 \tilde{P}(\omega)
\]

or
\[
\Delta \tilde{E}(\omega) + \frac{\omega^2}{c^2} \tilde{E}(\omega) - \operatorname{graddiv} \tilde{E}(\omega) = -\mu_0 \varepsilon_0 \tilde{P}(\omega)
\]

- and for auxiliary fields
  \[
  \tilde{D}(\omega) = \varepsilon_0 \tilde{E}(\omega) + \tilde{P}(\omega)
  \]

The general response function and the respective susceptibility given above simplifies for certain properties of the medium:

\textbf{Different types of media}

\textbf{A)} linear, homogenous, isotropic, non-dispersive media (most simple but very unphysical case)

- homogenous \( \xrightarrow{} \chi_{ij} (r,\omega) = \chi_{ij} (\omega) \)
- isotropic \( \xrightarrow{} \chi_{ij} (r,\omega) = \chi(r,\omega) \delta_{ij} \)
- non-dispersive \( \xrightarrow{} \chi_{ij} (r,\omega) = \chi_r (r) \rightarrow \text{instantaneous: } R_j (r,t) = \chi_r (r) \delta(t) \)
  (Attention: This is unphysical!)

\( \chi_r (r,\omega) \rightarrow \chi \) is a scalar constant

\textbf{Maxwell:}

\[
\nabla \cdot \mathbf{D} = 0 \rightarrow \nabla \cdot \tilde{E}(\omega) = 0 \text{ for } \varepsilon(\omega) \neq 0
\]

\[
\Delta \tilde{E}(\omega) + \frac{\omega^2}{c^2} \tilde{E}(\omega) = 0 \rightarrow \Delta \tilde{E}(r,t) - \frac{\varepsilon}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{E}(r,t) = 0
\]

- approximation is valid only for a certain frequency range, because all media are dispersive
- based on an unphysical material model

\textbf{B)} linear, homogenous, isotropic, \textbf{dispersive} media \( \Leftrightarrow \chi(\omega) \)

\[
\tilde{P}(\omega) = e_0 \chi(\omega) \tilde{E}(\omega) \Leftrightarrow P(r,t) = e_0 \chi(E(r,t)) \text{ (unphysical!)}
\]

\[
\tilde{D}(\omega) = e_0 \varepsilon(\omega) \tilde{E}(\omega) \Leftrightarrow D(r,t) = e_0 \varepsilon(E(r,t)) \rightarrow \varepsilon = 1 + \chi
\]

\textbf{Maxwell:}

\[
\nabla \cdot \mathbf{D} = 0 \rightarrow \nabla \cdot \tilde{E}(\omega) = 0 \text{ for } \varepsilon(\omega) \neq 0
\]

\[
\Delta \tilde{E}(\omega) + \frac{\omega^2}{c^2} \tilde{E}(\omega) = 0 \rightarrow \Delta \tilde{E}(r,t) - \frac{\varepsilon}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{E}(r,t) = 0
\]

\textbf{Helmholtz equation}
This description is sufficient for many materials.

C) linear, inhomogeneous, isotropic, dispersive media \( \rightarrow \chi(r,\omega) \)

\[
P(r,\omega) = \varepsilon_0 \chi(r,\omega) \mathbf{E}(r,\omega),
\]

\[
\mathbf{D}(r,\omega) = \varepsilon_0 \varepsilon(r,\omega) \mathbf{E}(r,\omega).
\]

\[
\text{div } \mathbf{D}(r,\omega) = 0
\]

\[
\text{div } \mathbf{D}(r,\omega) = \varepsilon_0 \varepsilon(r,\omega) \text{div } \mathbf{E}(r,\omega) + \varepsilon_0 \varepsilon(r,\omega) \text{grad } \varepsilon(r,\omega) = 0,
\]

\[
\rightarrow \text{div } \mathbf{E}(r,\omega) = -\frac{\text{grad } \varepsilon(r,\omega)}{\varepsilon(r,\omega)} \mathbf{E}(r,\omega).
\]

\[
\Delta \mathbf{E}(r,\omega) + \frac{\omega^2}{c^2} \varepsilon(r,\omega) \mathbf{E}(r,\omega) = -\text{grad } \left[ \frac{\text{grad } \varepsilon(r,\omega)}{\varepsilon(r,\omega)} \mathbf{E}(r,\omega) \right].
\]

All field components couple.

D) linear, homogeneous, anisotropic, dispersive media \( \rightarrow \chi_i(\omega) \)

\[
\mathbf{P}(r,\omega) = \varepsilon_0 \sum_i \chi_i(\omega) \mathbf{E}_i(r,\omega)
\]

\[
\mathbf{D}(r,\omega) = \varepsilon_0 \sum_i \varepsilon_i(\omega) \mathbf{E}_i(r,\omega).
\]

\[
\rightarrow \text{see chapter on crystal optics}
\]

This is the worst case for a medium with linear response.

Before we start writing down the actual material model equations, let us summarize what we want to do:

**What kind of light-matter interaction do we want to consider?**

**I) Interaction of light with bound electrons and the lattice**

The contributions of bound electrons and lattice vibrations in dielectrics and semiconductors give rise to the polarization \( P \). The lattice vibrations (phonons) are the ionic part of the material model. Because of the large mass of the ions \( (10^3 \times \text{mass of electron}) \) the resulting oscillation frequencies will be small. Generally speaking, phonons are responsible for thermal properties of the medium. However, some phonon modes may contribute to optical properties, but they have small dispersion (weak dependence on frequency \( \omega \)).

Fully understanding the electronic transitions of bound electrons requires quantum theoretical treatment, which allows an accurate computation of the transition frequencies. However, a (phenomenological) classical treatment of the oscillation of bound electrons is possible and useful.

**II) Interaction of light with free electrons**

The contribution of free electrons in metals and excited semiconductors gives rise to a current density \( j \). We assume a so-called (interaction-)free electron gas, where the electron charges are neutralized by the background ions. Only collisions with ions and related damping of the electron motion will be considered.

We will look at the contributions from I) and II) separately, and join the results later.

2.2.2 Dielectric polarization and susceptibility

Let us first focus on bound charges (ions, electrons). In the so-called Drude model, the electric field \( E(r,t) \) gives rise to a displacement \( s(r,t) \) of charged particles from their equilibrium positions. In the easiest approach this can be modeled by a driven harmonic oscillator:

\[
\frac{\partial^2}{\partial t^2} s(r,t) + \frac{\omega_0^2}{c^2} s(r,t) + \omega_0^2 s(r,t) = \frac{q}{m} \mathbf{E}(r,t)
\]

- resonance frequency \( \omega_0 \)
- damping \( \rightarrow g \)
- charge \( \rightarrow q \)
- mass \( \rightarrow m \)

The induced electric dipole moment due to the displacement of charged particles is given by

\[
\mathbf{p}(r,t) = qs(r,t),
\]

We further assume that all bound charges of the same type behave identical, i.e., we treat an ensemble of non-coupled, driven, and damped harmonic oscillators. Then, the dipole density (polarization) is given by

\[
\mathbf{P}(r,t) = N \mathbf{p}(r,t) = N q s(r,t)
\]

Hence, the governing equation for the polarization \( \mathbf{P}(r,t) \) reads as

\[
\frac{\partial^2}{\partial t^2} \mathbf{P}(r,t) + \frac{\omega_0^2}{c^2} \mathbf{P}(r,t) + \omega_0^2 \mathbf{P}(r,t) = \frac{q^2 N}{m} \mathbf{E}(r,t) = \frac{e_0}{m} \mathbf{E}(r,t)
\]

with oscillator strength \( f = \frac{1}{e_0^2 N} \), for \( q = -e \) (electrons)

This equation is easy to solve in Fourier domain:

\[
-\omega^2 \mathbf{P}(r,\omega) - i \omega \mathbf{p}(r,\omega) + \omega_0^2 \mathbf{P}(r,\omega) = \frac{e_0}{m} f \mathbf{E}(r,\omega)
\]

\[
\rightarrow \mathbf{P}(r,\omega) = \frac{e_0 f}{(\omega_0^2 - \omega^2) - i \omega \omega_0} \mathbf{E}(r,\omega)
\]
with $\mathbf{P}(r, \omega) = \varepsilon_0 \chi(\omega) \mathbf{E}(r, \omega) \Rightarrow \chi(\omega) = \frac{f}{(\omega_0^2 - \omega^2) - i\gamma\omega}$

In general we have several different types of oscillators in a medium, i.e., several different resonance frequencies. Nevertheless, since in a good approximation they do not influence each other, all these different oscillators contribute individually to the polarization. Hence the model can be constructed by simply summing up all contributions.

- several resonance frequencies

$$\mathbf{P}(r, \omega) = \varepsilon_0 \sum_j \left[ \frac{f_j}{(\omega_0^2_j - \omega^2) - i\gamma\omega} \right] \mathbf{E}(r, \omega) = \varepsilon_0 \chi(\omega) \mathbf{E}(r, \omega)$$

$$\chi(\omega) = \sum_j \left[ \frac{f_j}{(\omega_0^2_j - \omega^2) - i\gamma\omega} \right]$$

- $\chi(\omega)$ is the complex, frequency dependent susceptibility

$$\mathbf{E}(r, \omega) = \varepsilon_0 \varepsilon(\omega) \mathbf{E}(r, \omega) + \varepsilon_0 \chi(\omega) \mathbf{E}(r, \omega) = \varepsilon_0 \varepsilon(\omega) \mathbf{E}(r, \omega)$$

- $\varepsilon(\omega)$ is the complex frequency dependent dielectric function

Example: (plotted for eta and kappa with $\varepsilon(\omega) = \left[ \eta(\omega) + i\kappa(\omega) \right]^2$)

2.2.3 Conductive current and conductivity

Let us now describe the response of a free electron gas with positively charged background (no interaction). Again we use the model of a driven harmonic oscillator, but this time with resonance frequency $\omega_0 = 0$. This corresponds to the case of zero restoring force.

$$\frac{\partial^2}{\partial t^2} \mathbf{J}(r, t) + \frac{\partial}{\partial t} \mathbf{J}(r, t) = \frac{e^2 N}{m} \mathbf{E}(r, t)$$

The resulting induced current density is given by

$$\mathbf{J}(r, t) = -Ne \frac{\partial}{\partial t} \mathbf{E}(r, t)$$

and the governing dynamic equation reads as

$$\frac{\partial}{\partial t} \mathbf{J}(r, t) + g \mathbf{J}(r, t) = \frac{e^2 N}{m} \mathbf{E}(r, t)$$

with plasma frequency $\omega_p^2 = \frac{1}{\varepsilon_0} \frac{e^2 N}{m}$

Again we solve this equation in Fourier domain:

$$-i\omega \mathbf{J}(r, \omega) + g \mathbf{J}(r, \omega) = \frac{e_0 \omega_p^2}{g - i\omega} \mathbf{E}(r, \omega)$$

$$\mathbf{J}(r, \omega) = \frac{e_0 \omega_p^2}{g - i\omega} \mathbf{E}(r, \omega)$$

Here we introduced the complex frequency dependent conductivity

$$\sigma(\omega) = \frac{e_0 \omega_p^2}{g - i\omega} = -\frac{e_0 \frac{\varepsilon_0 \omega_p^2}{g - i\omega}}{-\omega_p^2 - i\gamma\omega}$$

**Remarks on plasma frequency**

We consider a cloud of electrons and positive ions described by the total charge density $\rho$ in their self-consistent field $\mathbf{E}$. Then we find according to Maxwell:

$$\varepsilon_0 \varepsilon_0 \mathbf{E}(r, t) = \rho(r, t)$$

For cold electrons, and because the total charge is zero, we can use our damped oscillator model from before to describe the current density (only electrons move):

$$\frac{\partial}{\partial t} \mathbf{J} + g \mathbf{J} = \varepsilon_0 \omega_p^2 \mathbf{E}(r, t)$$

Now we apply divergence operator and plug in from above (red terms):

$$\mathbf{div} \frac{\partial}{\partial t} \mathbf{J} + g \mathbf{div} \mathbf{J} = \varepsilon_0 \omega_p^2 \mathbf{div} \mathbf{E}(r, t) = \omega_p^2 \rho(r, t)$$

With the continuity equation for the charge density (from Maxwell's equations)

$$\frac{\partial}{\partial t} \rho + \mathbf{div} \mathbf{J} = 0,$$

We can substitute the divergence of the current density and find:
\[
\frac{\partial^2}{\partial t^2} \rho - g \frac{\partial}{\partial t} \rho = \omega_p^2 \rho
\]

\[
\frac{\partial^2}{\partial t^2} \rho + g \frac{\partial}{\partial t} \rho + \omega_p^2 \rho = 0,
\]

\[\Rightarrow \text{harmonic oscillator equation}\]

Hence, the plasma frequency \(\omega_p\) is the eigen-frequency of such a charge density.

### 2.2.4 The generalized complex dielectric function

In the sections above we have derived expressions for both polarization (bound charges) and conductive current density (free charges). Let us now plug our \(\mathcal{J}(\mathbf{r},\omega)\) and \(\mathbf{P}(\mathbf{r},\omega)\) into the wave equation (in Fourier domain)

\[
\text{rotrot} \, \tilde{\mathbf{E}}(\mathbf{r},\omega) = \frac{\omega^2}{c^2} \tilde{\mathbf{E}}(\mathbf{r},\omega) = \mu_0 \varepsilon_0 \mathbf{P}(\mathbf{r},\omega) + i \omega \mu_0 \mathcal{J}(\mathbf{r},\omega)
\]

\[
= \left[ \mu_0 \varepsilon_0 \omega^2 \chi(\omega) + i \omega \mu_0 \sigma(\omega) \right] \tilde{\mathbf{E}}(\mathbf{r},\omega)
\]

Hence we can collect all terms proportional to \(\tilde{\mathbf{E}}(\mathbf{r},\omega)\) and write

\[
\text{rotrot} \, \tilde{\mathbf{E}}(\mathbf{r},\omega) = \frac{\omega^2}{c^2} \left[ 1 + \chi(\omega) + \frac{i}{\omega \varepsilon_0} \sigma(\omega) \right] \tilde{\mathbf{E}}(\mathbf{r},\omega)
\]

\[
\text{rotrot} \, \tilde{\mathbf{E}}(\mathbf{r},\omega) = \frac{\omega^2}{c^2} \varepsilon(\omega) \tilde{\mathbf{E}}(\mathbf{r},\omega)
\]

Here, we introduced the generalized complex dielectric function

\[
\varepsilon(\omega) = 1 + \chi(\omega) + \frac{i}{\omega \varepsilon_0} \sigma(\omega) = \varepsilon'(\omega) + i \varepsilon''(\omega)
\]

So, in general we have

\[
\varepsilon(\omega) = 1 + \sum_j \left( \frac{f_j}{(\omega_j - \omega)^2 - i \gamma_j \omega} \right) + \frac{\omega_p^2}{-\omega^2 - i \omega \gamma}
\]

because (from before)

\[
\chi(\omega) = \sum_j \left( \frac{f_j}{(\omega_j - \omega)^2 - i \gamma_j \omega} \right), \quad \sigma(\omega) = -i \frac{\varepsilon_0 \omega_0^2}{-\omega^2 - i \omega \gamma}
\]

\(\varepsilon(\omega)\) contains contributions from vacuum, phonons (lattice vibrations), bound and free electrons.

---

### Some special cases for materials in the IR und VIS:

#### A) Dielectrics (insulators) near phonon resonance (IR)

If we are interested in dielectrics (insulators) near phonon resonance in the infrared spectral range we can simplify the dielectric function as follows:

\[
\varepsilon(\omega) = 1 + \sum_j \left( \frac{f_j}{(\omega_j - \omega)^2 - i \gamma_j \omega} \right) + \frac{f}{(\omega_0^2 - \omega^2)^2 - 2i \omega \gamma}
\]

\[
\varepsilon(\omega) \approx \varepsilon_{\infty} + \frac{f}{(\omega_0^2 - \omega^2)^2 - 2i \omega \gamma}
\]

The contribution of vacuum and electronic transitions show almost no frequency dependence (dispersion) in this regime and can be expressed as a constant \(\varepsilon_{\infty}\). Let us study the real and the imaginary part of the resulting \(\varepsilon(\omega)\) separately:

\[
\varepsilon_{\infty} \rightarrow \text{vacuum and electronic transitions}
\]

\[
\Rightarrow \varepsilon(\omega) = \Re \varepsilon(\omega) + i \Im \varepsilon(\omega) = \varepsilon'(\omega) + i \varepsilon''(\omega)
\]

\[
\varepsilon'(\omega) = \varepsilon_{\infty} + \frac{f}{(\omega_0^2 - \omega^2)^2 + 2 \omega^2 \gamma^2}
\]

\[
\varepsilon''(\omega) = \frac{g \omega}{(\omega_0^2 - \omega^2)^2 + 2 \omega^2 \gamma^2}
\]

- resonance frequency \(\omega_0\)
- width of resonance peak \(g\)
- static dielectric constant in the limit \(\omega \rightarrow 0:\) \(\varepsilon'' = \varepsilon_{\infty} + \frac{f}{\omega_0^2}\)
- so-called longitudinal frequency \(\omega_L:\) \(\varepsilon'(\omega) = \varepsilon''(\omega) = 0\)
- \(\varepsilon'(\omega) \neq 0:\) absorption and dispersion appear always together

Example: single resonance
- near resonance we find $\varepsilon'(\omega) < 0$ (damping without absorption if $\varepsilon'' = 0$)
- normal dispersion $\rightarrow \partial \varepsilon'(\omega) / \partial \omega > 0$
- anomalous dispersion $\rightarrow \partial \varepsilon'(\omega) / \partial \omega < 0$

Example: sharp resonance for undamped oscillator $g \rightarrow 0$

- relation between resonance frequency $\omega_0$ and longitudinal frequency $\omega_L$ (Lyddane-Sachs-Teller relation)

$$
\varepsilon'(\omega_L) = \varepsilon_\infty + \frac{f}{(\omega_0^2 - \omega_L^2)} = 0, \quad f = (\varepsilon_0^2 - \varepsilon_\infty) \omega_0^2 \quad \text{(from above)}
$$

$$
\omega_L = \omega_0 \sqrt{\frac{\varepsilon_\infty}{\varepsilon_0}}
$$

B) Dielectric media in visible spectral range

Dielectric media in visible spectral range can be described by a so-called double resonance model (phonon in IR and electronic transition in UV).

$$
\varepsilon(\omega) = \varepsilon_\infty + \frac{f_p}{(\omega_0^2 - \omega^2) - \frac{1}{4} \varepsilon_0^2} + \frac{f_e}{(\omega_0^2 - \omega^2) - \frac{1}{4} \varepsilon_0^2} \omega_p \ll \omega \ll \omega_0
$$

$\varepsilon_\infty \rightarrow$ contribution of vacuum and other (far away) resonances

The generalization of this approach in the transparent spectral range leads to so-called Sellmeier formula:

$$
\varepsilon'(\omega) - 1 = \sum_j \frac{f_j \omega_j^2}{(\omega_j^2 - \omega^2)}
$$

- describes many media very well (dispersion of absorption is neglected)
- oscillator strengths and resonance frequencies are fit parameters

C) Metals in visible spectral range

If we want to describe metals in visible spectral range we find

$$
\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + \varepsilon_0^2}\quad \omega_p \gg \omega
$$

$$
\varepsilon'(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + \varepsilon_0^2}, \quad \varepsilon''(\omega) = \frac{g \omega_p^2}{\omega(\omega^2 + \varepsilon_0^2)}
$$

Metals show a large negative real part of the dielectric function $\varepsilon'(\omega)$
2.2.5 Material models in time domain

Let us now transform our results of the material models back to time domain. In Fourier domain we found for homogeneous and isotropic media:

\[ \chi(\omega) = \varepsilon_0 \varepsilon''(\omega) \]

The response function (or Green's function) \( R(t) \) is then given by

\[ R(t) = \int_{-\infty}^{\infty} \chi(\omega) \exp(-i\omega t) d\omega = \int_{-\infty}^{\infty} R(t) \exp(i\omega t) dt \]

To prove this, we can use the convolution theorem

\[ P(r, t) = \int_{-\infty}^{\infty} P(r, \omega) \exp(-i\omega t) d\omega = \varepsilon_0 \int_{-\infty}^{\infty} \chi(\omega) E(r, \omega) \exp(-i\omega t) d\omega \]

Now we switch the order of integration, and identify the response function \( R \) (red terms):

\[ R(t) = \int_{-\infty}^{\infty} \chi(\omega) \frac{1}{2\pi} \int_{-\infty}^{\infty} E(r, t') \exp(i\omega t') dt' \exp(-i\omega t) d\omega \]

For a “delta” excitation in the electric field we find the response or Greens function as the polarization:

\[ E(r, t) = \varepsilon_0 \delta(t) \rightarrow P(r, t) = \varepsilon_0 R(t) \delta(t) \rightarrow \text{Green's function} \]

For instantaneous (or non-dispersive) media we find:

\[ R(t) = \chi \delta(t) \rightarrow P(r, t) = \varepsilon_0 \chi E(r, t) \quad \text{(unphysical)} \]

**Examples**

A) dielectric media

\[ R_\varepsilon(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \chi(\omega) \exp(-i\omega t) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{\varepsilon_0^2 - \omega^2} \exp(-i\omega t) d\omega, \]

- Using the residual theorem we can find:

\[ R(t) = \begin{cases} \frac{g}{\varepsilon_0^2} \exp(-g t) & t \geq 0 \\ 0 & t < 0 \end{cases} \]

\[ P(r, t) = \varepsilon_0 \chi \exp(-g(t-t')) \exp[\Omega(t-t')] E(r, t') dt' \]

B) metals

\[ R_j(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sigma(\omega) \exp(-i\omega t) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\varepsilon_0 \varepsilon''}{g - \varepsilon''} \exp(-i\omega t) d\omega, \]

- Using again the residual theorem we can find:

\[ R(t) = \begin{cases} \frac{\sigma_0^2}{g} \exp(-g t) & t \geq 0 \\ 0 & t < 0 \end{cases} \]

\[ j(r, t) = \varepsilon_0 \varepsilon'' \int_{-\infty}^{\infty} \exp(-g(t-t')) E(r, t') dt' \]

2.3 The Poynting vector and energy balance

2.3.1 Time averaged Poynting vector

The energy flux of the electromagnetic field is given by the Poynting vector \( S \). In practice, we always measure the energy flux through a surface (detector), \( S \cdot n \), where \( n \) is the normal vector of surface. To be more precise, the Poynting vector \( S(r, t) = E_x(r, t) \times H_y(r, t) \) gives the momentary energy flux. Note that we have to use the real electric and magnetic fields, because a product of fields occurs.

In optics we have to consider the following time scales:

- optical cycle \( \rightarrow T_\circ = 2\pi/\varepsilon_0 \leq 10^{-14}\text{s} \)
- pulse duration \( \rightarrow T_p \quad \text{in general} \quad T_p >> T_\circ \)
duration of measurement $\rightarrow T_m$ in general $T_m \gg T_0$

Hence, in general the detector does not recognize the fast oscillations of the optical field $\sim e^{i\omega t}$ (optical cycles) and delivers a time averaged value. For the situation described above, the electro-magnetic fields factorize in slowly varying envelopes and fast carrier oscillations:

$$\frac{1}{2} \left[ \mathbf{E}(r,t) \exp(-i\omega_0 t) + c.c. \right] = \mathbf{E}_0(r,t)$$

For such pulses, the momentary Poynting vector reads:

$$\mathbf{S}(r,t) = \mathbf{E}_0(r,t) \times \mathbf{H}_0(r,t)$$

$$ = \frac{1}{4} \left[ \mathbf{E}(r,t) \times \mathbf{H}^*(r,t) + \mathbf{E}^*(r,t) \times \mathbf{H}(r,t) \right]$$

$$+ \frac{1}{4} \left[ \mathbf{E}(r,t) \times \mathbf{H}^*(r,t) \exp(-2i\omega_0 t) + \mathbf{E}^*(r,t) \times \mathbf{H}(r,t) \exp(2i\omega_0 t) \right]$$

$$= \frac{1}{2} \Re\left[ \mathbf{E}(r,t) \times \mathbf{H}^*(r,t) \right] + \frac{1}{2} \Re\left[ \mathbf{E}^*(r,t) \times \mathbf{H}(r,t) \right] \cos(2\omega_0 t)$$

$$+ \frac{1}{2} \Im\left[ \mathbf{E}^*(r,t) \times \mathbf{H}(r,t) \right] \sin(2\omega_0 t).$$

We find that the momentary Poynting vector has some slow contributions which change over time scales of the pulse envelope $T_p$, and some fast contributions $\sim \cos(2\omega_0 t)$, $\sim \sin(2\omega_0 t)$ changing over time scales of the optical cycle $T_0$. Now, a measurement of the Poynting vector over a time interval $T_m$ leads to a time average of $\mathbf{S}(r,t)$

$$\langle \mathbf{S}(r,t) \rangle = \frac{1}{T_m} \int_{-T_m/2}^{T_m/2} \mathbf{S}(r,t') dt'$$

The fast oscillating terms $\sim \cos 2\omega_0 t$ and $\sim \sin 2\omega_0 t$ cancel by the integration since the pulse envelope does not change much over one optical cycle. Hence we get only a contribution from the slow term:

$$\langle \mathbf{S}(r,t) \rangle \approx \frac{1}{2} \left[ \mathbf{E}^*(r,t') \times \mathbf{H}(r,t') \right] dt'$$

Let us now have a look at the special (but important) case of stationary (monochromatic) fields. Then, the pulse envelope does not depend on time at all (infinitely long pulses):

$$\mathbf{E}(r,t') = \mathbf{E}(r), \quad \mathbf{H}(r,t') = \mathbf{H}(r)$$

$$\langle \mathbf{S}(r,t) \rangle = \frac{1}{2} \frac{\Re\left[ \mathbf{E}(r) \times \mathbf{H}^*(r) \right]}{2}$$

This is the definition for the optical intensity $I = |\mathbf{S}(r,t)|$. We see that an intensity measurement destroys information on the phase.

2.3.2 Time averaged energy balance

Let us motivate a little bit further the concept of the Poynting vector. It appears naturally in the Poynting theorem, the equation for the energy balance of the electromagnetic field. The Poynting theorem can be derived directly from Maxwell’s equations. We multiply the two curl equations by $\mathbf{H}$ resp. $\mathbf{E}$: (note that we use real fields):

$$\mathbf{H} \cdot \mathbf{rot} \mathbf{E} + \mu_0 \mathbf{H} \cdot \frac{\partial}{\partial t} \mathbf{H} = 0$$

$$-\varepsilon_0 \mathbf{E} \cdot \frac{\partial}{\partial t} \mathbf{E} + \mathbf{E} \cdot \mathbf{rot} \mathbf{H} = \mathbf{E} \cdot (\mathbf{j} + \frac{\partial}{\partial t} \mathbf{P})$$

Next, we subtract the two equations and get

$$\mathbf{H} \cdot \mathbf{rot} \mathbf{E} - \mathbf{E} \cdot \mathbf{rot} \mathbf{H} = \varepsilon_0 \mathbf{E} \cdot \frac{\partial}{\partial t} \mathbf{E} + \mu_0 \mathbf{H} \cdot \frac{\partial}{\partial t} \mathbf{H} = -\mathbf{E} \cdot (\mathbf{j} + \frac{\partial}{\partial t} \mathbf{P}).$$

This equation can be simplified by using the following vector identity:

$$\text{div} (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \cdot \mathbf{rot} \mathbf{E} - \mathbf{E} \cdot \mathbf{rot} \mathbf{H}$$

Finally, with $\mathbf{E} \cdot \frac{\partial}{\partial t} \mathbf{E} = \frac{1}{2} \frac{\partial}{\partial t} \mathbf{E}^2$ we find Poynting’s theorem

$$\frac{1}{2} \varepsilon_0 \frac{\partial}{\partial t} \mathbf{E}^2 + \frac{1}{2} \mu_0 \frac{\partial}{\partial t} \mathbf{H}^2 + \text{div} (\mathbf{E} \times \mathbf{H}) = -\mathbf{E} \cdot (\mathbf{j} + \frac{\partial}{\partial t} \mathbf{P})$$

This equation has the general form of a balance equation, here it represents the energy balance. Apart from the appearance of the Poynting vector (energy flux), we can identify the vacuum energy density $u = \frac{1}{2} \varepsilon_0 \mathbf{E}^2 + \frac{1}{2} \mu_0 \mathbf{H}^2$.

The right-hand-side of Poynting's theorem contains the so-called source terms.

where $u = \frac{1}{2} \varepsilon_0 \mathbf{E}^2 + \frac{1}{2} \mu_0 \mathbf{H}^2 \rightarrow$ vacuum energy density

In the case of stationary fields and isotropic media (simple but important)
\[
\begin{align*}
E_x(r,t) &= \frac{1}{2} \left[ E(r) \exp(-i\omega t) + c.c. \right] \\
H_z(r,t) &= \frac{1}{2} \left[ H(r) \exp(-i\omega t) + c.c. \right]
\end{align*}
\]

Time averaging of the left hand side of Poynting’s theorem (*) yields:
\[
\left\{ \frac{1}{2} c \frac{\partial}{\partial t} E_x^2(r,t) + \frac{1}{2} M \frac{\partial}{\partial t} H_z^2(r,t) + \text{div} \left[ E_x(r,t) \times H_z(r,t) \right] \right\} = \frac{1}{2} \left\{ \frac{\partial}{\partial t} \left[ E(r) \times H^*(r) \right] \right\} = \text{div} \langle S(r,t) \rangle.
\]

Note that the time derivative removes stationary terms in \( E_x(r,t) \) and \( H_z(r,t) \).

Time averaging of the right hand side of Poynting’s theorem yields (source terms):
\[
\left\{ \frac{1}{2} \sigma(\omega) E(r) e^{-i\omega t} - i \omega_0 \varepsilon_0 \chi(\omega_0) E(r) e^{-i\omega t} + c.c. \right\} = \frac{1}{4} \left\{ \varepsilon(\omega_0) E(r) e^{-i\omega t} + c.c. \right\}
\]

Now we use our generalized dielectric function:
\[
\frac{1}{4} \left\{ -i \omega_0 \varepsilon_0 \left( \chi(\omega_0) + i \frac{\sigma(\omega_0)}{\omega_0 \varepsilon_0} \right) E(r) \exp(-i\omega t) + c.c. \right\} \left\{ E(r) \exp(-i\omega t) + c.c. \right\}
\]

\[
= \frac{1}{4} \left( \varepsilon(\omega_0) - 1 \right) E(r) E^*(r) + c.c.
\]

Again, all fast oscillating terms \( \exp(\pm i\omega_0 t) \) cancel due to the time average. Finally, splitting \( \varepsilon(\omega_0) \) into real and imaginary part yields
\[
= \frac{1}{4} \left( \varepsilon'(\omega_0) - 1 + i \varepsilon''(\omega_0) \right) E(r) E^*(r) + c.c. = \frac{1}{2} \omega_0 \varepsilon_0 \varepsilon''(\omega_0) E(r) E^*(r).
\]

Hence, the divergence of the time averaged Poynting vector is related to the imaginary part of the generalized dielectric function:
\[
\text{div} \langle S \rangle = -\frac{1}{2} \omega_0 \varepsilon_0 \varepsilon''(\omega_0) E(r) E^*(r).
\]

This shows that a nonzero imaginary part of epsilon \( \varepsilon''(\omega) \neq 0 \) causes a drain of energy flux. In particular, we always have \( \varepsilon''(\omega) > 0 \), otherwise there would be gain of energy. In particular near resonances we have \( \varepsilon''(\omega) \neq 0 \) and therefore absorption.

Further insight into the meaning of \( \text{div} \langle S \rangle \) gives the so-called divergence theorem. If the energy of the electro-magnetic field is flowing through some volume, and we wish to know how much energy flows out of a certain region within that volume, then we need to add up the sources inside the region and subtract the sinks. The energy flux is represented by the (time averaged) Poynting vector, and the Poynting vector’s divergence at a given point describes the strength of the source or sink there. So, integrating the Poynting vector’s divergence over the interior of the region equals the integral of the Poynting vector over the region’s boundary.

\[
\int_V \text{div} \langle S \rangle dV = \int_A \langle S \rangle \cdot n dA
\]

### 2.4 Normal modes in homogeneous isotropic media

Using the linear material models which we discussed in the previous chapters we can now look for solutions to the wave equation.

Because it is convenient to use the generalized complex dielectric function
\[
\varepsilon(\omega) = 1 + \chi(\omega) + \frac{1 - \varepsilon'(\omega)}{\omega_0 \varepsilon_0} \sigma(\omega) = \varepsilon'(\omega) + i \varepsilon''(\omega)
\]

We will do our analysis in Fourier domain. In particular, we will focus on the most simple solution to the wave equation in Fourier domain, the so-called normal modes. We will see later that it is possible to construct general solutions from the normal modes. The wave equation in Fourier domain reads

\[
\text{rotrot } \tilde{E}(r,\omega) = \frac{\omega^2}{c^2} \varepsilon(\omega) \tilde{E}(r,\omega)
\]

According to Maxwell the solutions have to fulfill additionally
\[
\varepsilon_0 [1 + \chi(\omega)] \text{div } \tilde{E}(r,\omega) = 0
\]

In general, this additional condition implies that the electric field is free of divergence:
\[
1 + \chi(\omega) \neq 0 \rightarrow \text{div } \tilde{E}(r,\omega) = 0 \quad \text{(normal case)}
\]
Let us for a moment assume that we already know that we can find plane wave solutions of the form:

$$\mathbf{E}(r, \omega) = \mathbf{E}(\omega) \exp(i \mathbf{k} \cdot \mathbf{r}),$$

where \( \mathbf{k} \) is an unknown complex wave-vector.

