Grundkonzepte der Optik
Sommersemester 2014

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

Table of content

0. Introduction ................................................................. 4
1. Ray optics - geometrical optics ........................................ 15
  1.1 Introduction ......................................................... 15
  1.2 Postulates .............................................................. 15
  1.3 Simple rules for propagation of light ......................... 16
  1.4 Simple optical components ..................................... 16
  1.5 Ray tracing in inhomogeneous media (graded-index - GRIN optics) ........................................... 20
  1.5.1 Ray equation .................................................... 20
  1.5.2 The eikonal equation .......................................... 22
  1.6 Matrix optics ........................................................ 22
  1.6.1 The ray-transfer-matrix ....................................... 23
  1.6.2 Matrices of optical elements ............................... 23
  1.6.3 Cascaded elements ............................................ 24
2. Optical fields in dispersive and isotropic media ................ 25
  2.1 Maxwell’s equations ............................................. 25
  2.1.1 Adaption to optics ............................................ 25
  2.1.2 Temporal dependence of the fields ....................... 28
  2.1.3 Maxwell’s equations in Fourier domain ................. 29
  2.1.4 From Maxwell’s equations to the wave equation ...... 29
  2.1.5 Decoupling of the vectorial wave equation ............. 31
  2.2 Optical properties of matter .................................... 32
  2.2.1 Basics .............................................................. 32
  2.2.2 Dielectric polarization and susceptibility ............... 35
  2.2.3 Conductive current and conductivity ..................... 37
  2.2.4 The generalized complex dielectric function .......... 38
  2.2.5 Material models in time domain ........................... 42
  2.3 The Poynting vector and energy balance .................... 44
  2.3.1 Time averaged Poynting vector ........................... 44
  2.3.2 Time averaged energy balance ......................... 45
  2.4 Normal modes in homogeneous isotropic media .......... 49
  2.4.1 Transversal waves ............................................ 50
  2.4.2 Longitudinal waves .......................................... 51
  2.4.3 Plane wave solutions in different frequency regimes ... 52
  2.4.4 Time averaged Poynting vector of plane waves ....... 58
  2.5 The Kramers-Kronig relation ................................... 58
  2.6 Beams and pulses - analogy of diffraction and dispersion ... 61
  2.7 Diffraction of monochromatic beams in homogeneous isotropic media ........................................ 63
  2.7.1 Arbitrarily narrow beams (general case) ................. 64
  2.7.2 Fresnel- (paraxial) approximation ....................... 70
  2.7.3 The paraxial wave equation .............................. 75
  2.8 Propagation of Gaussian beams ............................... 76
  2.8.1 Propagation in paraxial approximation ................... 77
  2.8.2 Propagation of Gauss beams with q-parameter formalism ... 82
  2.8.3 Gaussian optics ............................................... 82
  2.8.4 Gaussian modes in a resonator ............................ 85
3. Diffraction theory ..................................................... 91
  3.1 Interaction with plane masks ..................................... 95
  3.2 Propagation using different approximations ................ 106
  3.2.1 The general case - small aperture ........................ 106
  3.2.2 Fresnel approximation (paraxial approximation) ....... 106
  3.2.3 Paraxial Fraunhofer approximation (far field approximation) ........................................ 107
  3.2.4 Non-paraxial Fraunhofer approximation ................ 109
  3.3 Fraunhofer diffraction at plane masks (paraxial) .......... 114
  3.4 Remarks on Fresnel diffraction ............................... 114
4. Fourier optics - optical filtering .................................. 116
  4.1 Imaging of arbitrary optical field with thin lens .......... 116
  4.1.1 Transfer function of a thin lens ........................... 116
  4.1.2 Optical imaging using the 2f-setup ....................... 117
  4.2 Optical filtering and image processing ...................... 119
  4.2.1 The 4f-setup ................................................... 119
  4.2.2 Examples of aperture functions .......................... 122
  4.2.3 Optical resolution ............................................ 123
5. The polarization of electromagnetic waves ........................ 126
  5.1 Introduction ......................................................... 126
  5.2 Polarization of normal modes in isotropic media .......... 126
  5.3 Polarization states .................................................. 127
6. Principles of optics in crystals ..................................... 129
  6.1 Susceptibility and dielectric tensor .......................... 129
  6.2 The optical classification of crystals ....................... 131
  6.3 The index ellipsoid ............................................... 132
  6.4 Normal modes in anisotropic media .......................... 133
  6.4.1 Normal modes propagating in principal directions ....... 134
  6.4.2 Normal modes for arbitrary propagation direction .... 135
  6.4.3 Normal surfaces of normal modes ....................... 140
  6.4.4 Special case: uniaxial crystals ............................ 142
7. Optical fields in isotropic, dispersive and piecewise homogeneous media .......................... 145
  7.1 Basics ................................................................. 145
  7.1.1 Definition of the problem .................................... 146
  7.1.2 Decoupling of the vectorial wave equation ............. 146
  7.1.3 Interfaces and symmetries ................................... 147
  7.1.4 Transition conditions ........................................ 148
  7.2 Fields in a layer system → matrix method ................. 148
  7.2.1 Fields in one homogeneous layer ......................... 148
  7.2.2 The fields in a system of layers .......................... 150
  7.3 Reflection – transmission problem for layer systems ....... 152
  7.3.1 General layer systems ....................................... 152
  7.3.2 Single interface ............................................... 159
  7.3.3 Periodic multi-layer systems - Bragg-mirrors - 1D photonic crystals ................................. 166
  7.3.4 Fabry-Perot-resonators ...................................... 173
  7.4 Guided waves in layer systems ................................. 179
  7.4.1 Field structure of guided waves .......................... 179
  7.4.2 Dispersion relation for guided waves .................... 180
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) What is light?
- electromagnetic wave \( (c = 3 \times 10^8 \text{ m/s}) \)
- amplitude and phase → complex description
- polarization, coherence

### Spectrum of Electromagnetic Radiation

<table>
<thead>
<tr>
<th>Region</th>
<th>Wavelength [nm]</th>
<th>Wavelength [nm (nm=10^-9m)]</th>
<th>Frequency [Hz] (THz=10^12Hz)</th>
<th>Energy [eV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radio</td>
<td>&gt; 10^9</td>
<td>&gt; 10^-1</td>
<td>&lt; 3 \times 10^9</td>
<td>&lt; 10^5</td>
</tr>
<tr>
<td>Microwave</td>
<td>10^8 - 10^5</td>
<td>10^-1 - 10^-4</td>
<td>3 \times 10^9 - 3 \times 10^12</td>
<td>10^5 - 0.01</td>
</tr>
<tr>
<td>Infrared</td>
<td>10^6 - 700</td>
<td>7 \times 10^-7 - 4 \times 10^-7</td>
<td>4.3 \times 10^14 - 7.5 \times 10^14</td>
<td>0.01 - 2</td>
</tr>
<tr>
<td>Visible</td>
<td>700 - 400</td>
<td>4 \times 10^-7 - 4 \times 10^-7</td>
<td>4.3 \times 10^14 - 7.5 \times 10^14</td>
<td>2 - 3</td>
</tr>
<tr>
<td>Ultraviolet</td>
<td>400 - 1</td>
<td>4 \times 10^-7 - 10^-9</td>
<td>7.5 \times 10^14 - 3 \times 10^17</td>
<td>3 - 10^1</td>
</tr>
<tr>
<td>X-Rays</td>
<td>1 - 0.01</td>
<td>10^-9 - 10^-11</td>
<td>3 \times 10^17 - 3 \times 10^19</td>
<td>10^2 - 10^5</td>
</tr>
<tr>
<td>Gamma Rays</td>
<td>&lt; 0.01</td>
<td>&lt; 10^-11</td>
<td>&gt; 3 \times 10^19</td>
<td>&gt; 10^5</td>
</tr>
</tbody>
</table>
B) 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

C) Propagation of light through matter
- light-matter interaction (G: Licht-Materie-Wechselwirkung)
  - dispersion → diffraction → absorption → scattering
    - ↓ frequency → spatial → center of → wavelength
      - spectrum → frequency → spectrum

- matter is the medium of propagation → the properties of the medium (natural or artificial) determine the propagation of light
- light is the means to study the matter (spectroscopy) → measurement methods (interferometer)
- design media with desired properties: glasses, polymers, semiconductors, compounded media (effective media, photonic crystals, meta-materials)
D) **Light can modify matter**
- light induces physical, chemical and biological processes
- used for lithography, material processing, or modification of biological objects (bio-photonics)

E) **Optical telecommunication**
- transmitting data (Terabit/s in one fiber) over transatlantic distances

*Hole "drilled" with a fs laser at Institute of Applied Physics, FSU Jena.*

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

Figure 1. Rotation of an intracellular object inside *Elodea densa* 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) Light sensors and light sources

- new light sources to reduce energy consumption
- new projection techniques
H) Micro- and nano-optics
- ultra small camera

I) Relativistic optics

Figure 3. Two relativistic lasers: (a) Helios circa 1980 at LBNL was the first relativistic laser with \( \Delta \nu \sim 1 \) at a millihertz repetition rate. (Courtesy of LBNL) (b) The \( \lambda^3 \) laser at the University of Michigan is has an \( \Delta \nu \sim 1 \) at a kilohertz repetition rate.

Figure 4. Relativistic rectification in plasma: (a) a high-intensity pulse before it enters the plasma and (b) the \( \Delta \nu \) 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.

Insect inspired camera system develop at Fraunhofer Institute IOF Jena
J) Schematic of optics

- geometrical optics
  - $\lambda \ll$ size of objects $\rightarrow$ daily experience
  - 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

K) Literature

- Fundamental
     in German: "Grundlagen der Photonik" Wiley (2008)
  2. Hecht, 'Optic', Addison-Wesley (2001)
     in German: "Optik", Oldenbourg (2005)
  7. Sommerfeld, 'Optik'

- Advanced
  1. W. Silvast, 'Laser Fundamentals',
  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

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 $n(r) < 0$)

\[ n = \frac{c}{c_n} \quad c_n \text{ - speed of light in the medium} \]

C) optical path length \( \sim \) delay
   i) homogeneous media
   ii) inhomogeneous media

\[ \int n(r) ds \]

D) Fermat’s principle

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

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

1.3 Simple rules for propagation of light
A) Homogeneous media
   - $n = \text{const.} \to \text{minimum delay = minimum distance}$
   - 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 $\to$ reflected ray plus refracted ray
   - The reflected ray obeys b).
   - The refracted ray lies in the plane of incidence.

\[ \frac{n_1 \sin \theta_1}{n_2 \sin \theta_2} = \text{constant} \]

- 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 $\rightarrow$ caustic
- Parallel, paraxial rays converge to the focal point $f = (-R)/2$
- Convention: $R < 0$ - concave mirror; $R > 0$ - convex mirror.
- 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 $P_2$

\[ \frac{1}{z_1} + \frac{1}{z_2} = \frac{2}{(-R)} \] (imaging formula)

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

C) Spherical interface (paraxial)
- Paraxial imaging
\[ \theta_2 = \frac{n_1}{n_2} \theta_1 - \frac{n_2 - n_1}{n_2} \frac{y}{R} \] (*)

\[ \frac{n_1}{z_1} - \frac{n_2}{z_2} \approx \frac{n_2 - n_1}{R} \] (imaging formula)

\[ m = -\frac{n_1}{n_2} \frac{z_1}{z_2} \] (magnification)