Then, the divergence condition implies that those waves are transversal if \( \mathbf{k} \perp \mathbf{E}(\omega) \leftrightarrow \text{transverse wave} \). 

The corresponding stationary field in time domain is given by

$$\mathbf{E}(r, t) = \mathbf{E}(\omega) \exp \left[ -i (\mathbf{k} \cdot \mathbf{r} - \omega t) \right]$$

$$\Rightarrow \text{monochromatic plane wave} \Rightarrow \text{normal mode}$$

This is a monochromatic plane wave, the simplest solution we can expect, a so-called normal mode.

If we split the complex wave vector into real and imaginary part \( \mathbf{k} = \mathbf{k}' + i \mathbf{k}'' \), we can define:

- planes of constant phase \( k' r = \text{const.} \)
- planes of constant amplitude \( k'' r = \text{const.} \)

In the following we will call the solutions

A) if those planes are identical \( \Rightarrow \text{homogeneous waves} \)
B) if those planes are perpendicular \( \Rightarrow \text{evanescent waves} \)
C) otherwise \( \Rightarrow \text{inhomogeneous waves} \)

We will see that in dielectrics \( \epsilon(\omega) = 0 \) we can find a second, exotic type of wave solutions: At \( \omega = \omega_L \rightarrow \epsilon(\omega_L) = 0 \), so-called longitudinal waves \( \mathbf{k} \parallel \mathbf{E}(\omega) \) appear.

### 2.4.1 Transversal waves

As pointed out above, for \( \omega \neq \omega_L \) the electric field becomes free of divergence:

$$\epsilon(\omega) \text{div} \mathbf{E}(r, \omega) = 0 \quad \Rightarrow \quad \text{div} \mathbf{E}(r, \omega) = 0$$

Then, the wave equation reduces to the Helmholtz equation:

$$\Delta \mathbf{E}(r, \omega) + \frac{\omega^2}{c^2} \epsilon(\omega) \mathbf{E}(r, \omega) = 0.$$ 

Hence, we have three scalar equations for \( \mathbf{E}(r, \omega) \) (from Helmholtz), and together with the divergence condition we are left with two independent field components. We will now construct solutions using the plane wave ansatz:

$$\mathbf{E}(r, \omega) = \mathbf{E}(\omega) \exp(i \mathbf{k} \cdot \mathbf{r})$$

Immediately we see that the wave is transversal:

$$0 = \text{div} \mathbf{E}(r, \omega) = i \mathbf{k} \cdot \mathbf{E}(r, \omega) \Rightarrow \mathbf{k} \perp \mathbf{E}(\omega).$$

Hence, we have to solve

$$\left[ -\mathbf{k}^2 + \frac{\omega^2}{c^2} \epsilon(\omega) \right] \mathbf{E}(\omega) = 0 \quad \text{and} \quad \mathbf{k} \cdot \mathbf{E}(\omega) = 0.$$ 

which leads to the following dispersion relation

$$\mathbf{k}^2 = k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2} \epsilon(\omega)$$

We see that the so-called wave-number \( k(\omega) = \frac{\omega}{c} \sqrt{\epsilon(\omega)} \) is a function of the frequency. We can conclude that transversal plane waves are solutions to Maxwell’s equations in homogeneous, isotropic media, only if the dispersion relation \( k(\omega) \) is fulfilled.

In general, \( \mathbf{k} = k' + i k'' \) is complex. Therefore it is sometimes useful to introduce the complex refractive index (if \( k'' \neq 0 \)):

$$\hat{n}(\omega) = \frac{\epsilon(\omega)^{1/2}}{\epsilon(\omega)^{1/2} + n(\omega)} = \frac{\epsilon(\omega)^{1/2}}{\epsilon(\omega)^{1/2} + n(\omega) + i n'(\omega)}.$$ 

With the knowledge of the electric field we can compute the magnetic field if desired:

$$\mathbf{H}(r, \omega) = \frac{i}{\omega \mu_0} \text{rot} \mathbf{E}(r, \omega) = \frac{1}{\omega \mu_0} \left[ \mathbf{k} \times \mathbf{E}(\omega) \right] \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r})$$

$$\Rightarrow \mathbf{H}(r, \omega) = \hat{n}(\omega) \mathbf{E}(\omega) \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r}), \quad \text{with} \quad \hat{n}(\omega) = \frac{1}{\omega \mu_0} \left[ \mathbf{k} \times \mathbf{E}(\omega) \right]$$

### 2.4.2 Longitudinal waves

Let us now have a look at the rather exotic case of longitudinal waves. Those waves can only exist for \( \epsilon(\omega) = 0 \) in dielectrics at the longitudinal frequency \( \omega = \omega_L \). In this case, we cannot conclude that \( \text{div} \mathbf{E}(r, \omega) = 0 \), and the wave equation reads (the l.h.s. vanishes because \( \epsilon(\omega) = 0 \)):

$$\Rightarrow \text{rot} \mathbf{E}(r, \omega) = 0$$

If we try our plane wave ansatz and assume \( \mathbf{k} \) real, it is useful to split the electric field into transversal and longitudinal components with respect to the wave vector:

$$\mathbf{E}(r, \omega) = \mathbf{E}_T(\omega) \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r}) = \mathbf{E}_T(\omega) \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r}) + \mathbf{E}_L(\omega) \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r}),$$

with, \( \mathbf{k} \perp \mathbf{E}_L(\omega) \) and \( \mathbf{k} \parallel \mathbf{E}_T(\omega) \)

With \( \text{rot} \left[ \mathbf{E}(\omega) \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r}) \right] = \mathbf{k} \times \mathbf{E}(\omega) \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r}) \) we get from the wave equation:

$$\mathbf{k} \times \left[ \mathbf{k} \times \mathbf{E}(\omega) \right] = 0$$
Now we plug in the electric field (longitudinal and transversal):

\[
k \times [k \times (\mathbf{E}_l + \mathbf{E}_t)] \exp(ikr) = 0,
\]

\[
k \times [k \times \mathbf{E}_t] \exp(ikr) = 0,
\]

\[
k \times [k \times 
\mathbf{E}_l] \exp(ikr) = 0,
\]

\[
\Rightarrow k^2 \mathbf{E}_l = 0 \Rightarrow \mathbf{E}(r, \omega_k) = \mathbf{E}(\omega_k) \exp(ikr)
\]

2.4.3 Plane wave solutions in different frequency regimes

The dispersion relation for plane wave solutions \( k^2 = k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2} \varepsilon(\omega) \) dictates the (complex) wavenumber only. Thus, different solutions for the complex wave vector \( k = k' + ik'' \) are possible. In addition, the generalized dielectric function is complex. In this chapter we will discuss possible scenarios and resulting plane wave solutions.

A) Positive real valued epsilon \( \varepsilon(\omega) = \varepsilon'(\omega) > 0 \)

This is the regime favorable for optics. We have transparency, and the frequency is far from resonances. The dispersion relation gives

\[
k^2 = k'^2 - k''^2 + 2i k' k'' = \frac{\omega^2}{c^2} \varepsilon'(\omega) = \frac{\omega^2}{c^2} n^2(\omega) \Rightarrow k' \cdot k'' = 0
\]

There are two possibilities to fulfill this condition, either \( k'' = 0 \) or \( k' \perp k'' \).

A.1) real valued wave-vector \( k'' = 0 \)

- In this case the wave vector is real and we find the dispersion relation

\[
k(\omega) = \frac{\omega}{c n(\omega)} = \frac{\omega}{c n} \Rightarrow \frac{\omega}{c} = \frac{2\pi}{\lambda}
\]

- Because \( k'' = 0 \) these waves are homogeneous (trivial, because the amplitude is constant).

Example 1: single resonance in dielectric material

- for lattice vibrations (phonons)

\[
\varepsilon(\omega) = \varepsilon'(\omega) = \varepsilon_n + \frac{f}{\omega_0^2 - \omega^2}
\]

- We can invert the dispersion relation \( k(\omega) = \frac{\omega}{c} \sqrt{\varepsilon(\omega)} \rightarrow \omega(k) : \)

Example 2: free electrons

- for plasma and metal

- Again the imaginary part of \( \varepsilon(\omega) \) is neglected

\[
\varepsilon(\omega) = \varepsilon'(\omega) = 1 - \frac{\omega_k^2}{\omega^2}
\]

- We again invert the dispersion relation \( k(\omega) = \frac{\omega}{c} \sqrt{\varepsilon(\omega)} \rightarrow \omega(k) : \)

A.2) complex valued wave-vector \( k' \perp k'' \)

- The second possibility to fulfill the dispersion relation leads to a complex wave-vector and so-called evanescent waves. We find

\[
k^2 = k'^2 - k''^2 = \frac{\omega^2}{c^2} \varepsilon(\omega) \text{ and therefore } k''^2 = k'^2 - k^2
\]

- This means that
We will discuss the importance of evanescent waves in the next chapter, where we will study the propagation of arbitrary initial field distributions. What is interesting to note here is that evanescent waves can have arbitrary large \(k' > k''\), whereas the homogeneous waves of i) \((k'' = 0)\) obey \(k'' = k'\). If we plug our findings into the plane wave ansatz we get: for the evanescent waves:

\[
E(r, \omega) = E(\omega) \exp \left\{ \mathbf{i} \left[ k'(\omega) \mathbf{r} \right] \right\} \exp(-k''(\omega)\mathbf{r})
\]

The planes defined by the equation \(k''(\omega)\mathbf{r} = \text{const.}\) are the so-called planes of constant amplitude, those defined by \(k'(\omega)\mathbf{r} = \text{const.}\) are the planes of constant phase. Because of \(\mathbf{k}' \perp \mathbf{k}''\), these planes are perpendicular to each other.

The factor \(\exp(-k''(\omega)\mathbf{r})\) leads to exponential growth of evanescent waves in homogeneous space. Therefore, evanescent waves are not normal modes of homogeneous space and can only exist at interfaces.

### B) Negative real valued epsilon \(\varepsilon(\omega) = \varepsilon'(\omega) < 0\)

This situation (negative but real \(\varepsilon(\omega)\)) can occur near resonances in dielectrics \((\omega_e < \omega < \omega_r)\) or below the plasma frequency \((\omega < \omega_p)\) in metals. Then the dispersion relation gives

\[
k^2 = k'^2 - k''^2 + 2\mathbf{k}' \cdot \mathbf{k}'' = \frac{\omega^2}{c^2} \varepsilon'(\omega) < 0
\]

As in the previous case A), the imaginary term has to vanish and \(\mathbf{k}' \cdot \mathbf{k}'' = 0\). Again this can be achieved by two possibilities.

#### B.1) \(k' = 0\)

\[
k''^2 = \frac{\omega^2}{c^2} \varepsilon'(\omega)
\]

\(\Rightarrow E(r, \omega) \sim \exp(-k''r) \Rightarrow \text{strong damping}\)

As above, these evanescent waves exist only at interfaces (like for \(\varepsilon(\omega) = \varepsilon'(\omega) > 0\)). The interesting point is that here we find evanescent waves for all values of \(k''^2\). In particular, case i) \((k'' = 0)\) is included. Hence, we can conclude that for \(\varepsilon(\omega) = \varepsilon'(\omega) < 0\) we find only evanescent waves!

### C) Complex valued epsilon \(\varepsilon(\omega)\)

This is the general case, which is in particular relevant near resonances. From our (optical) point of view only weak absorption is interesting. Therefore, in the following we will always assume \(\varepsilon''(\omega) << \varepsilon'(\omega)\). As we can see in the following sketch, we can have \(\varepsilon'(\omega) > 0, \varepsilon''(\omega) > 0\), or \(\varepsilon'(\omega) < 0, \varepsilon''(\omega) > 0\).

Let us further consider only the important special case of quasi-homogeneous plane waves, i.e., \(\mathbf{k}'\) and \(\mathbf{k}''\) are almost parallel. Then, it is convenient to use the complex refractive index

\[
k(\omega) = \frac{\omega}{c} \sqrt{\varepsilon(\omega)} = \frac{\omega}{c} \hat{\varepsilon}(\omega) = \frac{\omega}{c} \left[ n(\omega) + \mathbf{i} \kappa(\omega) \right],
\]

\[
|k'| = \frac{\omega}{c} n(\omega), \quad |k''| = \frac{\omega}{c} \kappa(\omega).
\]

The dispersion relation in terms of the complex refractive index gives

\[
k^2 = k^2 = \frac{\omega^2}{c^2} \varepsilon(\omega) = \frac{\omega^2}{c^2} \left[ n(\omega) + \mathbf{i} \kappa(\omega) \right]^2
\]
Here we have
\[ \varepsilon(\omega) = \varepsilon'(\omega) + i \varepsilon''(\omega) = n^2(\omega) - \kappa^2(\omega) + 2i n(\omega) \kappa(\omega), \]
and therefore \( \varepsilon'(\omega) = n'(\omega) - \kappa'(\omega) \)
\[ \varepsilon'(\omega) = 2n(\omega) \kappa(\omega) \]
\[ n'(\omega) = \varepsilon' \left[ \operatorname{sgn}(\varepsilon') \sqrt{1 + \left(\varepsilon''/\varepsilon'\right)^2} + 1 \right], \]
\[ \kappa'(\omega) = \varepsilon' \left[ \operatorname{sgn}(\varepsilon') \sqrt{1 + \left(\varepsilon''/\varepsilon'\right)^2} - 1 \right]. \]

Two important limiting cases of quasi-homogeneous plane waves:

**Region 1** \( \varepsilon', \varepsilon^* > 0, \varepsilon^* < \varepsilon' \), *(dielectric media)*
\[ n(\omega) \approx \sqrt{\varepsilon'(\omega)}, \quad \kappa(\omega) \approx \frac{1}{2} \frac{\varepsilon'(\omega)}{\sqrt{\kappa'(\omega)}} \]

In this regime propagation dominates \( (n(\omega) \gg \kappa(\omega)) \), and we have weak absorption:
\[ \mathbf{k}^2 - \mathbf{k}'^2 = \frac{\omega^2}{c^2} \varepsilon'(\omega), \quad 2 \mathbf{k}' \cdot \mathbf{k}'' = \frac{\omega^2}{c^2} \varepsilon''(\omega). \]
\[ |\mathbf{k}'| = \frac{\omega}{c} n(\omega) \approx \frac{\omega}{c} \sqrt{\varepsilon'(\omega)}, \quad |\mathbf{k}''| = \frac{\omega}{c} \kappa(\omega) \approx \frac{1}{2} \frac{\omega}{c} \frac{\varepsilon'(\omega)}{\sqrt{\varepsilon'(\omega)}} \]
\[ - \mathbf{k}' \cdot \mathbf{k}'' \approx |\mathbf{k}'| |\mathbf{k}''| \]
\[ - \mathbf{k}' \text{ and } \mathbf{k}'' \text{ almost parallel } \Rightarrow \text{homogeneous waves} \]
\[ \Rightarrow \text{in homogeneous, isotropic media, next to resonances, we find damped, homogeneous plane waves, } \mathbf{k}' \| \mathbf{k}'' \| \mathbf{e}_k \text{ with } \mathbf{e}_k \text{ being the unit vector along } \mathbf{k} \]
\[ \mathbf{E}(r, t) = \mathbf{E}(\omega) \exp(i \mathbf{k} \cdot \mathbf{r}) = \mathbf{E}(\omega) \exp \left[ i \left( \frac{\omega}{c} n(\omega) \left( \mathbf{e}_k \cdot \mathbf{r} \right) \right) \exp \left[ - \frac{\omega}{c} \kappa(\omega) \left( \mathbf{e}_k \cdot \mathbf{r} \right) \right] \right. \]

**Region 2** \( \varepsilon' < 0, \varepsilon^* > 0, \varepsilon^* < |\varepsilon'| \), *(metals and dielectric media in so-called Reststrahl domain)*
\[ n(\omega) \approx \frac{1}{2} \frac{\varepsilon''(\omega)}{\sqrt{\varepsilon'(\omega)}}, \quad \kappa(\omega) \approx \frac{1}{2} \frac{\varepsilon'(\omega)}{\sqrt{\varepsilon'(\omega)}}. \]

In this regime damping dominates \( (n(\omega) \ll \kappa(\omega)) \), we find a very small refractive index. Interestingly, propagation (nonzero \( n \)) is only possible due to absorption (see time averaged Poynting vector below).

---

Summary of normal modes

- **a)** undamped homogeneous waves and evanescent waves
- **b)** evanescent waves
- **c)** weakly damped quasi-homogeneous waves
- **d)** strongly damped quasi-homogeneous waves

2.4.4 Time averaged Poynting vector of plane waves

\[ \left\langle S(\mathbf{r}, t) \right\rangle = \frac{1}{2 \pi c} \int_{-\infty}^{\infty} \Re \left[ \mathbf{E}(\mathbf{r}, t) \times \mathbf{H}^* (\mathbf{r}, t) \right] dt, \]

For plane waves we find:
\[ \mathbf{E}(\mathbf{r}, t) = \mathbf{E} \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r} - \omega t) = \mathbf{E} \exp(\mathbf{i} \mathbf{k}' \cdot \mathbf{r} - \omega t) \]
\[ \mathbf{H}(\mathbf{r}, t) = \frac{1}{\omega \mu_0} \mathbf{k} \times \mathbf{E}(\mathbf{r}, t) \]
assuming a stationary case \( \mathbf{E}(t) = \overline{\mathbf{E}}(\omega) \)
\[ \left\langle S(\mathbf{r}, t) \right\rangle = \frac{1}{2} \frac{1}{\omega \mu_0} \exp[-2 \mathbf{k}' \cdot \mathbf{r}] |\mathbf{E}|^2 = \frac{1}{2} \frac{1}{\omega \mu_0} |\mathbf{E}|^2 \exp \left[ - \frac{\omega}{c} \kappa(\mathbf{n}_k \cdot \mathbf{r}) \right] |\mathbf{E}|^2. \]

2.5 Beams and pulses - analogy of diffraction and dispersion

In this chapter we will analyze the propagation of light. In particular, we will answer the question how an arbitrary beam (spatial) or pulse (temporal) will change during propagation in isotropic, homogeneous, dispersive media. Relevant (linear) physical effects are diffraction and dispersion. Both phenomena can be understood very easily in the Fourier domain. Temporal effects, i.e. the dispersion of pulses, will be treated in temporal Fourier domain (temporal frequency domain). Spatial effects, i.e. the diffraction of...
beams, will be treated in the spatial Fourier domain (spatial frequency
domain). We will see that:
• **Pulses with finite spatial width** (i.e. pulsed beams) are superposition of
  normal modes (in frequency- and spatial frequency domain).
• **Spatio-temporally localized optical excitations delocalize** during
  propagation because of different phase evolution for different frequencies
  and spatial frequencies (different propagation directions of normal modes).

Let us have a look at the different possibilities (beam, pulse, pulsed beam)

**A) beam ➔ finite transverse width ➔ diffraction**

A beam is a continuous superposition of stationary plane waves (normal
modes) with different propagation directions

\[ A \rightarrow \begin{bmatrix} k^1 \end{bmatrix} + \begin{bmatrix} k^2 \end{bmatrix} + \begin{bmatrix} k^3 \end{bmatrix} + \begin{bmatrix} k^4 \end{bmatrix} + \begin{bmatrix} k^5 \end{bmatrix} + \ldots \]

\[ E(r,t) = \int_{-\infty}^{\infty} \hat{E}(k) \exp[i(k \cdot r - \omega t)] d^3k \]

**B) pulse ➔ finite duration ➔ dispersion**

A pulse is a continuous superposition of stationary plane waves (normal
modes) with different frequencies

\[ A \rightarrow \begin{bmatrix} \omega^1 \end{bmatrix} + \begin{bmatrix} \omega^2 \end{bmatrix} + \begin{bmatrix} \omega^3 \end{bmatrix} + \ldots \]

\[ E(r,t) = \int_{-\infty}^{\infty} \hat{E}(\omega) \exp[i(k \cdot r - \omega t)] d\omega \]

**C) pulsed beams ➔ finite transverse width and finite duration ➔
diffraction and dispersion**

A pulsed beam is a continuous superposition of stationary plane waves
(normal modes) with different frequency and different propagation direction

\[ E(r,t) = \int_{-\infty}^{\infty} \hat{E}(k,\omega) \exp[i(k \cdot r - \omega t)] d^3k d\omega \]

**2.5.1 Diffraction of monochromatic beams in homogeneous isotropic media**

Let us have a look at the propagation of monochromatic beams first. In this
situation, we have to deal with diffraction only. We will see later that pulses
and dispersion can be treated in a very similar way. Treating diffraction in the
framework of wave-optical theory (or even Maxwell) allows us to treat
rigorously many important optical systems and effects, i.e., optical imaging
and resolution, filtering, microscopy, gratings, ...

In this chapter, we assume stationary fields and therefore \( \omega = \text{const} \). For
technical convenience and because it is sufficient for many important
problems, we will make the following assumptions and approximations:
• \( \varepsilon(\omega) = \varepsilon'(\omega) > 0 \), ➔ optical transparent regime ➔ normal modes are
  stationary homogeneous and evanescent plane waves
• **scalar approximation**
  \[ \hat{E}(r,\omega) \rightarrow \hat{E}_s(r,\omega) e^{-i\omega t} \rightarrow \hat{E}_s(r,\omega) \rightarrow \mathbf{u}(r,\omega). \]
  – exact for one-dimensional beams and linear polarization
  – approximation in two-dimensional case

In homogeneous isotropic media we have to solve the Helmholtz equation

\[ \Delta \hat{E}(r,\omega) + \frac{\omega^2}{c^2} \varepsilon(\omega) \hat{E}(r,\omega) = 0. \]

In scalar approximation and for fixed frequency \( \omega \) it reads
\[ \Delta u(r,\omega) + \frac{\omega^2}{c^2} \varepsilon(\omega) u(r,\omega) = 0, \]

scalar Helmholtz equation

\[ \Delta u(r,\omega) + k^2(\omega) u(r,\omega) = 0. \]

In the last step we inserted the dispersion relation (wave number \( k \)). In the following we often omit the fixed frequency \( \omega \).

### 2.5.1.1 Arbitrarily narrow beams \( \Rightarrow \) the general case

Let us consider the following fundamental problem. We want to compute from a given field distribution in the plane \( z = 0 \) the complete field in the half-space \( z > 0 \) (\( z \) is our “propagation direction”)

The governing equation is the scalar Helmholtz equation

\[ \Delta u(r,\omega) + k^2(\omega) u(r,\omega) = 0 \]

To solve this equation and to calculate the dynamics of the fields, we can switch again to the Fourier domain.

We take the Fourier transform

\[ u(r,\omega) = \int_{-\infty}^{\infty} U(k,\omega) \exp\left[i k(\omega)r\right] d^3k \]

which can be interpreted as a superposition of normal modes with different propagation directions and wavenumbers \( k(\omega) \) (here the absolute value of the wave-vector \( k \)). Naively, we could expect that we just constructed a general solution to our problem, but the solution is not correct because of dispersion relation:

\[ k^2 = k^2_x + k^2_y + k^2_z = \frac{\omega^2}{c^2} \varepsilon(\omega) \]

\[ \rightarrow \text{only two components of } k \text{ are independent}, \text{ e.g., } k_x, k_y. \]

Our naming convention is in the following: \( k_x = \alpha, \quad k_y = \beta, \quad k_z = \gamma. \)

Then, the dispersion relation reads:

\[ k^2(\omega) = \alpha^2 + \beta^2 + \gamma^2 \]

Thus, to solve our problem we need only a two-dimensional Fourier transform, with respect to transverse directions to the “propagation direction \( z \)”: \( u(r) = \int \int_{-\infty}^{\infty} U(\alpha,\beta,z) \exp\left[i (\alpha x + \beta y)\right] d\alpha d\beta. \]

In analogy to the frequency \( \omega \) we call \( \alpha, \beta \) spatial frequencies. If we plug this expression into the scalar Helmholtz equation

\[ \Delta u(r) + k^2(\omega) u(r) = 0 \]

This way we can transfer the Helmholtz equation in two spatial dimensions into Fourier space

\[ \left( \frac{d^2}{dz^2} + k^2 - \alpha^2 - \beta^2 \right) U(\alpha,\beta,z) = 0, \]

\[ \left( \frac{d^2}{dz^2} + \gamma^2 \right) U(\alpha,\beta,z) = 0. \]

This equation is easily solved and yields the general solution \( U(\alpha,\beta,z) = U_1(\alpha,\beta) \exp\left[i \gamma(\alpha,\beta)z\right] + U_2(\alpha,\beta) \exp[-i \gamma(\alpha,\beta)z] \).

depending on \( \gamma(\alpha,\beta) = \sqrt{k^2(\omega) - \alpha^2 - \beta^2} \).

We can identify two types of solutions:

A) \( \gamma^2 \geq 0 \): \( \alpha^2 + \beta^2 \leq k^2 \), i.e., \( k \text{ real} \rightarrow \text{homogeneous waves} \)

\[ \gamma^2 \geq 0 \rightarrow \alpha^2 + \beta^2 \leq k^2, \text{ i.e., } k \text{ real} \rightarrow \text{homogeneous waves} \]

B) \( \gamma^2 < 0 \): \( \alpha^2 + \beta^2 > k^2 \), i.e., \( k \text{ complex}, \text{ because } \gamma = k_z \text{ imaginary} \)

Then, we have \( k = k_r + ik^*, \text{ with } k_r = \alpha e_\alpha + \beta e_\beta \text{ and } k^* = \gamma e_\gamma. \)

\( \rightarrow k' \perp k'' \rightarrow \text{evanescent waves} \)
We see immediately that in the half-space $z > 0$ the solution $\sim \exp(-i\gamma z)$ grows exponentially. Because this does not make sense, we have to set $U^+(\alpha, \beta) = 0$. In fact, we will see later that $U^+(\alpha, \beta)$ corresponds to backward running waves, i.e., light propagating in the opposite direction. We therefore find the solution:

$$U(\alpha, \beta; z) = U^+(\alpha, \beta) \exp[\pm i\gamma(\alpha, \beta)z]$$

Furthermore the following boundary condition holds: $U(\alpha, \beta; 0) = U^+_0(\alpha, \beta)$. In spatial space, we can find the optical field for $z > 0$ by inverse Fourier transform:

$$u(r) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\alpha, \beta; z) \exp[i(ax + by)] d\alpha d\beta.$$ 

For homogeneous waves (real $\gamma$) the red term above causes a certain phase shift for the respective plane wave during propagation. Hence, we can formulate the following result:

**Diffraction is due to different phase shifts in propagation direction for different spatial frequencies $\alpha, \beta$.**

The initial spatial frequency spectrum or angular spectrum at $z = 0$

$$U^+_0(\alpha, \beta) = \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_0(x, y) \exp[-i(ax + by)] dxdy,$$

(boundary condition) follows from $u_0(x, y) = u(x, y, 0)$.

As mentioned above the wave-vector components $\alpha, \beta$ are the so-called spatial frequencies. Another common terminology is “direction cosine” for the quantities $\alpha / k, \beta / k$, because of the direct link to the angle of the respective plane wave with the optical $z$-axis.

We can formulate a general scheme to describe diffraction:

1. initial field: $u_0(x, y)$
2. initial spectrum: $U^+_0(\alpha, \beta)$ by Fourier transform
3. propagation: by multiplication with $\exp[i\gamma(\alpha, \beta)z]$
4. new spectrum: $U(\alpha, \beta; z) = U^+_0(\alpha, \beta) \exp[i\gamma(\alpha, \beta)z]$
5. new field distribution: $u(x, y, z)$ by Fourier back transform

This scheme allows for two interpretations:

1) The resulting field distribution is the Fourier transform of the propagated spectrum

$$u(r) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\alpha, \beta; z) \exp[i(ax + by)] d\alpha d\beta.$$ 

2) The resulting field distribution is a superposition of homogeneous and evanescent plane waves ('plane-wave spectrum') which obey the dispersion relation.

$$u(r) = \int_{-\infty}^{\infty} U^+_0(\alpha, \beta) \exp[i(ax + by + \gamma(\alpha, \beta)z)] d\alpha d\beta.$$ 

Let us now discuss the complex transfer function $H(\alpha, \beta; z) = \exp[i\gamma(\alpha, \beta)z]$, which describes the beam propagation in Fourier space. For $z = \text{const.}$ (finite propagation distance) it looks like:

**A) homogeneous waves** $\rightarrow \alpha^2 + \beta^2 \leq k^2$

$$\Rightarrow \exp[i\gamma(\alpha, \beta)z] = 1, \quad \text{arg}(\exp[i\gamma(\alpha, \beta)z]) \neq 0$$

Obviously, $H(\alpha, \beta; z) = \exp[i\gamma(\alpha, \beta)z]$ acts differently on homogeneous and evanescent waves:
Upon propagation the homogeneous waves are multiplied by the phase factor
\[ \exp \left[ i \sqrt{k^2 - \alpha^2 - \beta^2} z \right] \]

**B) evanescent waves** \( \rightarrow \alpha^2 + \beta^2 > k^2 \)
\[ \exp \left[ i \gamma(\alpha, \beta) z \right] = \exp \left[ -i \sqrt{\alpha^2 + \beta^2 - k^2} z \right], \quad \arg \left( \exp \left[ i \gamma(\alpha, \beta) z \right] \right) = 0 \]

Upon propagation the evanescent waves are multiplied by an amplitude factor <1
\[ \exp \left[ -i \sqrt{\alpha^2 + \beta^2 - k^2} z \right] < 1 \]
This means that their contribution gets damped with increasing propagation distance \( z \).

Now the question is: When do we get evanescent waves? Obviously, the answer lies in the boundary condition: Whenever \( u_0(x, y) \) yields an angular spectrum \( U_0(\alpha, \beta) \neq 0 \) for \( \alpha^2 + \beta^2 > k^2 \) we get evanescent waves.

**Example:**
Let us consider the following one-dimensional initial condition which corresponds to an aperture of a slit:

\[
u_0(x) = \begin{cases} 1 & \text{for } |x| = \frac{a}{2}, \\ 0 & \text{otherwise} \end{cases}
\]

\[ U_0(\alpha) = FT[u_0(x)] \sim \frac{\sin \left( \frac{a\alpha}{2} \right)}{\left( \frac{a\alpha}{2} \right)} = \sin \left( \frac{a\alpha}{2} \right) \]

General result
We have seen in the example above that evanescent waves appear for structures < wavelength in the initial condition. Information about those small structures gets lost for \( z \gg \lambda \).

**Conclusion**
In homogeneous media, only information about structural details with \( |\Delta x|, |\Delta y| > \lambda / n \) are transmitted over macroscopic distances. Homogeneous media act like a low-pass filter for light.

**Summary of beam propagation scheme**
\[
u_0(x, y) \rightarrow U_0(\alpha, \beta) \rightarrow U(\alpha, \beta; z) = H(\alpha, \beta; z)U_0(\alpha, \beta) \rightarrow u(x, y, z)
\]

**Remark: diffraction free beams**
With our understanding of diffraction it is straightforward to construct so-called diffraction free beams, i.e., beams that do not change their amplitude distribution during propagation. Translated to Fourier space this means that all spatial frequency components get the same phase shift \( \rightarrow \)

\[
U(\alpha, \beta; z) = U_0(\alpha, \beta) \exp \left[ i \gamma(\alpha, \beta) z \right] = U_0(\alpha, \beta) \exp \left[ i \gamma z \right] \rightarrow u(x, y, z) = \exp \left[ i \gamma z \right] u_0(x, y)
\]
it is straightforward to see that the relevant spatial frequencies lie on a circular ring in the \((\alpha, \beta)\) plane. For constant spectral amplitude on this ring Fourier back-transform yields:

\[ u_0(x, y) = J_0(\rho r) \] (see exercises)

The propagation of the spectrum in Fresnel approximation works in complete analogue to the general case, we just use the modified transfer function to describe the propagation:

\[ U_\alpha(\alpha, \beta; z) = H_\alpha(\alpha, \beta; z)U_\alpha(\alpha, \beta) \]

For a coarse initial field distribution \(u_0(x, y, z)\) the angular spectrum \(U_\alpha(x, y)\) is nonzero for \(\alpha^2 + \beta^2 \ll k^2\) only. Then, only paraxial plane waves are relevant for transmitting information.

### Description in real space

Of course it is possible to formulate beam propagation in Fresnel (paraxial) approximation in position space as well:

\[
H(x, y, z) = \int_{-\infty}^{\infty} H_\alpha(\alpha, \beta; z) U_\alpha(\alpha, \beta) \exp\left[i(\alpha x + \beta y)\right] d\alpha d\beta
\]

with the spatial response function from the convolution theorem

\[ h_\alpha(x, y; z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H_\alpha(\alpha, \beta; z) \exp\left[i(\alpha x + \beta y)\right] d\alpha d\beta \]

This Fourier integral can be solved and we find:

\[
h_\alpha(x, y; z) = \exp(iz)\left\{ -\frac{ik}{2\pi} \exp\left[\frac{i}{2z}(x^2 + y^2)\right]\right\} = \frac{ik}{2\pi} \exp\left[\frac{ikz}{2z^2}(x^2 + y^2)\right].
\]

The response function is a spherical wave in paraxial approximation.

To sum up, in position space paraxial beam propagation is given by:
\[ u_p(x, y, z) = \frac{-i k}{2\pi} \exp(ikz) \int_{-\infty}^{\infty} u_p(x', y') \exp \left( \frac{i k}{2\pi} \left[ (x - x')^2 + (y - y')^2 \right] \right) dx' dy'. \]

Of course, the two descriptions in position space and in the spatial Fourier domain are completely equivalent.