(Proof: exercise)
- if paraxiality is violated \( \rightarrow \) aberration
- rays coming from one point of the object do not intersect in one point of the image (caustic)

D) Spherical thin lens (paraxial)

- two spherical interfaces \((R_1, R_2, \Delta)\) apply (*) two times and assume \(y=\text{const (}\Delta\text{ small)}\)

\[ \theta_2 = \theta_1 - \frac{y}{f} \] 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)

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

computation:

\[ L = \int n[r(s)] ds \]

variation of the path: \(r(s) + \delta r(s)\)
\[ \delta L = \int_a^b \delta n \, ds + \int_a^b n \delta ds \]

\[ \delta n = \text{grad} \cdot \delta \mathbf{r} \]

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

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

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

\[ \delta L = \int_a^b \left( \text{grad} \cdot \delta \mathbf{r} + n \frac{dr}{ds} \frac{d\delta r}{ds} \right) \, ds \]

\[ = \int_a^b \left( \text{grad} n - \frac{d}{ds} \left( n \frac{dr}{ds} \right) \right) \, d\mathbf{r} ds \]

\[ \delta L = 0 \] for arbitrary variation

\[ \text{grad} n = \frac{d}{ds} \left( n \frac{dr}{ds} \right) \] ray equation

Possible solutions:

A) trajectory

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

- solve for \( x(z), y(z) \)

- paraxial rays \( \rightarrow \) \( ds \approx dz \)

\[ \frac{d}{dz} \left[ n(x, y, z) \frac{dx}{dz} \right] \approx \frac{dn}{dx} \]

\[ \frac{d}{dz} \left[ n(x, y, z) \frac{dy}{dz} \right] \approx \frac{dn}{dy} \]

B) homogeneous media

- straight lines

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

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

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

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

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

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

for \( n(y)-n_0\ll1: \)

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

1.5.2 The eikonal equation

- bridge between geometrical optics and wave

- eikonal \( S(r) = \) constant \( \rightarrow \) planes perpendicular to rays

- from \( S(r) \) we can determine direction of rays \( \sim \) \( \text{grad} S(r) \) (like potential)

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

Remark: it is possible to derive Fermat’s principle from eikonal equation

- geometrical optics: Fermat’s or eikonal equation

\[ S(r_0) = S(r_\alpha) = \int_{r_0}^{r_\alpha} \text{grad} S(r) \, ds = \int_{r_0}^{r_\alpha} n(r) \, ds \]

eikonal \( \rightarrow \) optical path length \( \sim \) 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 \( \rightarrow \) \( 2 \times 2 \) 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

in paraxial approximation:

\[ y_2 = A y_1 + B \theta_1 \]
\[ \theta_2 = C y_1 + D \theta_1 \]

A=0: same \( y_1 \) \( \rightarrow \) 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

\[
M = \begin{bmatrix} 1 & 0 \\ 0 & n_1/n_2 \end{bmatrix}
\]

C) refraction on spherical interface

\[
M = \begin{bmatrix} 1 & 0 \\ 0 & n_1/n_2 \end{bmatrix}
\]

\[ n_1 n_2 \]

D) thin lens

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

E) reflection on planar mirror

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

F) reflection on spherical mirror (compare to lens)

\[
M = \begin{bmatrix} 1 & 0 \\ 2/R & 1 \end{bmatrix}
\]

1.6.3 Cascaded elements

\[
\begin{bmatrix} y_{N+1} \\ \theta_{N+1} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} \rightarrow M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} M = M_N \ldots M_2 M
\]
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 which is 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 electromagnetic 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.

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

- electric field (G: elektrisches Feld) \( \mathbf{E}(r,t) \) [V/m]
- magnetic flux density (magnetic induction) \( \mathbf{B}(r,t) \) [Vs/m²] or [tesla]
- electric flux density (electric displacement field) \( \mathbf{D}(r,t) \) [As/m²]
- magnetic field (G: magnetisches Feld) \( \mathbf{H}(r,t) \) [A/m]
- external charge density \( \rho_{\text{ext}}(r,t) \) [As/m³]
- macroscopic current density \( \mathbf{j}_{\text{macro}}(r,t) \) [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 \( \mathbf{D}(r,t) \) and the magnetic field \( \mathbf{H}(r,t) \). These two fields are related to the electric field \( \mathbf{E}(r,t) \) and magnetic flux density \( \mathbf{B}(r,t) \) by two other new fields.

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

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

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 \( \mathbf{D}(r,t) \) and the
magnetic field $\mathbf{H}(r,t)$ to the electric field $\mathbf{E}(r,t)$ and the magnetic flux density $\mathbf{B}(r,t)$. This is achieved by modeling the material properties by introducing the material equations.

- The effect of the medium gives rise to polarization $\mathbf{P}(r,t) = f[\mathbf{E}]$ and magnetization $\mathbf{M}(r,t) = f[\mathbf{B}]$. In order to solve Maxwell’s equations we need material models describing these quantities.

- In optics, we generally deal with non-magnetizable media $\mathbf{M}(r,t) = 0$ (exceptions are metamaterials with $\mathbf{M}(r,t) \neq 0$).

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 (G: Dichte freier Ladungsträger)
  $\rho_{\text{ext}}(r,t) \text{[As/m}^3\text{]}$

- macroscopic current density (G: makroskopische Stromdichte)
  $\mathbf{j}_{\text{macro}}(r,t) = \mathbf{j}_{\text{cond}}(r,t) + \mathbf{j}_{\text{conv}}(r,t) \text{[A/m}^2\text{]}$

- conductive current density (G: Konduktionsstromdichte)
  $\mathbf{j}_{\text{cond}}(r,t) = f[\mathbf{E}]$

- convective current density (G: Konvektionsstromdichte)
  $\mathbf{j}_{\text{conv}}(r,t) = \rho_{\text{ext}}(r,t) \mathbf{v}(r,t)$

In optics, we generally have no free charges which change at speeds comparable to the frequency of light:

$\rho_{\text{ext}}(r,t) = 0 \rightarrow \mathbf{j}_{\text{conv}}(r,t) = 0$

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

  $\text{rot} \mathbf{E}(r,t) = -\mu_0 \frac{\partial \mathbf{H}(r,t)}{\partial t} - \varepsilon_0 \partial_t \mathbf{E}(r,t) = -\text{div} \mathbf{P}(r,t)$

  $\text{rot} \mathbf{H}(r,t) = \mathbf{j}(r,t) + \frac{\partial \mathbf{P}(r,t)}{\partial t} + \varepsilon_0 \frac{\partial \mathbf{E}(r,t)}{\partial t} \quad \text{div} \mathbf{H}(r,t) = 0$

  - In optics, the medium (or more precisely the mathematical material model) determines the dependence of the polarization on the electric field $\mathbf{P}(\mathbf{E})$ and the dependence of the (conductive) current density on the electric field $\mathbf{j}(\mathbf{E})$.
  - 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 (G: Dichte gebundener Ladungsträger)
  $\rho_{b}(r,t) = -\text{div} \mathbf{P}(r,t)$

- and a bound current density (G: Stromdichte gebundener Ladungsträger)
  $\mathbf{j}_{b}(r,t) = \frac{\partial \mathbf{P}(r,t)}{\partial t}$

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

Complex field formalism (G: komplexer Feld-Formalismus):

- 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 $\cos(x)$ and $\sin(x)$.

- convention in this lecture
  - real physical field: $\mathbf{E}_{r}(r,t)$
  - complex mathematical representation: $\mathbf{E}(r,t)$

- They are related by
  $\mathbf{E}_{r}(r,t) = \frac{1}{2} \left[ \mathbf{E}(r,t) + \mathbf{E}^*(r,t) \right] = \text{Re}[\mathbf{E}(r,t)]$

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

- This means in general: For calculations we use the complex fields $[\mathbf{E}(r,t)]$ 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! This is 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
B) polychromatic light ↔ 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)

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

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

Remark: The position of the sign in the exponent and the factor \(1/2\pi\) can be defined differently in different textbooks.

### 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:

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

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

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

\[
\begin{align*}
\nabla \times E(r,\omega) &= i\omega \mu_0 H(r,\omega) \\
\nabla \times H(r,\omega) &= j(r,\omega) - i\omega \varepsilon_0 P(r,\omega) - i\omega \varepsilon_0 E(r,\omega) \\
\n\nabla \cdot \varepsilon E(r,\omega) &= -\nabla \cdot \mu J(r,\omega)
\end{align*}
\]

### B) Frequency domain derivation

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

\[
\begin{align*}
\nabla \times \mathcal{E}(r,\omega) &= -\omega^2 \mu_0 \mathcal{J}(r,\omega) \\
\n\nabla \times \mathcal{H}(r,\omega) &= j(r,\omega) - i\omega \varepsilon_0 \mathcal{P}(r,\omega) - i\omega \varepsilon_0 \mathcal{E}(r,\omega)
\end{align*}
\]

and

\[
\nabla \cdot \varepsilon_0 \mathcal{E}(r,\omega) + \mu_0 \mathcal{P}(r,\omega) = 0
\]

- magnetic field:

#### A) Time domain derivation

We start from applying the curl operator \(\nabla \times\) a second time on \(\nabla \times E(r,t)\) and substitute \(\nabla \times H\) with the other Maxwell equation

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

And find the wave equation for the electric field

\[
\begin{align*}
\nabla \times E(r,t) + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathcal{E}(r,t) &= -\mu_0 \frac{\partial}{\partial t} \mathcal{J}(r,t) - \mu_0 \frac{\partial^2 \mathcal{P}(r,t)}{\partial t^2}
\end{align*}
\]

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:

\[
\nabla \cdot \varepsilon_0 E(r,t) + \mu_0 J(r,t) = 0
\]

Once we have solved the wave equation, we know the electric field. From that we can easily compute the magnetic field:

\[
\frac{\partial}{\partial t} \mathcal{H}(r,t) = -\frac{1}{\mu_0} \nabla \times E(r,t)
\]

Remarks:

- An analog procedure is possible for \(\mathcal{H}\), i.e., we can derive a wave equation for the magnetic field.
- Generally, the wave equation for \(\mathcal{E}\) is more convenient, because the material model defines \(\mathcal{P}(\mathcal{E})\).
- However, for inhomogeneous media \(\mathcal{H}\) can be the better choice for the numerical solution of the partial differential equation since it forms a hermitian operator.
- analog procedure possible for \(\mathcal{H} \rightarrow \mathcal{E}\)
- generally, wave equation for \(\mathcal{E}\) is more convenient, because \(\mathcal{P}(\mathcal{E})\) given
- for inhomogeneous media \(\mathcal{H}\) can be better choice
\[ \mathbf{H}(r, \omega) = - \frac{i}{\omega \mu_0} \text{rot} \mathbf{E}(r, \omega) \]

- 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

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

### 2.1.5 Decoupling of the vectorial wave equation

So far we have seen that for the general problem of electromagnetic waves all 3 field components of the electric or the magnetic field are coupled. Hence we have to solve a vectorial wave equation for the general problem. However, it would be desirable to express problems also by scalar equation since they are much easier to solve. For problems with **translational invariance** in at least one direction, as e.g. for homogeneous infinite media, layers or interfaces, this can be achieved since the vectorial components of the fields can be decoupled.