**Remark on the validity of the scalar approximation**

\[ \hat{E}(r, \omega) = \int \hat{E}(\alpha, \beta, \omega) \exp(\alpha x + \beta y + \gamma z) d\alpha d\beta \]

\[ \text{div}\hat{E}(r, \omega) = 0 \rightarrow \alpha \hat{E}_x + \beta \hat{E}_y + \gamma \hat{E}_z = 0 \]

A) One-dimensional beams

- Translational invariance in y-direction: \( \beta = 0 \)
- And linear polarization in y-direction: \( \hat{E}_y \rightarrow U \)

\( \rightarrow \) scalar approximation is exact since divergence condition is strictly fulfilled

B) Two-dimensional beams

- Finite beam which is localized in the \( x, y \)-plane: \( \alpha, \beta \neq 0 \)
- And linear polarization, w.l.o.g. in y-direction: \( \hat{E}_y \rightarrow U \)

\( \rightarrow \) divergence condition: \( \beta \hat{E}_y + \gamma \hat{E}_z = 0 \)

\[ \hat{E}_z(\alpha, \beta, \omega) = -\frac{\beta}{\gamma} \hat{E}_y(\alpha, \beta, \omega) = -\frac{\beta}{\sqrt{k^2 - \alpha^2 - \beta^2}} \hat{E}_y(\alpha, \beta, \omega) \approx 0 \]

In paraxial approximation \((\alpha^2 + \beta^2) \ll k^2\) the scalar approximation is automatically justified.

### 2.5.1.3 The paraxial wave equation

In paraxial approximation the propagated spectrum is given by

\[ U_p(\alpha, \beta; z) = H_p(\alpha, \beta; z) U_o(\alpha, \beta) = \exp(ikz) \exp \left( \frac{-i(\alpha^2 + \beta^2)z}{2k} \right) U_o(\alpha, \beta) \]

Let us introduce the slowly varying spectrum \( V(\alpha, \beta; z) \):

\[ U_p(\alpha, \beta; z) = \exp(ikz) V(\alpha, \beta; z) \rightarrow V(\alpha, \beta; z) = \exp \left( \frac{-i(\alpha^2 + \beta^2)z}{2k} \right) V_o(\alpha, \beta). \]

Differentiation of \( V \) with respect to \( z \) gives:

\[ \frac{d}{dz} V(\alpha, \beta; z) = \frac{1}{2k} (\alpha^2 + \beta^2) V(\alpha, \beta; z) \]

\[ \frac{d}{dz} \int_{-\infty}^{\infty} V(\alpha, \beta; z) \exp(\alpha x + \beta y + \gamma z) d\alpha d\beta \]

\[ = \frac{1}{2k} \int_{-\infty}^{\infty} (\alpha^2 + \beta^2) V(\alpha, \beta; z) \exp(\alpha x + \beta y + \gamma z) d\alpha d\beta \]

\[ \rightarrow \frac{d}{dz} V(\alpha, \beta; z) = \frac{1}{2k} \left( \frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \beta^2} \right) \int_{-\infty}^{\infty} V(\alpha, \beta; z) \exp(\alpha x + \beta y + \gamma z) d\alpha d\beta \]

\[ \rightarrow \frac{d}{dz} v(x, y, z) + \frac{1}{2k} \Delta v(x, y, z) = 0 \]

**Remark**

The slowly varying envelope \( v(x, y, z) \) (Fourier transform of the slowly varying spectrum) relates to the scalar field as \( u_p(x, y, z) = v(x, y, z) \exp(ikz) \).

### Extension of the wave equation to weakly inhomogeneous media

(Paraxial wave equation - SVEA)

There is an alternative, more general way to derive the paraxial wave equation, the so-called slowly varying envelope approximation. This approximation even allows us to treat inhomogeneous media. We start with scalar Helmholtz equation (for inhomogeneous media already an approximation assuming weak spatial fluctuations in \( c(r, \omega) \)):

\[ \Delta u(x, y, z) + k^2(r, \omega) u(x, y, z) = 0 \quad \text{with} \quad k^2(r, \omega) = \frac{\omega^2}{c(r, \omega)} \]

We use the ansatz \( u(x, y, z) = v(x, y, z) \exp(ikz) \) with \( k_0 = \langle k \rangle \) being the average wavenumber. With the SVEA condition

\[ \langle k \rangle v \gg |\frac{\partial v}{\partial z}| \]

we can simplify the scalar Helmholtz equation as follows:

\[ \frac{\partial^2}{\partial x^2} v(x, y, z) + \frac{\partial^2}{\partial y^2} v(x, y, z) + \Delta^2 v(x, y, z) + \left[ k^2(r, \omega) - k_0^2 \right] v(x, y, z) = 0, \]

\[ \rightarrow \frac{\partial}{\partial z} v(x, y, z) + \frac{1}{2k_0} \Delta^2 v(x, y, z) + \left[ k^2(r, \omega) - k_0^2 \right] v(x, y, z) = 0 \]

This is the paraxial wave equation for inhomogeneous media (weak index contrast).
**Beam propagation scheme**

![Beam propagation scheme diagram]

Relation between transfer and response function

\[ h(x, y; z) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} H(\alpha, \beta; z) e^{i(\alpha x + \beta y)} d\alpha d\beta \]

Transfer functions for homogeneous space

\[ H(\alpha, \beta; z) = \exp \left[ i \gamma (\alpha, \beta) z \right] = \exp \left[ i \sqrt{k^2 - \alpha^2 - \beta^2} z \right] \text{ exact solution} \]

\[ H_r(\alpha, \beta; z) = \exp [ikz] \exp \left[ i \frac{\alpha^2 + \beta^2}{2k} z \right] \text{ Fresnel approximation} \]

with \( k = k(\omega) = \frac{\omega}{c} n(\omega) \)

### 2.5.2 Propagation of Gaussian beams

The propagation of Gaussian beams is an important special case. First of all, the transversal fundamental mode of many lasers has Gaussian shape. Second, in paraxial approximation it is possible to compute the Gaussian beam evolution analytically.

![Fundamental Gaussian beam in focus](image)

The general form of a Gaussian beam is elliptic, with curved phase.

\[ u_g(x, y) = v_0(x, y) = A_0 \exp \left[ -\frac{x^2}{w_x^2} - \frac{y^2}{w_y^2} \right] \exp \left[ i \phi(x, y) \right]. \]
Here, we will restrict ourselves to rotational symmetry → $w_x^2 = w_y^2 = w_0^2$ and (initially) 'flat' phase $\varphi(x, y) = 0$, which corresponds to a beam in the focus. The Gaussian beam in the focal plane (flat phase) is characterized by amplitude $A$ and width $w_0$: $u_0(x^2 + y^2 = w_0^2) = A_0 \exp(-1) = A_0/e$. In practice, the so-called 'full width at half maximum' (FWHM) is often used instead of $w_0$.

$$w_{\text{FWHM}}^2 = -\ln 2 \Rightarrow w_{\text{FWHM}}^2 = 2 \ln 2 w_0^2 \approx 1.386 w_0^2$$

2.5.2.1 Propagation in paraxial approximation

Let us now compute the propagation of a Gaussian beam starting from the focus in paraxial approximation:

1) Field at $z = 0$:

$$u_0(x, y) = v_0(x, y) = A_0 \exp\left(-\frac{x^2 + y^2}{w_0^2}\right).$$

2) Angular spectrum at $z = 0$:

$$U_0(\alpha, \beta) = V_0(\alpha, \beta) = \frac{1}{(2\pi)^2} A_0 \iint \exp\left(-\frac{x^2 + y^2}{w_0^2}\right) \exp[-i(\alpha x + \beta y)] dy dx$$

$$= \frac{A_0}{4\pi} w_0^2 \exp\left(-\frac{\alpha^2 + \beta^2}{4/w_0^2}\right) = \frac{A_0}{4\pi} w_0^2 \exp\left(-\frac{\alpha^2 + \beta^2}{w_0^2}\right).$$

We see that the angular spectrum has a Gaussian profile as well and that the width in position space and Fourier space are linked by $w_x \times w_0 = 2$.

$$U_0$$

$\beta$

$\alpha$

Angular spectrum in the focal plane

C) Check if paraxial approximation is fulfilled:

We can say that $U_0(\alpha, \beta) \approx 0$ for $(\alpha^2 + \beta^2) \geq 16/w_0^2$, because $\exp(-4) \approx 0.02$.

For paraxial approximation we need $k^2 \gg (\alpha^2 + \beta^2)$

$$\Rightarrow k^2 \gg 16/w_0^2$$

$$w_0^2 \gg \frac{16}{(2\pi)^2}\left(\frac{2\lambda}{\pi n}\right)^2 \approx \left(\frac{\Lambda}{n}\right)^2,$$

$$\Rightarrow$$

paraxial approximation works for $w_0 \gg 10\frac{\lambda}{n} \equiv 10\lambda_n$

D) Propagation of the angular spectrum:

$$U(\alpha, \beta; z) = V(\alpha, \beta; z) \exp(i k z)$$

$$V(\alpha, \beta; z) = U_0(\alpha, \beta) \exp\left(-\frac{\alpha^2 + \beta^2}{2k z}\right)$$

$$= \frac{A_0}{4\pi} w_0^2 \exp\left(-\frac{\alpha^2 + \beta^2}{4}\right) \exp\left(-\frac{\alpha^2 + \beta^2}{2k z}\right).$$

E) Fourier back-transformation to position space

$$v(x, y, z) = \frac{A_0}{4\pi} w_0^2 \iint \exp\left[-\frac{(\alpha^2 + \beta^2)}{4} - \frac{1}{2k} \right] \exp\left[i(\alpha x + \beta y)\right] d\alpha d\beta$$

$$= \frac{A_0}{1 + \frac{4\lambda^2}{kw_0^2}} \exp\left[-\frac{x^2 + y^2}{w_0^2 \left(1 + \frac{4\lambda^2}{kw_0^2}\right)}\right]$$

$$= \frac{A_0}{1 + \frac{\lambda^2}{w_0^2}} \exp\left[-\frac{x^2 + y^2}{w_0^2 \left(1 + \frac{\lambda^2}{w_0^2}\right)}\right].$$

With the Rayleigh length $z_0$ which determines the propagation of a Gaussian beam:

$$z_0 = \frac{kw_0^2}{2} = \frac{\pi}{\lambda_n} w_0^2$$

Note that we use the slowly varying envelope $v$!

Conclusion:

- Gaussian beam keeps its shape, but amplitude, width, and phase change upon propagation
- Two important parameters: propagation length $z$ and Rayleigh length $z_0$.
Some books use the “diffraction length” \( L_\text{b} = 2z_0 \), a measure for the “focus depth” of the Gaussian beam. E.g.: \( w_0 \gg 10\lambda_\text{m} \rightarrow L_\text{b} \gg 600\lambda_\text{m} \).

From our computation above we know that the Gaussian beam evolves like:

\[
v(x,y,z) = A_0 \frac{1}{1 + (z/z_0)} \exp \left( -\frac{x^2 + y^2}{w_0^2 (1 + (z/z_0)^2)} \right)
\]

For practical use, we can write this expression in terms of \( z \)-dependent amplitude, width, etc.:

\[
v(x,y,z) = A_0 \frac{1-z}{z_0} \exp \left( -\frac{x^2 + y^2}{w_0^2 (1 + (z/z_0)^2)} \right) \exp \left( \frac{k}{2z} \left( \frac{x^2 + y^2}{z_0^2} \right) \right) \exp \left[ i\varphi(z) \right].
\]

Here we used that \( w_0^2 = 2z_0/k \). The \((x,y)\)-independent phase \( \varphi(z) \) is given by \( \tan \varphi = -z/z_0 \), the so-called Gouy phase shift.

In conclusion, we see that the propagation of a Gaussian beam is given by a \( z \)-dependent amplitude, width, phase curvature and phase shift:

\[
v(x,y,z) = A(z) \exp \left( -\frac{x^2 + y^2}{w(z)} \right) \exp \left( \frac{k}{2z} \left( \frac{x^2 + y^2}{R(z)} \right) \right) \exp \left[ i\varphi(z) \right].
\]

**Discussion:**

The amplitude is given as:

\[
A(z) = A_0 \frac{1}{\sqrt{1 + \left( \frac{z}{z_0} \right)^2}} = A_0 \frac{1}{\sqrt{1 + \left( \frac{2z}{I_\text{b}} \right)^2}}.
\]

Hence, we get for the Intensity profile \( I \propto A^2 \):
The beam radius \( W(z) \) has its minimum value \( W_0 \) at the waist \( (z = 0) \), reaches \( \sqrt{2}W_0 \) at \( z = \pm z_0 \), and increases linearly with \( z \) for large \( z \).

The radius of the phase curvature is given by

\[
R(z) = z \left[ 1 + \left( \frac{z_0}{z} \right)^2 \right] = z \left[ 1 + \left( \frac{L_0}{2z} \right)^2 \right]
\]

The flat phase in the focus \((z=0)\) corresponds to an infinite radius of curvature. The strongest curvature (minimum radius) appears at the Rayleigh distance from the focus. The \((x,y)\)-independent Gouy phase is given by

\[
\tan \varphi = \frac{z}{z_0} = \frac{2z}{L_0}
\]

\(2.5.2.2\) Propagation of Gauss beams with \( q \)-parameter formalism

In the previous chapter we gave the expressions for Gaussian beam propagation, i.e., we know how amplitude, width, and phase change with the propagation variable \( z \). However, the complex beam parameter \( q(z) = z - \pm z_0 \) \( q \)-parameter allows an even simpler computation of the evolution of a Gaussian beam. In fact, if we take the inverse of the "\( q \)-parameter",

\[
\frac{1}{q(z)} = \frac{1}{z - \pm z_0} = \frac{z}{z^2 \pm z_0^2} + \frac{1}{z_0^2} = \frac{1}{z_0} \frac{1}{1 + \frac{z_0}{z}}
\]

we can observe that real and imaginary part are directly linked to radius of phase curvature and beam width:

\[
\frac{1}{q(z)} = \frac{1}{R(z)} + \frac{\lambda_n}{\pi w^2(z)}
\]

because \( z_0 = \frac{k w_0^2}{2} = \frac{\pi}{\lambda_n w_0^2} \).

Thus, the \( q \)-parameter describes beam propagation for all \( z \)!

Example: propagation in homogeneous space by \( z = d \)

A) initial conditions:

\[
\frac{1}{q(0)} = \frac{1}{R(0)} + \frac{\lambda_n}{\pi w^2(0)}
\]

B) propagation (by definition of \( q \) parameter) \( q(d) = q(0) + d \)

C) \( q \)-parameter at \( z = d \) determines new width and radius of curvature

\[
\frac{1}{q(d)} = \frac{1}{q(0) + d} \pm \frac{1}{R(d)} + \frac{\lambda_n}{\pi w^2(d)}
\]
2.5.3 Gaussian optics

We have seen in the previous chapter that the complex q-parameter formalism makes a simple description of beam propagation possible. The question is whether it is possible to treat optical elements (lens, mirror, etc.) as well.

**Aim:** for given $R_0, w_0$ (i.e. $q_0$) $\rightarrow R_n, w_n$ (i.e. $q_n$) after passing through $n$ optical elements

We will evaluate the q-parameter at certain propagation distances, i.e., we will have values at discrete positions: $q(z) \rightarrow \hat{q}$.

Surprising property: We can use ABCD-Matrices from ray optics! This is remarkable because here we are doing wave-optics (but with Gaussian beams).

How did it work in geometrical optics?

A) propagation through one optical element:

$$\hat{M} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$ 

B) propagation through multiple elements:

$$\hat{M} = \hat{M}_n\hat{M}_{n-1}...\hat{M}_1 = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$ 

C) matrix connects distances to the optical axis y and inclination angles $\Theta$ before and after the element

$$\begin{bmatrix} y_2 \\ \Theta_2 \end{bmatrix} = \hat{M} \begin{bmatrix} y_1 \\ \Theta_1 \end{bmatrix}.$$
**Link to Gaussian beams**

Let us consider the distance to the intersection of the ray with the optical axis, as it was defined in chapter 1.6.1 on "The ray-transfer-matrix":

\[
z_1 = \frac{y_1}{\Theta_1} \Rightarrow z_2 = \frac{y_2}{\Theta_2} = \frac{A y_1 + B \Theta_1}{C y_1 + D \Theta_1} = \frac{A z_1 + B}{C z_1 + D}
\]

The distances \( z_1, z_2 \) are connected by matrix elements, but not by normal matrix vector multiplication. It turns out that we can pass to Gaussian optics by replacing \( z \) by the complex beam parameter \( q \). The propagation of \( q \)-parameters through an optical element is given by:

\[
q_i = \frac{A q_0 + B}{C q_0 + D}
\]

\( \rightarrow \) propagation through \( N \) elements:

\[
q_n = \frac{A q_0 + B}{C q_0 + D}
\]

with the matrix \( \hat{M} = \hat{M}_0 \hat{M}_1 \ldots \hat{M}_n = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \).

\( \rightarrow \) This works for all ABCD matrices given in chapter 1.6 for ray optics!!!

Here: we will check two important examples:

i) propagation in free space by \( z = d \):

\( \rightarrow \) propagation (by definition of \( q \)-parameter) \( q(d) = q(0) + d \)

\[
\hat{M} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}
\]

\[
q_i = \frac{A q_0 + B}{C q_0 + D} = \frac{q_0 + d}{0 + 1} = q_0 + d
\]

ii) thin lens with focal length \( f \)

What does a thin lens do to a Gaussian beam \( \exp \left( -\frac{1}{2} \frac{x^2 + y^2}{w_0^2} \right) \) in paraxial approximation?

- no change of the width
- but change of phase curvature \( R_f : \times \exp \left( \frac{k}{2} \frac{(x^2 + y^2)}{R_f} \right) \)

How can we see that?

Trick:

We start from the focus which is produced by the lens with \( z_0 = z_f = \frac{\pi w_f^2}{\lambda_s} \)

and \( w_f \) is the focal width. The radius of curvature evolves as:

\[
R(z) = z \left[ 1 + \left( \frac{z_f}{z} \right)^2 \right] \approx z \text{ for } z \gg z_f
\]

We can invert the propagation from the focal position to the lens at the distance of the focal length \( f \) and obtain \( R_f = -f \)

\[
f < 0 \quad \quad f > 0 \quad \quad \hat{M} = \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix}
\]

\( \rightarrow \) double concave lens \( \rightarrow \) defocusing \( \rightarrow \) double convex lens \( \rightarrow \) focusing

\[
q_i = \frac{A q_0 + B}{C q_0 + D} = \frac{q_0}{q_0 + f + 1}
\]

\[
1 = \frac{q_0}{q_0 + f + 1} = \frac{1}{f + \frac{\pi w_0^2}{\lambda_s}} \quad \text{for} \quad q_0 = -i \frac{\pi w_0^2}{\lambda_s}
\]

Be careful: Gaussian optics describes the evolution of the beam's width and phase curvature only!

\( \rightarrow \) Changes of amplitude and reflection are not included!
2.5.4 Gaussian modes in a resonator

In this chapter we will use our knowledge about paraxial Gaussian beam propagation to derive stability conditions for resonators. An optical cavity or optical resonator is an arrangement of mirrors that forms a standing wave cavity resonator for light waves. Optical cavities are a major component of lasers, surrounding the gain medium and providing feedback of the laser light (see He-Ne laser experiment in Labworks).

2.5.4.1 Transversal fundamental modes (rotational symmetry)

The general idea to get a stable light configuration in a resonator is that mirrors and wave fronts (planes of constant phase) coincide. Then, radiation patterns are reproduced on every round-trip of the light through the resonator. Those patterns are the so-called modes of the resonator.

In paraxial approximation and Gaussian beams this condition is easily fulfilled: The radii of mirror and wave front have to be identical!

In this lecture we use the following conventions (different to Labworks script, see remark below!):

- \( z_{1,2} \) is the position of mirror '1', '2'; \( z=0 \) is the position of the focus!
- \( d \) is the distance between the two mirrors \( z_2 - z_1 = d \)
- because \( R(z) = z + \frac{z^2}{2} \) radius of wave front <0 for \( z <0 \)
- from Chapter 1: beam hits concave mirror \( \rightarrow \) radius \( R_i(i=1,2) < 0 \).
- beam hits convex mirror \( \rightarrow \) radius \( R_i(i=1,2) > 0 \).

Examples:

A) \( R(z_1), R(z_2) > 0 \); \( R_1 > 0, R_2 < 0 \); \( z_1 > 0, z_2 > 0 \)

B) \( R(z_1) < 0, R(z_2) > 0 \); \( R_1, R_2 < 0 \); \( z_1 < 0, z_2 > 0 \)

According to our reasoning above, the conditions for stability are:

\[
R_1 = R(z_1), \quad R_2 = -R(z_2)
\]

\[
\implies R_1 = z_1 + \frac{z_1^2}{z_1}, \quad -R_2 = z_2 + \frac{z_2^2}{z_2}
\]

In both expressions we find the Rayleigh length \( z_0 \), which we eliminate:

\[
z_1(R_1 - z_1) = -z_2(R_2 + z_2)
\]

with \( z_1 = z_1 + d \) we find \( z_1 = -\frac{d(R_2 + d)}{R_1 + R_2 + 2d} \).

Now we can choose \( R_1, R_2, d \) and compute modes in the resonator. However, we have to make sure that those modes exist. In our calculations above we have eliminated the Rayleigh length \( z_0 \), a real and positive quantity. Hence, we have to check that the so-called stability condition \( z_0^2 > 0 \) is fulfilled!

\[
\implies z_0^2 = R_1(z_1 - z_1^2) = -\frac{d(R_1 + d)(R_2 + d)}{(R_1 + R_2 + 2d)} > 0
\]

The denominator \( (R_1 + R_2 + 2d)^2 \) is always positive we need to fulfill
\[ -d(R_1 + d)(R_2 + d)(R_1 + R_2 + d) > 0 \]

If we introduce the so-called resonator parameters

\[ g_1 = \left(1 + \frac{d}{R_1}\right), \quad g_2 = \left(1 + \frac{d}{R_2}\right) \]

We can re-express the stability condition as

\[ -d(R_1 + d)(R_2 + d)(R_1 + R_2 + d) = dg_1g_2R_1R_2 \left(1 - \frac{g_1g_2}{R_1R_2}\right)^2 > 0 \]

This inequality is fulfilled for

\[ 0 < g_1g_2 < 1 \quad \text{or} \quad 0 < \left(1 + \frac{d}{R_1}\right) \left(1 + \frac{d}{R_2}\right) < 1 \]

This final form of the stability condition can be visualized: The range of stability of a resonator lies between coordinate axes and hyperbolas:

Resonator stability diagram. A spherical-mirror resonator is stable if the parameters \( g_1 = 1 + \frac{d}{R_1} \) and \( g_2 = 1 + \frac{d}{R_2} \) lie in the unshaded regions bounded by the lines \( g_1 = 0 \) and \( g_2 = 0 \), and the hyperbola \( g_2 = 1/g_1 \). \( R \) is negative for a concave mirror and positive for a convex mirror. Various special configurations are indicated by letters. All symmetrical resonators lie along the line \( g_2 = g_1 \).

Resonator stability diagram. A spherical-mirror resonator is stable if the parameters \( g_1 = 1 + \frac{d}{R_1} \) and \( g_2 = 1 + \frac{d}{R_2} \) lie in the unshaded regions bounded by the lines \( g_1 = 0 \) and \( g_2 = 0 \), and the hyperbola \( g_2 = 1/g_1 \). \( R \) is negative for a concave mirror and positive for a convex mirror. Various special configurations are indicated by letters. All symmetrical resonators lie along the line \( g_2 = g_1 \).

Examples for a stable and an unstable resonator:

A) \( R_1, R_2 < 0 \); \( R_1 > d \); \( \cap \; 0 \leq g_1 \leq 1 \); \( \cap \; 0 \leq g_2 \leq 1 \); \( \cap \; 0 \leq g_1g_2 \leq 1 \) \( \cap \) stable

B) \( R_1, R_2 < 0 \); \( |R_1| < d \); \( |R_2| > d \); \( \cap \; g_1 \leq 0 \); \( \cap \; 0 \leq g_2 \leq 1 \); \( \cap \; g_1g_2 \leq 0 \) \( \cap \) unstable

Remark connection to He-Ne-Labwork script (and Wikipedia):

In Labworks (he_ne_laser.pdf) a slightly different convention is used:

- Direction of z-axis reversed for the two mirrors
- beam hits concave mirror \( \rightarrow \) radius \( R(i=1,2) > 0 \).
- beam hits convex mirror \( \rightarrow \) radius \( R(i=1,2) < 0 \).
- \( z_{1,2} \) is the distance of mirror ‘1’,‘2’ to the focus!
- \( d \) is the distance between the two mirrors \( \rightarrow \) \( z_2 + z_1 = d \)

Examples:

A) \( R(z_1) < 0 \); \( R(z_2) > 0 \); \( R_1 < 0, R_2 > 0 \); \( z_1 < 0, z_2 > 0 \)

B) \( R(z_1) > 0 \); \( R(z_2) > 0 \); \( R_1, R_2 > 0 \); \( z_1 > 0, z_2 > 0 \)
Then the conditions for stability are:

\[ R_1 = R(z_1), \quad R_2 = R(z_2) \]

With analog calculation as above we find with for the resonator parameters

\[ g_1 = \left(1 - \frac{d}{R_1}\right), \quad g_2 = \left(1 - \frac{d}{R_2}\right) \]

the same stability condition

\[ g_1 g_2 (1 - g_1 g_2) (R_1 R_2)^2 > 0, \quad 0 < g_1 g_2 < 1 \]

2.5.4.2 Higher order resonator modes

For the derivation of the above stability condition we needed the wave fronts only. Hence, there may exist other modes with same wave fronts but different intensity distribution. For the fundamental mode we have:

\[ v_g(x, y, z) = A \frac{w_0}{w(z)} \exp \left[ -\frac{x^2 + y^2}{w(z)} \right] \exp \left[ i \frac{k}{2} \frac{x^2 + y^2}{R(z)} \right] \exp \left[i \phi(z)\right]. \]

Higher order modes: (x, y-dependence of phase is the same)

\[ u_{lm}(x, y, z) = A_{lm} \frac{w_0}{w(z)} G_l \left[ \sqrt{2 \frac{x}{w(z)}} \right] G_m \left[ \sqrt{2 \frac{y}{w(z)}} \right] \times \]

\[ \exp \left[ i \frac{k}{2} \frac{x^2 + y^2}{R(z)} \right] \exp \left[i k z\right] \exp \left[i (l + m + 1) \phi(z)\right]. \]

\[ G_l(\xi) = H_l(\xi) \exp \left(-\frac{\xi^2}{2}\right) \]

The functions \( G_l \) are given by the so-called Hermite polynomials:

\( H_l(\xi) \quad (H_0 = 1, H_1 = 2\xi \quad \text{and} \quad H_{l+1} = 2\xi H_l - 2l H_{l-1}) \).

2.5.5 Pulse propagation

2.5.5.1 Pulses with finite transverse width (pulsed beams)

In the previous chapters we have treated propagation of monochromatic beams, where the frequency \( \omega \) was fix and therefore the wavenumber \( k(\omega) \) was constant as well. This is the typical situation when we deal with continuous-wave (cw) lasers.

However, for many applications (spectroscopy, nonlinear optics, telecommunication, material processing) we need to consider the propagation of pulses. In this situation, we have typical envelope length \( T_0 \) of \( 10^{-13} \) s (100 fs) \( \leq T_0 \leq 10^{-10} \) s (100 ps).

Let us compute the spectrum of the (Gaussian) pulse:

\[ f(t) = \exp \left(-i \omega t\right) \exp \left(-\frac{t^2}{T_0^2}\right) \]

\[ F(\omega) \sim \exp \left[-\frac{(\omega - \omega_0)^2}{4 / T_0^2}\right] \Rightarrow \omega_0^2 = \frac{4}{T_0^2} \Rightarrow \omega T_0 = 2 \]

\( \Rightarrow \) spectral width: \( 4 \cdot 10^{10} \text{s}^{-1} \leq \omega_0 \leq 4 \cdot 10^{13} \text{s}^{-1} \)
center frequency of visible light: $\omega_0 = 2\pi v - 4 \cdot 10^{15}$ s$^{-1}$

→ optical cycle: $T_s = 2\pi / \omega_0 \approx 1.6$ fs

Hence, we have the following order of magnitudes:

$$\omega_s \ll \omega_0 \rightarrow \omega - \omega_0 = \omega_0 \ll \omega_s$$

In this situation it can be helpful to replace the complicated frequency dependence of the wave number (dispersion relation) by a Taylor expansion at $\omega = \omega_0$. In general, a parabolic (or cubic) approximation will be sufficient:

$$k(\omega) \approx k(\omega_0) + \frac{\partial k}{\partial \omega_0} (\omega - \omega_0) + \frac{1}{2} \frac{\partial^2 k}{\partial \omega_0^2} (\omega - \omega_0)^2 + \ldots$$

The following terminology is commonly used in the literature:

- phase velocity $v_{ph}$
- group velocity $v_g$
- group velocity dispersion (GVD) $D_g$

We reduce the complicated dispersion relation to three parameters:

- phase velocity $v_{ph}$
- group velocity $v_g$
- group velocity dispersion (GVD) $D_g$

Discussion:

A) group velocity and group index

→ group velocity is the velocity of the center of the pulse (see below)

$$k(\omega) = k(\omega_0) + \frac{\partial k}{\partial \omega_0} (\omega - \omega_0) + \frac{1}{2} \frac{\partial^2 k}{\partial \omega_0^2} (\omega - \omega_0)^2 + \ldots$$

$$\frac{1}{v_g} = \frac{\partial k}{\partial \omega_0} (\omega_0) = \frac{n(\omega_0)}{c} \quad \text{phase velocity}$$

$$D_g = \frac{\partial^2 k}{\partial \omega_0^2} (\omega_0) \quad \text{group velocity dispersion (GVD)}$$

→ GVD changes pulse shape upon propagation (see below)

$$D = D_g = \frac{\partial^2 k}{\partial \omega_0^2} (\omega_0) = \frac{\partial}{\partial \omega_0} \left( \frac{1}{v_g} \right)$$

$$\Delta \mathbf{P}(\mathbf{r}, \omega) + \frac{\omega^2}{c^2} \mathbf{e}(\omega) \cdot \mathbf{P}(\mathbf{r}, \omega) = 0$$

In contrast to monochromatic beam propagation, we now have for each $\omega$ one Fourier component of the optical field:

$$\mathbf{D} \mathbf{E}(\mathbf{r}, \omega) + \frac{\omega^2}{c^2} \mathbf{E}(\omega) \cdot \mathbf{E}(\mathbf{r}, \omega) = 0$$

$$\frac{\partial}{\partial \omega_0} \left( \frac{1}{v_g} \right) = -\frac{1}{v_g^2} \frac{\partial v_g}{\partial \omega_0}$$

$$D = \frac{\partial}{\partial \omega_0} \left( \frac{1}{v_g} \right)$$

$$D = \frac{\partial}{\partial \omega_0} \left( \frac{1}{v_g} \right) = -\frac{1}{v_g^2} \frac{\partial v_g}{\partial \omega_0}$$

$$D = -\frac{1}{v_g^2} \frac{\partial v_g}{\partial \omega_0}$$

Let us now discuss the propagation of pulsed beams. We start with the scalar Helmholtz equation, with the full dispersion (no Taylor expansion yet):

$$\Delta \mathbf{P}(\mathbf{r}, \omega) + \frac{\omega^2}{c^2} \mathbf{e}(\omega) \cdot \mathbf{P}(\mathbf{r}, \omega) = 0$$

Alternatively in telecommunication one often uses

$$D = \frac{\partial}{\partial \lambda} \left( \frac{1}{v_g} \right) = -\frac{2\pi}{\lambda^2} \frac{\partial D_g}{\partial \omega_0}$$

Let us now discuss the propagation of pulsed beams. We start with the scalar Helmholtz equation, with the full dispersion (no Taylor expansion yet):

$$\Delta \mathbf{P}(\mathbf{r}, \omega) + \frac{\omega^2}{c^2} \mathbf{e}(\omega) \cdot \mathbf{P}(\mathbf{r}, \omega) = 0$$

In contrast to monochromatic beam propagation, we now have for each $\omega$ one Fourier component of the optical field:

$$\mathbf{D} \mathbf{E}(\mathbf{r}, \omega) + \frac{\omega^2}{c^2} \mathbf{E}(\omega) \cdot \mathbf{E}(\mathbf{r}, \omega) = 0$$

$$\text{dispersion relation: } k^2(\omega) = \frac{\omega^2}{c^2} \mathbf{e}(\omega)$$
Hence, we need to consider the propagation of the Fourier spectrum (Fourier transform in space and time):

\[ U(\alpha, \beta, \omega; z) = U_0(\alpha, \beta, \omega) \exp\left[ i \gamma(\alpha, \beta, \omega) \right] \] 

with \( \gamma(\alpha, \beta, \omega) = \sqrt{k^2(\omega) - \alpha^2 - \beta^2} \)

The initial spectrum at \( z = 0 \) is \( U_0(\alpha, \beta, \omega) \)

\[ U_0(\alpha, \beta, \omega) = \frac{1}{(2\pi)^3} \iint u_0(x, y, t) \exp\left[ -i(\alpha x + \beta y - \omega t) \right] dx dy dt \]

Let us further assume Fresnel (paraxial) approximation is justified \( (k^2(\omega) >> \alpha^2 + \beta^2) \)

\[ U(\alpha, \beta, \omega; z) \approx U_0(\alpha, \beta, \omega) \exp\left[ i k(\omega) z \right] \exp\left[ -i \frac{\alpha^2 + \beta^2}{2k(\omega)} z \right] \]