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 (\( \partial / \partial y = 0 \)) and the operators in the wave equation simplify.

\[
\text{rot rot } \mathbf{E} = \text{grad div } \mathbf{E} - \Delta \mathbf{E} = \begin{bmatrix}
0 \\
\frac{\partial}{\partial z} (\frac{\partial E_z}{\partial y} + \frac{\partial E_y}{\partial z}) & \frac{\partial}{\partial x} (\frac{\partial E_x}{\partial y} + \frac{\partial E_y}{\partial x}) & \frac{\partial}{\partial y} (\frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial y}) \\
0 & \frac{\partial}{\partial z} (\frac{\partial E_z}{\partial x} + \frac{\partial E_x}{\partial z}) & \frac{\partial}{\partial x} (\frac{\partial E_x}{\partial x} + \frac{\partial E_x}{\partial x}) \\
0 & 0 & \frac{\partial}{\partial y} (\frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial y})
\end{bmatrix} - \begin{bmatrix}
\Delta^{(2)} E_z \\
\Delta^{(2)} E_x \\
\Delta^{(2)} E_y
\end{bmatrix}
\]

The decoupling becomes visible when the three components of the general vectorial field are decomposed in the following way.

- **decomposition of electric field**
  
  \[ \mathbf{E} = \mathbf{E}_\perp + \mathbf{E}_\parallel \Rightarrow \mathbf{E}_\perp = \begin{pmatrix} 0 \\ \mathbf{E}_z \\ 0 \end{pmatrix}, \quad \mathbf{E}_\parallel = \begin{pmatrix} \mathbf{E}_x \\ 0 \\ \mathbf{E}_y \end{pmatrix} \]
  
  with Nabla operator \( \nabla^{(2)} = \begin{bmatrix} \partial / \partial x \\ 0 \\ \partial / \partial z \end{bmatrix} \) and Laplace \( \Delta^{(2)} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \)

Hence we obtain two wave equations for the \( \mathbf{E}_\perp \) and \( \mathbf{E}_\parallel \) fields.

- **gives two decoupled wave equations**

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

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

These two wave equations are independent as long as the material response, which is expressed by \( \mathbf{j} \) and \( \mathbf{P} \), does not couple the respective field components by some anisotropic response.

**Properties**

- propagation of perpendicularly polarized fields \( \mathbf{E}_\perp \) and \( \mathbf{E}_\parallel \) can be treated separately
- propagation of \( \mathbf{E}_\parallel \) is described by scalar equation
- similarly the other field components can be described by a scalar equation for \( \mathbf{H} \)
- 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 we will use the same model but without restoring force, leading eventually to a current density. 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 \( \mathbf{P}(\mathbf{E}) \) and \( \mathbf{j}(\mathbf{E}) \). 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
- In time domain: we introduce the response function $(G: \text{Responsfunktion})$
  \[ P(t) = \varepsilon_0 \sum_{j=x,y,z} R_{ij}(t-t')E_j(t')dt' \]
  with $R$ being a 2nd rank tensor $i=x,y,z$ and summing over $j=x,y,z$

- In frequency domain: we introduce the susceptibility $(G: \text{Suszeptibilität})$
  \[ P(\omega) = \varepsilon_0 \sum_{j=x,y,z} \chi_{ij}(\omega)E_j(\omega) \]
  response function and susceptibility are linked via Fourier transform (convolution theorem)
  \[ R_{ij}(t) = \frac{1}{2\pi} \int \chi_{ij}(\omega) \exp(-i\omega t) d\omega \]

- Obviously, things look friendlier in frequency domain. Using the wave equation from before and assuming that there are no currents ($j = 0$) we find

  \[
  \begin{align*}
  \text{rotrot } & P(t) - \frac{\omega^2}{c^2} P(t) = \mu_0 \omega^2 P(t) \\
  \text{or } \quad & \Delta P(t) + \frac{\omega^2}{c^2} P(t) - \text{graddive}(t) = -\mu_0 \omega^2 P(t) \\
  \text{and for auxiliary fields } \quad & D(t) = \varepsilon_0 E(t) + P(t)
  \end{align*}
  \]

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

### Simplification of the wave equation for different types of media

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

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

$\chi_{ij}(r,\omega) \rightarrow \chi$ is a scalar constant

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

B) linear, homogeneous, isotropic, dispersive media $\Rightarrow \chi(\omega)$

- Maxwell: $\text{div } D = 0 \rightarrow \text{div } E(\omega) = 0$ for $\varepsilon(\omega) \neq 0$

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

Helmholtz equation

- This description is sufficient for many materials.

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

- Maxwell: $\text{div } D(\omega) = 0 \land \text{div } E(\omega) = 0$ for $\varepsilon(\omega) \neq 0$

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

- This description is sufficient for many materials.
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 model material. Because of the large mass of the ions ($10^3 \times$ 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 $\mathbf{E}(r, t)$ gives rise to a displacement $\mathbf{s}(r, t)$ of charged particles from their equilibrium positions. In the easiest approach this can be modeled by a driven harmonic oscillator:

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

- resonance frequency (electronic transition) $\Rightarrow \omega_0$
- damping $\Rightarrow g$
- charge $\Rightarrow q$
- mass $\Rightarrow m$

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

$$p(r, t) = q \mathbf{s}(r, t),$$

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

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

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

$$\frac{\partial^2}{\partial t^2} \mathbf{P}(r, t) + g \frac{\partial}{\partial t} \mathbf{P}(r, t) + \omega_0^2 \mathbf{P}(r, t) = \frac{q^2 N}{m} \mathbf{E}(r, t) = e_0 f / \mathbf{E}(r, t)$$

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

This equation is easy to solve in Fourier domain:

$$-\omega^2 \tilde{\mathbf{P}}(r, \omega) - i g \omega \tilde{\mathbf{P}}(r, \omega) + \omega_0^2 \tilde{\mathbf{P}}(r, \omega) = \frac{e_0 f}{\omega_0^2 - \omega^2 - i g \omega} \mathbf{E}(r, \omega)$$

$$\Rightarrow \tilde{\mathbf{P}}(r, \omega) = \frac{e_0 f}{\omega_0^2 - \omega^2 - i g \omega} \mathbf{E}(r, \omega)$$

with $\tilde{\mathbf{P}}(r, \omega) = e_0 \chi(\omega) \mathbf{E}(r, \omega) \Rightarrow \chi(\omega) = \frac{f}{\omega_0^2 - \omega^2 - i g \omega}$

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

$$\mathbf{P}(r, \omega) = e_0 \sum_j \left[ \frac{f_j}{(\omega_0^2 - \omega_j^2) - i g \omega_j} \right] \mathbf{E}(r, \omega) = e_0 \chi(\omega) \mathbf{E}(r, \omega)$$
Again we solve this equation in Fourier domain:
\[-\omega_0^2 \mathbf{j}(\mathbf{r}, \omega) + g \mathbf{j}(\mathbf{r}, \omega) = \varepsilon_0 \varepsilon_p^2 \mathbf{E}(\mathbf{r}, \omega)\]
\[\rightarrow \mathbf{j}(\mathbf{r}, \omega) = \frac{\varepsilon_0 \omega_0^2}{g - 1 \omega^2} \mathbf{E}(\mathbf{r}, \omega) = \sigma(\omega) \mathbf{E}(\mathbf{r}, \omega).\]

Here we introduced the complex frequency dependent conductivity
\[
\sigma(\omega) = \frac{\varepsilon_0 \omega_0^2}{g - 1 \omega^2} = -\frac{1}{g - 1 \omega^2} \frac{\varepsilon_0 \omega_0^2}{\omega^2 - 1 g \omega^0}
\]

**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:
\[\varepsilon_0 \nabla \cdot \mathbf{E} = \rho(\mathbf{r}, t)\]

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):
\[\frac{\partial}{\partial t} \mathbf{j} + g \mathbf{j} = \varepsilon_0 \omega_p^2 \mathbf{E}(\mathbf{r}, t)\]

Now we apply divergence operator and plug in from above (red terms):
\[\nabla \cdot \frac{\partial}{\partial t} \mathbf{j} + g \nabla \cdot \mathbf{j} = \varepsilon_0 \omega_p^2 \nabla \cdot \mathbf{E}(\mathbf{r}, t) = \omega_p^2 \rho(\mathbf{r}, t)\]

With the continuity equation for the charge density (from Maxwell's equations)
\[\frac{\partial}{\partial t} \rho + \nabla \cdot \mathbf{j} = 0,\]

We can substitute the divergence of the current density and find:
\[-\frac{\partial}{\partial t} \rho - \frac{g}{\omega_p^2} \rho = \omega_p^2 \rho\]
\[\frac{\partial^2}{\partial t^2} \rho + \frac{g}{\omega_p^2} \rho + \omega_p^2 = 0 \rightarrow \text{harmonic oscillator equation}\]

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 \( \mathbf{j}(\mathbf{r}, \omega) \) and \( \mathbf{P}(\mathbf{r}, \omega) \) into the wave equation (in Fourier domain)
rotrot \( \mathbf{E}(\mathbf{r}, \omega) - \frac{\omega^2}{c^2} \mathbf{E}(\mathbf{r}, \omega) = \mu_0 \omega^2 \mathbf{P}(\mathbf{r}, \omega) + i \omega \mu_0 \mathbf{j}(\mathbf{r}, \omega) \)

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

Hence we can collect all terms proportional to \( \omega \) and write

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

Here, we introduced the generalized complex dielectric function

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

So, in general we have

\[ \varepsilon(\omega) = 1 + \sum_j \left\{ \frac{f_j}{(\omega_j^2 - \omega^2) - i \omega \gamma_j} \right\} + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i \gamma_0 \omega} \]

because (from before)

\[ \chi(\omega) = \sum_j \left\{ \frac{f_j}{(\omega_j^2 - \omega^2) - i \omega \gamma_j} \right\}, \quad \sigma(\omega) = -i \frac{\varepsilon_0 \omega_0^2}{\omega_0^2 - \omega^2 - i \gamma_0 \omega} \]

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

**Some special cases for materials in the infrared and visible spectral range:**

**A) Dielectrics (insulators) in the infrared (IR) spectral range near phonon resonance**

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

\[ \varepsilon(\omega) = 1 + \sum_j \left\{ \frac{f_j}{(\omega_j^2 - \omega^2) - i \omega \gamma_j} \right\} + \frac{f}{(\omega_0^2 - \omega^2) - i \omega \gamma_0} \]

with \( \omega_0 \ll \omega_j \) and \( \omega \sim \omega_0 \)

\[ \Rightarrow \varepsilon(\omega) = \varepsilon_\infty + \frac{f}{(\omega_0^2 - \omega^2) - i \omega \gamma_0} \]

The contribution from electronic transitions shows almost no frequency dependence (dispersion) in this frequency range far away from the electronic resonances. Hence it can be expressed together with the vacuum contribution as a constant \( \varepsilon_\infty \).