We see that propagation of pulsed beams in Fresnel approximation in Fourier space is described by the following propagation function (transfer function):

\[ H_{\text{tr}}(\alpha, \beta, \omega; z) = \exp\left[ i k(\omega) z \right] \exp\left[ -i \frac{\alpha^2 + \beta^2}{2k(\omega)} z \right] \]

Now let us consider the Taylor expansion of \( k(\omega) \) from above. If the pulse is not too short, we can replace

\[ k(\omega) \approx k(\omega_0) + \frac{\partial k}{\partial \omega}(\omega - \omega_0) + \frac{1}{2} \frac{\partial^2 k}{\partial \omega^2}(\omega - \omega_0)^2 + .. \]

Moreover, in the second term \( \exp\left[ -i \frac{\alpha^2 + \beta^2}{2k(\omega)} z \right] \) (which is already small due to paraxiality) we can use \( k(\omega_0) \approx k_0 \) (breaks down for \( T_0 \lesssim 15 \) fs). By doing so, we get the so-called parabolic approximation:

\[ H_{\text{tr}}(\alpha, \beta, \omega; z) \approx \exp\left[ i k_0 z \right] \exp\left[ \frac{1}{v_g}(\omega - \omega_0) z \right] \times \exp\left[ i \frac{D}{2}(\omega - \omega_0)^2 z \right] \exp\left[ -i \frac{\alpha^2 + \beta^2}{2k_0} z \right] \]

\[ = \exp\left[ i k_0 z \right] \exp\left[ i z \left( \frac{\omega - \omega_0}{v_g} + \frac{1}{2} \frac{D}{k_0} \right) \frac{1}{2} \left( \alpha^2 + \beta^2 \right) \right] \]

with \( \omega = \omega - \omega_0 \)

In the last line of the above equation we have introduced a new variant of the propagation function. \( \overline{H}_{\text{tr}}(\alpha, \beta, \omega, z) \) is the propagation function for the slowly varying envelope \( v(x, y, t) \):

\[ u(x, y, t, z) = \exp\left[ i k_0(z - \omega t) \right] \iint U_0(\alpha, \beta, \omega) \overline{H}_{\text{tr}}(\alpha, \beta, \omega, z) \exp\left[ i \left( \alpha x + \beta y - \omega t \right) \right] d\alpha d\beta d\omega \]

\[ v(x, y, t) = v(x, y, t) \exp\left[ i k_0(z - \omega t) \right] \]

In order to complete the formalism, we also need the initial spectrum of the slowly varying envelope

\[ v_0(x, y, t) = v_0(x, y, t) \exp\left[ -i \omega_0 t \right] \]

\[ \rightarrow V_0(\alpha, \beta, \omega_0) = \frac{1}{(2\pi)^3} \iint v_0(x, y, t) \exp\left[ -i \left( \alpha x + \beta y - \omega t \right) \right] dx dy dt \]
Thus, the propagation of the slowly varying envelope is given by:

\[
v(x, y, z, t) = \int_0^\infty V_0(\alpha, \beta, \omega, z) \bar{H}_{\text{rep}}(\alpha, \beta, \omega; z) \exp\left[\frac{i}{v_g} \left(\alpha x + \beta y - \omega_0 t\right)\right] d\alpha d\beta d\omega
\]

The next step is to introduce a co-moving reference frame with

\[
\bar{H}_{\text{rep}}(\alpha, \beta, \omega; z) = \exp\left[\frac{-i}{v_g} \left(\omega_0 z\right)\right] H_{\text{rep}}(\alpha, \beta, \omega; z)
\]

\[
v(x, y, z, t) = \int_0^\infty V_0(\alpha, \beta, \bar{\omega}, z) \bar{H}_{\text{rep}}(\alpha, \beta, \bar{\omega}; z) \exp\left[\frac{i}{v_g} \left(\alpha x + \beta y - \bar{\omega}_0 (t - z/v_g)\right)\right] d\alpha d\beta d\bar{\omega}
\]

The last line above involves the propagation function \( \bar{H}_{\text{rep}}(\alpha, \beta, \bar{\omega}; z) \), the propagation function for the slowly varying envelope in the co-moving frame of the pulse:

\[
\tau = t - \frac{z}{v_g}
\]

This frame is called co-moving because \( \bar{H}_{\text{rep}}(\alpha, \beta, \bar{\omega}; z) \) is now purely quadratic in \( \bar{\omega} \), i.e., the pulse does not "move" anymore. In contrast, the linear \( \omega \)-dependence in Fourier space had given a shift in the time domain. Thus, the slowly varying envelope in the co-moving frame evolves as:

\[
\bar{v}(x, y, z, \tau) = \int_0^\infty V_0(\alpha, \beta, \bar{\omega}, z) \exp\left[\frac{i}{v_g} \left(\alpha \frac{z}{v_g} - \frac{\omega_0^2 + \beta^2}{k_0}\right)\right] \times \exp\left[\frac{i}{v_g} \left(\alpha x + \beta y - \omega_0 \tau\right)\right] d\alpha d\beta d\bar{\omega}
\]

The optical field \( u \) reads in the co-moving frame:

\[
\cdots \dot{u}(x, y, z, \tau) = \bar{v}(x, y, z, \tau) \exp\left[\frac{i}{v_g} \left(k_0 z - \omega_0 \tau\right)\right] = \bar{v}(x, y, z, \tau) \exp\left[\frac{i}{v_g} \left(k_0 z - \omega_0 \tau\right)\right]
\]

Finally, let us derive the propagation equation for the slowly varying envelope in the co-moving frame:

\[
\frac{\partial \bar{v}(x, y, z, \tau)}{\partial \tau} = \frac{v_0}{v_g} \nabla \cdot \nabla \bar{v}(x, y, z, \tau)
\]

\[
\frac{\partial \bar{v}(x, y, z, \tau)}{\partial \tau} = \frac{1}{2} \left(D_0 \omega^2 - \frac{\alpha^2 + \beta^2}{k_0}\right) \bar{v}(x, y, z, \tau)
\]

As before in the case of monochromatic beams, we use Fourier back-transformation to get the differential equation in time-position domain.
B) the evolution equation for slowly varying envelope in the co-moving frame reads
\[ \frac{\partial \tilde{v}(z, \tau)}{\partial z} - \frac{D}{2} \frac{\partial^2 \tilde{v}(z, \tau)}{\partial \tau^2} = 0 \]

**Analogy of diffraction and dispersion**

**DIFFRACTION**
\[ \frac{\partial^2 v(x, y, z)}{\partial x^2} + \frac{1}{2k} A^{2D} v(x, y, z) = 0 \]

\[ (\alpha, \beta) \leftrightarrow \hat{\omega} \]
\[ (x, y) \leftrightarrow \tau \]
\[ V \leftrightarrow \frac{\partial}{\partial \tau} \]
\[ \frac{1}{k_0} \leftrightarrow -D \text{ but } D \lesssim 0 \text{ can vary} \]

**Examples of pulse propagation**

2.5.5.3 **Example 1: Gaussian pulse without chirp**

use analogies to spatial diffraction

1. **Initial pulse shape**

pulse without chirp \( \rightarrow \) corresponds to Gaussian pulse in the waist with flat phase

\[ u_0(t) = A_0 \exp \left( -\frac{t^2}{T_0} \right) \exp \left( -i \omega_0 t \right) \]

\[ v_0(\tau) = A_0 \exp \left( -\frac{\tau^2}{T_0} \right) \]

2. **Initial pulse spectrum**

\[ V_0(\omega) = A_0 \frac{T_0}{2\sqrt{\pi}} \exp \left( -\frac{\omega^2 T_0^2}{4} \right) \]

**Dispersion length:** \( L_D = 2|z_0| \)

Use results from propagation of Gaussian beams:

\[ z_0 \text{ describes Gaussian pulse} \]

\[ z_0 = \frac{k}{2} v_0^2 \rightarrow z_0 = -\frac{1}{2} \frac{T_0^2}{D} \geq 0 \]

hence anomalous GVD is equivalent to 'normal' diffraction

3. **Evolution of the amplitude**

\[ \tilde{v}(z, \tau) = A_0 \sqrt{\frac{T_0}{T(z)}} \exp \left[ -\frac{\tau^2}{T(z)} \right] \exp \left[ -\frac{i}{2D} \frac{\tau^2}{R(z)} \right] \exp [i \varphi(z)] \]

with

\[ A(z) = A_0 \sqrt{\frac{1}{1 + \left( \frac{z}{z_0} \right)^2}} \]

\[ T(z) = T_0 \sqrt{1 + \left( \frac{z}{z_0} \right)^2} \]

\[ A^2(z)T(z) = \text{const.} \]

'Phase curvature' is not fitting to the description of pulses \( \rightarrow \) introduction of new parameter **Chirp**

\[ \frac{\partial^2 \Phi(x, y, z)}{\partial x^2 + \frac{\partial^2 \Phi(x, y, z)}{\partial y^2}} = \frac{k}{R(z)} \]

\[ \Phi(t) = -\omega t \rightarrow -\frac{\partial \Phi(t)}{\partial t} = \omega \]

\( \rightarrow \) arbitrary time dependence of phase

\[ -\frac{\partial \Phi(t)}{\partial t} = \omega(t) \text{ and } -\frac{\partial^2 \Phi(t)}{\partial t^2} = \frac{\partial \omega(t)}{\partial t} \neq 0 \rightarrow \text{chirp} \]

\( \rightarrow \) parabolic approximation \( \rightarrow \) 'chirp' constant \( \rightarrow \) dimensionless chirp parameter (often just chirp)

\[ C = -\frac{T_0^2}{2} \frac{\partial^2 \Phi(t)}{\partial t^2} \]

\( \rightarrow \) variable frequency of the pulse in time
integration leads to:
\[-\Phi'(\tau) = \omega(\tau) = \omega_0 + 2C \frac{\tau}{T_0^2}, \quad -\Phi(\tau) = \omega_0 \tau + C \frac{\tau^2}{T_0^2} \]

C > 0 \rightarrow \text{up-chirp}

C < 0 \rightarrow \text{down-chirp}

leading front

trailing tail

phase curvature \( R(z) \rightarrow \text{Chirp } C(z) \)

**Complete phase:**

\[ \Phi(\tau) = -\omega_0 \left( \tau + \frac{z}{v_g} \right) - \frac{\tau^2}{2DR(z)} = -\omega_0 \left( \tau + \frac{z}{v_g} \right) - C(z) \frac{\tau^2}{T_0^2} \]

\[ \rightarrow C(z) = \frac{T_0^2}{2DR(z)} = -\frac{z_0}{R(z)} \]

with

\[ R(z) = \frac{z^2 + z_0^2}{z} \]

\[ C(z) = -\frac{z_0 z}{z^2 + z_0^2} = -\frac{z}{z_0 \left(1 + \frac{z}{z_0} \right)} \]

\[ \rightarrow C(0) = 0, \quad C\left(\left|z_0\right|\right) = \frac{1}{2} \text{sgn}(z_0), \quad C(z \rightarrow \infty) = -\frac{z_0}{z} \text{ with } z_0 = -\frac{T_0^2}{2D} \]

Attention: Chirp depends on sign of \( z_0 \) and hence of \( D \).

Complete field:

\[ u(z, \tau) = A \sqrt{\frac{T_0}{T(z)}} \exp\left[-\frac{\tau^2}{T(z)}\right] \exp\left[-iC(z)\frac{\tau^2}{T_0^2}\right] \exp[i\phi(z)] \exp\left[i(k_0 z - \omega_0 \tau)\right] \]

Dynamics of a pulse is equivalent to that of a beam.

important parameter \( \rightarrow \) dispersion parameter \( z_0 = -\frac{T_0^2}{2D} \)

1) \( z \ll \left|z_0\right| \): no effect
2) \( z = \left|z_0\right| \): similar to beam diffraction
3) \( z \gg \left|z_0\right| \): asymptotic dependence

\[ \left|v_\delta(\omega)\right|^2 \approx \frac{4(1 + C_0^2)}{T_0^2} \]

\( \Rightarrow \) spectral width of chirped pulse is larger than that of unchirped pulse
Aim: calculation of pulse width and chirp in dependence on z for given initial conditions

Gaussian beam → q -parameter → similar for Gaussian pulse

Use analogy: however it is limited to homogeneous space

\( q(z) = q(0) + z \)

Remember beams:

\[
\frac{1}{q(z)} = \frac{1}{R(z)} + \frac{k}{w^2(z)}.
\]

\( k \rightarrow -\frac{1}{D}, \quad w^2(z) \rightarrow R^2(z), \quad R(z) \rightarrow 2DC(z) \)

\[
\frac{1}{q(z)} = \frac{2DC(z)}{T^2_0} - \frac{1}{T^2(z)}
\]

\[
\frac{1}{q(z)} = \frac{2D}{T^2_0} C(z) - \frac{k}{T^2(z)}
\]

\( T_0 \) → pulse width at \( z = 0 \), which is not necessarily in the 'focus' or waist hence at \( z = 0 \):

\[
\frac{1}{q(0)} = \frac{2D}{T^2_0} [C(0) - 1] \quad C_0 = C(0).
\]

Idea:

a) \( q(z) = q(0) + z \) and \( \frac{1}{q(0)} = \frac{2D}{T^2_0} [C(0) - 1] \) insert into \( \frac{1}{q(z)} \)

b) \( \frac{1}{q(z)} = \frac{2D}{T^2_0} C(z) - \frac{k}{T^2(z)} \)

set a) and b) equal \( T(z), C(z) \)

generally: 2 equations for \( C_0, T_0, C(z), T(z) \) → 3 values predetermined here: \( z = d \)

1) Determination of q parameter at input

\[
q(0) = T^2_0 \left( C_0 + \frac{1}{d} \right) \]

2) Evolution of q parameter

\[
q(d) = q(0) + d = \frac{T^2_0 \left( C_0 + \frac{1}{d} \right)}{2D(1 + C_0^2)} + d = \frac{1}{2D(1 + C_0^2)} \left[ 2Dd \left( 1 + C_0^2 \right) + C_0 T^2_0 + \frac{1}{d} T^2_0 \right]
\]

3) Inversion of general equation (*) for q(d)

\[
\frac{1}{q(d)} = \frac{2D}{T^2_0} C(d) - \frac{k}{T^2(d)}
\]

\[
q(d) = \frac{T^2_0 T^2(d)}{2D \left[ C^2(d)T^4(d) + T^2_0 \right]} \left[ C(d)T^2(d) + \frac{1}{d} T^2_0 \right]
\]

4) Set two equations equal

\[
\frac{2Dd \left( 1 + C_0^2 \right) + C_0 T^2_0 + \frac{1}{d} T^2_0}{2D(1 + C_0^2)} = \frac{T^2_0 T^2(d)}{2D \left[ C^2(d)T^4(d) + T^2_0 \right]}
\]

a) real part

\[
\frac{2Dd \left( 1 + C_0^2 \right) + C_0 T^2_0}{(1 + C_0^2)} = \frac{C(d)T^2(d)}{C^2(d)T^4(d) + T^2_0}
\]

b) imaginary part

\[
\frac{1}{(1 + C_0^2)} = \frac{T^2_0 T^2(d)}{C^2(d)T^4(d) + T^2_0}
\]

If we predetermine 3 parameters \( C_0, T_0, C(d) \), we can determine the other 2 unknown parameters \( d, T(d) \).

Important case: Where is the pulse compressed to the smallest length?

given: \( C_0, T_0, C(d) = 0 \) (focus)

unknown: \( d, T(d) \)

a) real part must be zero

\[
2Dd \left( 1 + C_0^2 \right) + C_0 T^2_0 = 0
\]

\[
d = -\frac{1}{2D(1 + C_0^2)} = \frac{1}{2} \text{sgn}(D) \frac{C_0}{(1 + C_0^2)} L_p
\]

b) \( T^2(d) = \frac{T^2_0}{(1 + C_0^2)} \)

Resulting properties:

1) A pulse can be compressed when the product of initial chirp and dispersion is negative \( \rightarrow C_0 D < 0 \).

2) The eventual compression increases with initial chirp.
Physical interpretation
If e.g. $C_0 < 0$ and $D > 0 \rightarrow \tilde{\epsilon}_v / \tilde{\epsilon}_a < 0$ is 'red' faster than 'blue'.

Compression of a chirped pulse in a medium with normal dispersion. The low frequency (marked R for red) occurs after the high frequency (marked B for blue) in the initial pulse, but it catches up since it travels faster. Upon further propagation, the pulse spreads again as the R component arrives earlier than the B component.

1) First the 'red tail' of the pulse catches up with the 'blue front' until $C(z) = 0$ (waist), i.e. the pulse is compressed, no chirp remains.
2) Then $C(z) > 0$ and red is in front. Subsequently the 'red front' is faster than the 'blue tail', i.e. the pulse gets wider.

$$C(z) = -\frac{z}{z_0 \left(1 + \frac{z_0}{z} \right)} = -\frac{z}{z_0 + \frac{z_0^2}{2D}}$$

2.6 (The Kramers-Kronig relation, covered by lecture Structure of Matter)
The topic of "Kramers-Kronig relation" is not covered in the course "Fundamentals of modern optics". This topic will be covered rather by the course "Structure of matter". The following part of the script which is devoted to this topic is just included in the script for consistency.

It is possible to derive a very general relation between $\varepsilon'(\omega)$ (dispersion) and $\varepsilon''(\omega)$ (absorption). This means in practice that we can compute $\varepsilon''(\omega)$ from $\varepsilon'(\omega)$ and vice versa. For example, if we have access to the absorption spectrum of a medium, we can calculate the dispersion relation.

The Kramers-Kronig relation follows from reality and causality of response function $R$. That the response function is real valued is a direct consequence from Maxwell's equations which are real valued as well, and causality is also a very fundamental property: The polarization must not depend on some future electric field. As we have seen in the previous chapter, in time-domain the polarization and the electric field are related as.

$$P_i (r, t) = \varepsilon_0 \int_{-\infty}^{\infty} R(t-t')E_i (r, t') dt' \leftrightarrow P_i (r, t) = \varepsilon_0 \int_{-\infty}^{\infty} R(t)E_i (r, t-t) dt$$

Reality of the response function implies:

$$R(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \chi'(\omega)e^{i\omega t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \chi^*(\omega)e^{i\omega t} \rightarrow \chi(\omega) = \chi^*(-\omega)$$

Causality of the response function implies:

$$R(\omega) = \theta(\omega) y(\omega) \rightarrow \begin{cases} 1 & \text{for } \omega > 0 \\ 0 & \text{for } \omega < 0 \end{cases}$$

In the following, we will make use of the Fourier transform of Heaviside distribution:

$$2\pi \delta(\omega) = \int_{-\infty}^{\infty} dt \theta(t)e^{i\omega t} = P \frac{i}{\omega} + \pi \delta(\omega)$$

In Fourier space, the Heaviside distribution consists of the Dirac delta distribution

$$\int_{-\infty}^{\infty} d\omega \delta(\omega) f(\omega) = f(\omega_0) \rightarrow \text{Dirac delta distribution}$$

and an expression involving a Cauchy principal value:

$$P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} f(\omega) = \lim_{\alpha \rightarrow 0} \left[ \int_{-\infty}^{-\alpha} d\omega \frac{i}{\omega} f(\omega) + \int_{\alpha}^{\infty} d\omega \frac{i}{\omega} f(\omega) \right]$$

As we have seen above, causality implies that the response function has to contain a multiplicative Heaviside function. Hence, in Fourier space (susceptibility) we expect a convolution:

$$\chi(\omega) = \int_{-\infty}^{\infty} d\tau R(\tau)e^{i\omega \tau} = \int_{-\infty}^{\infty} d\tau \theta(\tau) y(\tau)e^{i\omega \tau} \rightarrow \tilde{\chi}(\omega) = \frac{1}{2\pi} \frac{1}{\omega} + \frac{1}{2} \frac{P}{\omega}$$

In order to derive the Kramers-Kronig relation we can use a small trick (this trick saves us using complex integration in the derivation). Because of the Heaviside function, we can choose the function $y(\tau)$ for $\tau < 0$ arbitrarily without altering the susceptibility! In particular, we can choose:
a) \( y(-\tau) = y(\tau) \) even function

b) \( y(-\tau) = -y(\tau) \) odd function

In this case \( y(-\tau) = y(\tau) \) is a real valued and even function. We can exploit this property and show that

\[
\mathcal{P}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \, y(\tau) e^{i\omega \tau} = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \, y(\tau) e^{-i\omega \tau} = \mathcal{P}^*(\omega) \text{ is real as well}
\]

Hence, we can conclude from equation (*) above that

\[
\chi^*(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \, \frac{i \mathcal{P}(\omega)}{\omega - \omega_0} + \frac{\mathcal{P}(\omega)}{2} = \mathcal{P}(\omega)
\]

Now we have expressions for \( \chi(\omega), \chi^*(\omega) \) and can compute real and imaginary part of the susceptibility:

\[
\chi(\omega) + \chi^*(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \, \frac{i \mathcal{P}(\omega)}{\omega - \omega_0} + \frac{\mathcal{P}(\omega)}{2} - \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \, \frac{i \mathcal{P}(\omega)}{\omega - \omega_0} + \frac{\mathcal{P}(\omega)}{2} = \mathcal{P}(\omega)
\]

\[
\chi(\omega) - \chi^*(\omega) = \ldots = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \, \frac{i \mathcal{P}(\omega)}{\omega - \omega_0}
\]

Plugging the last two equations together we find the first Kramers-Kronig relation:

\[
\rightarrow \Re \chi(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \, \frac{\mathcal{P}(\omega)}{\omega - \omega_0} 1. \text{K-K relation}
\]

Knowledge of the real part of the susceptibility (dispersion) allows us to compute the imaginary part (absorption).

b) \( y(-\tau) = -y(\tau) \)

The second K-K relation can be found in a similar procedure when we assume that \( y(-\tau) = -y(\tau) \) is a real odd function. We can show that in this case

\[
\mathcal{P}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \, y(\tau) e^{i\omega \tau} = -\frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \, y(\tau) e^{-i\omega \tau} = -\mathcal{P}^*(\omega) \text{ is purely imaginary}
\]

With equation (*) we then find that

\[
\chi^*(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \, \frac{i \mathcal{P}(\omega)}{\omega - \omega_0} \frac{\mathcal{P}(\omega)}{2} \text{ (see *)} \quad \text{and}
\]

Again we can then compute real and imaginary part of the susceptibility.
3. Diffraction theory

3.1 Interaction with plane masks

In this chapter we will use our knowledge on beam propagation to analyze diffraction effects. In particular, we will treat the interaction of light with thin and plane masks/apertures. Therefore we would like to understand how a given transversally localized field distribution propagates in a half-space. There are different approximations commonly used to describe light propagation behind an amplitude mask:

A) If we use geometrical optics we get a simple shadow.

B) We can use scalar diffraction theory with approximated interaction, i.e., a so-called aperture is described by a complex transfer function $t(x, y)$ with $t(x, y) = 0$ for $|y| > a$ (aperture) here we consider the description based on scalar diffraction theory. Then we can split our diffraction problem into three processes:

i) propagation from light source to aperture

$\rightarrow$ not important, generally plane wave (no diffraction)

ii) multiply field distribution of illuminating wave by transfer function $u_{i}(x, y, z) = t(x, y)u_{i}(x, y, z_{a})$

iii) propagation of modified field distribution behind the aperture through homogeneous space

$u(x, y, z) = \int_{-\infty}^{\infty} H(\alpha, \beta; z_{a}) U_{i}(\alpha, \beta; z_{a}) \exp\left[\frac{i}{\lambda}(\alpha x + \beta y)\right] d\alpha d\beta$

or

$u(x, y, z) = \int_{-\infty}^{\infty} h(x-x', y-y', z_{a}) u_{i}(x', y', z_{a}) d\alpha d\beta$

with $h = \frac{1}{(2\pi) \cdot \text{FT}^{-1}[H]}$

In the following we will use the notation $z_{a} = z - z_{a}$. According to our choice of the propagation function $H$, resp. $h$, we can compute this propagation either exactly or in a paraxial approximation (Fresnel). In the following, we will see that a further approximation for very large $z_{a}$ is possible, the so-called Fraunhofer approximation.

3.2 Propagation using different approximations

3.2.1 The general case - small aperture

We know from before that for arbitrary fields (arbitrary wide angular spectrum) we have to use the general propagation function

$H(\alpha, \beta; z_{a}) = \exp\left[\frac{i}{\lambda}(\alpha \gamma(\alpha, \beta) z_{a})\right]$

where $\gamma^{2} = k^{2}(\omega) - \alpha^{2} - \beta^{2}$.

Then we have no constraints with respect to spatial frequencies $\alpha$, $\beta$. We get homogeneous and evanescent waves and can treat arbitrary small structures in the aperture by:

$u(x, y, z) = \int_{-\infty}^{\infty} U_{i}(\alpha, \beta) H(\alpha, \beta; z_{a}) \exp\left[\frac{i}{\lambda}(\alpha x + \beta y)\right] d\alpha d\beta$

where $U_{i}(\alpha, \beta) = \text{FT}[u_{i}(x, y)]$

Derivation of the response function

We start from the Weyl-representation of a spherical wave:

$\frac{1}{r} \exp\left[\frac{i}{\lambda}kr\right] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{r} \exp\left[\frac{i}{\lambda}(\alpha x + \beta y + \gamma z)\right] d\alpha d\beta$

Now we can compute the response function $h$, which we did not do in the previous chapter, where we computed only $h_{F}$ (Fresnel approximation). The following trick shows that

$-2\pi \frac{\partial}{\partial z_{a}} \int_{-\infty}^{\infty} \frac{1}{r} \exp\left[\frac{i}{\lambda}kr\right] = \int_{-\infty}^{\infty} \exp\left[\frac{i}{\lambda}(\alpha x + \beta y + \gamma z)\right] d\alpha d\beta = \text{FT}^{-1}[H] = (2\pi)^{2} h$

and therefore

$h(x, y, z) = -\frac{1}{2\pi} \frac{\partial}{\partial z_{a}} \frac{1}{r} \exp\left[\frac{i}{\lambda}kr\right]$ with $r = \sqrt{x^{2} + y^{2} + z^{2}}$.

The resulting expression in position space for the propagation of monochromatic beams is also called ‘Rayleigh-formula’:

$u(x, y, z) = \int_{-\infty}^{\infty} h(x-x', y-y', z_{a}) u_{i}(x', y', z_{a}) d\alpha d\beta$

3.2.2 Fresnel approximation (paraxial approximation)

From the previous chapter we know that we can apply the Fresnel approximation if $\alpha^{2} + \beta^{2} << k z_{a}$ which is valid for a limited angular spectrum, and therefore a large size of the structures inside the aperture. Then

$H_{F}(\alpha, \beta; z_{a}) = \exp\left[\frac{i}{\lambda}kz_{a}\right] \exp\left[-\frac{1}{2k}(\alpha^{2} + \beta^{2})\right]$
3.2.3 Paraxial Fraunhofer approximation (far field approximation)

A further simplification of the beam propagation is possible for many diffraction problems. Let us assume a narrow angular spectrum \( \alpha^2 + \beta^2 \ll k^2 \)

and the additional condition for the so-called Fresnel number \( N_r \)

\[
N_r \lesssim 0.1 \text{ with } N_r = \frac{a^2}{\lambda z_n}
\]

where \( a \) is the (largest) size of the aperture (like the "beam width").

Hence the approximation, which we derive in the following, is only valid in the so-called 'far field', which means far away from the aperture.

To understand the influence of this new condition on the Fresnel number, we have a look at beam propagation in paraxial approximation:

\[
u(x, y, z_n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_x(\alpha, \beta, z_n) U_x(\alpha, \beta) \exp\left[ i(kx + \beta y) \right] d\alpha d\beta
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_x(x', y', z_n) u_x(x', y') dxdy'
\]

\[
= -\frac{i}{\lambda z_n} \exp(ikz_n) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_x(x', y') \exp\left[ i\frac{k}{2z_n} \left( (x-x')^2 + (y-y')^2 \right) \right] dx'dy'
\]

In this situation it is easier to treat the beam propagation in position space, because

\[
u_x(x, y) = t(x, y) u_x(x, y) \text{, and } t(x, y) = 0 \text{ for } |x|, |y| > a \text{ (aperture)}
\]

\[
\rightarrow u_x(x, y) = 0 \text{ for } |x|, |y| > a
\]

This means that we need to integrate only over the aperture in the above integral:

\[
u_x(x, y, z_n) = -\frac{i}{\lambda z_n} \exp(ikz_n) \int_{-a}^{a} \int_{-a}^{a} u_x(x', y') \exp\left[ i\frac{k}{2z_n} \left( (x-x')^2 + (y-y')^2 \right) \right] dx'dy'
\]

Now, let us have a closer look at the exponential expression:

\[
\exp\left[ i\frac{k}{2z_n} \left( (x-x')^2 + (y-y')^2 \right) \right]
\]

\[
= \exp\left[ i\frac{k}{2z_n} \left( x^2 - 2xx' + x'^2 + y^2 - 2yy' + y'^2 \right) \right]
\]

\[
= \exp\left[ i\frac{k}{2z_n} \left( x^2 + y^2 \right) \right] \exp\left[ -i \frac{k}{2z_n} \left( xx' + yy' \right) \right] \exp\left[ i\frac{k}{2z_n} \left( x'^2 + y'^2 \right) \right]
\]

So far, nothing happened, we just sorted the factors differently. But here comes the trick:

Because of the integration range, we have \( x', y' < a \) and therefore

\[
\rightarrow \frac{k}{2z_n} \left( x^2 + y^2 \right) \ll \frac{ka^2}{z_n} = 2\pi N_r
\]

\[
\rightarrow \text{for } N_r \lesssim 0.1 \rightarrow \exp\left[ -i \frac{k}{2z_n} \left( xx' + yy' \right) \right] \approx 1
\]

This means that we can neglect the quadratic phase term in \( x', y' \) and we get for the far field:

\[
u_{fr}(x, y, z_n) = -\frac{i}{\lambda z_n} \exp(ikz_n) \exp\left[ i\frac{k}{2z_n} \left( x^2 + y^2 \right) \right]
\]

\[
\times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_x(x', y') \exp\left[ -i \left( \frac{k}{z_n} \right) \left( x + y \right) \right] dx'dy'
\]

\[
= -\frac{(2\pi)^2}{\lambda z_n} \exp(ikz_n) U_x(kx/z_n, ky/z_n) \exp\left[ i\frac{k}{2z_n} \left( x^2 + y^2 \right) \right]
\]

This is the far-field in paraxial Fraunhofer approximation. Surprisingly, the intensity distribution of the far field in position space is just given by the Fourier transform of the field distribution at the aperture

\[
I_{fr}(x, y, z_n) \sim \frac{1}{(\lambda z_n)^2} \left| U_x(kx/z_n, ky/z_n) \right|^2
\]

Interpretation

For a plane in the far field at \( z = z_n \) in each point \( x, y \) only one angular frequency \( \alpha = kx/z_n, \beta = ky/z_n \) with spectral amplitude \( U_x(kx/z_n, ky/z_n) \) contributes to the field distribution. This is in contrast to the previously considered cases, where all angular frequencies contributed to the response in a single position point.

In summary, we have shown that in (paraxial) Fraunhofer approximation the propagated field, or diffraction pattern, is very simple to calculate. We just
need to Fourier transform the field at the aperture. In order to apply this approximation we have to check that:

A) \[ \alpha^2 + \beta^2 < k^2 \] smallest features \( \Delta x, \Delta y \gg \lambda \) narrow angular spectrum (paraxiality)

B) \[ N_s = \frac{\lambda}{z_a} < 1 \] largest feature \( a \) determines \( z_a > \frac{1}{\alpha^2} \) far field

Example: \( \Delta x, \Delta y = 10 \lambda, \quad a = 100 \lambda, \quad \lambda = 1 \mu m \) \( \rightarrow z_a > 10^3 \lambda = 1 \text{ cm} \)

3.2.4 Non-paraxial Fraunhofer approximation

The concept that the angular components of the input spectrum separate in the far field due to diffraction works also beyond the paraxial approximation. If we have arbitrary angular frequencies in our spectrum, all \( \alpha^2 + \beta^2 \leq k^2 \) contribute to the far field distribution. Evanescent waves decay for \( k z_a > 1 \rightarrow z_a > \lambda \).

Form the previous chapter we know that the diffraction pattern in the far field in paraxial Fraunhofer approximation is given as:

\[ I(x,y,z_a) \sim |u(x,y,z_a)|^2 \sim \frac{1}{(\lambda z_a)^2} |U_r(k \frac{x}{z_a}, k \frac{y}{z_a})|^2 \]

The diffraction pattern is proportional to the spectrum behind the mask at \( \alpha = k \frac{x}{z_a}, \beta = k \frac{y}{z_a} \).

Let us calculate the field for inclined incidence of the excitation. The field behind the mask is given by:

\[ u_r(x,y,z_a) = u(x,y,z_a) t(x,y) = A \exp \left[ i \left( k x + k y + k z_a \right) \right] t(x,y) \]

The Fourier transform gives the spectrum:

\[ U_r(k \frac{x}{z_a}, k \frac{y}{z_a}) = \frac{A}{(2\pi)^2} \exp \left[ i k z_a \right] \int t(x',y') \exp \left[ -i \left( k \frac{x}{z_a} - k x' \right) - i \left( k \frac{y}{z_a} - k y' \right) \right] d x' d y' \]

\[ = A \exp \left[ i k z_a \right] T \left( k \frac{x}{z_a} - k x, k \frac{y}{z_a} - k y \right) \]

Hence, the intensity distribution of the diffraction pattern is given as:

\[ I(x,y,z_a) \sim \frac{1}{(\lambda z_a)^2} T \left( k \frac{x}{z_a} - k x, k \frac{y}{z_a} - k y \right) \]

This is the absolute square of the Fourier transform of the aperture function. An inclination of the illuminating plane wave just shifts the pattern (in paraxial approximation).

Examples

A) Rectangular aperture illuminated by normal plane wave

\[ t(x,y) = \begin{cases} 1 & \text{for } |x| \leq a, |y| \leq b \\ 0 & \text{elsewhere} \end{cases} \]

\[ I(x,y,z_a) \sim \text{sinc}^2 \left( k a \frac{x}{z_a} \right) \text{sinc}^2 \left( k b \frac{y}{z_a} \right) \]
Fraunhofer diffraction pattern from a rectangular aperture. The central lobe of the pattern has half-angular widths $\theta_x = \lambda / D_x$ and $\theta_y = \lambda / D_y$.

B) Circular aperture (pinhole) illuminated by normal plane wave

$$t(x, y) = \begin{cases} 1 & \text{for } x^2 + y^2 \leq a^2 \\ 0 & \text{elsewhere} \end{cases}$$

$$I(x, y, z_b) \sim \left[ \frac{J_1 \left( \frac{ka}{z_b} \sqrt{x^2 + y^2} \right)}{\frac{ka}{z_b} \sqrt{x^2 + y^2}} \right]^2 \rightarrow \text{Airy disk}$$

The Fraunhofer diffraction pattern from a circular aperture produces the Airy pattern with the radius of the central disk subtending an angle $\theta = 1.22 \lambda / D$.

The first zero of the Bessel function (size of Airy disk):

$$\frac{ka}{z_b} = 1.22 \pi \rightarrow \frac{\rho}{z_b} = \frac{0.61 \lambda}{a} \quad \text{with} \quad \rho^2 = x^2 + y^2$$

So-called angle of aperture: $\Theta = \frac{2\rho}{z_b} = \frac{1.22 \lambda}{a}$

C) One-dimensional periodic structure illuminated by normal plane wave

For periodic arrangements of slits we can gain deeper insight in the structure of the diffraction pattern. Let us assume a period slit aperture with: period $b$ and a size of each slit $2a$:

$$t_i(x)$$

$$t_s(x)$$

Then, we can express the mask function $t$ as:

$$t(x) = \sum_{n=0}^{\infty} t_i(x - nb) \text{ with } t_i(x) = \begin{cases} 1 & \text{ for } |x| \leq a \\ 0 & \text{ elsewhere} \end{cases}$$

The Fourier transform of the mask then reads

$$T \left( k \frac{x}{z_b} \right) \sim \sum_{n=0}^{\infty} T_s(x' - nb) \exp \left( -\frac{1}{2}k \frac{x}{z_b} \right) dx'$$

With the new variable $x' - nb = X'$ we can simplify further:

$$T \left( k \frac{x}{z_b} \right) \sim \sum_{n=0}^{\infty} T_s(X') \exp \left( -\frac{1}{2}k \frac{x}{z_b} \right) \exp \left( -i k \frac{x}{z_b} nb \right) dX'$$

We see that the Fourier transform $T_s$ of the elementary slit $t_s$ appears. The second factor has its origin in the periodic arrangement. With some math we can identify this second expression as a geometrical series and perform the summation:

$$\sum_{n=0}^{\infty} \exp (-i \delta n) = \frac{\sin (N \delta)}{\sin (\delta)}$$

Thus we finally write:

$$T \left( k \frac{x}{z_b} \right) \sim T_s \left( k \frac{x}{z_b} \right) \frac{\sin (N \frac{x}{z_b} b)}{\sin \left( \frac{x}{z_b} a \right)}$$

For the particular case of a simple grating of slit apertures we have

$$T_s \left( k \frac{x}{z_b} \right) = \text{sinc} \left( k \frac{x}{z_b} \right)$$
The width of a maximum in the far-field diffraction pattern $x_N$ (smallest length scale in the pattern) is determined by $N^* b$, the total size of the mask (largest length scale of the mask).

These observations are consistent with the general property of the Fourier-transform: small scales in position space give rise to a broad angular spectrum.

### 3.4 Remarks on Fresnel diffraction

**Fresnel number** $\rightarrow N_F = \frac{a}{\lambda z_a}$

- $N_F \approx 10$ (a large, $\lambda z_a$ small, $z_b < 1/30 z_a$) → shadow
- $N_F \leq 0.1$ ($z_b > 3z_a$) → Fraunhofer → FT of aperture
- $10 \geq N_F \geq 0.1$ ($1/30 z_a < z_b < 3z_a$) → Fresnel diffraction

Fresnel diffraction from a slit of width $D = 2a$.

(a) Shaded area is the geometrical shadow of the aperture. The dashed line is the width of the Fraunhofer diffraction beam. (b) Diffraction pattern at four axial positions marked by the arrows in (a) and corresponding to the Fresnel numbers $N_F = 10, 1, 0.5$ and 0.1. The shaded area represents the geometrical shadow of the slit. The dashed lines at $x = 1/3 D$ represent the width of the Fraunhofer pattern in the far field. Where the dashed lines coincide with the edges of the geometrical shadow, the Fresnel number $N_F = a^2 / \lambda d = 0.5$.
4. Fourier optics - optical filtering

From previous chapters we know how to propagate the optical field through homogeneous space, and we also know the transfer function of a thin lens. Thus, we have all tools at hand to describe optical imaging. Here, we will use again paraxial approximation, which is in general sufficient for optical systems.

We will see in the following that with the right setup of our imaging system we can generate the Fourier transform of the object on a much shorter distance than by far field diffraction in the Fraunhofer approximation.

The general idea of Fourier optics is the following:
1) An imaging system generates the Fourier transform of the object.
2) A spatial filter (e.g. an aperture) in the Fourier plane manipulates the field.
3) Another imaging system performs the Fourier back-transform and hence results in a manipulated image.

Mathematical concept:
- propagation in free space \( \rightarrow \) in Fourier domain
- interaction with lens or filter \( \rightarrow \) in position space

4.1 Imaging of arbitrary optical field with thin lens

4.1.1 Transfer function of a thin lens

A thin lens changes only the phase of the optical field, since due to its infinitesimal thickness, no diffraction occurs. By definition, it transforms a spherical wave into a plane wave. If we write down this definition in paraxial approximation we get

\[
\psi_{lens}(x,y) = \frac{1}{\lambda f} \exp(i k f) \exp\left(\frac{i k}{2 f} (x^2 + y^2)\right)
\]

And therefore the transfer function for a thin lens is given as (see chapter 2.6.3):

\[
T_l(x,y) = \frac{1}{\lambda f} \exp\left(-\frac{i k}{2 f} (x^2 + y^2)\right)
\]

In Fourier domain we find consequently

\[
T_l(\alpha,\beta) = \frac{-i}{\lambda f} \frac{2\pi}{(2\pi)^2} \exp\left(\frac{i k f}{2k} (\alpha^2 + \beta^2)\right)
\]

4.1.2 Optical imaging

Let us now consider optical imaging. We place our object in the first focus of a thin lens, with a field distribution \( u_0(x,y) \), and follow the usual light propagation recipe.

A) Spectrum in object plane

\[
U_0(\alpha,\beta) = \text{FT}\left[u_0(x,y)\right]
\]

B) Propagation from object to lens (lens positioned at distance \( f \))

\[
U_-(\alpha,\beta; f) = H_f(\alpha,\beta; f) U_0(\alpha,\beta)
\]

\[
U_-(\alpha,\beta; f) = \exp(i k f) \exp\left(-\frac{i k f}{2k} (\alpha^2 + \beta^2)\right) U_0(\alpha,\beta)
\]

C) Interaction with lens (multiplication in position space or convolution in Fourier domain)

\[
u_+(x,y,f) = t_l(x,y) u_0(x,y,f)
\]

\[
U_+(\alpha,\beta; f) = T_l(\alpha,\beta) \ast U_-(\alpha,\beta; f)
\]

\[
= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \exp(i k f) \int_{-\infty}^{\infty} \exp\left(-\frac{i k f}{2k} (\alpha - \alpha')^2 + (\beta - \beta')^2\right)
\]

\[
\cdot \exp\left(-\frac{i k f}{2k} (\alpha^2 + \beta^2)\right) U_0(\alpha',\beta') d\alpha' d\beta'
\]

D) Propagation from lens to image plane

\[
U(\alpha,\beta; 2f) = H_f(\alpha,\beta; f) U_+(\alpha,\beta; f)
\]
$$U(\alpha, \beta; 2f) = -\frac{i}{2} \frac{jf}{(2\pi)^2} \exp(2ikf') \int_{-\infty}^{\infty} \exp \left\{ -\frac{i}{2k} \left[ (\alpha - \alpha')^2 + (\beta - \beta')^2 \right] \right\} d\alpha' d\beta'$$

$$\cdot \exp \left[ -\frac{i}{2k} \left( \alpha^2 + \beta^2 \right) f \right] \exp \left[ -\frac{i}{2k} \left( \alpha^2 + \beta^2 \right) f \right] U_o(\alpha', \beta') d\alpha' d\beta'$$

Quadratic terms $$\left[ -\frac{i}{2k} \left( \alpha^2 + \beta^2 \right) f \right]$$ and $$\left[ -\frac{i}{2k} \left( \alpha^2 + \beta^2 \right) f \right]$$ in the exponent cancel with quadratic terms from $$\frac{1}{k} \left( \alpha - \alpha' \right)^2 + \left( \beta - \beta' \right)^2$$.

$$U(\alpha, \beta; 2f) = -\frac{i}{2} \frac{jf}{(2\pi)^2} \exp(2ikf') \int_{-\infty}^{\infty} \exp \left\{ -\frac{i}{k} (\alpha + \beta) \right\} U_o(\alpha', \beta') d\alpha' d\beta'$$

$$= -\frac{i}{2} \frac{jf}{(2\pi)^2} \exp(2ikf) U_o \left( -\frac{f}{k} \alpha, -\frac{f}{k} \beta \right)$$

We see that the spectrum in the image plane is given by the optical field in the object plane.

**E) Fourier back transform in image plane**

$$u(x, y, 2f) = \mathcal{F}^{-1}[U(\alpha, \beta; 2f)]$$

$$= -\frac{i}{2} \frac{jf}{(2\pi)^2} \exp(2ikf') \int_{-\infty}^{\infty} \exp \left[ \frac{i}{k} (\alpha x + \beta y) \right] d\alpha' d\beta'$$

With the coordinate transformation

$$x' = -\frac{f}{k} \alpha, \quad y' = -\frac{f}{k} \beta \quad \rightarrow \quad d\alpha = -\frac{2\pi}{\lambda f} dx', \quad d\beta = -\frac{2\pi}{\lambda f} dy'$$

we get:

$$\rightarrow u(x, y, 2f) = -\frac{i}{2} \frac{1}{\lambda f} \exp(2ikf') \int_{-\infty}^{\infty} u_{o}(x', y') \exp \left[ \frac{1}{k} (xx' + yy') \right] dx' dy'$$

$$u(x, y, 2f) = -\frac{i}{2} \frac{(2\pi)^2}{\lambda f} \exp(2ikf') U_o \left( \frac{k}{\lambda f}, \frac{k}{\lambda f}, \frac{k}{\lambda f} \right)$$

The image in the second focal plane is the Fourier transform of the optical field in the object plane. → like far field in Fraunhofer approximation, but $$z_o \leftrightarrow f$$. This finding allows us to perform an optical Fourier transform, and in the Fourier plane it is possible to manipulate the spectrum.

### 4.2 Optical filtering and image processing

#### 4.2.1 The 4f-setup

For image manipulation (filtering) it would be advantageous if we could perform a Fourier back-transform by means of an optical imaging setup as well. It turns out that this leads to the so-called 4f-setup:

We know that the image in Fourier plane is the FT of the optical field in object plane.

$$\Rightarrow u(x, y, 2f) = AU_o \left( \frac{k}{f}, \frac{k}{f}, \frac{k}{f} \right)$$

The filtering (manipulation) happens in the second focal plane (Fourier plane after 2f) by applying a transmission mask $$p(x, y)$$. In order to retrieve the filtered image we use a second lens:

$$u(x, y, 2f) = AU_o \left( \frac{k}{f}, \frac{k}{f}, \frac{k}{f} \right)$$

We have to compute the imaging with the second lens after manipulation. Our final goal is the transmission function $$H_x(\alpha, \beta; 4f)$$ of the complete imaging system:

$$u(-x, -y, 4f) = \int_{-\infty}^{\infty} H_x(\alpha, \beta; 4f) U_o(\alpha, \beta) \exp \left[ \frac{i}{\lambda} (\alpha x + \beta y) \right] d\alpha d\beta$$
Note: We will see in the following calculation that the second lens does a Fourier transform $\exp[-i(\alpha x + \beta y)]$. In order to obtain a proper back transform we have to pass to mirrored coordinates $x \rightarrow -x, y \rightarrow -y$. The transmission mask $p(x,y)$ contains all constrains of the system (e.g. a limited aperture) and optical filtering (which we can design).

**A) Field behind transmission mask**

$$u_1(x,y,2f) = u(x,y,2f)p(x,y) \sim AU_0 \left(\frac{k}{f}x, \frac{k}{f}y\right)p(x,y)$$

**B) Second lens $\Rightarrow$ Fourier back-transform of field distribution**

$$u(x,y,4f) = -i(\frac{2\pi}{\lambda f})^2 \exp\left(2ikf\right)U_0 \left(\frac{k}{f}x, \frac{k}{f}y; 2f\right)$$

Now we can make the link to the initial spectrum in the object plane $U_0$:

$$u(x,y,4f) \sim \int \int u_1(x',y',2f) \exp\left[-i\frac{k}{f}(xx' + yy')\right] dx'dy'$$

$$\sim \int \int U_0 \left(\frac{k}{f}x', \frac{k}{f}y'\right)p(x',y') \exp\left[-i\frac{k}{f}(xx' + yy')\right] dx'dy'$$

Here we do not care about the amplitudes and just write $\sim$:

To get the anticipated form we need to perform a coordinate transformation:

$$\alpha = \frac{k}{f}x', \beta = \frac{k}{f}y'$$

Then we can write:

$$u(x,y,4f) \sim \int \int U_0(\alpha,\beta) p\left(\frac{f}{k}\alpha, \frac{f}{k}\beta\right) \exp\left[-i(\alpha x + \beta y)\right] d\alpha d\beta$$

By passing to mirrored coordinates $x \rightarrow -x, y \rightarrow -y$ we get

$$u(-x,-y,4f) \sim \int \int p\left(\frac{f}{k}\alpha, \frac{f}{k}\beta\right)U_0(\alpha,\beta) \exp\left[i(\alpha x + \beta y)\right] d\alpha d\beta$$

Hence we can identify the transmission function of the system

$$H_k(\alpha,\beta;4f) \sim p\left(\frac{f}{k}\alpha, \frac{f}{k}\beta\right)$$

**Summary**

- Fourier amplitudes get multiplied by transmission mask
- transmission mask $\rightarrow$ transmission function

**Example 1: The ideal image (infinite aperture)**

Be careful, we use paraxial approximation $\rightarrow$ limited angular spectrum.

In position space we can formulate propagation through a 4f-system by using the response function $h_k(x,y)$

$$u(-x,-y,4f) = \int \int h_k(x-x',y-y')u_0(x',y')dx'dy'$$

As usual, the response function is given as:

$$h_k(x,y) = \frac{1}{(2\pi)^2} \int \int H_k(\alpha,\beta;4f) \exp\left[i(\alpha x + \beta y)\right] d\alpha d\beta$$

From above we have $H_k(\alpha,\beta;4f) \sim p\left(\frac{f}{k}\alpha, \frac{f}{k}\beta\right)$

$$h_k(x,y) \sim \int \int p\left(\frac{f}{k}\alpha, \frac{f}{k}\beta\right) \exp\left[i(\alpha x + \beta y)\right] d\alpha d\beta$$

We introduce the coordinate transform $\left(\bar{x} = \frac{f}{k}\alpha, \bar{y} = \frac{f}{k}\beta\right)$

$$h_k(x,y) \sim \int \int p(\bar{x},\bar{y}) \exp\left[i\frac{k}{f}(xx' + yy')\right] d\bar{x}d\bar{y} \sim P\left[\frac{k}{f}x, \frac{k}{f}y\right]$$

$$u(-x,-y,4f) \sim \int \int P\left[\frac{k}{f}(x-x'), \frac{k}{f}(y-y')\right]u_0(x',y')dx'dy'$$

The response-function is proportional to the Fourier transform of the transmission mask.

**4.2.2 Examples of aperture functions**

- coordinates of image $\rightarrow$ mirrored coordinates of object
- In position space we can formulate propagation through a 4f-system by using the response function $h_k(x,y)$

![Diagram](image-url)
Example 2: Finite aperture

\[
p(x, y) = \begin{cases} 
1 & \text{for } x^2 + y^2 \leq \left( \frac{D}{2} \right)^2 \\
0 & \text{elsewhere}
\end{cases}
\]

Translated to position space, the smallest transmitted structural information is given by:

\[
\Delta r_{\text{min}} = \frac{2\pi}{\rho_{\text{max}}} = \frac{2\lambda f}{nD}
\]

A more precise definition of the optical resolution can be derived the following way:

\[
\alpha_{\text{min}} = 1.22 \frac{\lambda}{D}
\]

Further examples for 4f filtering:

We can define the limit of optical resolution:

Two objects in the object plane can be independently resolved in the image plane as long as the intensity maximum of one of the objects is not closer to the other object than its first intensity minimum:

\[
\frac{kD}{2f} \Delta r_{\text{min}} = 1.22\pi
\]

Hence we find:

\[
\Delta r_{\text{min}} = \frac{1.22\lambda f}{nD}
\]
5. The polarization of electromagnetic waves

5.1 Introduction

We are interested in the temporal evolution of the electric field vector $\mathbf{E}(r,t)$. In the previous chapters we mostly used a scalar description, assuming linearly polarized light. However, in general one has to consider the vectorial nature, i.e. the polarization state, of the electric field vector.

We know that the normal modes of homogeneous isotropic dielectric media are plane waves $\mathbf{E}(r,t) = \mathbf{E}_0 \exp \left[ i \left( \mathbf{k} \cdot \mathbf{r} - \omega t \right) \right]$.

If we assume propagation in $z$ direction ($\mathbf{k}$-vector points in $z$-direction), $\nabla \cdot \mathbf{E}(r,t) = 0$ implies that we can have two nonzero transversal field components $E_x, E_y$.

The orientation and shape of the area which the (real) electric field vector covers is in general an ellipse. There are two special cases:
- line (linear polarization)
- circle (circular polarization)

5.2 Polarization of normal modes in isotropic media

\[
\mathbf{k} = \begin{pmatrix} 0 \\ 0 \\ k \end{pmatrix} \rightarrow \text{propagation in } z \text{ direction}
\]

The evolution of the real electric field vector is given as

\[
\mathbf{E}_x(r,t) = \mathbb{R} \left\{ \mathbf{E}_0 \exp \left[ i (kz - \omega t) \right] \right\}
\]

Because the field is transversal we have two free complex field components

\[
\mathbf{E} = \begin{pmatrix} E_x \exp(i\phi_x) \\ E_y \exp(i\phi_y) \\ 0 \end{pmatrix} \text{ with } E_{x,y} \text{ and } \phi_{x,y} \text{ being real}
\]

Then the real electric field vector is given as

\[
\mathbf{E}_x(r,t) = \begin{pmatrix} E_x \cos(\omega t - kz - \phi_x) \\ E_y \cos(\omega t - kz - \phi_y) \\ 0 \end{pmatrix}
\]

Here, only the relative phase is interesting $\delta = \phi_y - \phi_x$

Conclusion:
Normal modes in isotropic, dispersive media are in general elliptically polarized; \( E_x, E_y \) and phase difference \( \delta = (\varphi_y - \varphi_x) \) are free parameters.

### 5.3 Polarization states

Let us have a look at different possible parameter settings:

A) **linear polarization**  \( \rightarrow \delta = n\pi \) (or \( E_x = 0 \) or \( E_y = 0 \))

B) **circular polarization**  \( \rightarrow E_x = E_y = E, \delta = \pm \pi / 2 \)

\( \delta = \pm \pi / 2 \) \( \rightarrow \) counterclockwise rotation

\[ \delta = -\pi / 2 \rightarrow \text{clockwise rotation} \]

These pictures are for an observer looking contrary to the propagation direction.

C) **elliptic polarization**  \( \rightarrow E_x \neq E_y \neq 0, \delta \neq n\pi \)

\( 0 < \delta < \pi \) \( \rightarrow \) counterclockwise

\( \pi < \delta < 2\pi \) \( \rightarrow \) clockwise

Example:

Remark

A linearly polarized wave can be written as a superposition of two counter-rotating circularly polarized waves. Example: Let's observe the temporal evolution at a fixed position \( k_z = 0 \) with \( \delta = \pm \pi / 2 \).

\[
\begin{pmatrix}
\cos(\omega t) & \cos(\omega t) & \cos(\omega t)
\end{pmatrix}
\begin{pmatrix}
E \\
\sin(\omega t) \\
0
\end{pmatrix}
+ \begin{pmatrix}
\cos(\omega t) & -\cos(\omega t) & 0
\end{pmatrix}
\begin{pmatrix}
E \\
0 \\
-\sin(\omega t)
\end{pmatrix}
= 2E \begin{pmatrix}
\cos(\omega t) \\
0 \\
0
\end{pmatrix}
\]
6. Principles of optics in crystals

In this chapter we will treat light propagation in anisotropic media (the worst case). Like in the isotropic case before we will seek for the normal modes, and in order to keep things simple we assume homogeneous media.

6.1 Susceptibility and dielectric tensor

before: isotropy (optical properties independent of direction)
now: anisotropy (optical properties depend on direction)

The common reason for anisotropy in many optical media (in particular crystals) is that the polarization $P$ depends on direction of electric field vector. The underlying reason is that in crystals the atoms have a periodic distribution with different symmetries in different directions.

Prominent examples for anisotropic materials are:
- Lithium Niobat $\rightarrow$ electro-optical material
- Quartz $\rightarrow$ polarizer
- liquid crystals $\rightarrow$ displays, NLO

In order to keep things as simple as possible we make the following assumptions:
- one frequency- (monochromatic), one angular frequency (plane wave)
- no absorption

From previous chapters we know that in isotropic media the normal modes are elliptically polarized, monochromatic plane waves. The question is how the normal modes of an anisotropic medium look like $\rightarrow$ ??

Before (isotropic):
\[ P(r,\omega) = \varepsilon_0 \chi(\omega) \vec{E}(r,\omega) \]
\[ D(r,\omega) = \varepsilon_0 \epsilon(\omega) \vec{E}(r,\omega) \]

In the following we will write $\vec{E} \rightarrow \vec{E}$, because we assume monochromatic light and the frequency $\omega$ is just a parameter. Now (anisotropic):
\[ P(r,\omega) = \varepsilon_0 \sum_{\alpha=1}^{3} \chi_{\alpha}(\omega) E_{\alpha}(r,\omega) \]
\[ D(r,\omega) = \varepsilon_0 \epsilon_{\alpha} \vec{E}(r,\omega) \]

The linear susceptibility tensor has $3x3=9$ tensor components. Direct consequences of this relation between polarization $P$ and electric field $E$ are:
- $P \neq E$: the polarization is not necessarily parallel to the electric field
- The tensor elements $\chi_{\alpha}$ depend on the structure of crystal. However, we do not need to know the microscopic structure because of the different length scales involved (optics $\sim 5 \cdot 10^{-7}$ m; crystal $\sim 5 \cdot 10^{-10}$ m), but the field is influenced by the symmetries of the crystal (see next subchapter 6.2).

- In complete analogy we find for the $D$ field:
\[ D(r,\omega) = \varepsilon_0 \sum_{\alpha=1}^{3} \epsilon_{\alpha}(\omega) E_{\alpha}(r,\omega) \]

As for the polarization we find:
- $D \neq E$

We introduce the following notation:
- $\chi = \chi_{\alpha}$ $\rightarrow$ susceptibility tensor
- $\epsilon = \epsilon_{\alpha}$ $\rightarrow$ dielectric tensor
- $\sigma = (\epsilon)^{-1} = \sigma_{\alpha}$ $\rightarrow$ inverse dielectric tensor

The following properties of the dielectric and inverse dielectric tensor are important:
- $\sigma_{\alpha}, \epsilon_{\alpha}$ are real in the transparent region (omit $\omega$), we have no losses (see our assumptions above)
- The tensors are symmetric (hermitian), only 6 components are independent $\epsilon_{\alpha} = \epsilon_{\beta}, \sigma_{\alpha} = \sigma_{\beta}$.
- It is known (see any book on linear algebra) that for such tensors a transformation to principal axes by rotation is possible (matrix is diagonalizable by orthogonal transformations).
- If we write down this for $\sigma_{\alpha}$, it means that we are looking for directions where $D \neq E$, i.e., our principal axes:
\[ \sigma_{\alpha}^{\alpha} D_{\alpha}^{\alpha} = \lambda D_{\alpha}^{\alpha} \]

This is a so-called eigenvalue problem, with eigenvalues $\lambda$. If we want to solve for the eigenvalues we get
\[ \det[\sigma_{\alpha}^{\alpha} - \lambda I_{\alpha}] = 0, \text{ with } I_{\alpha} = \delta_{\alpha} \]

This leads to a third order equation in $\lambda$, hence we expect three solutions (roots) $\lambda^{(\alpha)}$. The corresponding eigenvectors can be computed from
\[ \sum_{\alpha=1}^{3} \sigma_{\alpha} D_{\alpha}^{(\alpha)} = \lambda^{(\alpha)} D_{\alpha}^{(\alpha)} \]

The eigenvectors are orthogonal:
\[ D_{(\alpha)} D_{(\beta)}^{\alpha} = 0 \text{ for } \lambda^{(\alpha)} \neq \lambda^{(\beta)} \]

This property holds if and only if the eigenvalues are distinct. If the eigenvalues are not distinct, that is, if $\lambda^{(\alpha)} = \lambda^{(\beta)}$, then the eigenvectors are not orthogonal.
The directions of the principal axes (defined by the eigenvectors) correspond to the symmetry axes of the crystal. The diagonalized dielectric and inverse dielectric tensors are linked:

\[
\varepsilon_y = \varepsilon_i \delta_y, \quad \sigma_y = \sigma_i \delta_y = \frac{1}{\varepsilon_i} \delta_y
\]

The above reasoning shows that anisotropic media are characterized in general by three independent dielectric functions (in the principal coordinate system). It is easier to do all calculations in the principal coordinate system (coordinate system of the crystal) and back-transform the final results to the laboratory system.

### 6.2 The optical classification of crystals

Let us now give a brief overview over crystal classes and their optical properties:

**A) isotropic**
- three crystallographic equivalent orthogonal axes
- cubic crystals (diamond, Si, …)

\[
\varepsilon_i(\omega) = \varepsilon_j(\omega) = \varepsilon_k(\omega) \Rightarrow D_i = \varepsilon_i \varepsilon(\omega) E_i
\]

Cubic crystals behave like gas, amorphous solids, liquids, and have no anisotropy.

**B) uniaxial**
- two crystallographic equivalent directions
- trigonal (quartz, lithium niobate), tetragonal, hexagonal

\[
\varepsilon_i(\omega) = \varepsilon_j(\omega) \neq \varepsilon_k(\omega)
\]

**C) biaxial**
- no crystallographic equivalent directions

- orthorhombic, monoclinic, triclinic

### 6.3 The index ellipsoid

The index ellipsoid offers a simple geometrical interpretation of the inverse dielectric tensor \( \hat{\sigma} = [\hat{\varepsilon}]^{-1} \). The defining equation for the index ellipsoid is

\[
\sum_{i,j=1}^{3} \sigma_{ij} x_i x_j = 1
\]

which describes a surface in three dimensional space.

Remark on the physics of the index ellipsoid: The index ellipsoid defines a surface of constant electric energy density:

\[
\sum_{i,j=1}^{3} \sigma_{ij} D_i D_j = \varepsilon_0 \sum_{j=1}^{3} E_j D_j = 2 \varepsilon_0 \varepsilon_1
\]

In the principal coordinate system the defining equation of the index ellipsoid reads:

\[
\sigma_{1} x_1^2 + \sigma_{2} x_2^2 + \sigma_{3} x_3^2 = \frac{x_1^2}{\varepsilon_1} + \frac{x_2^2}{\varepsilon_2} + \frac{x_3^2}{\varepsilon_3} = 1
\]

Here the elements of the dielectric tensor can be related to refractive indexes

\[ n_i = \sqrt{\varepsilon_i} \]
Summary:
- anisotropic media → tensor instead of scalar
  → in principal system: $n_i = \sqrt{\varepsilon_i}$
- The index ellipsoid is degenerate for special cases:
  - isotropic crystal: sphere
  - uniaxial crystal: rotational symmetric with respect to z-axis and $n_1 = n_2$

6.4 Normal modes in anisotropic media

Let us now look for the normal modes in crystals. A normal mode is:
- a solution to the wave equation, which shows only a phase dynamics during propagation while amplitude and polarization remain constant
  → most simple solution: $\exp\left[i\mathbf{k}(\omega)\mathbf{r} - \omega t\right]$
- a solution where the spatial and temporal evolution of the phase are connected by a dispersion relation
  $\omega = \omega(k)$ or $k = k(\omega)$

Before – isotropic media

In isotropic media the normal modes are monochromatic plane waves

$$E(r,t) = E \exp\left[i\mathbf{k}(\omega)\mathbf{r} - \omega t\right]$$

with the dispersion relation

$$k^2(\omega) = \frac{\omega^2}{c^2} \varepsilon(\omega)$$

with $\varepsilon(\omega) > 0$ and real as well as $\mathbf{k} \cdot \mathbf{E} = \mathbf{k} \cdot \mathbf{D} = 0$. The normal modes are elliptically polarized, and the polarization is conserved during propagation.

Now – anisotropic media

What are the normal modes in anisotropic media?

6.4.1 Normal modes propagating in principal directions

Let us first calculate the normal modes for propagation in the direction of the principal axes of the index ellipsoid, which is the simple case.

We assume without loss of generality that the principal axes are in $x, y, z$ direction and the light propagates in z-direction ($\mathbf{k} \rightarrow \mathbf{k}_z$). Then, the fields are arbitrary in the $x, y$-plane

$$D_x, D_y \neq 0$$

and

$$D_z = \varepsilon_z \varepsilon_z E_z$$

In general we have $\mathbf{E} \neq \mathbf{D}$, but here $\mathbf{k} \cdot \mathbf{D} = 0 \Rightarrow \mathbf{k} \cdot \mathbf{E} = 0$, and the two polarization directions $x, y$ are decoupled:

$$D_x, \varepsilon_x \Rightarrow D_x \exp\left[i\mathbf{k}(\omega)\mathbf{r} - \omega t\right] = D_x \exp\left[i\phi_z(z)\right] \exp\left[-i\omega t\right]$$

with $k_z^2 = \frac{\omega^2}{c^2} \varepsilon_z(\omega)$

$$D_y, \varepsilon_y \Rightarrow D_y \exp\left[i\mathbf{k}(\omega)\mathbf{r} - \omega t\right] = D_y \exp\left[i\phi_z(z)\right] \exp\left[-i\omega t\right]$$

with $k_z^2 = \frac{\omega^2}{c^2} \varepsilon_z(\omega)$

We see that in contrast to isotropic media, normal modes can't be elliptically polarized, since the polarization direction would change during propagation. But, for polarization in the direction of a principal axis ($x$ or $y$) only the phase changes during propagation, thus we found our normal modes:

$$\mathbf{D}^{(a)} = \left[D_x \exp\left[i\mathbf{k}_z\mathbf{r} - \omega t\right]\right] \varepsilon_x \rightarrow k_x^2 = \frac{\omega^2}{c^2} n_x^2 = k_z^2 \rightarrow \text{normal mode a}$$

$$\mathbf{D}^{(b)} = \left[D_y \exp\left[i\mathbf{k}_z\mathbf{r} - \omega t\right]\right] \varepsilon_y \rightarrow k_y^2 = \frac{\omega^2}{c^2} n_y^2 = k_z^2 \rightarrow \text{normal mode b}$$
For light propagation in principle direction we find two perpendicular linearly polarized normal modes with $\mathbf{E} \parallel \mathbf{D}$.

### 6.4.2 Normal modes for arbitrary propagation direction

#### 6.4.2.1 Geometrical construction

Before we will do the derivation and actually calculate normal modes and dispersion relations, let us have a look at the results visualized in the index ellipsoid. It is actually possible to construct the normal modes geometrically:
- For a specific crystal and a given frequency $\omega$ we take (in principal axis system) the $\varepsilon_i$ and construct the index ellipsoid.
- We then fix the propagation direction $\mathbf{k} = \mathbf{u}$.
- We draw a plane through the origin of index ellipsoid which is perpendicular to $\mathbf{k}$.
- The resulting intersection is an ellipse, the so-called index ellipse.
- The half-lengths of the principle axes of this ellipse equal the refractive indices $n_a, n_b$, of the normal modes for the propagation direction $\mathbf{u} = \mathbf{k} / k$
- The directions of the principal axes of the index ellipse are the polarization direction of the normal modes $\mathbf{D}_a$ and $\mathbf{D}_b$.
- The electric field vectors of the normal modes $\mathbf{E}^{(a)}$ and $\mathbf{E}^{(b)}$ follow from
  $$ E_i^{(a)} = \frac{D_i^{(a)}}{\varepsilon_i n_a}, \quad E_i^{(b)} = \frac{D_i^{(b)}}{\varepsilon_i n_b} $$
- Thus, $\mathbf{D}^{(a,b)} \parallel \mathbf{E}^{(a,b)}$, and $\mathbf{E}^{(a,b)}$ are not perpendicular to $\mathbf{k}$.
- This has a direct consequence on the pointing vector:

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

hence $\mathbf{k}$ is not parallel to $\langle S \rangle$ because $\langle S \rangle \perp \mathbf{E}$

- If the index ellipse is a circle, the direction of this particular $\mathbf{k}$-vector defines the optical axis of the crystal.