Let us study the real and the imaginary part of the resulting \( \varepsilon(\omega) \) separately:

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

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

\[ \varepsilon'(\omega) = \varepsilon_\infty + \frac{f}{(\omega_0^2 - \omega^2) + i \omega \gamma_0} \]

\[ \varepsilon''(\omega) = \frac{gf}{(\omega_0^2 - \omega^2)^2 + g^2 \omega^2} \rightarrow \text{Lorentz curve} \]

**properties:**

- resonance frequency: \( \omega_0 \)
- width of resonance peak: \( g \)
- static dielectric constant in the limit \( \omega \rightarrow 0 \): \( \varepsilon_\infty = \varepsilon_\infty + \frac{f}{\omega_0^2} \)
- so called longitudinal frequency \( \omega_L : \varepsilon'(\omega) = \omega_L = 0 \)
- \( \varepsilon''(\omega) \neq 0 \): absorption and dispersion appear always together
- near resonance we find \( \varepsilon''(\omega) < 0 \) (damping, i.e. decay of field, without absorption if \( \varepsilon'' \approx 0 \))
- frequency range with normal dispersion: \( \partial \varepsilon'(\omega) / \partial \omega > 0 \)
- frequency range with anomalous dispersion: \( \partial \varepsilon'(\omega) / \partial \omega < 0 \)

Simplified example: sharp resonance for undamped oscillator \( g \rightarrow 0 \)
relation between resonance frequency $\omega_0$ and longitudinal frequency $\omega_L$ (Lyddane-Sachs-Teller relation)

$$\epsilon'(\omega_0) = \epsilon + \frac{f}{(\omega_0^2 - \omega_L^2)} = 0, f = (\epsilon - \epsilon') \omega_0^2 \text{ (from above)}$$

$$\omega_L = \sqrt{\frac{\epsilon}{\epsilon'}}$$

B) Dielectrics in the visible (VIS) spectral range

Dielectric media in visible (VIS) spectral range can be described by a so-called double resonance model, where a phonon resonance exists in the infrared (IR) and an electronic transition exists in the ultraviolet (UV).

$$\epsilon(\omega) = \epsilon_\infty + \frac{f_p}{(\omega_p^2 - \omega^2)} - \frac{f_e}{(\omega_e^2 - \omega^2)} \text{ with } \omega_p \ll \omega \ll \omega_e$$

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

![Graph of dielectric function](image)

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

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:

$$\overline{D}(r, \omega) = \epsilon_0 \epsilon(\omega) \overline{E}(r, \omega)$$

$$\overline{P}(r, \omega) = \epsilon_0 \chi(\omega) \overline{E}(r, \omega).$$

The response function (or Green's function) $R(t)$ in the time domain is then given by

$$R(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \chi(\omega) \exp(-i\omega t) d\omega \quad \chi(\omega) = \int_{-\infty}^{\infty} R(t) \exp(i\omega t) dt$$
To prove this, we can use the convolution theorem

\[ P(r,t) = \int_{-\infty}^{\infty} P(r,\omega) \exp(-i\omega t) d\omega = \varepsilon_{0} \int_{-\infty}^{\infty} \chi(\omega) E(r,\omega) \exp(-i\omega t) d\omega \]

\[ = \varepsilon_{0} \int_{0}^{\infty} \gamma(\omega) \frac{1}{2\pi} \int_{-\infty}^{\infty} E(r,t') \exp(i\omega t') dt' \exp(-i\omega t) d\omega \]

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

\[ = \varepsilon_{0} \int_{-\infty}^{\infty} R(t-t') E(r,t') dt' \]

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

\[ E(r,t') = \varepsilon_{0} \delta(t' - t_0) \rightarrow P(r,t) = \varepsilon_{0} R(t-t_0) e \rightarrow \text{Green’s function} \]

**Examples**

A) instantaneous media (unphysical simplification)

- For instantaneous (or non-dispersive) media, which cannot not really exist in nature, we would find:

\[ R(t) = \gamma \delta(t) \rightarrow P(r,t) = \varepsilon_{0} \gamma E(r,t) \quad \text{(unphysical!)} \]

B) dielectrics

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

- Using the residual theorem we find:

\[ R(t) = \begin{cases} \frac{\varepsilon_{0} \gamma}{\Omega} \exp \left( -\frac{t}{2\omega_{0}^{2}} \right) \sin \Omega t & t \geq 0 \\ 0 & t < 0 \end{cases} \]

with \( \Omega = \sqrt{\omega_{0}^{2} - \frac{g^{2}}{4}} \)

\[ P(r,t) = \frac{\varepsilon_{0} \gamma}{\Omega} \int \exp \left[ -\frac{g}{2} (t-t') \right] \sin \left[ \Omega(t-t') \right] E(r,t') dt' \]

C) metals

\[ R_{j}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sigma(\omega) \exp(-i\omega t) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\varepsilon_{0} \omega_{0}^{2}}{g - i\omega} \exp(-i\omega t) d\omega, \]

- Using again the residual theorem we find:

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

\[ j(r,t) = \varepsilon_{0} \omega_{0}^{2} \int \exp[-g(t-t')] E(r,t') dt' \]

**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 \( S(r,t) = E_{r}(r,t) \times H_{r}(r,t) \) 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: \( T_{0} = 2\pi / \omega_{0} \leq 10^{-14} \text{s} \)
- pulse duration: \( T_{p} \) in general \( T_{p} \gg T_{0} \)
- duration of measurement: \( T_{m} \) 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:

\[ \frac{1}{2} \int \left[ E(r,t) \exp(-i\omega t) + c.c. \right] = E_{r}(r,t) \]

For such pulses, the momentary Poynting vector reads:
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 $\sim \cos(2\omega_f t)$, $\sim \sin(2\omega_f t)$ 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 $\langle S(r,t) \rangle$.

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

The fast oscillating terms $\sim \cos 2\omega_f t$ and $\sim \sin 2\omega_f t$ 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.

$$\langle S(r,t) \rangle = \frac{1}{2} \int_{t-T_m/2}^{t+T_m/2} \Re[\mathbf{E}(r,t') \times \mathbf{H}'(r,t')] dt'$$

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).

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

$$\langle S(r,t) \rangle = \frac{1}{2} \Re[\mathbf{E}(r) \times \mathbf{H}'(r)]$$

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

$$I = |\langle S(r,t) \rangle| \rightarrow \text{measurement destroys phase information}$$

### 2.3.2 Time averaged energy balance

Let us motivate a little bit further the concept of the Poynting vector. Some interesting insight on the energy flow of light and hence also on the transport of information can be obtained from the Poynting theorem, which is 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):

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

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

$$\mathbf{j} \cdot \mathbf{rot} \mathbf{E} - \mathbf{E} \cdot \mathbf{rot} \mathbf{H} = \frac{\varepsilon_0}{2} \mathbf{E} \cdot \frac{\partial}{\partial t} \mathbf{E} + \mathbf{H} \cdot \frac{\partial}{\partial t} \mathbf{H}$$

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

$$\frac{1}{2} \mathbf{j} \cdot \mathbf{rot} \mathbf{E} + \frac{1}{2} \mathbf{H} \cdot \frac{\partial}{\partial t} \mathbf{H} + \mathbf{E} \cdot \frac{\partial}{\partial t} \mathbf{E} = -\mathbf{E} \cdot \left( \mathbf{j}_e + \frac{\partial}{\partial t} \mathbf{P} \right)$$

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 the Poynting’s theorem contains the so-called source terms.

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

In the case of stationary fields and isotropic media (simple but important)

$$\mathbf{E}_s(r,t) = \frac{1}{2} \mathbf{E}(r) \exp(-i\omega_f t) + c.c.$$  

$$\mathbf{H}_s(r,t) = \frac{1}{2} \mathbf{H}(r) \exp(-i\omega_f t) + c.c.$$  

Time averaging of the left hand side of Poynting’s theorem (*) yields:

$$\left\{ \frac{1}{2} \varepsilon_0 \frac{\partial}{\partial t} \mathbf{E}_s^2(r,t) + \frac{1}{2} \mu_0 \frac{\partial}{\partial t} |\mathbf{H}_s|^2 + \mathbf{E}_s \cdot \frac{\partial}{\partial t} \mathbf{E}_s \right\} = \frac{1}{2} \mathbf{E} \cdot \mathbf{rot} \mathbf{E}$$

Note that the time derivative removes stationary terms in $\mathbf{E}_s^2(r,t)$ and $|\mathbf{H}_s|^2$. Time averaging of the right hand side of Poynting’s theorem yields (source terms):

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

This equation has the form of a balance equation as above. 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 the Poynting’s theorem contains the so-called source terms.

$$u = \frac{1}{2} \varepsilon_0 |\mathbf{E}|^2 + \frac{1}{2} \mu_0 |\mathbf{H}|^2 \rightarrow \text{vacuum energy density}$$
\[-\left[ j_\omega (r,t) + \frac{\partial}{\partial t} P_\omega (r,t) \right] E_\omega (r,t) \right\]

\[ = -\frac{1}{4} \left( \sigma(\omega) E(r)e^{-i\omega t} - i\omega_\omega e\chi(\omega) E(r)e^{-i\omega t} c.c. \right) \left[ E(r)e^{i\omega t} + c.c. \right] \]

Now we use our generalized dielectric function:

\[ = -\frac{1}{4} \left[ -i\omega_\omega e\chi(\omega) + i \frac{\sigma(\omega)}{\omega_\omega e} \right] E(r) \exp(-i\omega t) + c.c. \left[ E(r) \exp(-i\omega t) + c.c. \right] \]

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

Again, all fast oscillating terms $\sim \exp(\pm 2i\omega t)$ cancel due to the time average.

Finally, splitting $\varepsilon(\omega)$ into real and imaginary part yields

\[ = \frac{1}{4} i\omega_\omega e\left[ \varepsilon'(\omega) - 1 + i\varepsilon''(\omega) \right] E(r)E(r)^* + c.c. = -\frac{1}{2} \omega_\omega e\varepsilon''(\omega) E(r)E(r)^*. \]

Hence, the divergence of the time averaged Poynting vector is related to the imaginary part of the generalized dielectric function:

\[ \rightarrow \text{div} \langle S \rangle = -\frac{1}{2} \omega_\omega e\varepsilon''(\omega) E(r)E(r)^*. \]

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.

\[ \int \text{div} \langle S \rangle dV = \int \langle S \rangle \cdot n dA \]
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 self-consistent solutions to the wave equation including the material response. It is convenient to use the generalized complex dielectric function for the derivative of the solution of the wave equation

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

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

\[ \nabla \times \nabla \times E(r, \omega) = \frac{\omega^2}{c^2} \varepsilon(\omega) E(r, \omega) \]

According to Maxwell the solutions have to fulfill additionally the divergence equation:

\[ \varepsilon_0 [1 + \chi(\omega)] \nabla \cdot E(r, \omega) = 0 \]

In general, this additional condition implies that the electric field is free of divergence:

\[ 1 + \chi(\omega) \neq 0 \Rightarrow \nabla \cdot E(r, \omega) = 0 \] (normal case)

Let us for a moment assume that we already know that we can find plane wave solutions of the following form in the frequency domain:

\[ E(r, \omega) = E(\omega) \exp(iKr), \quad k = \text{unknown complex wave-vector} \]

The corresponding stationary field in time domain is given by:

\[ \hat{E}(r, t) = E \exp \left[ i (Kr - \omega t) \right] \]

\[ \rightarrow \text{monochromatic plane wave} \rightarrow \text{normal mode} \]

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

Then, the divergence condition implies that those waves are transversal

\[ k \perp \hat{E}(\omega) \rightarrow \text{transverse wave} \]

If we split the complex wave vector into real and imaginary part \( k = k' + ik'' \), 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 \( (\sigma(\omega) = 0) \) we can find a second, exotic type of wave solutions: At \( \omega = \omega_L \rightarrow \varepsilon(\omega_L) = 0 \), so-called longitudinal waves \( k \parallel \hat{E}(\omega) \) appear.

2.4.1 Transversal waves

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

\[ \varepsilon_0 k(\omega) \nabla \cdot E(r, \omega) = 0 \Rightarrow \nabla \cdot E(r, \omega) = 0 \]

Then, the wave equation reduces to the Helmholtz equation:

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

Hence, we have three scalar equations for \( \hat{E}(r, \omega) \) (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:

\[ \hat{E}(r, \omega) = E(\omega) \exp(iKr) \]

Immediately we see that the wave is transversal:

\[ 0 = \nabla \cdot \hat{E}(r, \omega) = i k \cdot \hat{E}(r, \omega) \rightarrow k \perp \hat{E}(\omega) \]

Hence, we have to solve

\[ \begin{bmatrix} -k^2 + \frac{\omega^2}{c^2} \varepsilon(\omega) \end{bmatrix} E(\omega) = 0 \quad \text{and} \quad k \cdot \hat{E}(\omega) = 0. \]

which leads to the following dispersion relation

\[ k^2 = k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2} \varepsilon(\omega) \]

We see that the so-called wave-number \( k(\omega) = \frac{\omega}{c \sqrt{\varepsilon(\omega)}} \) 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, \( k = k' + ik'' \) is complex. Alternatively it is sometimes useful to introduce the complex refractive index (if \( k'' \neq 0 \)):

\[ n(\omega) = \frac{\omega}{c \sqrt{\varepsilon(\omega)}} = \frac{\omega}{c} n(\omega) = \frac{\omega}{c} \left[ \varepsilon(\omega) + 1k(\omega) \right] \]
However, instead of assuming that $\hat{n}(\omega)$ and $\sqrt{\varepsilon(\omega)}$ are just the same, one should clearly distinguish between the two. While $\varepsilon(\omega)$ is a property of the medium, $\hat{n}(\omega)$ is a property of a particular type of the electromagnetic field in the medium, i.e. a property of the infinitely extended monochromatic plane wave.

$$\hat{E}(r,\omega) = \hat{E}(\omega) \exp(\mathbf{i}kr)$$

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

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

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

### 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 $\varepsilon(\omega) = 0$ in dielectrics at the longitudinal frequency $\omega = \omega_L$. In this case, we cannot conclude that $\varepsilon(\omega) = 0$:

$$\rightarrow$$

As for the transversal waves we try the plane wave ansatz and assume $\mathbf{k}$ to be real.

$$\hat{E}(r,\omega) = \hat{E}(\omega) \exp(\mathbf{i}kr)$$

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

$$\mathbf{k} \times \left[ \mathbf{k} \times \hat{E}(r,\omega) \right] = 0$$

Now we decompose the electric field into transversal and longitudinal components with respect to the wave vector:

$$\hat{E}(r,\omega) = \hat{E}_\perp(\omega) \exp(\mathbf{i}kr) + \hat{E}_L(\omega) \exp(\mathbf{i}kr)$$

with, $\hat{E}_\perp(\omega) \perp \mathbf{k}$ and $\hat{E}_L(\omega) \parallel \mathbf{k}$

This decomposed field is inserted into the wave equation:

$$\mathbf{k} \times \left[ \mathbf{k} \times (\hat{E}_\perp(\omega) + \hat{E}_L(\omega)) \right] \exp(\mathbf{i}kr) = 0$$

$$\mathbf{k} \times \left[ \mathbf{k} \times \hat{E}_\perp(\omega) \right] \exp(\mathbf{i}kr) + \mathbf{k} \times \left[ \mathbf{k} \times \hat{E}_L(\omega) \right] \exp(\mathbf{i}kr) = 0$$

Since the cross product of $\mathbf{k}$ with the longitudinal field $\hat{E}_L(\omega)$ is trivially zero the remaining wave equation is:

$$k^2 \hat{E}_L = 0$$

Hence the transversal field $\hat{E}_\perp$ must vanish and the only remaining field component is the longitudinal field $\hat{E}_L(\omega)$:

$$\rightarrow \hat{E}(r,\omega) = \hat{E}_L(\omega) \exp(\mathbf{i}kr)$$

### 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 $k$ only. Thus, different solutions for the complex wave vector $\mathbf{k} = k' + \mathbf{i}k''$ are possible. In addition, the generalized dielectric function $\varepsilon(\omega)$ 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

$$k^2 = k_x'^2 - k'^2 + 2\mathbf{k}' \cdot \mathbf{k}'' = \frac{\omega^2}{c^2} \varepsilon'(\omega) = \frac{\omega^2}{c^2} n'(\omega) \quad \Rightarrow \quad k' \cdot k'' = 0$$

There are two possibilities to fulfill this condition, either $k'' = 0$ or $\mathbf{k} \perp \mathbf{k}''$.

#### A.1) Real valued wave-vector $k'' = 0$

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

$$k(\omega) = \frac{\omega}{c} n(\omega) = \frac{\omega}{c} \frac{2\pi}{\lambda} n(\omega)$$

- Because $k'' = 0$ these waves are homogeneous, i.e. planes of constant phase are parallel to the planes of constant amplitude. This is trivial, because the amplitude is constant.

**Example 1:** single resonance in dielectric material

- for lattice vibrations (phonons)
Now the imaginary part of $\varepsilon(\omega)$ is neglected, which mathematically corresponds to an undamped resonance

$$\varepsilon(\omega) = \varepsilon'(\omega) = \varepsilon_\infty + \frac{\varepsilon_1}{\omega_0 - \omega^2}$$

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

$$\varepsilon(\omega) = \varepsilon'(\omega) = 1 - \frac{\omega_0^2}{\omega^2}$$

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

$$k^2 = k'^2 - k''^2 = \frac{\omega_0^2}{c^2} \varepsilon(\omega)$$

and therefore $k''^2 = k'^2 - k^2$

- This means that

$$k''^2 \neq 0 \quad \text{and} \quad k'^2 > k^2$$

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 $k'^2 > k^2$, whereas the homogeneous waves of case A.1) ($k'' = 0$) obey $k'^2 = k^2$. If we plug our findings into the plane wave ansatz we get: for the evanescent waves:

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

- 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 $k' \perp k''$ 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 can't be physically justified normal modes of homogeneous space and can only exist in inhomogeneous space, where the exponential growth is suppressed, e.g. 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 ($\omega_0 < \omega < \omega_2$) or below the plasma frequency ($\omega < \omega_p$) in metals. Then the dispersion relation gives
As in the previous case A), the imaginary term has to vanish and \( k' k'' = 0 \). Again this can be achieved by two possibilities.

\[ B.1) \quad k' = 0 \]
\[ \rightarrow \quad k'' = \frac{\varepsilon''}{c^2} |\varepsilon'(\omega)| \rightarrow \bar{E}(r, \omega) \sim \exp(-k''r) \rightarrow \text{strong damping} \]

\[ B.2) \quad k' \perp k'' \rightarrow \text{evanescent waves} \]
\[ k^2 = k'^2 - k''^2 = -\frac{\varepsilon''}{c^2} |\varepsilon'(\omega)| \]
\[ k''^2 = \frac{\varepsilon''}{c^2} |\varepsilon'(\omega)| + k'^2. \]

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 \( k'' \). In particular, case B.1) \( (k' = 0) \) is included. Hence, we can conclude that for \( \varepsilon(\omega) = \varepsilon'(\omega) < 0 \) we find only evanescent waves!

\[ C) \text{Complex valued epsilon} \ \varepsilon(\omega) \]

This is the general case, which is relevant particularly near resonances. From our (optical) point of view only weak absorption is interesting. Therefore, in the following we will always assume \( \varepsilon'(\omega) \ll |\varepsilon'(\omega)| \). 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., \( k' \) and \( k'' \) are almost parallel. Then, it is convenient to use the complex refractive index

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

Since \( k' \) and \( k'' \) are almost parallel:

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

The dispersion relation in terms of the complex refractive index gives

\[ k^2 = k'^2 = \frac{\varepsilon'}{c^2} |\varepsilon'(\omega)| = \frac{\varepsilon'}{c^2} [n'(\omega) + i \kappa'(\omega)]^2 \]

Here we have

\[ \varepsilon(\omega) = \varepsilon'(\omega) + i \kappa(\omega) = n''(\omega) - \kappa''(\omega) + 2 i n'(\omega) \kappa(\omega), \]

and therefore

\[ \varepsilon'(\omega) = n''(\omega) - \kappa''(\omega) \]
\[ \varepsilon'(\omega) = 2 n'(\omega) \kappa(\omega) \]
\[ n''(\omega) = \frac{\varepsilon'}{2} \text{sgn}(\varepsilon') \sqrt{1 + \left(\frac{\varepsilon''}{\varepsilon'}\right)^2} + 1 \]
\[ \kappa''(\omega) = \frac{\varepsilon'}{2} \text{sgn}(\varepsilon') \sqrt{1 + \left(\frac{\varepsilon''}{\varepsilon'}\right)^2} - 1 \]

Two important limiting cases of quasi-homogeneous plane waves:

\[ C.1) \quad \varepsilon', \varepsilon'' > 0, \quad \varepsilon'' \ll \varepsilon', \quad \text{(Dielectric media)} \]
\[ n(\omega) \approx \sqrt{\varepsilon'(\omega)}, \quad \kappa(\omega) \approx \frac{1}{2} \frac{\varepsilon'(\omega)}{\sqrt{\varepsilon'(\omega)}} \]
In this regime propagation dominates \((n(\omega) \gg \kappa(\omega))\), and we have weak absorption:

\[ k'^2 - k''^2 = \frac{\omega^2}{c^2} \varepsilon'(\omega), \quad 2k' \cdot k'' = \frac{\omega^2}{c^2} \varepsilon''(\omega). \]

\[ |k'| = \frac{\omega}{c} n(\omega) \approx \frac{\omega}{c} \sqrt{\varepsilon'(\omega)}, \quad |k''| = \frac{\omega}{c} \kappa(\omega) \approx \frac{\omega}{c} \frac{\varepsilon''(\omega)}{2} \]

\[-k' \times k'' \approx |k'||k''|, \]

- \(k'\) and \(k''\) almost parallel \(\Rightarrow\) homogeneous waves

\(\Rightarrow\) in homogeneous, isotropic media, next to resonances, we find damped, homogeneous plane waves, \(k'\parallel k''\parallel e_k\), with \(e_k\) being the unit vector along \(k\).