#### 6.4.2.2 Mathematical derivation of dispersion relation

Let us now derive the dispersion relation for normal modes of the form

$$ \mathbf{E}(\mathbf{r}, t) = E \exp \left[ i (\mathbf{k}(\omega) \mathbf{r} - \omega t) \right] \\
\mathbf{D}(\mathbf{r}, t) = D \exp \left[ i (\mathbf{k}(\omega) \mathbf{r} - \omega t) \right] $$

In the isotropic case we found the dispersion relation

$$ k^2 = \frac{\omega^2}{c^2} \varepsilon(\omega) $$

where the absolute value of the $\mathbf{k}$-vector is independent of its direction. The fields of the normal modes are elliptically polarized.

In the anisotropic case the normal modes are again monochromatic plane waves $\sim \exp \left[ i (\mathbf{k}(\omega) \mathbf{r} - \omega t) \right]$, but the wavenumber depends on the direction $\mathbf{u}$ of propagation

$$ k = k(\omega, \mathbf{u}) $$

and the polarization of the normal modes is not elliptic.

In the following, we start again from Maxwell’s equations and plug in the plane wave ansatz. We will use the following notation for the directional dependence of $\mathbf{k}$:

$$ k = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \quad \text{with} \quad u_1^2 + u_2^2 + u_3^2 = 1 $$

Our aim is to derive $\omega = \omega(k_1, k_2, k_3)$ or $\omega = \omega(k, u_1, u_2, u_3)$ or $k = k(\omega, u_1, u_2, u_3)$. We start from Maxwell’s equations for the plane wave Ansatz:

$$ \mathbf{k} \cdot \mathbf{D} = 0 \quad \mathbf{k} \times \mathbf{E} = \omega \mathbf{u}_1 \mathbf{H} $$

hence $\mathbf{k}$ is not parallel to $\langle S \rangle$ because $\langle S \rangle \perp \mathbf{E}$

- If the index ellipse is a circle, the direction of this particular $\mathbf{k}$-vector defines the optical axis of the crystal.
\[ \mathbf{k} \cdot \mathbf{H} = 0 \quad \mathbf{k} \times \mathbf{H} = -\omega \mathbf{D} \]

Now we follow the usual derivation of the wave equation:

\[ -\left[ \mathbf{k} \times (\mathbf{k} \times \mathbf{E}) \right] = \frac{\omega^2}{c^2} \mathbf{D} \Rightarrow -\mathbf{k} (\mathbf{k} \cdot \mathbf{E}) + \mathbf{k} \mathbf{E} = \frac{\omega^2}{c^2} \mathbf{D} \]

- Here \( \mathbf{k} \cdot \mathbf{E} \) does not vanish as in the isotropic case!
- In the principal coordinate system and with \( D_i = e_i \mathbf{E}_i \), we find

\[ -k_i \sum_j k_j \mathbf{E}_j + k^2 \mathbf{E}_i = \frac{\omega^2}{c^2} \mathbf{E}_i, \]

\[ \left( \frac{\omega^2}{c^2} e_i - k^2 \right) \mathbf{E}_i = -k_i \sum_j k_j \mathbf{E}_j \]

Remark: for isotropic media the r.h.s. of this equation would vanish (\( \mathbf{k} \cdot \mathbf{E} = 0 \)).

Now, we have the following problem to solve:

\[ \sum_j \left( k_j \mathbf{E}_j \right) = 0 \]

We can write

\[ \sum_j k_j^2 \mathbf{E}_j = -k_i \sum_j k_j \mathbf{E}_j, \]

\[ \Rightarrow \sum_j \frac{u_j^2}{n_j^2 - e_i} = 1 \quad \text{final form of DR} \]

Result:

- For given \( e_i(\omega) \) and direction \((u_1, u_2)\), (because \( u_1^2 + u_2^2 + u_3^2 = 1 \)) we can compute the refractive index \( n(\omega, u_1, u_2) \) seen by the normal mode.

- Explicit calculation (multiplication by all denominators):

\[ u_1^2 \left( n^2 - e_i \right) \left( n^2 - e_i \right) + u_2^2 \left( n^2 - e_i \right) \left( n^2 - e_i \right) n^2 + u_3^2 \left( n^2 - e_i \right) \left( n^2 - e_i \right) n^2 = \]

\[ \left( n^2 - e_i \right) \left( n^2 - e_i \right) \left( n^2 - e_i \right) \]

- The resulting equation is quadratic in \( n \) since the \( n^6 \)-term cancels. Hence, we get two (positive) solutions \( n_a, n_b \), and therefore

\[ k_{sa} = \frac{\omega}{c} n_a, \quad k_{sb} = \frac{\omega}{c} n_b \]

for the two orthogonally polarized normal modes \( \mathbf{E}^{(a)} \) and \( \mathbf{E}^{(b)} \).

In particular, for the propagation in direction of the principal axis \( \Rightarrow u_1 = 1, \quad u_2 = u_3 = 0 \), see 6.4 we find:

\[ \Rightarrow \left( n^2 - e_i \right) \left( n^2 - e_i \right) n^2 = \left( n^2 - e_i \right) \left( n^2 - e_i \right) \left( n^2 - e_i \right) \]

\[ \Rightarrow \left( n^2 - e_i \right) \left( n^2 - e_i \right) \left( n^2 - e_i \right) e_i = 0 \]

\[ \Rightarrow n_a^2 = e_i, \quad n_b^2 = e_i \]

Finally, we can compute the fields of the normal modes. We know:

\[ \sum_j \left( k_j \mathbf{E}_j \right) = 0 \]

and hence

\[ \mathbf{E}_i = -k_i \sum_j k_j \mathbf{E}_j \]
where the sum does not depend on the index \( i \). Hence the last term of the equation must be constant

\[ \sum_{j} k_j E_j = \text{const.} \]  

Knowing that the last term is a constant we can calculate the relation of the individual field components from the first part of the equation

\[ E_1 : E_2 : E_3 = \frac{k_1}{\alpha^2 c^2 \epsilon_1 - k^2} : \frac{k_2}{\alpha^2 c^2 \epsilon_2 - k^2} : \frac{k_3}{\alpha^2 c^2 \epsilon_3 - k^2} \]

and with \( D_j = \epsilon_j \epsilon_i E_i \)

\[ D_1 : D_2 : D_3 = \frac{\epsilon_i k_1}{\alpha^2 c^2 \epsilon_1 - k^2} : \frac{\epsilon_i k_2}{\alpha^2 c^2 \epsilon_2 - k^2} : \frac{\epsilon_i k_3}{\alpha^2 c^2 \epsilon_3 - k^2} \]

How are the normal modes polarized?

- The ratio between the field components is real \( \rightarrow \) phase difference 0 \( \rightarrow \) linear polarization

How do we see the orthogonality \( \mathbf{D}^{(a)} \cdot \mathbf{D}^{(b)} = 0 \)? (be careful: \( \mathbf{E}^{(a)} \cdot \mathbf{E}^{(b)} \neq 0 \))

\[
\mathbf{D}^{(a)} \cdot \mathbf{D}^{(b)} \sim \sum_{i} \frac{k_i k_i c_i^2 u_i^2}{k_i^2 - \alpha^2 c_i^2 \epsilon_i}
\]

\[
= \frac{c^2}{\alpha^2} \frac{k^2}{k^2 - \alpha^2 c^2 \epsilon_i} \left[ k_2^2 \sum_{i} \frac{\epsilon_i u_i^2}{k_i^2 - \alpha^2 c_i^2 \epsilon_i} - k_3^2 \sum_{i} \frac{\epsilon_i u_i^2}{k_i^2 - \alpha^2 c_i^2 \epsilon_i} \right]
\]

Since the two red terms vanish due to the dispersion relation, it follows that \( \mathbf{D}^{(a)} \cdot \mathbf{D}^{(b)} = 0 \). The vanishing of the red terms can be seen when rewriting the dispersion relation:

\[
1 = \sum_{i} \frac{k_i^2 u_i^2}{k_i^2 - \alpha^2 c_i^2 \epsilon_i} = \sum_{i} \left( \frac{k_{i,b}^2 - \alpha^2 c_i^2 \epsilon_i}{k_{i,b}^2 - \alpha^2 c_i^2 \epsilon_i} \right) u_i^2 = \frac{\alpha^2}{c^2} \sum_{i} \frac{\epsilon_i u_i^2}{k_{i,b}^2 - \alpha^2 c_i^2 \epsilon_i}
\]

**6.4.3 Normal surfaces of normal modes**

In addition to the index ellipsoid there is a second geometrical interpretation, the so-called normal surfaces:

- If we plot the refractive indices (wave number or norm of the k-vector divided by \( k_i \)) of the normal modes in the \( k_i \)-space (normal surfaces), we get a centro-symmetric, two layer surface:

![Normal surfaces of normal modes](image)

**isotropic**: sphere

**uniaxial**: 2 points with \( n_s = n_b \) in the poles \( \Rightarrow \) connecting line defines the optical axis (for \( \epsilon_i = \epsilon_{\text{refr}} = \epsilon_{\text{refr}} \), \( \epsilon_i = \epsilon_s \) the z-axis is the optical axis)

**biaxial**: 4 points with \( n_s = n_b \) \( \Rightarrow \) connecting lines define two optical axes

How to read the figure:

- fix propagation direction (\( u_i, u_i \)) \( \rightarrow \) intersection with surfaces
- distances from origin to intersections with surfaces correspond to refractive indices of normal modes
- definition of optical axis \( \rightarrow \) \( n_s = n_b \)

**Summary**: there are two geometrical constructions:

A) index ellipsoid (visualization of dielectric tensor)

- fix propagation direction \( \sim \) index ellipse \( \sim \) half lengths of principal axes give \( n_s, n_b \) (refractive indices of the normal modes)
- optical axis \( \rightarrow \) index ellipse is a circle
- for uniaxial crystals the optical axis coincides with one principal axis

B) normal surfaces (visualization of dispersion relation)

- fix propagation direction \( \sim \) intersection with surfaces
- distances from origin give \( n_s, n_b \)
- optical axis connects points with \( n_s = n_b \)

**Conclusion**:
In anisotropic media and for a given propagation direction we find two normal modes, which are linearly polarized monochromatic plane waves with two different phase velocities \( c/n_a \), \( c/n_b \) and two orthogonal polarization directions \( \mathbf{D}^{(a)}, \mathbf{D}^{(b)} \).

### 6.4.4 Special case: uniaxial crystals

Let us now treat the special (simpler) case of uniaxial crystals. In biaxial crystals we do not find any other effects, just the description is more complicated. The main advantage of uniaxial crystals is that we have rotational symmetry in, e.g., \( x,y \)-direction and therefore all three-dimensional graphs (index-ellipsoid, normal surfaces) can be reduced to two dimensions, and we can sketch them more easily. As we have seen before, uniaxial crystals have trigonal, tetragonal, or hexagonal symmetry. Let us assume (without loss of generality) that the index ellipsoid is rotational symmetric around the \( z \)-axis, and we have

\[
\varepsilon_1 = \varepsilon_2 = \varepsilon_{ax}, \quad \varepsilon_3 = \varepsilon_a
\]

which we call ordinary and extraordinary refractive indices.

Then, we expect two normal modes:

A) ordinary wave \( \Rightarrow n_a \) independent of propagation direction

B) extraordinary wave \( \Rightarrow n_b \) depends on propagation direction

The \( z \)-axis is, according to definition, the optical axis with \( n_z = n_a \)

- The ordinary wave \( \mathbf{D}^{(or)} \) is polarized perpendicular to the \( z \)-axis and the \( \mathbf{k} \)-vector.
- The extraordinary wave \( \mathbf{D}^{(e)} \) is polarized perpendicular to the \( \mathbf{k} \)-vector and \( \mathbf{D}^{(or)} \).

Let us now derive the dispersion relation: From above we know the implicit form

\[
\sum_i \left( \frac{u_i^2}{n^2 - \varepsilon_i} \right) = \frac{1}{n^2}
\]

For uniaxial crystals this leads to

\[
\frac{u_1^2}{n^2 - \varepsilon_{ax}} + \frac{u_2^2}{n^2 - \varepsilon_{ax}} + \frac{u_3^2}{n^2 - \varepsilon_a} = \frac{1}{n^2}
\]

\[
n^2 \left( n^2 - \varepsilon_a \right) \left[ n^2 - \varepsilon_{ax} \right] \left( u_1^2 + u_2^2 \right) + n^2 \left[ n^2 - \varepsilon_{ax} \right] u_3^2 = \left[ n^2 - \varepsilon_a \right] \left[ n^2 - \varepsilon_{ax} \right]
\]

A) ordinary wave: independent of direction

\[
n_a^2 = \varepsilon_{ax} \Rightarrow k_a^2 = \frac{\omega^2}{c^2} n_a^2 = k_{ax}^2 \varepsilon_{ax}
\]

B) extraordinary wave (derivation is your exercise): dependent on direction

\[
\left( \frac{u_1^2 + u_2^2}{\varepsilon_a} \right) + \frac{u_3^2}{\varepsilon_{ax}} = \frac{1}{n_b^2}, \quad k_b^2 = \frac{\omega^2}{c^2} n_b^2 \left( u_1, u_2, u_3 \right)
\]

Hence for a given direction \( u_i \) one gets the two refractive indexes \( n_a, n_b \).

The geometrical interpretation as normal surfaces is straightforward and can be done, w.l.o.g., in the \( k_z, k_3 \) or \( y, z \) plane \( (u_z = 0) \). We have with \( k_i^2 = k_0^2 n_i^2 \)

A) ordinary wave

\[
k_0^2 = k_1^2 + k_2^2 + k_3^2 = k_{ax}^2 \varepsilon_{ax}
\]

B) extraordinary wave

\[
\frac{1}{\varepsilon_a} \left( k_1^2 + k_2^2 \right) + \frac{1}{\varepsilon_{ax}} k_3^2 = 1
\]

What about the fields? We know from before that

\[
\rightarrow \mathbf{D}_1 : \mathbf{D}_2 : \mathbf{D}_3 = \frac{\varepsilon_{ax} k_1}{\omega^2} : \frac{\varepsilon_{ax} k_3}{c^2} : \frac{\varepsilon_a k_3}{c^2}
\]

For the extraordinary wave all denominators are finite, and in particular \( k_1 = 0 \) implies \( \mathbf{D}_1^{(e)} = 0 \), hence \( \mathbf{D}^{(e)} \) is polarized in the \( y,z \) plane. Then, \( \mathbf{D}^{(or)} \perp \mathbf{D}^{(e)} \) implies that \( \mathbf{D}^{(or)} \) is polarized in \( x \)-direction.
In summary, we find for the polarizations of the fields:

A) ordinary: \( \mathbf{D} \) perpendicular to optical axis and \( \mathbf{k} \),
\[ \mathbf{D} \perp \mathbf{k}, \mathbf{D} \parallel \mathbf{E} \]

B) extraordinary: \( \mathbf{D} \) perpendicular to \( \mathbf{k} \) and in the plane \( \mathbf{k} \)-optical axis
\[ \mathbf{D} \perp \mathbf{k}, \mathbf{D} \parallel \mathbf{E} \], because
\[ D_1 = \varepsilon_{oo} E_1, D_3 = \varepsilon_{oo} E_3 \]

If we introduce an angle \( \Theta \), as in the figures below, to describe the propagation direction, a simple computation of \( n_x^2(\Theta) \) for the extraordinary wave is possible (exercise):
\[ n_x^2(\Theta) = \frac{\varepsilon_{oo} \sin^2 \Theta}{\varepsilon_{ee} \sin^2 \Theta + \varepsilon_e \cos^2 \Theta} \]

The following classification for uniaxial crystals is commonly used
\[ \varepsilon_{oo} > \varepsilon_e \rightarrow \text{negative uniaxial} \]
\[ \varepsilon_{ee} < \varepsilon_e \rightarrow \text{positive uniaxial} \]

7. Optical fields in isotropic, dispersive and piecewise homogeneous media

7.1 Basics

7.1.1 Definition of the problem
Up to now, we always treated homogeneous media. However, in the context of evanescent waves we already used the concept of an interface. This was already a first step in the direction we now want to pursue. When we treated interfaces so far we never considered effects of the interface, we just fixed the incident field on an interface and described its further propagation in the half-space.

In this chapter, we will go further and consider reflection and transmission properties of the following physical systems:
- interface
- layer (2 interfaces)
- system of layers

Aims
- We will study the interaction of monochromatic plane waves with arbitrary multilayer systems \( \rightarrow \) interferometers, dielectric mirrors, ...
- by superposition of such plane waves we can then describe interaction of spatio-temporal varying fields with multilayer systems
- We will see a new effect, the “trapping” of light in systems of layers \( \rightarrow \) new types of normal modes \( \rightarrow \) “guided” waves \( \rightarrow \) no diffraction

Approach
- take Maxwell’s transition condition for interfaces
- calculate field in inhomogeneous media \( \rightarrow \) matrix method
- solve reflection – transmission problem for interface, layer, and system of layers,
- apply the method to consider special cases like Fabry-Perot-interferometer, 1D photonic crystals, waveguide…

Background
- orthogonality of normal modes of hom. space: no interaction in homogeneous space
- inhomogeneity breaks this orthogonality \( \Rightarrow \) modes interact and exchange energy
- however, locally the concept of eigenmodes is still very useful and we will see that the interaction is limited to a small number of modes
7.1.2 Decoupling of the vectorial wave equation

Before we will start treating a single interface, it is worth looking again at the wave equation in homogeneous space in frequency domain

\[ \text{rot rot } \vec{E}(r, \omega) - \frac{\omega^2}{c^2} \vec{E}(r, \omega) = i \omega \mu_0 \vec{j}(r, \omega) + \mu_0 \omega^2 \vec{P}(r, \omega) \]

In general, for isotropic media all field components are coupled due to the rot rot operator. However, for problems with translational invariance in at least one direction (homogeneous infinite media, layers or interfaces) a simplification is possible. Let us assume, e.g. translational invariance of the system in \( y \) direction and propagation in the \( x-z \)-plane \( \partial / \partial y = 0 \)

\[ \text{rot rot } \vec{E} = \text{grad div } \vec{E} - \Delta \vec{E} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \Delta \vec{E} = \begin{bmatrix} \Delta^{(2)} \vec{E}_x \\ 0 \\ 0 \end{bmatrix} \]

Then, we can split the electric field as \( \vec{E} = \vec{E}_\perp + \vec{E}_\parallel \), with

\[
\vec{E}_\perp = \begin{bmatrix} E_x \\ E_y \\ 0 \end{bmatrix}, \quad \vec{E}_\parallel = \begin{bmatrix} 0 \\ 0 \\ E_z \end{bmatrix}, \quad \nabla^{(2)} = \begin{bmatrix} \partial^2 / \partial x^2 & \partial^2 / \partial x \partial z & \partial^2 / \partial z^2 \\ 0 & 0 & 0 \end{bmatrix}, \quad \Delta^{(2)} = \begin{bmatrix} \partial^2 / \partial x^2 & \partial^2 / \partial x \partial z & \partial^2 / \partial z^2 \\ 0 & 0 & 0 \end{bmatrix}
\]

\( \vec{E}_\perp \) is polarized perpendicular to the plane of propagation, \( \vec{E}_\parallel \) is polarized parallel to this plane.

Common notations are:

- perpendicular: \( \perp \rightarrow x \rightarrow \text{TE} \) (transversal electric)
- parallel: \( \parallel \rightarrow y \rightarrow \text{TM} \) (transversal magnetic)

Both components are decoupled and can be treated independently:

\[
\Delta^{(2)} \vec{E}_\perp + \frac{\omega^2}{c^2} \vec{E}_\perp(r, \omega) = -i \omega \mu_0 \vec{j}(r, \omega) - \mu_0 \omega^2 \vec{P}(r, \omega)
\]

\[
\Delta^{(2)} \vec{E}_\parallel + \frac{\omega^2}{c^2} \vec{E}_\parallel(r, \omega) - \nabla^{(2)} \vec{E}_\parallel = -i \omega \mu_0 \vec{j}(r, \omega) - \mu_0 \omega^2 \vec{P}(r, \omega)
\]

From

\[
\vec{H}(r, \omega) = -\frac{i}{\omega \mu_0} \text{rot } \vec{E}(r, \omega)
\]

we can conclude that the corresponding magnetic fields are

\[
\vec{E}_{\text{TE}} = \begin{bmatrix} E_x \\ 0 \\ 0 \end{bmatrix}, \quad \vec{H}_{\text{TE}} = \begin{bmatrix} H_x \\ 0 \end{bmatrix}, \quad \vec{E}_{\text{TM}} = \begin{bmatrix} 0 \\ E_y \\ 0 \end{bmatrix}, \quad \vec{H}_{\text{TM}} = \begin{bmatrix} 0 \\ H_y \\ 0 \end{bmatrix}
\]

7.1.3 Interfaces and symmetries

Up to now we treated plane waves of the form

\[
\vec{E}(r, \omega) = \exp\left[i(\vec{k} \cdot r - \omega t)\right]
\]

Here, homogeneous space implies \( \exp(\vec{k} \cdot r) \) and monochromaticity leads to \( \exp(-i \omega t) \)

Now, we will break homogenity in \( x \)-direction by considering an interface in \( y-z \) – plane which is infinite in \( y \) and \( z \)

W.l.o.g. we can assume \( \vec{k} = (k_x, 0, k_z) \) by choosing an appropriate coordinate system (plane of incidence is the \( x,z \) plane), and then the problem does not depend on the \( y \) - coordinate.

As pointed out before, we can split the fields in TE and TM polarization \( \vec{E} = \vec{E}_{\text{TE}} + \vec{E}_{\text{TM}} \) and treat them separately. We still have homogeneity in \( z \) - direction, and therefore we expect solutions \( -\exp(ik_z z) \).

The wave vector component \( k_z \) has to be continuous at the interface (follows strictly from continuity of transverse field components, see 7.1.4.). Therefore, we can write for the electric field:

\[
\vec{E}(x, z, t) = \vec{E}_{\text{TE}}(x) \exp\left[i(k_z z - \omega t)\right] + \vec{E}_{\text{TM}}(x) \exp\left[i(k_z z - \omega t)\right]
\]

7.1.4 Transition conditions

From Maxwell’s equations follow transition conditions for the field components. Here we will use that \( \vec{E}_i, \vec{H}_i \) (transverse components) are continuous at an interface between two media. This implies for the:

- **fields**
  - TE: \( E = E_z \) and \( H_z \) continuous
TM: $E_y$ and $H_z = H_z$ continuous

B) wave vectors
homogeneous in $z$-direction $\rightarrow$ phase $e^{ik_zz}$ $k_z$ continuous

### 7.2 Fields in a layer system $\rightarrow$ matrix method

We will now derive a quite powerful method to compute the electromagnetic fields in a system of layers with different dielectric properties.

#### 7.2.1 Fields in one homogeneous layer

Let us first compute the fields in one homogeneous layer of thickness $d$ and dielectric function $\varepsilon_\omega(z)$.

- **aim:** for given fields at $x=0$ $\rightarrow$ calculate fields at $x=d$
- **strategy:**
  - do computation with transverse field components (because they are continuous)
  - the normal components can be calculated later

We will assume monochromatic light (one Fourier component)

#### TE-polarization

We have to solve the wave equation (no $y$-dependence because of translational invariance):

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\omega^2}{c^2} \varepsilon_\omega \right] E_{TE}(x,z) = 0$$

We use the ansatz from above:

$$E_{TE}(x,z) = E_{TE}(x) \exp(ik_zz), \quad H_{TE}(x,z) = H_{TE}(x) \exp(ik_zz)$$

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\omega^2}{c^2} \varepsilon_\omega - k_z^2 \right] E_{TE}(x) = 0$$

with: $H_{TE}(x,z) = -\frac{i}{\omega \mu_0} \text{rot}E_{TE}(x,z)$

Now let us extract the equations for transversal fields $E_y, H_z$:

$$\left[\frac{\partial^2}{\partial x^2} + k_{TE}^2(k_z,\omega) \right] E(x) = 0 \quad \text{with} \quad k_{TE}^2(k_z,\omega) = \frac{\omega^2}{c^2} \varepsilon_\omega(\omega) - k_z^2$$

$$H_z(x) = -\frac{i}{\omega \mu_0} \frac{\partial}{\partial x} E(x)$$

This makes sense: The wave equation for the $y$-component of the electric field is of second order, so we need to specify the field and its first derivative as initial condition at $x=0$.

#### TM-polarization

analog for transversal components $H_y = H_y, \quad E_z$:

$$\left[\frac{\partial^2}{\partial x^2} + k_{TM}^2(k_z,\omega) \right] H(x) = 0$$

$$E_z(x) = \frac{i}{\omega \varepsilon_\omega} \frac{\partial}{\partial x} H(x)$$

Again, we succeed describing everything in transversal components.

We have the following problem to solve:

- calculate fields $(E, H)$ and derivatives $\frac{\partial}{\partial x} E(x), \frac{\partial}{\partial x} H(x)$ at $x=d$ for given values at $x=0$
- at the end: $H_{TM} \rightarrow E_{TM} \rightarrow E = E_{TM} + E_{TE}$

Because the equations for TE and TM have identical structure, we can treat them simultaneously. We rename

$E(x, z; \omega), H(x, z; \omega)$ and in the following we will omit $\omega$.

#### Generalized field

We know the general solution of this system (harmonic oscillator):

$$\begin{align*}
F(x) &= C_1 \exp(ik_x x) + C_2 \exp(-ik_x x) \\
G(x) &= \alpha_{TE} F(x) \\
H_{TM}(x) &= \alpha_{TM} F(x)
\end{align*}$$

with $\alpha_{TE} = 1, \alpha_{TM} = \frac{1}{\mu_0}$
The final solution of the initial value problem is therefore:

\[ F(x) = \cos(k_{\text{ix}}x)F(0) + \frac{1}{\alpha_i k_{\text{ix}}} \sin(k_{\text{ix}}x)G(0) \]

\[ G(x) = -\alpha_i k_{\text{ix}} \sin(k_{\text{ix}}x)F(0) + \cos(k_{\text{ix}}x)G(0) \]

We can write this formally as

\[
\begin{bmatrix}
F(x) \\
G(x)
\end{bmatrix}
= \hat{m}(x)
\begin{bmatrix}
F(0) \\
G(0)
\end{bmatrix}
\]

where the 2 x 2-matrix \( \hat{m} \) describes propagation of the fields:

\[
\hat{m}(x) = \begin{bmatrix}
\cos(k_{\text{ix}}x) & \frac{1}{k_{\text{ix}} \alpha_i} \\
-k_{\text{ix}} \alpha_i \sin(k_{\text{ix}}x) & \cos(k_{\text{ix}}x)
\end{bmatrix}
\]

- To compute the field at the end of the layer we set \( x = d \).
- We assume no absorption in the layer \( \Rightarrow \| \hat{m}(x) \| = 1 \).
- A system of layers is characterized by \( \varepsilon, d \).

If multiple layers are considered, the fields between them connect continuously since the field components used for the description of the fields are the continuous tangential components.

Hence, we can directly write the formalism for a multilayer system, it just requires matrix multiplication:

A) two layers

\[
\begin{bmatrix}
F(x) \\
G(x)
\end{bmatrix}_{d_1 + d_2} = \hat{m}_1(d_1) \hat{m}_2(d_2) \begin{bmatrix}
F(0) \\
G(0)
\end{bmatrix}
\]

B) \( N \) layers

\[
\begin{bmatrix}
F(x) \\
G(x)
\end{bmatrix}_{d + d_2 + \ldots + d_N} = \prod_{i=1}^{N} \hat{m}_i(d_i) \begin{bmatrix}
F(0) \\
G(0)
\end{bmatrix} = \hat{M}(x) \begin{bmatrix}
F(0) \\
G(0)
\end{bmatrix}
\]

All matrices \( \hat{m}_i \) have the same form, but different \( \alpha_i, d_i, k_{\text{ix}}^2 \).

Summary of matrix method:

- \( F(0) \) and \( G(0) \) given (\( E, H \) for TE, \( E_z, H \) for TM)
- \( k_{\text{ix}}, \alpha_i, d_i \) given \( \rightarrow \) matrix elements
- multiplication of matrices (in the right order) \( \rightarrow \) total matrix
- fields \( F(D) \) and \( G(D) \)
7.3 Reflection – transmission problem for layer systems

7.3.1 General layer systems

7.3.1.1 Reflection- and transmission coefficients → generalized Fresnel formulas

In the previous chapter, we have learned how to link the electromagnetic field on one side of an arbitrary multilayer system with the on the other side. We have seen that after splitting in TE/TM polarization, continuous (transversal) field components are sufficient to describe the whole field. What we will do now is to link those field components with accessible fields, i.e. incident, reflected, and transmitted fields. In particular, we want to solve the reflection transmission problem, which means that we have to compute reflected and transmitted fields for a given angle of incidence, frequency, layer system and polarization.

We introduce the wave vectors of incident \( k_s \), reflected \( k_r \) and transmitted \( k_t \) fields:

\[
\begin{align*}
&\begin{bmatrix} k_x \sin \theta \sin \varphi \end{bmatrix} \\
&\begin{bmatrix} 0 \end{bmatrix}
\end{align*},
\begin{align*}
&\begin{bmatrix} -k_x \sin \theta \sin \varphi \end{bmatrix} \\
&\begin{bmatrix} 0 \end{bmatrix}
\end{align*},
\begin{align*}
&\begin{bmatrix} k_x \sin \theta \sin \varphi \end{bmatrix} \\
&\begin{bmatrix} 0 \end{bmatrix}
\end{align*}
\]

with \( k_{\text{ss}} = \frac{\omega}{c} \sqrt{\varepsilon_s - k_s^2} = \sqrt{k_s^2(\omega) - k_s^2} \), \( k_{\text{ts}} = \frac{\omega}{c} \sqrt{\varepsilon_t - k_t^2} = \sqrt{k_t^2(\omega) - k_t^2} \),

where \( \varepsilon_s(\omega) \) and \( \varepsilon_t(\omega) \) are dielectric functions of substrate and cladding.

Remark on law of reflection and transmission (Snellius)

It is possible to derive Snellius law just from the fact that \( k_z \) is a conserved quantity:

1. \( k_s \sin \varphi_s = k_s \sin \varphi_r \) \( \lor \) \( k_s \sin \varphi_s = k_s \sin \varphi_t \) (reflection)
2. \( k_s \sin \varphi_s = k_s \sin \varphi_t \) \( \land \) \( n_s \sin \varphi_s = n_t \sin \varphi_t \) (Snellius)

Let us now rewrite the fields in order to solve the reflection transmission problem:

A) field in substrate

complex amplitudes \( F_s, F_n \):

\[
F_s(x, z) = \exp(i k_s z) \left[ F_s(x) \exp(i k_s x) + F_n(x) \exp(-i k_s x) \right]
\]

\[
G_n(x, z) = i \alpha_s k_{ss} \exp(i k_s z) \left[ F_s(x) \exp(i k_s x) - F_n(x) \exp(-i k_s x) \right]
\]

B) field in layer system

- matrix method

\[
F_z(x, z) = \exp(i k_z z) F(x)
\]

\[
G_z(x, z) = \exp(i k_z z) G(x)
\]

and the amplitudes \( F(x) \) and \( G(x) \) are given by

\[
\begin{bmatrix} F \\ G \end{bmatrix} = \hat{\mathbf{M}}(x) \begin{bmatrix} F \\ G \end{bmatrix}_0
\]

C) field in cladding

\[
F_c(x, z) = \exp(i k_z z) F_z \exp(i k_{ss} (x - D))
\]

\[
G_c(x, z) = i \alpha_c k_{ss} \exp(i k_s z) F_z \exp(i k_{ss} (x - D))
\]

Note that in the cladding we consider a forward (transmitted) wave only.

Reflection transmission problem

We want to compute \( F_n \) and \( F_z \) for given \( F_s \), \( k_s \sim \sin \varphi_s \). \( \varepsilon_s, d \). We know that \( F \) and \( G \) are continuous at the interfaces, in particular at \( x = 0 \) and \( x = D \).

We have:

\[
\begin{bmatrix} F \\ G \end{bmatrix}_D = \hat{\mathbf{M}}(D) \begin{bmatrix} F \\ G \end{bmatrix}_0
\]

Field in cladding at \( x = D \)

Field in substrate at \( x = 0 \)
On the other hand, we have expressions for the fields at x=0 and x=D from our decomposition in incident, reflected and transmitted field from above. Hence, we can write:

\[
\begin{pmatrix}
F_x \\
F_y
\end{pmatrix} = \begin{pmatrix}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{pmatrix} \begin{pmatrix}
F_x + F_x \\
F_y + F_y
\end{pmatrix}.
\]

We consider \( F_x \) as known, and \( F_x \) and \( F_y \) as unknown and get:

\[
F_x = \begin{pmatrix}
\alpha \alpha_k M_{22} - \alpha_k M_{11} \\
\alpha \alpha_k M_{22} + \alpha_k M_{11}
\end{pmatrix} + \begin{pmatrix}
M_{21} - \alpha \alpha_k M_{12} \\
M_{21} - \alpha \alpha_k M_{12}
\end{pmatrix} F_x \\
F_y = \begin{pmatrix}
2 \alpha \alpha_k (M_{12} - M_{11}) \\
2 \alpha \alpha_k (M_{12} - M_{11})
\end{pmatrix} F_y
\]

\[
F_x = \frac{2 \alpha \alpha_k (M_{12} - M_{11})}{\alpha \alpha_k (M_{12} - M_{11}) + i(M_{21} - \alpha \alpha_k M_{12})} F_x \\
F_y = \frac{2 \alpha \alpha_k (M_{12} - M_{11})}{\alpha \alpha_k (M_{12} - M_{11}) + i(M_{21} - \alpha \alpha_k M_{12})} F_y
\]

Those are the general formulas for reflected and transmitted amplitudes. The matrix elements depend on polarization direction \( \epsilon_{ij} \).