\[ E(r,\omega) = E(\omega) \exp(i kr) = E(\omega) \exp \left( i \frac{\omega}{c} n(\omega)(e_k r) \right) \exp \left[ -\frac{\omega}{c} \kappa(\omega)(e_k r) \right]. \]

C.2) \(\varepsilon' < 0, \varepsilon'' > 0, \varepsilon'' \ll |\varepsilon'|\). \((\text{metals and dielectric media in so-called Reststrahl domain})\)

\[ n(\omega) \approx \frac{1}{2} \sqrt{\varepsilon'(\omega)}, \quad \kappa(\omega) \approx \sqrt{\varepsilon''(\omega)} \]

In this regime damping dominates \((n(\omega) \ll \kappa(\omega))\) and 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

\[ \langle S(r,t) \rangle = \frac{1}{2} \frac{1}{T} \int_{-T}^{T} \mathbf{E}(r,t) \times \mathbf{H}^*(r,t) \, dt, \]

For plane waves we find:

\[ E(r,t) = E(\omega) \exp (i k r - i \omega t) = E(\omega) \exp (i k' r - i \omega t) \]

\[ H(r,t) = \frac{1}{\omega \mu_0} k \times E(r,t) \]

assuming a stationary case \(E(t) = E(\omega) \exp(-i \omega t)\)

\[ \Rightarrow \langle S(r,t) \rangle = \frac{1}{2} \frac{k'}{\omega \mu_0} \exp[-2k' \cdot r] |E|^2 = \frac{1}{2} n^2 \frac{\varepsilon_0}{\mu_0} \exp \left[ -2 \frac{\omega}{c} \kappa(e_k \cdot r) \right] |E|^2 \]

with \(e_k\) being the unit vector along \(k'\) and \(e_{k''}\) being the unit vector along \(k''\).

### 2.5 The Kramers-Kronig relation

In the previous sections we have assumed a very simple model for the description of the material's response to the excitation by the electromagnetic field. This model was based on quite strong assumptions, like a single charge which is attached to a rigid lattice etc. Hence, one could imagine that more complex matter could give rise to arbitrarily complex response functions if adequate models would be used for its description. However we can show from basic laws of physics, that several properties are common to all possible response functions, as long as a linear response to the excitation is assumed.

These fundamental properties of the response function are formulated mathematically by the Kramers-Kronig relation. It is a 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.

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

\[ P_i'(r,t) = \mathcal{E}_0 \int_{-\infty}^{t} R(t-t') E_i'(r,t') \, dt' \leftrightarrow P_i'(r,t) = \varepsilon_0 \int_{0}^{t} R'(\tau) E_i'(r,t-\tau) \, d\tau \]

Reality of the response function implies:
\[
R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \chi(\omega)e^{i\omega\tau} = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \chi^*(\omega)e^{-i\omega\tau} \quad \rightarrow \quad \chi(\omega) = \chi^*(-\omega)
\]

Causality of the response function implies:
\[
R(\tau) = \delta(\tau)v(\tau) \quad \text{with} \quad \begin{cases} 
1 & \text{for } \tau > 0 \\
0 & \text{for } \tau < 0 \\
\frac{1}{2} & \text{for } \tau = 0
\end{cases} \quad \text{Heaviside distribution}
\]

In the following, we will make use of the Fourier transform of Heaviside distribution:
\[
2\pi \delta(\omega) = \int_{-\infty}^{\infty} dt \delta(t)e^{i\omega t} = P \frac{i}{\omega} + \pi \delta(\omega) \quad \text{defined as integral only}
\]

In Fourier space, the Heaviside distribution consists of the Dirac delta distribution
\[
\int_{-\infty}^{\infty} d\omega \delta(\omega - \omega_0) f(\omega) = f(\omega_0) \quad \text{Dirac delta distribution}
\]

and the expression \( P(\delta/\omega) \) involving a Cauchy principal value:
\[
P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} f(\omega) = \lim_{\alpha \to 0} \int_{-\infty}^{-\alpha} d\omega \frac{i}{\omega} f(\omega) + \int_{\alpha}^{\infty} d\omega \frac{i}{\omega} f(\omega) \quad \text{Cauchy principle value}
\]

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:
\[
\chi(\omega) = \int_{-\infty}^{\infty} d\tau R(\tau) e^{i\omega\tau} = \int_{-\infty}^{\infty} d\tau \delta(\tau)v(\tau)e^{i\omega\tau} \\
= \int_{-\infty}^{\infty} d\omega \delta(\omega - \omega_0) \frac{\chi(\omega)}{\omega - \omega_0} \\
= \frac{1}{2\pi} P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} \frac{\chi(\omega)}{\omega - \omega_0} \quad (\star)
\]

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) \( y(-\tau) = y(\tau) \quad \text{even function} \)

b) \( y(-\tau) = -y(\tau) \quad \text{odd function} \)

In this case \( y(-\tau) = y(\tau) \) is a real valued and even function. We can exploit this property and show that
\[
\rightarrow \chi^*(\omega) = -\frac{1}{2\pi} P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} \frac{\chi(\omega)}{\omega - \omega_0} + \frac{\chi(\omega)}{\omega - \omega_0} = \chi(\omega)
\]

Hence, we can conclude from equation \((\star)\) above that
\[
\rightarrow \chi^*(\omega) = -\frac{1}{2\pi} P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} \frac{\chi(\omega)}{\omega - \omega_0} + \frac{\chi(\omega)}{\omega - \omega_0} = \chi(\omega)
\]

Here \( P \int \) is a so called principal value integral (G: Hauptwertintegral).

Now we have expressions for \( \chi(\omega), \chi^*(\omega) \) and can compute real and imaginary part of the susceptibility:
\[
\chi(\omega) + \chi^*(\omega) = \frac{1}{2\pi} P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} \frac{\chi(\omega)}{\omega - \omega_0} + \frac{\chi(\omega)}{\omega - \omega_0} - \frac{1}{2\pi} P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} \frac{\chi(\omega)}{\omega - \omega_0} + \frac{\chi(\omega)}{\omega - \omega_0} = \chi(\omega)
\]

Knowledge of the real part of the susceptibility (dispersion) allows us to compute the imaginary part (absorption).

b) \( y(-\tau) = -y(\tau) \quad \text{odd function} \)

The second K-K relation can be found by a similar procedure when we assume that \( y(-\tau) = -y(\tau) \) is a real odd function. We can show that in this case
\[
\rightarrow \chi^*(\omega) = -\frac{1}{2\pi} P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} \frac{\chi(\omega)}{\omega - \omega_0} + \frac{\chi(\omega)}{\omega - \omega_0} = -\chi(\omega)
\]

With equation \((\star)\) we then find that
\[
\rightarrow \chi^*(\omega) = -\frac{1}{2\pi} P \int_{-\infty}^{\infty} d\omega \frac{i}{\omega} \frac{\chi(\omega)}{\omega - \omega_0} + \frac{\chi(\omega)}{\omega - \omega_0} \quad (\text{see } (\star)) \quad \text{and}
\]

Again we can then compute real and imaginary part of the susceptibility.
\[ \chi(\omega) - \chi^*(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\mathcal{Y}(\omega) + \mathcal{Y}(\omega) - \mathcal{Y}(\omega) + \mathcal{Y}(\omega)}{\omega - \omega} d\omega \]

\[ \chi(\omega) + \chi^*(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2 \mathcal{Y}(\omega)}{\omega - \omega} d\omega \]

and finally obtain

\[ \mathcal{R} \chi(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{2 \mathcal{Y}(\omega)}{\omega - \omega} d\omega \]

2. K-K relation

The second Kramers-Kronig relation allows us to compute the real part of the susceptibility (dispersion) when we know its imaginary part (absorption).

The Kramers-Kronig relation can also be rewritten in terms of the dielectric function, where one applies also the symmetry relation for \( \omega \):

K-K relation for \( \varepsilon \):

\[ \chi(\omega) = \chi^*(\omega) \rightarrow \chi'(\omega) = \chi'(\omega) \]

\[ \chi(\omega) = \varepsilon(\omega) - 1 = [\varepsilon'(\omega) - 1] + 1 \varepsilon^*(\omega) \]

\[ \varepsilon'(\omega) - 1 = \frac{2}{\pi} \int_{0}^{\infty} \frac{\mathcal{Y}^* (\omega)}{\omega^2 - \omega^2} d\omega \]

\[ \varepsilon'(\omega) = \frac{2}{\omega^2} \int_{0}^{\infty} \left[ \varepsilon'(\omega) - 1 \right] d\omega \]

- dispersion and absorption are linked, e.g., we can measure absorption and compute dispersion

Example:

\[ \varepsilon^*(\omega) \sim \delta(\omega - \omega_o) \rightarrow \varepsilon'(\omega) - 1 \sim \frac{\omega_o}{\omega^2 - \omega_o^2} \rightarrow \text{Drude-Lorentz model} \]

2.6 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 \( \rightarrow \) finite transverse width \( \rightarrow \) diffraction

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

\[ \rightarrow \]

\[ k \rightarrow 2_w \]

B) pulse \( \rightarrow \) finite duration \( \rightarrow \) dispersion

A pulse is a continuous superposition of stationary plane waves (normal modes) with different frequencies.
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
\[ E(r,t) = \int_{-\infty}^{\infty} \hat{E}(\omega) \exp\left[ \text{i} (\mathbf{k} \cdot \mathbf{r} - \omega t) \right] d\omega. \]

2.7 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 their 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:
- \( \varepsilon(\omega) = \text{const}. > 0, \) → optical transparent regime → normal modes are
stationary homogeneous and evanescent plane waves
- scalar approximation

\[ E(r,\omega) = E_y(r,\omega) e_y \rightarrow E_y'(r,\omega) \rightarrow u(r,\omega). \]
- exact for one-dimensional beams and linear polarization
- approximation in two-dimensional case
In homogeneous isotropic media we have to solve the Helmholtz equation
\[ \Delta E(r,\omega) + \frac{\omega^2}{c^2} \varepsilon(\omega) E(r,\omega) = 0. \]
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 \( k(\omega) \)). In the
following we often even omit the argument of the fixed frequency \( \omega \).

2.7.1 Arbitrarily narrow beams (general case)
Let us consider the following fundamental problem. We want to compute from
a given field distribution \( u(x,y,0) \) in the plane \( z=0 \) the complete field
\( u(x,y,z) \) in the half-space \( z > 0 \), where \( z \) 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
\[ u(r,\omega) = \int_{-\infty}^{\infty} U(\mathbf{k},\omega) \exp[\text{i} \mathbf{k}(\omega) \cdot \mathbf{r}] d^3k \]
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 \( \mathbf{k} \)). Naively, we could expect that we just constructed a
general solution to our problem, but the solution is not correct because of the
dispersion relation:
\[ k^2 = \alpha^2 + \beta^2 + \gamma^2 = \frac{\omega^2}{c^2} \varepsilon(\omega) \]

→ only two components of \( \mathbf{k} \) are independent, 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 \( z \)“:

\[
u(r) = \iint_{-\infty}^{\infty} U(\alpha, \beta, z) \exp \left[ \i (ax + \beta y) \right] d\alpha d\beta.
\]

In analogy to the frequency \( \omega \) we call \( \alpha, \beta \) spatial frequencies.

Now 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

\[
U(\alpha, \beta; z) = U_{\gamma}(\alpha, \beta) \exp \left[ \i \gamma(\alpha, \beta) z \right] + U_{\gamma}^* (\alpha, \beta) \exp \left[ -\i \gamma(\alpha, \beta) z \right],
\]

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

We can identify two types of solutions:

A) Homogeneous waves

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

B) Evanescent waves

\( \gamma^2 < 0, \) \( \rightarrow \) \( \alpha^2 + \beta^2 > k^2 \), i.e., \( k \) complex, because \( \gamma = k \), imaginary. Then, we have \( k = k + \i k'' \), with \( k' = \alpha e_i + \beta e_r \), and \( k'' = \gamma e_r \).

\( \rightarrow k' \perp k'' \rightarrow \) 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, this component of the solution must vanish \( U_{\gamma}(\alpha, \beta) = 0 \). In fact, we will see later that \( U_{\gamma}(\alpha, \beta) \) corresponds to backward running waves, i.e., light propagating in the opposite direction. We therefore find the solution:

\[
U(\alpha, \beta; z) = U_{\gamma}(\alpha, \beta) \exp \left[ \i \gamma(\alpha, \beta) z \right] = U(\alpha, \beta; 0) \exp \left[ \i \gamma(\alpha, \beta) z \right] = U_{\gamma}(\alpha, \beta) \exp \left[ \i \gamma(\alpha, \beta) z \right]
\]

Furthermore the following boundary condition holds:

\[ U(\alpha, \beta; 0) = U_{\gamma}(\alpha, \beta). \]

In spatial space, we can find the optical field for \( z > 0 \) by inverse Fourier transform:

\[
u(r) = \iint_{-\infty}^{\infty} U(\alpha, \beta; z) \exp \left[ \i (ax + \beta y) \right] d\alpha d\beta.
\]

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 the different normal modes according to their different spatial frequencies \( \alpha, \beta \).**

The initial spatial frequency spectrum or angular spectrum at \( z = 0 \) forms the initial condition of the initial value problem and follows from \( u_{\gamma}(x, y) = u(x, y, 0) \) by Fourier transform:

\[
U_{\gamma}(\alpha, \beta) = \left( \frac{1}{2\pi} \right)^2 \iint_{-\infty}^{\infty} u_{\gamma}(x, y) \exp \left[ -\i (ax + \beta y) \right] dxdy.
\]

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. For example \( \cos x \) gives the angle of the plane wave’s propagation direction with the \( x \)-axis.

**Scheme for calculation of beam diffraction**

We can formulate a general scheme to describe the diffraction of beams:

1. initial field: \( u_0(x, y) \)
2. initial spectrum: \( U_0(\alpha, \beta) \) by Fourier transform
3. propagation: by multiplication with \( \exp(\imath \gamma (\alpha, \beta) z) \)
4. new spectrum: \( U(\alpha, \beta; z) = U_0(\alpha, \beta) \exp(\imath \gamma (\alpha, \beta) z) \)
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} U(\alpha, \beta; z) \exp(\imath (\alpha x + \beta y) \) 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(\imath (\alpha x + \beta y + \gamma (\alpha, \beta) z) \) d\alpha d\beta.
\]

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

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

\[
\exp(\imath \gamma (\alpha, \beta) z) \rightarrow 1, \quad \arg(\exp(\imath \gamma (\alpha, \beta) z)) \neq 0
\]

\( \rightarrow \) Upon propagation the homogeneous waves are multiplied by the phase factor

\[
\exp(\imath \sqrt{k^2 - \alpha^2 - \beta^2} z)
\]

**B) evanescent waves** \( \rightarrow \alpha^2 + \beta^2 > k^2 \)

\[
\exp(\imath \gamma (\alpha, \beta) z) \rightarrow \exp(\imath \sqrt{\alpha^2 + \beta^2 - k^2} z), \quad \arg(\exp(\imath \gamma (\alpha, \beta) z)) = 0
\]

\( \rightarrow \) Upon propagation the evanescent waves are multiplied by an amplitude factor <1

\[
\exp(\imath \sqrt{\alpha^2 + \beta^2 - k^2} z) < 1
\]

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: Slit**

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

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

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

Obviously, \( H(\alpha, \beta; z) = \exp(\imath \gamma (\alpha, \beta) z) \) acts differently on homogeneous and evanescent waves:
- All spatial frequencies ($-\infty \to \infty$) are excited.
- Important spectral information is contained in the interval $2/\alpha = \pi$.
- Largest important spectral frequency for a structure with width $a$ is $2/2\alpha = \pi$.
- Evanescent waves appear for $k \alpha > \pi$.
- To represent the relevant information by homogeneous waves the following condition must be fulfilled: $2\pi/a < k = 2\pi/\lambda n \to a > \lambda/\gamma$.

**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 having length scales of $|\Delta x|, |\Delta y| > \lambda/n$ are transmitted over macroscopic distances. 

Homogeneous media act like a low-pass filter for light.

**Summary of beam propagation scheme**

$$u_0(x,y) \rightarrow U_0(\alpha, \beta) \rightarrow U(\alpha,\beta; z) = H(\alpha,\beta; z) U_0(\alpha, \beta) \rightarrow u(x,y,z)$$

with the transfer function $H(\alpha,\beta; z) = \exp[i\gamma(\alpha, \beta) z]$.

**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 have to get the same phase shift during the propagation

$$U(\alpha,\beta; z) = U_0(\alpha, \beta) \exp[i\gamma(\alpha, \beta) z] = U_0(\alpha, \beta) \exp[iCz]$$

$\gamma(\alpha, \beta)$ is the Fourier transform of $u_0(x,y)$. Since in general $\gamma(\alpha, \beta) \neq \text{const}$ the excitation $u_0(x,y)$ must have a shape such that its Fourier transform has only components where the transfer function is of equivalent value

$$U_0(\alpha, \beta) \equiv 0 \text{ only for } \gamma(\alpha, \beta) = \sqrt{k^2 - \alpha^2 - \beta^2} = C$$

It is straightforward to see that the excited spatial frequencies must lie on a circular ring in the $(\alpha, \beta)$ plane.

$$\alpha^2 + \beta^2 = \rho^2$$

For constant spectral amplitude on this ring the Fourier back-transform yields (see exercises):

$$u_0(x,y) = J_0(\rho r)$$

**2.7.2 Fresnel- (paraxial) approximation**

The beam propagation formalism developed in the previous chapter can be simplified for the important special case of a narrowband angular spectrum

$$U_0(\alpha, \beta) \equiv 0 \text{ for } \alpha^2 + \beta^2 \ll k^2.$$ 

In this situation the beam consists of plane waves having only small inclination with respect to the optical $z$-axis (paraxial (Fresnel) approximation). Then, we can simplify the expression for $\gamma(\alpha, \beta)$ by a Taylor expansion to:

$$\gamma(\alpha, \beta) = \sqrt{k^2 - \alpha^2 - \beta^2} \approx k \left(1 - \frac{\alpha^2 + \beta^2}{2k^2}\right) \approx k - \frac{\alpha^2 + \beta^2}{2k}$$
The resulting expression for the transfer function in Fresnel approximation reads:

\[ H = \exp(i\gamma(\alpha,\beta)z) \approx \exp(ikz)\exp\left(-\frac{1}{2} \frac{\alpha^2 + \beta^2}{k^2} z^2\right) = H_0(\alpha,\beta;z) \]

We can see that this \( H_0(\alpha,\beta;z) \) is always real valued. Hence it does not account for the physics of evanescent waves. However, if we remember that for the derivation of the \( H_0(\alpha,\beta;z) \) as an approximation of \( H(\alpha,\beta;z) \) we had assumed that the spatial frequency spectrum is narrow (paraxial waves) which had excluded the excitation of evanescent waves already from the beginning to justify the paraxial approximation.

The assumption of a narrow frequency spectrum corresponds to the requirement that all structural details \( |\Delta x|, |\Delta y| \) of the field distribution in the excitation plane (at \( z = 0 \)) must be much larger than the wavelength:

\[ |\Delta x|, |\Delta y| > 10\lambda / n \approx \lambda / n \]

This requirement applies also to the phase of the excitation. Hence it is not sufficient that only the structural details of the intensity have a large scale. The underlying phase of the excitation field must fulfill this condition as well.

The propagation of the spectrum in Fresnel approximation works in complete analogy to the general case. We just use the modified transfer function to describe the propagation:

\[ U_0(\alpha,\beta;z) = H_0(\alpha,\beta;z)U_0(\alpha,\beta) \]

Summary of Fresnel approximation

For a coarse initial field distribution \( u_0(x,y,z) \) the angular spectrum \( U_0(\alpha,\beta) \) is nonzero for \( \alpha^2 + \beta^2 \ll k^2 \) only. Then, only paraxial plane waves are relevant for transmitting information and the transfer function of homogeneous space can be approximated by \( H_0(\alpha,\beta;z) \).

Description in real space

It is also possible to formulate beam propagation in Fresnel (paraxial) approximation in position space:

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

\[ = \iint_{-\infty}^{\infty} H_\alpha(\alpha,\beta;z)U_0(\alpha,\beta) \exp\left[i(\alpha x + \beta y)\right] d\alpha d\beta \]

\[ = \iint_{-\infty}^{\infty} h_\alpha(x-x',y-y';z)u_0(x',y') dx'dy' \]

The spatial response function \( h_\alpha(x,y;z) \) follows from the convolution theorem and is the Fourier transform of \( H_\alpha(\alpha,\beta;z) \):

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

\[ = \left( \frac{1}{2\pi} \right)^2 \exp(i\kappa z) \iint_{-\infty}^{\infty} \exp\left[-\frac{1}{2} \frac{\alpha^2 + \beta^2}{k^2} z^2\right] \exp\left[i(\alpha x + \beta y)\right] d\alpha d\beta. \]

This Fourier integral can be solved and we find:

\[ h_\alpha(x,y;z) = \exp(i\kappa z) \left[ \frac{1}{2\pi} \frac{k}{2\pi} \exp\left[i \frac{k}{2\pi} \left(x^2 + y^2\right)\right] \right] = \frac{i\kappa}{2\pi} \exp\left[i \kappa z \left(1 + \frac{(x^2 + y^2)}{2z^2}\right)\right]. \]

The response function corresponds to a spherical wave in paraxial approximation. Similar to Huygens principle, where from each point in the object plane a spherical wave is emitted towards the image plane, here paraxial approximations of spherical waves are emitted.