Let us now transform back to the physical fields, and write the solution for the results of the reflection transmission problem for TE and TM polarization:

**A) TE-polarization**

\[
F = E = E_y, \quad \alpha_{\text{TE}} = 1
\]

i) reflected field

\[
E_{\text{TE}} = R_{\text{TE}} E_x
\]

with the reflection coefficient

\[
R_{\text{TE}} = \begin{pmatrix}
M_{22} - M_{11} \\
M_{22} + M_{11}
\end{pmatrix} \begin{pmatrix}
M_{21} + M_{12} \\
M_{21} - M_{12}
\end{pmatrix}
\]

ii) transmitted field

\[
E_{\text{TE}} = T_{\text{TE}} E_x
\]

with the transmission coefficient

\[
T_{\text{TE}} = \frac{2 \alpha \alpha_k}{\alpha \alpha_k (M_{12} - M_{11}) + i(M_{21} - \alpha \alpha_k M_{12})} = \frac{2 \alpha \alpha_k}{\alpha \alpha_k},
\]

->We get complex coefficients for reflection and transmission, which determine the amplitude and phase of the reflected and transmitted light.

**B) TM-polarization**

\[
F = H = H_y, \quad \alpha_{\text{TM}} = \frac{1}{\epsilon}
\]

In the case of TM polarization we have the problem that an analog calculation to TE would lead to \( H_{\epsilon,} / H_{y} \), i.e., relations between the magnetic field. However, we want \( E_{\text{TM}} / E_{x} \) ! Therefore, we have to convert the \( H \)-field to the \( E \)-field:

As can be seen in the figure, we can express the amplitude of the \( E_{\text{TM}} \) field in terms of the \( E \)-component:

\[
E_{\text{TE}} = -\sin \varphi = \frac{k_x}{k}
\]

\[
\wedge E_{\text{TM}} = -\frac{k_y}{k_x} E_{\text{TE}}
\]

With Maxwell we can link \( E \) to \( H \):

\[
E = \frac{1}{\omega \epsilon_0 c} (k \times H) \rightarrow E_x = \frac{1}{\omega \epsilon_0 c} k_x H_y \rightarrow E_{\text{TM}} = -\frac{k_x}{\epsilon_0 c} H_y = -\frac{1}{\epsilon_0 c} H_y
\]

result:

\[
E_{\text{TM}} = \frac{\epsilon_{\text{TM}}}{\epsilon_{\text{TE}}} H_{\text{TM}} = \frac{\epsilon_{\text{TM}}}{\epsilon_{\text{TE}}} H_{\text{TE}} \rightarrow \sqrt{\epsilon_{\text{TM}}/\epsilon_{\text{TE}}} \text{ relevant for transmission only}
\]

Hence we find the following for TM polarization:

\[
E_{\text{TM}} = R_{\text{TM}} E_{\text{TE}}
\]

with the reflection coefficient

\[
R_{\text{TM}} = \begin{pmatrix}
\epsilon \epsilon_k M_{22} - \epsilon_k M_{11} \\
\epsilon \epsilon_k M_{22} + \epsilon_k M_{11}
\end{pmatrix} \begin{pmatrix}
M_{21} + \epsilon \epsilon_k M_{12} \\
M_{21} - \epsilon \epsilon_k M_{12}
\end{pmatrix}
\]

\[
E_{\text{TM}} = T_{\text{TM}} E_{\text{TE}}
\]

with the transmission coefficient
In summary, we have found different complex coefficients for reflection and transmission for TE and TM polarization. The resulting generalized Fresnel formulas for multilayer systems are

\[ R_{\text{TM}} = \frac{2k_n}{N_{\text{TM}}} \left( c_n k_{\text{TM}} M_{22}^{\text{TM}} + d_n k_{\text{TM}} M_{11}^{\text{TM}} \right) - \frac{1}{2} \left( c_n k_{\text{TM}} M_{21}^{\text{TM}} + d_n k_{\text{TM}} M_{12}^{\text{TM}} \right) \]

\[ T_{\text{TM}} = \frac{2k_n}{N_{\text{TM}}} \left( c_n k_{\text{TM}} M_{22}^{\text{TM}} + d_n k_{\text{TM}} M_{11}^{\text{TM}} \right) + \frac{1}{2} \left( c_n k_{\text{TM}} M_{21}^{\text{TM}} + d_n k_{\text{TM}} M_{12}^{\text{TM}} \right) \]

\[ R_{\text{TE}} = \frac{2k_n}{N_{\text{TE}}} \left( c_n k_{\text{TE}} M_{22}^{\text{TE}} + d_n k_{\text{TE}} M_{11}^{\text{TE}} \right) - \frac{1}{2} \left( c_n k_{\text{TE}} M_{21}^{\text{TE}} + d_n k_{\text{TE}} M_{12}^{\text{TE}} \right) \]

\[ T_{\text{TE}} = \frac{2k_n}{N_{\text{TE}}} \left( c_n k_{\text{TE}} M_{22}^{\text{TE}} + d_n k_{\text{TE}} M_{11}^{\text{TE}} \right) + \frac{1}{2} \left( c_n k_{\text{TE}} M_{21}^{\text{TE}} + d_n k_{\text{TE}} M_{12}^{\text{TE}} \right) \]

### 7.3.1.2 Reflectivity and transmissivity

In the previous chapter we have computed the coefficients of reflection and transmission, which relate the electric fields in TE and TM polarization of incident, reflected and transmitted wave. However, in many situations it is more important to know the relation of energy fluxes, the so called reflectivity and transmissivity. In order to get information on these quantities we have to compute the energy flux perpendicular to the interface:

- flux through a surface with \( x = \text{const} \)

\[ \langle S \rangle e_n = \frac{1}{2} \Re \left( E \times H^* \right) e_n \]

With \( H^* = \frac{1}{\omega \mu_0} (k^* \times E^*) \)

we find

\[ \langle S \rangle e_n = \frac{1}{2 \omega \mu_0} \Re (k^* e_n) |E|^2 = \frac{1}{2 \omega \mu_0} \Re (k e_n) |E|^2. \]

Since in an absorption free medium the energy flux is conserved, in an absorption free layer the energy flux is also conserved.

In the substrate, \( k_n = \sqrt{\mu_n} \) is supposed to be real, because we have our incident wave coming from there. The total energy flux from the substrate to the layer system is given as

\[ \langle S \rangle e_n = \frac{1}{2 \omega \mu_0} \Re (k e_n) |E|^2 \]

In contrast, in the cladding may be complex!

The energy flux from the layer system into the cladding is

\[ \langle S \rangle e_n = \frac{1}{2 \omega \mu_0} \Re (k e_n) |E|^2 \]

We now will compute the global reflectivity \( \rho \) and transmissivity \( \tau \) of a layer system. Of course, we will decompose into TE and TM polarizations and relate to the reflectivities \( \rho_{\text{TE,TM}} \) and transmissivities \( \tau_{\text{TE,TM}} \). We know:

\[ E_x = E_{x}^{\text{TE}} + E_{x}^{\text{TM}}, \quad E_x = E_{x}^{\text{TE}} + E_{x}^{\text{TM}} \]

\[ |E_{z}^{\text{TE}}|^2 = |E_{z}^{\text{TE}}|^2 + |E_{z}^{\text{TM}}|^2 + \Re \left( k_{z}^{\text{TE}} |E_{z}^{\text{TE}}|^2 \right) + \Re \left( \frac{k_{z}^{\text{TM}}}{k_n} |E_{z}^{\text{TM}}|^2 \right) \]

\[ \left( |E_{z}^{\text{TE}}|^2 + |E_{z}^{\text{TM}}|^2 + \Re \left( k_{z}^{\text{TE}} |E_{z}^{\text{TE}}|^2 \right) + \Re \left( \frac{k_{z}^{\text{TM}}}{k_n} |E_{z}^{\text{TM}}|^2 \right) \right) \]

Here, we just substituted the reflected and transmitted field amplitudes by incident amplitudes times Fresnel coefficients. Now, we decompose the incident field as follows:

\[ E_{z}^{\text{TE}} = E_{z}^{\text{TE}}(E_{z}^{\text{TE}}), \quad E_{z}^{\text{TM}} = E_{z}^{\text{TM}}(E_{z}^{\text{TM}}) \]

Then, we can divide by the (arbitrary) amplitude \( |E_{z}^{\text{TE}}|^2 \) and write

\[ 1 = \left| \frac{R_{\text{TE}}^{\text{TE}}}{k_{z}^{\text{TE}}} + \Re \left( \frac{k_{z}^{\text{TE}}}{k_n} |E_{z}^{\text{TE}}|^2 \right) \cos^2 \delta + \Re \left( \frac{k_{z}^{\text{TM}}}{k_n} |E_{z}^{\text{TM}}|^2 \right) \sin^2 \delta \right| \]

\[ 1 = \left| R_{\text{TM}}^{\text{TM}} |E_{z}^{\text{TM}}|^2 \cos^2 \delta + R_{\text{TM}}^{\text{TM}} |E_{z}^{\text{TM}}|^2 \sin^2 \delta + \Re \left( \frac{k_{z}^{\text{TE}}}{k_n} |E_{z}^{\text{TE}}|^2 \right) \sin^2 \delta \right| \]

The red and blue terms can be identified as \( |I| = \rho + \tau \)

The global reflectivity and transmittivity are therefore given as
\[ \rho = \rho_{\text{TE}} \cos^2 \delta + \rho_{\text{TM}} \sin^2 \delta \]
\[ \tau = \tau_{\text{TE}} \cos^2 \delta + \tau_{\text{TM}} \sin^2 \delta \]

with the reflectivities
\[ \rho_{\text{TE, TM}} = \left| R_{\text{TE, TM}} \right|^2, \quad \tau_{\text{TE, TM}} = \frac{\mathcal{R}(k_{\text{ex}})}{k_{\text{ex}}} \left| R_{\text{TE, TM}} \right|^2. \]

for the two polarization states TE and TM.

### 7.3.2 Single interface

#### 7.3.2.1 (classical) Fresnel formulas

Let us now consider the important example of the most simple layer system, namely the single interface. The relevant wave vectors are (as usual):

\[
\begin{align*}
\mathbf{k}_i &= \left( \begin{array}{c} k_{\text{ex}} \\ 0 \\ k_x \\ k_z \end{array} \right), \\
\mathbf{k}_r &= \left( \begin{array}{c} k_{\text{ex}} \\ 0 \\ k_x \\ k_z \end{array} \right), \\
\mathbf{k}_t &= \left( \begin{array}{c} 0 \\ 0 \\ k_x \\ k_z \end{array} \right).
\end{align*}
\]

The continuous component of the wave vector, expressed in terms of the angle of incidence, is

\[ k_x = \frac{\omega}{c} \sqrt{\epsilon_x} \sin \phi_t = \frac{\omega}{c} n_x \sin \phi_t. \]

Then, the discontinuous component is given as

\[ \rightarrow k_{\text{ex}} = \sqrt{c^2 - k_x^2 - k_z^2} = \frac{\omega}{c} \sqrt{\epsilon_x - \epsilon_x \sin^2 \phi_t} = \frac{\omega}{c} \sqrt{n_x^2 - n_x^2 \sin^2 \phi_t} \]
\[ \leftarrow k_{\text{ex}} = \frac{\omega}{c} n_x \cos \phi_t, \quad k_{\text{ex}} = \frac{\omega}{c} \sqrt{n_x^2 - n_x^2 \sin^2 \phi_t} = \frac{\omega}{c} n_x \cos \phi_t. \]

As above, we can assume that \( k_{\text{ex}} \) is always real, because otherwise we have no incident wave. \( k_{\text{ex}} \) is real for \( n_x > n_x \sin \phi_t \), but imaginary for \( n_x < n_x \sin \phi_t \) (total internal reflection).

The matrix for a single interface is the unit matrix
\[ \mathbf{M} = \hat{\mathbf{m}}(d = 0) = \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \]

and it is easy to compute coefficients for reflection and transmission, and reflectivity and transmissivity. Using the formulas from above we find:

**A) TE-Polarization**

\[
\begin{align*}
R_{\text{TE}} &= \frac{k_{\text{ex}} M_{22} - k_{\text{ex}} M_{11}}{k_{\text{ex}} M_{22} + k_{\text{ex}} M_{11}} - \mathbf{i} \left( M_{21} + k_{\text{ex}} M_{12} \right), \\
T_{\text{TE}} &= \frac{2k_{\text{ex}}}{N_{\text{TE}}},
\end{align*}
\]

**B) TM-Polarization**

\[
\begin{align*}
R_{\text{TM}} &= \frac{e_x k_x M_{22} - e_x k_x M_{11}}{e_x k_x M_{22} + e_x k_x M_{11}} - \mathbf{i} \left( e_x k_x M_{21} + e_x k_x M_{12} \right), \\
T_{\text{TM}} &= \frac{2\sqrt{\epsilon_x} k_{\text{ex}}}{N_{\text{TM}}},
\end{align*}
\]

with:
\[ \hat{\mathbf{m}}(d = 0) = \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \]
\[
\rho_{TM} = \left| R_{TM} \right|^2 = \frac{k_{xx} e_c - k_{xx} e_a}{k_{xx} e_c + k_{xx} e_a},
\]
\[
\tau_{TM} = \frac{\Im(k_{xx})}{k_{xx}} \left| R_{TM} \right|^2 = \frac{4k_{xx} \Im(k_{xx}) e_c e_a}{k_{xx} e_c + k_{xx} e_a}
\]
\[
\implies \rho_{TM} + \tau_{TM} = 1
\]

Remark

It may seem that we have a problem for \( \phi_i = 0 \). For \( \phi_i = 0 \), TE and TM polarization should be equivalent, because the fields are always polarized parallel to the interface. However, formally we have \( R_{TM} = -R_{TM} \), \( T_{TM} = T_{TM} \). The “strange” behavior of the coefficient of reflection can be explained by the following figures:

![Figure](image)

### 7.3.2.2 Total internal reflection (TIR) for \( \varepsilon_S > \varepsilon_C \)

Let us now consider the special case when all incident light is reflected from the interface. This means that the reflectivity is unity.

\[
\rho_{TE} = \left| R_{TE} \right|^2 = \frac{k_{xx} e_c - k_{xx} e_a}{k_{xx} e_c + k_{xx} e_a},
\]
\[
\rho_{TM} = \left| R_{TM} \right|^2 = \frac{k_{xx} e_c - k_{xx} e_a}{k_{xx} e_c + k_{xx} e_a}
\]

With \( k_{xx} = \frac{\omega}{c} \sqrt{n_c^2 - n_s^2 \sin^2 \phi_i} \) we can compute the smallest angle of incidence with \( \rho_{TE, TM} = 1 \):

\[
k_{xx} = 0 \quad \implies \quad n_s = n_s \sin \phi_{tot}
\]
\[
\sin \phi_{tot} = \frac{n_s}{n_s}
\]

For angles of incidence larger than this limit angle, \( \phi_i > \phi_{tot} \), we have

\[
k_{xx} = \frac{\omega}{c} \sqrt{n_c^2 - n_s^2 \sin^2 \phi_i} = \frac{\omega}{c} \sqrt{\frac{\varepsilon_s}{\varepsilon_c} \varepsilon_c} \quad \text{imaginary}
\]
\[
\implies \Im(k_{xx}) = 0 \to \text{TIR}
\]

Obviously, we find the same angle of TIR for TE and TM polarization. The energy fluxes are given as (here TE, same result for TM):

\[
\rho_{TE} = \frac{k_{xx} - i \mu_{ec}}{k_{xx} + i \mu_{ec}} = 1, \quad \tau_{TE} = \frac{4k_{xx} \Im(k_{xx})}{k_{xx} + i \mu_{ec}} = 0.
\]

Remark

For metals in visible range (below the plasma frequency) we have always TIR, because: \( \Im(\varepsilon_c) < 0 \to k_{xx} = \frac{\omega}{c} \sqrt{\varepsilon_c - n_s^2 \sin^2 \phi_i} \) always imaginary!

In the case of TIR the modulus of the coefficient of reflection is one, but the coefficient itself is complex \( \to \) nontrivial phase shift for reflected light:

#### A) TE-polarization

\[
R_{TE} = 1 \cdot \exp(i \Theta) = \frac{k_{xx} - i \mu_{ec}}{k_{xx} + i \mu_{ec}} = \frac{Z}{Z'} = \frac{\exp(i \alpha)}{\exp(-i \alpha)} = \exp(2i \alpha)
\]
\[
\implies \tan \alpha = \tan \frac{\Theta}{2} = -\frac{k_{xx}}{k_{xx}} = \frac{-n_s^2 \sin^2 \phi_i - n_s^2}{n_s \cos \phi_i} = -\frac{\sin^2 \phi_i}{\cos \phi_i}.
\]

#### B) TM-polarization

\[
R_{TM} = 1 \cdot \exp(i \Theta) = \frac{k_{xx} e_c - i \mu_{ec} e_a}{k_{xx} e_c + i \mu_{ec} e_a} = \frac{Z}{Z'} = \frac{\exp(i \alpha)}{\exp(-i \alpha)} = \exp(2i \alpha)
\]
\[
\implies \tan \alpha = \tan \frac{\Theta}{2} = -\frac{k_{xx}}{k_{xx}} = \frac{\varepsilon_c}{\varepsilon_c} = \tan \frac{\Theta}{2}.
\]

In conclusion, we have seen that the phase shifts of the reflected light at TIR is different for TE and TM polarization, and because \( \varepsilon_s > \varepsilon_c \),

\[
\left| \Theta_T \right| > \left| \Theta_M \right|
\]

As a consequence, incident linearly polarized light gets generally elliptically polarized after TIR \( \to \) Fresnel prism

Remark

The field in the cladding is evanescent \( \sim \exp(i k_{xx} x) = \exp(-i \mu_{ec} x) \).

\( \implies \) The averaged energy flux in the cladding normal to the interface vanishes.

\[
\langle S \rangle_k = \frac{1}{20 \mu_{ec}} \Re(k[E]^2) = \frac{1}{20 \mu_{ec}} \Im(k[E]^2) = 0.
\]

#### 7.3.2.3 The Brewster angle

There exists another special angle with particular reflection properties. For TM-polarization, for incident light at the Brewster angle \( \phi_B \) we find \( R_{TM} = 0 \):
\[ \rho_{\text{TM}} = \frac{k_{\text{ck}}^2 \varepsilon_{\text{a}} - k_{\text{ck}}^2 \varepsilon_{\text{s}}^2}{k_{\text{ck}}^2 \varepsilon_{\text{a}} + k_{\text{ck}}^2 \varepsilon_{\text{s}}^2} = 0, \]
\[ \Rightarrow k_{\text{ck}}^2 \varepsilon_{\text{a}} = k_{\text{ck}}^2 \varepsilon_{\text{s}}^2, \]
\[ \varepsilon_{\text{a}}^2 (\varepsilon_{\text{a}} - \sin^2 \varphi_{\text{a}} \varepsilon_{\text{a}}) = \varepsilon_{\text{s}}^2 (\varepsilon_{\text{c}} - \sin^2 \varphi_{\text{a}} \varepsilon_{\text{c}}) \]
\[ \sin^2 \varphi_{\text{a}} = \frac{\varepsilon_{\text{a}} \varepsilon_{\text{c}} (\varepsilon_{\text{a}} - \varepsilon_{\text{s}})}{\varepsilon_{\text{a}} (\varepsilon_{\text{a}} - \varepsilon_{\text{c}})} = \frac{\varepsilon_{\text{c}}}{\varepsilon_{\text{a}} + \varepsilon_{\text{c}}} \]
\[ \cos^2 \varphi_{\text{a}} = 1 - \sin^2 \varphi_{\text{a}} = 1 - \frac{\varepsilon_{\text{c}}}{\varepsilon_{\text{a}} + \varepsilon_{\text{c}}} = \frac{\varepsilon_{\text{a}}}{\varepsilon_{\text{a}} + \varepsilon_{\text{c}}} \]

With the last two lines we can write the final result for the Brewster angle:
\[ \tan \varphi_{\text{B}} = \frac{\varepsilon_{\text{a}}}{\varepsilon_{\text{c}}} \]

The Brewster angle exists only for TM polarization, but for any \( n_s \leq n_a \).

There is a simple physical interpretation, why there is no reflection at the interface for the Brewster angle.

\[ \tan \varphi_{\text{B}} = \frac{\sin \varphi_{\text{B}}}{\cos \varphi_{\text{B}}} = \frac{n_a}{n_s} \]
\[ \Rightarrow n_s \sin \varphi_{\text{B}} = n_a \cos \varphi_{\text{B}} = n_s \sin \left( \frac{\pi}{2} - \varphi_{\text{B}} \right) \]

At the same time the angle of the transmitted light is always
\[ n_s \sin \varphi_{\text{T}} = n_a \sin \varphi_{\text{T}} \cap \varphi_{\text{B}} = \frac{\pi}{2} - \varphi_{\text{T}} \]

Hence, at Brewster angle reflected and transmitted wave propagate in perpendicular directions. If we interpret the reflected light as an emission from oscillating dipoles in the cladding, no reflected wave can occur for TM polarization (no radiation in the direction of dipole oscillation).

In summary, we have the following results for reflectivity and transmittivity at a single interface with \( \varepsilon_{\text{a}} > \varepsilon_{\text{c}} \).

### 7.3.2.4 The Goos-Hänchen-Shift

The Goos-Hänchen shift is a direct consequence of the nontrivial phase shift of the reflected light at TIR. It appears when beams undergo total internal reflection at an interface. The reflected beam appears to be shifted along the interface. As a result it seems as if the beam penetrates the cladding and reflection occurs at a plane parallel to the interface at a certain depth, the so-called penetration depth. For sake of simplicity we will treat here TE-polarization only.

Let us start with an incident plane wave in TE polarization:
Thus, the reflected beam appears shifted by $d = -\Theta'$ (Goos–Hänchen Shift).

Let us finally compute the shift $d = -\Theta'$. We know from before that the phase shift for TIR is given as:

$$
\tan \alpha = \tan \frac{\Theta}{2} = -\frac{\mu_n}{k_n} \frac{1}{\gamma_s}
$$

$$
\Theta = -2 \arctan\left(\frac{\sqrt{\alpha^2 - \frac{\omega}{n_e^2} - \alpha^2}}{\sqrt{\alpha^2 - \frac{\omega}{n_e^2} - \alpha^2}}\right) = -2 \arctan\frac{\mu_n}{k_n}
$$

$$
\Theta' = \frac{d \Theta}{d \alpha} = -2 \times \frac{1}{1 + \mu_k^2} \times \frac{2 \alpha k_n - 2 \alpha k_n}{2 \mu_k k_n} = -2 \frac{\alpha (k_n^2 + \mu_n^2)}{2 \mu_k k_n^2 + \mu_n^2} = -2 \frac{\alpha}{\mu_k k_n^2 + \mu_n^2}
$$

$$
\Theta|_{\alpha_0} = -d = -2 x_{\text{ex}} \tan \phi_{10}
$$

with $x_{\text{ex}} = \frac{1}{\sqrt{\alpha_0 \gamma_s - \omega n_e^2}}$ and $\tan \phi_{10} = \frac{\alpha}{k} = \frac{\alpha}{\gamma}$

$$
\rightarrow x_{\text{ex}} \quad \text{depth of penetration}
$$

### 7.3.3 Periodic multi-layer systems - Bragg-mirrors - 1D photonic crystals

In the previous chapters we have learned how to treat (finite) arbitrary multi-layer systems. Interesting effects occur when those multi-layer systems become periodic. Periodic structures are important in physics (lattices, crystals, atomic chains, waveguide arrays...), and we can gain insight in general features of such systems by looking at optical properties of periodic (dielectric) multi-layer systems, so-called Bragg-mirrors. The reflectivity of such mirrors is almost 100% in certain frequency ranges; the more layers the closer we get to this ideal value. Bragg mirrors are important for resonators (laser, interferometer).

In our theoretical approach, we will assume these layer systems as infinite, i.e. consisting of an infinite number of layers, and we treat them as so-called one-dimensional photonic crystals. We will discuss effects like band gaps, dispersion and diffraction in such periodic media, and gain understanding of the basics of Bragg reflection and the physics of photonic crystals.

In order to keep things simple we will treat:
- semi-infinite periodic multi-layer systems $\left[ x > 0, (e_i, d_i), (e_2, d_2) \right]$
- TE-polarization only
- monochromatic light

At the interface between substrate and Bragg-mirror $(x = 0)$ we have incident and reflected electric field:

$$E_i + E_r = E_0 \quad \text{and} \quad i k_\alpha (E_i - E_r) = \frac{\partial E_0}{\partial x} = E'_0$$

or

$$E_i = \frac{E_0}{2} = \frac{i E'_0}{2k_\alpha} \quad \text{and} \quad E_r = \frac{E_0}{2} + \frac{i E'_0}{2k_\alpha}$$

In chapter 7.2, we developed a matrix formalism involving the generalized fields $F$ and $G$. Because here we treat TE polarization only, we can use directly the electric field amplitude and its derivative with respect to $x$, because $E = F$ and

$$i \omega \mu_0 H_z = G = \frac{\partial E}{\partial x} = E'$$

Let us now calculate the field in the multi-layer system. From before, we know how to treat finite systems with the matrix method. Here, we want to treat an infinite periodic medium (like a one-dimensional crystal). For our example of 2 periodically repeated layers, we have:

$$e(x) = e(x + \Lambda) \quad \text{with the period} \quad \Lambda = d_1 + d_2$$

For infinite periodic media, we can make use of the Bloch-theorem to find the generalized normal modes (Bloch modes or Bloch waves). We seek for solutions like:

$$E(x, z; \omega) = \exp \left[ i \left( k_x (k_x, \omega) x + k_z z \right) \right] E_{k_x} (x)$$

with $E_{k_x} (x + \Lambda) = E_{k_x} (x)$ being a periodic function. In other words, we are looking for solutions which have the same amplitude after one period of the medium, but we allow for a different phase $\sim \exp \left[ i \left( k_x (k_x, \omega) x \right) \right]$. Here $k_x$ is the (yet) unknown Bloch vector. Because in this easy example we deal with a one-dimensional problem, the Bloch vector is actually a scalar.

In the following, we will find a dispersion relation for the Bloch-waves $k_x (k_x, \omega)$, in complete analogy to the DR for plane waves $k^2 = \frac{\omega^2}{c^2} - k^2$ in homogeneous media. In order to make the difference to the homogeneous case more obvious, we change the notation for the Bloch vector: $k_x \to K$.

According to the Bloch-theorem (our ansatz) we have a relation between $E$ and $E'$ when we advance by one period of the multi-layer system (from period $N$ to period $N + 1$):

$$\left( \begin{array}{c} E \\ E' \end{array} \right)_{[N+1]\Lambda} = \exp \left( i K \Lambda \right) \left( \begin{array}{c} E \\ E' \end{array} \right)_{N\Lambda}$$

On the other hand, we know from our matrix method:

$$\left( \begin{array}{c} E \\ E' \end{array} \right)_{[N+1]\Lambda} = M \left( \begin{array}{c} E \\ E' \end{array} \right)_{N\Lambda}$$

with $M = \mathbf{m} (d_2) \mathbf{m} (d_1) \rightarrow M_\gamma = \Sigma m^{(2)} \gamma m^{(1)} \gamma$

If the Bloch wave is a solution to our problem, we can set the two expressions equal:

$$\left( M - \exp \left( i K \Lambda \right) \mathbf{i} \right) \left( \begin{array}{c} E \\ E' \end{array} \right)_{N\Lambda} = 0$$

and with $\mu = \exp \left( i K \Lambda \right)$ we have to solve an eigenvalue problem.

$$\left( \mathbf{M} - \mu \mathbf{K} \mathbf{i} \right) \left( \begin{array}{c} E \\ E' \end{array} \right)_{N\Lambda} = 0$$

This eigenvalue problem determines the Bloch vector $K$ and will finally give our dispersion relation.

As usual, we use the solvability condition $\det \left( \mathbf{M} - \mu \mathbf{K} \mathbf{i} \right) = 0$ to compute the dispersion relation expressed in $\mu = \exp \left( i K \Lambda \right)$. Hence we still need to compute $K$ afterwards.

$$\mu \pm = \exp \left( i K \Lambda \right) = \pm \left( \frac{M_{11} + M_{22}}{2} + \sqrt{\left( \frac{M_{11} + M_{22}}{2} \right)^2 - 1} \right)$$

Note that we used $\det \left( \mathbf{M} \right) = 1$, which explains why the off-diagonal elements of the matrix do not appear in the formula. Moreover, because of $\det \left( \mathbf{M} \right) = 1$ we have $\mu, \overline{\mu} = 1$.

The corresponding eigenvectors (field and its derivative at $x = N\Lambda$) can be computed from

$$\left( \mathbf{M} - \exp \left( i K \Lambda \right) \mathbf{i} \right) \left( \begin{array}{c} E \\ E' \end{array} \right)_{N\Lambda} = 0$$

$$\Rightarrow \left( \begin{array}{cc} M_{11} - \mu & M_{12} \\ M_{21} - \mu & M_{22} \end{array} \right) \left( \begin{array}{c} E \\ E' \end{array} \right)_{N\Lambda} = 0$$
\[
\left( M_{11} - \mu \right) E + M_{12} E' = 0
\]
\[
\begin{pmatrix} E \\ E' \end{pmatrix} = \begin{pmatrix} 1 \\ \left( \mu - M_{11} / M_{12} \right) \end{pmatrix} E_{M}\lambda.
\]

If field values of the Bloch mode, i.e., the function \( E_{M}(x + \Lambda) = E_{M}(x) \), inside the layers are desired, they can be computed by using the matrix formalism and the above eigenvector \( E_{M} \).

We are interested in the reflection properties of an infinite Bragg mirror. Reasoning in terms of the electric field and derivative at the interface \( (x = 0) \), \( E_0 \) and \( E'_0 \), we can express the reflectivity of the Bragg mirror as
\[
\rho = \left| \frac{E_0^*}{E_0} \right|
\]
with \( E_0 = \frac{E_{10}}{2} + \frac{iE'_{10}}{2k_{n1}} \) and \( E_1 = \frac{E_{20}}{2} - \frac{iE'_{20}}{2k_{n2}} \) from before.

\[
\rho = \left| \frac{k_{n1} E_0 + iE'_{0}}{k_{n2} E_0 - iE'_{0}} \right|^2
\]

With our knowledge of the eigenvector from above we can compute the reflectivity:
\[
E_0^* = \frac{\mu - M_{11}}{M_{12}} E_0 \rightarrow \rho = \left| \frac{k_{n1} E_0 + iE'_{0}}{k_{n2} E_0 - iE'_{0}} \right|^2 = \left| \frac{k_{n1}^2 - \frac{i\mu - M_{11}}{M_{12}}}{k_{n2}^2 - \frac{i\mu - M_{11}}{M_{12}}} \right|^2
\]

According to this formula two scenarios are possible:

**A) total internal reflection \( \rho = 1 \)**

Hence \( \mu \) has to be real which results in the condition
\[
\left| \frac{(M_{11} + M_{22})}{2} \right| \leq 1
\]
with \( [\text{for our example } (n_1, n_2, d_1, d_2)] \)
\[
M_{11} = \cos(k_{nx} d_1) \cos(k_{nx} d_2) - \frac{k_{nx}}{k_{nx}} \sin(k_{nx} d_1) \sin(k_{nx} d_2)
\]
\[
M_{22} = \cos(k_{nx} d_1) \cos(k_{nx} d_2) - \frac{k_{nx}}{k_{nx}} \sin(k_{nx} d_1) \sin(k_{nx} d_2).
\]

This defines the so-called band gap, i.e., frequencies of excitation for which no propagating solutions exist.

**B) propagating normal modes**

Hence \( \mu \) must be complex which results in the condition
\[
\left| \frac{(M_{11} + M_{22})}{2} \right| < 1
\]

We can compute the explicit dispersion relation:
\[
\mu = \exp(\pm k \Lambda) = \frac{(M_{11} + M_{22})}{2} \pm \sqrt{\left( \frac{(M_{11} + M_{22})}{2} \right)^2 - 1}
\]

In infinite periodic media, only if the Bloch vector \( K \) fulfills this DR the Bloch wave is a solution to Maxwell’s equations. This is in complete analogy to plane waves in homogeneous media with the DR \( \omega = c k \).

**Interpretation**

- For the case of total internal reflection (\( \mu \) real, \( \left| \frac{(M_{11} + M_{22})}{2} \right| \geq 1 \)) the Bloch vector \( K \) is complex, \( \mu = \exp(\pm k \Lambda) = \exp(\pm \Re(K) \Lambda) \exp(-\Im(K) \Lambda) \).

Hence \( \Re(K(k_\omega, \omega) \Lambda) = n \pi \) and
\[
\Im(K(k_\omega, \omega) \Lambda) = -\ln \left( (-1)^n \left| \frac{(M_{11} + M_{22})}{2} \pm \sqrt{\left( \frac{(M_{11} + M_{22})}{2} \right)^2 - 1} \right| \right).
\]

The \( \pm \) accounts for exponentially damped and growing solution, as we usually expect in the case of complex wave vectors and evanescent waves.

- There is an infinite number of so-called band gaps or forbidden bands, because \( n = 1, 2, \ldots \). Those band gaps are interesting for Bragg mirrors and Bragg waveguides.

- The limits of the bands are given by \( \Re(K(k_\omega, \omega) \Lambda) = n \pi \) and \( \Im(K(k_\omega, \omega) \Lambda) = 0 \) and \( K(k_\omega, \omega) = n \pi / \Lambda \).

- Outside the band gaps, i.e., inside the bands, we find propagating solutions which have different properties than the normal modes in homogeneous media (different dispersion relation). We can exploit the strong curvature, i.e., frequency dependence, of DR for, e.g., dispersion compensation or diffraction free propagation.
**Special case: normal incidence**

In general there is a complex interplay of angle of incidence and frequency of light on the properties of multilayer systems. Therefore let us have a look at the simpler case of normal incidence ($k_z = 0$). Then, the band gaps correspond to "forbidden" frequency ranges. In a graphical representation of the dispersion relation for $k_z = 0$ it is common to use the following dimensionless quantities

$$\frac{\omega}{cG} \quad \text{and} \quad \frac{K}{G}$$

with the scaling constant $G = \frac{2\pi}{\Lambda}$.

Examples for normal incidence

- $n_1 = 1.4$, $d_1 = 0.5\Lambda$
- $n_2 = 3.4$, $d_2 = 0.5\Lambda$

It is common to use the reduced band structure - Brillouin zone, where the information for all possible Bloch vectors is mapped onto the Bloch vectors up to $(k / G) = 0.5$.

Because of $e^{ik\Lambda}$ we need only $-\pi \leq K\Lambda \leq \pi \rightarrow \frac{|K|}{G} \leq 0.5$ to describe the dispersion relation.

Inside the band gap, we find damped solutions:

- $n_0 = 1.0$, $n_1 = 1.4$
- $n_0 = n_1 = 1.4$

Let us quantify the damping. In our example $(n_1,n_2,d_1,d_2)$ we have

$$\frac{M_{11} + M_{22}}{2} = \cos\left(\frac{\omega}{c} n_1 d_1\right) \cos\left(\frac{\omega}{c} n_2 d_2\right) - \frac{1}{2} \left(\frac{n_2}{n_1} + \frac{n_1}{n_2}\right) \sin\left(\frac{\omega}{c} n_1 d_1\right) \sin\left(\frac{\omega}{c} n_2 d_2\right)$$

In the middle of the first band gap (optimum configuration for high reflection) we have $\frac{\omega}{c} n_1 d_1 = \frac{\omega}{c} n_2 d_2 = \frac{\pi}{2}$, with $\omega_n$ being the Bragg frequency, and therefore

$$\frac{M_{11} + M_{22}}{2} = -\frac{1}{2} \left(\frac{n_2}{n_1} + \frac{n_1}{n_2}\right) < -1.$$
expression for $\Im(K)$ and assume small index contrast $|n_2 - n_1| \ll (n_2 + n_1)$ we find

$$\Lambda \Im(K)_{\text{max}} \approx 2 \frac{n_2 - n_1}{n_2 + n_1}$$

(do derivation as an exercise)

$\triangleright$ Damping is proportional to index contrast $|n_2 - n_1|

The spectral width of the gap $\left|\frac{(M_{11} + M_{22})}{2} \right| \geq 1$ is then

$$\Delta \omega_{\text{gap}} \approx \frac{2 \omega_2}{\pi} \Lambda \Im(K)_{\text{max}}$$

(do derivation as an exercise)

$\triangleright$ Spectral width is proportional to index contrast as well.

### 7.3.4 Fabry-Perot-resonators

In this chapter we will treat a special multi-layer system, a so-called Fabry-Perot-Resonator. In this case, one single layer, the so-called cavity, is distinguished from the others and we are interested in the forward and backward propagating fields inside. The other layers function as mirrors, and may be periodic multilayer-systems or metal films. Fabry-Perot-resonators are very important in optics, as they appear as:

- Fabry-Perot- interferometer
- laser with plane mirrors $\rightarrow$ Fabry-Perot-Resonator with active medium inside the cavity
- nonlinear optics $\rightarrow$ high intensities inside the cavity $\rightarrow$ nonlinear optical effects for low intensity incident light:
  - bistability,
  - modulational instability
  - pattern formation, solitons

Here, we want to compute the transmission properties of the resonator for arbitrary plane mirrors. For simplicity, we will restrict ourselves to TE-polarization. The figure shows our setup with two mirrors at $x=0$ and $x=D$, characterized by coefficients of reflection and transmission $R_0$, $T_0$, $R_D$, $T_D$.

- $E_z, E_x$ and $E_y$ $\rightarrow$ amplitudes of incident, reflected and transmitted external fields in substrate and cladding.
- $E_+, E_-$ $\rightarrow$ amplitudes of forward and backward running internal fields inside the cavity.

Using the known coefficients of reflection and transmission, we can link the field amplitudes. Our aim is to eliminate $E_+$ and $E_-$. 

**A)** At the lower mirror inside the cavity:

$$T_0E_+ + R_0E_-(0) = E_z(0)$$

**B)** At the upper mirror outside the cavity:

$$E_+ = T_0E_z(D)$$

And with $E_+(D) = E_z(0) \exp(\mathbf{i}k_{xx}D)$

$$E_-(0) = \frac{R_0}{T_0} E_x \exp(-\mathbf{i}k_{xx}D)$$

**C)** At the upper mirror inside the cavity:

$$E_+(D) = R_0E_z(D) = \frac{R_0}{T_0} E_x$$

and with $E_+(0) = E_z(D) \exp(\mathbf{i}k_{xx}D)$

$$E_+(0) = \frac{R_0}{T_0} E_x \exp(\mathbf{i}k_{xx}D)$$

**D)** we substitute $E_+(0)$ and $E_-(0)$ in A)

$$T_0E_+ + R_0E_-(0) = E_z(0)$$

$$T_0E_+ + R_0 \frac{R_0}{T_0} E_x \exp(\mathbf{i}k_{xx}D) = \frac{E_2}{T_0} \exp(-\mathbf{i}k_{xx}D)$$

$$\Rightarrow E_+ = \frac{1}{T_0} \left\{ \exp(-\mathbf{i}k_{xx}D) - R_0 R_0 \exp(\mathbf{i}k_{xx}D) \right\} E_x.$$
Thus, the coefficient of transmission for the whole FP-resonator expressed in coefficients of the mirrors \((R_0, T_0, R_D, T_D)\), cavity properties and angle of incidence \((\theta, k)\) reads:

\[
T_\text{FP} = \frac{E_\text{out}}{E_\text{in}} = \frac{T_0 T_D \exp(1k_D D)}{1 - R_0 R_D \exp(2ik_D D)}
\]

This is the general transmission function of a lossless Fabry-Perot resonator. In general, the mirror coefficients are complex and the fields get certain phase shifts \(\phi_0, \phi_D\) induced by the coefficients of reflection \(R_0, R_D\). Obviously, only the phase shifts induced by the coefficients of reflection \(R_0, R_D\) are important for the transmissivity of the FP resonator \(\tau_\text{FP} = |T_\text{FP}|^2\).

For given \( |R_0|, |R_D|, |T_0|, |T_D| \), and \(\phi_0, \phi_D\), the general transmissivity of a lossless Fabry-Perot resonator reads:

\[
|T_\text{FP}|^2 = |T|^2 = \frac{|T_0|^2 |T_D|^2}{1 + |R_0|^2 |R_D|^2 - 2|R_0||R_D| \cos \left(2k_D D + \phi_0 + \phi_D\right)}
\]

\[
\tau = \frac{k_\text{ex}}{k_\text{in}} |T|^2 = \frac{k_\text{in}}{k_\text{ex}} \left(\frac{|T_0|^2 |T_D|^2}{1 + |R_0|^2 |R_D|^2 - 2|R_0||R_D| \cos \delta}\right)
\]

Here we introduced the phase-shift \(\delta\) which the field acquires in one round-trip in the cavity.

\section*{Discussion}

Depending on whether the two mirrors have identical properties we distinguish between symmetric and asymmetric FP-resonators.

\subsection*{a) asymmetric FP-resonator}

\[
\tau = \frac{k_\text{ex}}{k_\text{in}} |T|^2 = \frac{|T_0|^2 |T_D|^2}{1 + |R_0|^2 |R_D|^2 - 2|R_0||R_D| \cos \delta}
\]

Because we assume no losses we can use energy conservation at each mirror to eliminate \(T_{0,R}\):

\[
|T_0|^2 |T_D|^2 = \frac{k_\text{ex}}{k_\text{in}} \left(1 - |R_0|^2\right) \frac{k_\text{in}}{k_\text{ex}} \left(1 - |R_D|^2\right) \rightarrow
\]

\[
|T_0|^2 |T_D|^2 = \frac{k_\text{ex}}{k_\text{in}} \left(1 - |R_0|^2\right) \left(1 - |R_D|^2\right)
\]

\[
\text{Note:} \tau \text{ and } \rho \text{ for a lossless mirror are the same for both sides of the mirror.}
\]

\[
\text{For lossy mirrors only } \tau \text{ is the same, } \rho \text{ is then side-dependent.}
\]

Plugging everything in we get

\[
\tau = \frac{1 + |R_0|^2 |R_D|^2 - 2|R_0||R_D| \cos \frac{\delta}{2}}{1 - |R_0|^2 |R_D|^2 - 2|R_0||R_D| \cos \delta}
\]

\[
\tau = \left\{\frac{(1 - |R_0||R_D|)^2}{1 - |R_0|^2} + 4|R_0||R_D| \sin^2 \frac{\delta}{2}\right\}^{-1}
\]

and with

\[
\rho_0 = |R_0|^2, \rho_D = |R_D|^2
\]

\[
\rho_m = |R_0|^2 = |R_D|^2 = \rho_0 = \rho_D, \quad \phi_0 = \phi_D = \phi_0
\]

\[
\tau = \left\{\frac{(1 - \rho_m)^2}{(1 - \rho_m)^2 + 4\rho_m \sin^2 \frac{\delta}{2}}\right\}^{-1}
\]

\[
\tau = \left\{1 + F \sin^2 \frac{\delta}{2}\right\}^{-1}
\]

\[
\text{with } F = \frac{4\rho_m}{(1 - \rho_m)^2}, \quad \frac{\delta}{2} = k_D D + \phi
\]

The \textit{Airy-formula} (see also Labworks script, where phase shifts \(\phi\) due to the mirrors are not considered) gives the transmissivity of a symmetric, lossless Fabry-Perot-resonator. Only for this case we can get the maximum transmissivity \(\tau = 1\) for \(\delta/2 = n\pi\).
Remarks and conclusions

- We can do an analog calculation for TM-polarization \( \rightarrow R^{TM} \) resp. \( \rho^{TM} \).
- Resonances of the cavity with \( T \) occur for
  \[
  \frac{1}{2} f_k D_m + \phi = m\pi \quad \text{with} \quad k_k = \frac{2\pi}{\lambda} \sqrt{n_r^2 - n_s^2 \sin^2 \varphi_1} \]
  \[
  \Rightarrow T = \frac{m\pi - \phi}{k_k} = \frac{\lambda}{2} \sqrt{n_r^2 - n_s^2 \sin^2 \varphi_1} = \frac{\lambda}{2} \text{cavity} \]

- Transmission properties of a given resonator depend on \( \varphi_1 \) and \( \lambda \).
- Minimum transmission is given as
  \[
  \tau_{\text{min}} = \frac{1}{1 + T} \]

- It is favorable to have large \( F \rightarrow \text{e.g.} \):
  \[
  100 = F = \frac{4\rho_m}{(1 - \rho_m)^2} \]
  \[
  \rho_m = 1 - \tau_m \sim \frac{4(1 - \tau_m)}{\tau_m} \approx 4 \quad \Rightarrow \tau_m = 0.2, \quad \rho_m = 0.8. \]

- Pulses and beams can be treated efficiently in Fourier domain:
  \( \rightarrow \) e.g. TE: \( E_x(\alpha, \beta, \omega) = T_{xx}(\alpha, \beta, \omega)E_x(\alpha, \beta, \omega) \)
  Fourier back transformation: \( E_y(x, y, t) = FT^{-1}[E_x(\alpha, \beta, \omega)] \)

- Beams, because they always contain a certain range of incident angles \( \varphi_1 \), produce interference rings in the farfield output (or image of lens, like in labworks).

- A quantity often used to characterize a resonator is the finesse:
  \[
  \Phi = \frac{\text{distance between resonance}}{\text{full width at half maximum of resonance}} = \frac{\Delta}{\varepsilon} \]

with \( \varepsilon \) the FWHM and \( \Delta = \pi \) the distance in rad between two resonances.

\[
\left\{ 1 + F \sin^2 \left( \frac{m\pi \pm \varepsilon}{2} \right) \right\}^{-1} \approx \frac{1}{2} \]

For narrow resonances (small line width \( \varepsilon \)) we can write

\[
\left\{ 1 + F \left( \frac{\varepsilon}{2} \right)^2 \right\}^{-1} \approx \frac{1}{2} \Rightarrow F \left( \frac{\varepsilon}{2} \right)^2 = 1 \quad \Rightarrow \varepsilon = \frac{2}{\sqrt{F}}
\]

- The line width \( \varepsilon \) (FWHM) is inversely proportional to the finesse \( \Phi \).
- The Airy-formula can be expressed in terms of the finesse

\[
\tau = \left\{ 1 + \left( \frac{2\pi}{\lambda} \sin \frac{\delta}{2} \right)^2 \right\}^{-1}.
\]

- A Fabry-Perot-resonator can be used as a spectroscope. Then, we can ask for its resolution (here: normal incidence). Resonances (maximum transmission) occur at:

\[
kD + \phi = m\pi \quad \text{reduced transmission by factor 1/2}
\]

\[
\Rightarrow kD + \phi \pm \frac{\Delta k}{2} = m\pi \pm \frac{\varepsilon}{2} = m\pi \pm \frac{\pi}{2\Phi}
\]

\[
\Rightarrow |\Delta k| = \frac{2\pi}{\lambda} n_r |\Delta \lambda| = \frac{\pi}{\Phi D}
\]

\[
\Rightarrow \frac{\Delta \lambda}{\lambda} = \frac{\lambda}{n_r \Phi D} - \frac{\lambda}{n_r \Phi D} = \frac{\varepsilon}{\Phi D}
\]

Example: \( \lambda = 5 \cdot 10^{-7} \text{m}, \quad \Phi = 30, \quad n_r D = 4 \cdot 10^{-5} \text{m} \quad \Rightarrow \Delta \lambda = 2 \cdot 10^{-12} \text{m} \)

- The field amplitudes (here forward field) inside the cavity are given as:

\[
E_x = T_x E_x(D) \quad \text{Because} \quad |F|^2 - 1/(1 - \rho) - \Phi \quad \text{these intra-cavity fields can be very high \( \rho \) important for nonlinear effects}
\]

- Lifetime of photons in cavity: Via the "uncertainty relation":

\[
\Delta t \Delta \omega = \text{const.} \approx 1 \quad \text{it is possible to define a lifetime of photons inside the cavity:}
\]

\[
|\Delta t| = \frac{1}{\varepsilon} n_r \Delta \omega = \frac{\pi}{\Phi D}
\]

\[
\Delta \omega = \frac{c}{n_r \Phi D} \quad \Rightarrow \quad T_c = \frac{1}{\Delta t} = \frac{n_r \Phi D}{c \pi} = \frac{D}{\varepsilon}.
\]
7.4 Guided waves in layer systems

Finally, we want to explore our layer systems as waveguides. For many applications it is interesting to have waves which propagate without diffraction. This is crucial for integrated optics, where we want to guide light in very small (micrometric or smaller) dielectric layers (film, fiber), or optical communication technology where some light encoded information is transported over long distances. Moreover, waveguides are important in nonlinear optics, due to confinement and long propagation distances nonlinear effects become important. Here, we will treat wave-guiding in one dimension only because we restrict ourselves to layer-systems, but general concepts developed can be transferred to other settings.

### 7.4.1 Field structure of guided waves

Let us first do some general consideration about the field structure of guided waves. We want to find guided waves in a layer system. In such systems, till now, we have solved the reflection and transmission problem: For given \( k_\alpha,\epsilon_{i\alpha} \to \text{calculate} \ E_{i\alpha} \).

Inside each layer we have plane waves \( \exp[\pm ik_x x + k_z z - \omega t] \).

The question is, how can we trap (or guide) waves within a finite layer system? A possible hint gives the effect of total internal reflection, where the transmitted field is:

\[
E_{i\alpha}(x,z) = E_x \exp(\pm ik_x x) \exp(-\mu_{c\alpha} x)
\]

Obviously, in the case of TIR we have no energy flux in the cladding medium. If TIR is the key mechanism to guide light, can we have TIR on two sides (vs. cladding and substrate? And what about a single interface?

Guided waves have the following field structure:

- plane wave in propagation direction \( \sim \exp(ik_z z) \)
- evanescent waves in substrate and cladding
  \[
  \sim \exp[-\mu_{c\alpha}(x-D)] \quad \text{cladding}
  \]
  \[
  \sim \exp(\mu_{c\alpha} x) \quad \text{substrate}
  \]
  \[
  \mu_{c\alpha} = \sqrt{k_z^2 - \frac{\omega^2}{c^2} \epsilon_{c\alpha}(\omega)} > 0
  \]

Note that this 2. condition is not obvious at this stage. It appears due to transition conditions at the boundaries to substrate and cladding.

In summary, the \( z \)-component of the wave vector of guided waves has to fulfill:

\[
\max \left( \frac{\omega}{c \cdot \eta_{c\alpha}} \right) < k_z < \max \left( \frac{\omega}{c \cdot \eta_{\alpha}} \right)
\]

The field structure in substrate and cladding is given as:

- \( E_x(x,z) = E_x \exp(\pm ik_x x) \exp(-\mu_{c\alpha} x) \)
- \( E_x(x,z) = E_x \exp(\pm ik_x x) \exp(\mu_{c\alpha} x) \)

If we go back to our usual reflection transmission problem, we have here reflected and transmitted (evanescent) field for zero incident field \( E_i \to 0 \).

We will see in the following how this can be exploited to derive the dispersion relation for guided waves.

### 7.4.2 Dispersion relation for guided waves

As we have seen above, we have \( E_x, E_x \neq 0 \) for \( E_i \to 0 \), and this can be exploited. The coefficients for reflection and transmission are

\[
R^{TE,TM} = \begin{pmatrix} E_x \\ E_z \end{pmatrix}^{TE,TM}, \quad T^{TE,TM} = \begin{pmatrix} E_x \\ E_z \end{pmatrix}^{TE,TM} \quad E_i \to 0,
\]

and we find \( R, T \to \infty \) in the case of guided waves. In this sense, guided waves are resonances of the system. Let us compare to a driven harmonic oscillator

\[
x = \frac{F}{\omega^2 - \omega_0^2}, \quad x \text{- action, } F \text{- cause}
\]

\( \omega = \omega_0 \) in the case of resonance we get action for infinitesimal cause.

Hence, we can get the dispersion relation of guided waves by looking for a vanishing denominator in \( R, T \). This reasoning is a general principle in physics:
The poles of the response function (or Greens function) are the resonances of the system.

We know the coefficients of reflection and transmission for a layer system from before, and they have the same denominator:

\[ R = \frac{F_k}{F_0} = \frac{(\alpha_k k_{M_2} + \alpha_0 k_{M_1}) - i (M_{M_2} + \alpha k_k k_{M_1})}{\alpha k k_{M_2} + \alpha_0 k_{M_1} + i (M_{M_2} - \alpha_k k_{M_1})} \]

The pole is then given as:

\[ \alpha k k_{M_2} + \alpha_0 k_{M_1} + i (M_{M_2} - \alpha_k k_{M_1}) = 0 \]

With (because we have evanescent waves in substrate and cladding)

\[ k_{ax} = i\mu_a = i\sqrt{k_x^2 - \frac{\omega^2}{\epsilon_k} \epsilon_\alpha(\omega)}, \quad k_{cx} = i\mu_c = i\sqrt{k_x^2 - \frac{\omega^2}{\epsilon_k} \epsilon_\alpha(\omega)} \]

We can write the general dispersion relation in an arbitrary layer system

\[ M_{11,TE,TM} + \alpha_{ax} \mu_{ax} M_{12,TE,TM} + \frac{1}{\alpha_{cx} \mu_c} M_{21,TE,TM} + \alpha_{cx} M_{22,TE,TM} \]

Here, as usual, we have \( \alpha_{ax} = 1 \), \( \alpha_{cx} = \frac{1}{\epsilon} \)

In addition to the material dispersion in the layers

\[ k_x^2(\omega) = \frac{\omega^2}{\epsilon_k} \epsilon_\alpha(\omega) = k_{ax}^2 + k_{cx}^2 \]

we get the waveguide dispersion relation \( k_x(\omega, \text{geometry}) \)

The waveguide dispersion relation gives a discrete set of solutions, so-called waveguide modes. For given \( \epsilon, d, \omega \) we get \( k_x(\omega) \).

In the case of guided waves it is easy to compute the field (mode profile) inside the layer system:

- take \( k_x \) from dispersion relation
- in the substrate we have:

\[ F(x) = F \exp(i k_x x), \quad G(x) = \alpha \frac{\partial}{\partial x} \{ F \exp(i k_x x) \} \]

Hence, for given \( F(0) \) we get \( G(0) = \alpha k_F(0) \quad (F(0) \to \text{free parameter}) \)

\[ \begin{bmatrix} F \\ G \end{bmatrix} = \begin{bmatrix} M(x) & F \\ G & M(0) \end{bmatrix} \begin{bmatrix} 1 \\ \alpha k_F \end{bmatrix} F(0). \]

**Analogy of optics to the stationary Schrödinger equation in QM**

**Optics** (e.g. TE-polarization) **Quantum mechanics**

The guided waves ↔ discrete energy eigenvalues

\( k_x^2 > \frac{\omega^2}{c^2} \epsilon_{ax} \) \( \leftrightarrow \quad E < V_{\text{ext}} \)

Tunnel effect

\( k_x^2 > \frac{\omega^2}{c^2} \epsilon_{\text{film}} \) \( \leftrightarrow \quad E < V_{\text{barriere}} \)

**7.4.3 Guided waves at interface - surface polariton**

Let us first have a look at the most simple case where the guiding layer structure is just an interface.

Our condition for guided waves ist that on both sides of the interface we have evanescent waves: \( k_x^2 > \frac{\omega^2}{c^2} \epsilon_{ax} \), because

\[ \mu_{ax} = \sqrt{k_x^2 - \frac{\omega^2}{c^2} \epsilon_{ax}} > 0 \]

The general dispersion relation we derived before reads

\[ M_{11,TE,TM} + \alpha_{ax} M_{12,TE,TM} + \frac{1}{\alpha_{cx} \mu_c} M_{21,TE,TM} + \alpha_{cx} M_{22,TE,TM} \]

and with the matrix for a single interface:

\[ \tilde{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

\[ \begin{bmatrix} d^2 \\ dx^2 + \frac{\omega^2}{c^2} \epsilon(x) \end{bmatrix} E(x) = k_F^2 E(x) \quad \leftrightarrow \quad \begin{bmatrix} d^2 \\ dx^2 - \frac{2m}{\hbar^2} \end{bmatrix} \psi(x) = \frac{2m}{\hbar^2} E(x) \]

guided waves ↔ discrete energy eigenvalues

\( k_x^2 > \frac{\omega^2}{c^2} \epsilon_{ax} \max \{ \epsilon_{ax} \} \) \( \leftrightarrow \quad E < V_{\text{ext}} \)

Tunnel effect

\( k_x^2 > \frac{\omega^2}{c^2} \epsilon_{\text{film}} \) \( \leftrightarrow \quad E < V_{\text{barriere}} \)
we get the dispersion relation
\[ 1 + \frac{\alpha_s H_s}{\alpha_c H_c} = 0 \]

A) **TE-polarization** (\( \alpha = 1 \))
\[ \mu_s + \mu_e = 0 \rightarrow \text{no solution because } \mu_e, \mu_s > 0 \]

B) **TM-polarization** (\( \alpha = 1/\varepsilon \))
\[ \frac{\mu_s + \mu_e}{\varepsilon_c + \varepsilon_s} = 0 \]
with \( \mu_e, \mu_s > 0 \cap \varepsilon_c \cdot \varepsilon_s < 0 \),
→ one of the media has to have negative \( \varepsilon \) (dielectric near resonance or metal)

In dielectrics \( \omega_{0(T)} < \omega < \omega_L \) we can find surface-phonon-polaritons.
In metals \( \omega < \omega_p \) we can find surface-plasmon-polaritons.

**Remark**
Surface polaritons occur in TM polarization only, similar to the phenomenon of Brewster-angle (no reflection for \( \frac{k_{\omega}}{\varepsilon_c} = \frac{k_s}{\varepsilon_s} = 0 \)).

Let us now compute the explicit dispersion relation for surface polaritons.
W.o.l.g. we assume \( \varepsilon_c(\omega) < 0 \), and because \( \varepsilon_c(\omega) \) is near a resonance it will show a much stronger \( \omega \) dependence than \( \varepsilon_s \).

\[ (\mu_s, \varepsilon_s)^2 = (\mu_c, \varepsilon_c)^2 \]
\[ \varepsilon_s^2(\omega) \left\{ k_s^2 - \frac{\omega^2}{c^2} \varepsilon_s^2 \right\} = \varepsilon_c^2 \left\{ k_c^2 - \frac{\omega^2}{c^2} \varepsilon_c(\omega) \right\} \]
\[ k_s(\omega) = \frac{\omega}{c} \sqrt{\varepsilon_s(\omega) \varepsilon_c} \]

There is a second condition for existence of surface polaritons: \( \varepsilon_s + \varepsilon_s(\omega) < 0 \)

**Conclusion:**

Surface polaritons may have very small effective wavelengths in z-direction:
\[ \varepsilon_c = 1, |\varepsilon_s(\omega)| \approx 1 \rightarrow k_s(\omega) \approx \frac{2\pi}{\lambda} \sqrt{\varepsilon_s(\omega) \varepsilon_c} = \frac{2\pi}{\lambda_{eff}} \rightarrow \lambda_{eff} \ll \lambda \]

**7.4.4 Guided waves in a layer – film waveguide**
The prototype of a waveguide is the film waveguide, where the waveguide consists of one guiding layer with \( \frac{\omega^2}{c^2} \varepsilon_s(\omega) > k_s^2(\omega) \).
Such film waveguides are the basis of integrated optics. Typical parameters are:

\[ d \approx \text{a few } \alpha \cdot d \approx \]
\[ \Delta \varepsilon \approx 10^{-3} - 10^{-4} \]

Fabrication of film waveguides can be achieved by coating, diffusion or ion implantation. The matrix of a single layer (film) is given as:

\[
M^{TE,TM} = \begin{pmatrix}
\cos(k_x d) & -\frac{1}{\mu_{\text{ex}}} \sin(k_x d) \\
-\frac{1}{\mu_{\text{ex}}} \sin(k_x d) & \cos(k_x d)
\end{pmatrix}
\]

From this matrix we can compute the dispersion relation for guided modes:

\[
M_{11} + \alpha_{\text{ex}} \mu_{\text{ex}} M_{12} + \frac{1}{\alpha_{\text{ex}} \mu_{\text{ex}}} M_{21} + \frac{\alpha_{\text{ex}} \mu_{\text{ex}}}{\alpha_{\text{ex}} \mu_{\text{ex}}} M_{22} = 0
\]

\[
\cos(k_x d) + \frac{\alpha_{\text{ex}} \mu_{\text{ex}}}{\alpha_{\text{ex}} \mu_{\text{ex}}} \sin(k_x d) - \frac{\alpha_{\text{ex}} k_x}{\alpha_{\text{ex}} \mu_{\text{ex}}} \sin(k_x d) + \frac{\alpha_{\text{ex}} \mu_{\text{ex}}}{\alpha_{\text{ex}} \mu_{\text{ex}}} \cos(k_x d) = 0
\]

\[
\sin(k_x d) \frac{\cos(k_x d)}{\cos(k_x d)} = \tan(k_x d) = \frac{1 + \alpha_{\text{ex}} \mu_{\text{ex}}}{\alpha_{\text{ex}} \mu_{\text{ex}} - \alpha_{\text{ex}} \mu_{\text{ex}}} = \frac{\alpha_{\text{ex}} k_x}{\alpha_{\text{ex}} \mu_{\text{ex}} + \alpha_{\text{ex}} \mu_{\text{ex}}}
\]

Here: TE-Polarisation (\( \alpha = 1 \))

\[
\tan(k_x d) = \frac{k_x \left( \mu_{\text{ex}} + \mu_{\text{ex}} \right)}{k_x^2 - \mu_{\text{ex}}^2 \mu_{\text{ex}}}
\]

This waveguide dispersion relation is an implicit equation for \( k_x \). For given frequency \( \omega \) and thickness \( d \) we get several solutions with index \( k_x \).

Here is an example for fixed frequency \( \omega \), the effective index \( n_{\text{eff}} = k_x / \left( \frac{\omega}{c} \right) \) versus the thickness \( d \):
We can see in the figure that for large thickness \( d \) we have many modes. If we decrease \( d \), more and more modes vanish at a certain cut-off thickness.

**Definition of cut-off:**

A guided mode vanishes \( \rightarrow \) cut-off (here w.o.l.g. \( \varepsilon_c < \varepsilon_s \))

The idea of the cut-off is that a mode is not guided anymore. Guiding means evanescent fields in the substrate and cladding, so cut-off means

\[
\mu_s = \sqrt{k_s^2 - \omega^2/c^2 \varepsilon_s} = 0 \rightarrow \text{no guiding} \quad \Rightarrow \quad k_s^2 = \frac{\omega^2}{c^2} \varepsilon_s
\]

We can plug this cut-off condition in the DR:

\[
\tan (k_s d) = \frac{k_{fs}}{k_{fs}^2 - \mu_s^2} \mu_s^2
\]

\[
\tan \left( \frac{\omega}{c} \sqrt{\varepsilon_r - \varepsilon_s} d \right) = \frac{\sqrt{\varepsilon_r - \varepsilon_s} \sqrt{\varepsilon_s}}{\varepsilon_r - \varepsilon_s} \approx \frac{\varepsilon_s}{\varepsilon_r - \varepsilon_s}
\]

\[
(\omega d)_{co}^{TE} = \frac{c}{\sqrt{\varepsilon_r - \varepsilon_s}} \left\{ \arctan \left( \frac{\varepsilon_s}{\varepsilon_r - \varepsilon_s} + \nu \pi \right) \right\}
\]

\[
(\omega d)_{co}^{TM} = \frac{c}{\sqrt{\varepsilon_r - \varepsilon_s}} \left\{ \arctan \left( a + \nu \pi \right) \right\}
\]

with parameter of asymmetry \( a: \varepsilon_s \approx \varepsilon_c \rightarrow a \rightarrow 0 \)

\( \varepsilon_s \approx \varepsilon_c \rightarrow a \approx \infty \)

\[
\rightarrow \text{we can define a cut-off frequency} \quad \text{for} \quad k_s (\omega) \quad \text{when we keep} \quad d \quad \text{fix}
\]

\[
\rightarrow \text{we can define a cut-off thickness} \quad \text{for} \quad k_s (d) \quad \text{when we keep} \quad \omega \quad \text{fix}
\]

In a symmetric waveguide the fundamental mode \( (\nu = 0) \) has cut-off \( \neq 0 \).

If we plot the dispersion curves for each mode we get a graphical representation of the dispersion relation:

7.4.5 how to excite guided waves

Finally, we want to address the question how we can excite guided waves. In principle, there are two possibilities, we can adapt the field profile or the wave vector \( (k_s) \):

A) adaption of field \( \rightarrow \) front face coupling

Then, inside the waveguide we have (without radiative modes):

\[
E(x, z) = \sum a_v E_v(x) \exp(ik_{sv} z)
\]

\[
\Rightarrow E(x, 0) = \sum a_v E_v(x) \quad \| E_v(x) \|
\]

with:

\[
P_v = \frac{k_{sv}}{2 \omega \mu_0} \int_0^\infty |E_v(x)|^2 dx
\]

\[
a_v = \frac{k_{sv}}{2 \omega \mu_0 P_v} \int_0^\infty E_{sv}(x) E_v(x) dx,
\]

\( \rightarrow \) mode \( \nu \) couples to the incident field \( E_{in} (0) \) with amplitude \( a_v \).

\( \rightarrow \) Gauss-beam couples very good to the fundamental mode

B) adaption of wave vector \( \rightarrow \) coupling through the interface
We know that $k_z$ is continuous at interface. The condition for the existence of guided modes is

$$k_z > \frac{\omega}{c} \sqrt{\varepsilon_{ss}}$$

but dispersion relation for waves in bulk media dictates

$$k_z = \sqrt{\frac{\omega^2}{c^2} \varepsilon_{ss} - k_{zz}^2} < \frac{\omega}{c} \sqrt{\varepsilon_{ss}}$$

We got a problem!

There are two solutions:

i) coupling by prism

we bring a medium with $\varepsilon_p > \varepsilon_t$ (prism) near the waveguide

$$\sim k_z < \frac{\omega}{c} \sqrt{\varepsilon_t} \sim k_z < \frac{\omega}{c} \sqrt{\varepsilon_p} \sim k_{zs} = \sqrt{\frac{\omega^2}{c^2} \varepsilon_p - k_z^2} > 0.$$  

ii) coupling by grating

grating on waveguide (modulated thickness of layer $d$):

$$d(z) = d + \zeta(z)$$

$$\zeta(z) = A \sin(\frac{g z}{\varepsilon}) \quad \text{mit} \quad g = \frac{2\pi}{P} \cdot \text{p-per.}$$

coupling works for $m$'th diffraction order:

$$k_{zp} = k_z + mg$$

$$= \frac{\omega}{c} n_i \sin \phi + mg.$$  

→ light can couple to the waveguide via optical tunneling: ATR (‘attenuated total reflection’)

i) coupling by grating