To sum up, in position space paraxial beam propagation is given by:

\[ u_\alpha(x,y,z) = \frac{i\kappa}{2\pi} \exp(i\kappa z) \iint_{-\infty}^{\infty} u_0(x',y') \exp\left[i \frac{k}{2\pi} \left(x-x'\right)^2 + \left(y-y'\right)^2\right] dx'dy'. \]

Of course, the two descriptions in position space and in the spatial Fourier domain are completely equivalent.
The correspondence between real and frequency space

Relation between transfer and response function:

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

Transfer functions for homogeneous space

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

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

with \( k = k(\omega) = k_0 n(\omega) = \frac{\omega}{c} n(\omega) \)

Remark on the validity of the scalar approximation

\[ \mathbf{E}(\mathbf{r}, \omega) = \iint \mathbf{E}^\dagger(\alpha, \beta, \omega) k^2(\alpha, \beta, \gamma) \exp\left[i(\alpha x + \beta y + \gamma z)\right] d\alpha d\beta \]

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

A) One-dimensional beams

- translational invariance in y-direction: \( \beta = 0 \)
- and linear polarization in y-direction: \( \hat{E}_x \rightarrow U \)
\[ \rightarrow \text{scalar approximation is exact since divergence condition is strictly fulfilled} \]

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: \( \mathbf{E}_x = 0, \mathbf{E}_y \rightarrow U \)
\[ \rightarrow \text{divergence condition: } \beta \hat{E}_x + \gamma \hat{E}_z = 0 \]

\[ \mathbf{E}_z(\alpha, \beta, \omega) = -\frac{\beta}{\gamma} \hat{E}_y(\alpha, \beta, \omega) = -\frac{\beta}{\sqrt{k^2 - \alpha^2 - \beta^2}} \hat{E}_y(\alpha, \beta, \omega) \approx 0 \]
In paraxial approximation \((\alpha^2 + \beta^2 \ll k^2)\) the scalar approximation is automatically justified.

### 2.7.3 The paraxial wave equation

In paraxial approximation the propagated spectrum is given by

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

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

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

\[
U_\epsilon(\alpha, \beta; z) = \exp(\text{i} k z) V(\alpha, \beta; z) \rightarrow V(\alpha, \beta; z) = \exp\left(-\frac{\text{i} \alpha^2 + \beta^2}{2k} z\right)V_\alpha(\alpha, \beta).
\]

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

\[
\frac{\text{i}}{2k} \frac{\partial}{\partial z} V(\alpha, \beta; z) = \frac{1}{2k} \left(\alpha^2 + \beta^2\right) V(\alpha, \beta; z)
\]

Fourier transformation back to position space leads to the so-called paraxial wave equation:

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

\[
\rightarrow \frac{\text{i}}{2k} \frac{\partial}{\partial z} v(x, y, z) + \frac{1}{2k} \Delta(\alpha^2 + \beta^2) v(x, y, z) = 0
\]

This is the paraxial wave equation for inhomogeneous media (weak index contrast).

### 2.8 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.

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

\[
u_\epsilon(x, y, z) = v_\epsilon(x, y) = A \exp\left[-\frac{x^2}{w_x^2} + \frac{y^2}{w_y^2}\right] \exp\left[\text{i} \phi(x, y)\right].
\]

Here, we will restrict ourselves to rotational symmetry \(w_x^2 = w_y^2 = w_0^2\) and (initially) ‘flat’ phase \(\phi(x, y) = 0\), which corresponds to a beam in the focus. The Gaussian beam in the focal plane (flat phase) is characterized by
amplitude \( A \) and width \( w_0 \): \( u_0(x^2 + y^2) = A_0 \exp(-1) \). In practice, the so-called 'full width at half maximum' (FWHM) is often used instead of \( w_0 \).

\[
\left| u_0(x^2 + y^2) \right|^2 = \exp\left( -\frac{w_{\text{FWHM}}^2}{2w_0^2} \right) \approx \frac{1}{2} \\
-w_{\text{FWHM}}^2 = -\ln(2) \rightarrow w_{\text{FWHM}}^2 = 2\ln(2)w_0^2 \approx 1.386w_0^2
\]

**2.8.1 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 \( z = 0 \):

\[
U_0(\alpha, \beta) = V_0(\alpha, \beta) = \frac{1}{(2\pi)} A_0 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left( -\frac{x^2 + y^2}{w_0^2} \right) \exp[-i(\alpha x + \beta y)] \, dx \, dy \\
= A_0 \frac{w_0^2}{4\pi} \exp \left( -\frac{\alpha^2 + \beta^2}{4/w_0^2} \right) = A_0 \frac{w_0^2}{4\pi} \exp \left( -\frac{\alpha^2 + \beta^2}{w_0^2} \right).
\]

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 \( w_x \times w_\beta = 2 \)

\[
\text{Angular spectrum in the focal plane}
\]

C) Check if paraxial approximation is fulfilled:

We can say that \( U_0(\alpha, \beta) \approx 0 \) for \( \left( \alpha^2 + \beta^2 \right) \geq 16/w_0^2 \), because \( \exp(-4) \approx 0.02 \). For paraxial approximation we need \( k^2 \gg \left( \alpha^2 + \beta^2 \right) \)

\[
k^2 \gg 16/w_0^2
\]

\[
\approx \frac{16}{w_0^2} \Rightarrow \frac{2\lambda}{\pi n} \approx \left( \frac{\lambda}{n} \right)^2
\]

\[
\Rightarrow \text{paraxial approximation works for } w_0 \gg 10\frac{\lambda}{n} = 10\lambda_n
\]

D) Propagation of the angular spectrum:

\[
U(\alpha, \beta; z) = V(\alpha, \beta; z) \exp(\pm ikz) \\
= A_0 \frac{w_0^2}{4\pi} \exp \left[ -\frac{w_0^2}{4} \left( \frac{\alpha^2 + \beta^2}{2k} \right) \right] \exp(-i(\alpha x + \beta y))
\]

E) Fourier back-transformation to position space

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

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

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" \( L_\text{d} = 2z_0 \), a measure for the "focus depth" of the Gaussian beam. E.g.: \( w_0 \gg 10\lambda_n \Rightarrow L_\text{d} \gg 600\lambda_n \).
From our computation above we know that the Gaussian beam evolves like:

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

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

\[ v(x, y, z) = A_b \frac{1 - \frac{iz}{z_0}}{1 + \left( \frac{z}{z_0} \right)^2} \exp \left[ -\frac{x^2 + y^2}{w_0^2 \left( 1 + \left( \frac{z}{z_0} \right)^2 \right)} \right] \exp \left[ \frac{4k}{2} \frac{1}{z} \left( \frac{x^2 + y^2}{z^2} \right) \right] \exp \left[ i \phi(z) \right]. \]

Here we used that \( w_0^2 = 2z_0 / k \). The \( (x, y) \)-independent phase \( \phi(z) \) is given by \( \tan \phi = -z / z_0 \), 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:

\[ v(x, y, z) = A(z) \exp \left[ \frac{x^2 + y^2}{w^2(z)} \right] \exp \left[ \frac{4k}{2} \frac{1}{R(z)} \right] \exp \left[ i \phi(z) \right]. \]

**Discussion**

The amplitude is given as:

\[ A(z) = A_b \frac{1}{\sqrt{1 + \left( \frac{z}{z_0} \right)^2}} = A_b \frac{1}{\sqrt{1 + \frac{2z^2}{w_0^2}}} \]

Hence, we get for the Intensity profile \( I \sim A^2 \):

\[ w(z) = w_0 \frac{1 + \left( \frac{z}{z_0} \right)^2}{1 + \left( \frac{2z}{L_b} \right)^2}. \]

The normalized beam intensity \( I / I_0 \) at different axial distances: (a) \( z = 0 \); (b) \( z = z_0 \); (c) \( z = 2z_0 \).

The on-axis intensity \( (x=y=0) \) evolves like

The normalized beam intensity \( I / I_0 \) at points on the beam axis \( (\rho = 0) \) as a function of \( z \).

The beam width evolves like:

\[ W(z) = W_0 \sqrt{1 + \left( \frac{z}{z_0} \right)^2} = W_0 \sqrt{1 + \frac{2z^2}{L_b^2}}. \]

The normalized beam intensity \( I / I_0 \) as a function of the radial distance \( \rho \) at different axial distances: (a) \( z = 0 \); (b) \( z = z_0 \); (c) \( z = 2z_0 \).

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

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

The radius of curvature \(R_z\) of the wavefronts of a Gaussian beam. The dashed line is the radius of curvature of a spherical wave.

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

$$\tan \phi = -\frac{z}{z_0} = -\frac{2z}{L_\phi}$$

The Gouy phase is not important for many applications because it is ‘flat’. However, in resonators and in the context of nonlinear optics it can play an important role (i.e., harmonic generation in focused geometries). The wavefronts (planes of constant phase) of a Gaussian beam are given by

$$\Phi(x, y, z) = \left\{ k \left[ z + \frac{x^2 + y^2}{2R(z)} \right] + \phi(z) \right\} = \text{const.}$$

### 2.8.2 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

$$q(z) = z - iz_0$$

allows an even simpler computation of the evolution of a Gaussian beam. In fact, if we take the inverse of the “q-parameter”,

$$\frac{1}{q(z)} = \frac{1}{z - iz_0} = -\frac{z}{z^2 + z_0^2} + \frac{1}{z^2 + z_0^2} = \frac{1}{z(1 + \frac{i}{z})} + \frac{1}{z_0(1 + \frac{i}{z_0})}$$

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

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

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

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

A) initial conditions: \(\frac{1}{q(0)} = \frac{1}{R(0)} + \frac{1}{\lambda_\phi \pi w^2(0)}\)

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

$$\frac{1}{q(d)} = \frac{1}{q(0) + d} = \frac{1}{R(d)} + \frac{1}{\lambda_\phi \pi w^2(d)}$$

### 2.8.3 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 \(R_n, w_n\) (i.e. \(q_n\)) \(\rightarrow R_n, w_n\) (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: $q(z_i) \rightarrow q_i$.

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:

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

B) propagation through multiple elements:

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

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

$$\begin{pmatrix} y_2 \\ \Theta_2 \end{pmatrix} = \mathbf{\tilde{M}} \begin{pmatrix} y_1 \\ \Theta_1 \end{pmatrix}.$$  

**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":

$$z_i = \frac{y_i}{\Theta_i} \Rightarrow z_2 = \frac{y_2}{\Theta_2} = \frac{A y_1 + B \Theta_1}{C y_1 + D \Theta_1} = \frac{A z_1 + B}{C z_1 + D}.$$  

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:

$$q_i = \frac{A q_0 + B}{C q_0 + D}.$$  

→ propagation through $N$ elements:

$$q_n = \frac{A q_0 + B}{C q_0 + D}.$$  

with the matrix $\mathbf{\tilde{M}} = \mathbf{\tilde{M}}_n \mathbf{\tilde{M}}_{n-1} \cdots \mathbf{\tilde{M}}_1 = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

→ 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 $z = d$:

→ propagation (by definition of $q$-parameter) $q(d) = q(0) + d$

$$\mathbf{\tilde{M}} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}.$$  

$$q_i = \frac{A q_0 + B}{C q_0 + D} = q_0 + d.$$  

ii) thin lens with focal length $f$

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

- no change of the width
- but change of phase curvature $R_i \cdot \exp\left[\frac{k}{2} \frac{(x^2 + y^2)}{R_i}\right]$  

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:
\[ R(z) = z \left[ 1 + \left( \frac{z_f}{z} \right)^2 \right] \approx z \quad \text{for} \quad z \gg z_f \]
We can invert the propagation from the focal position to the lens at the
distance of the focal length \( f \) and obtain \( R_f = -f \)
\[ f < 0 \quad f > 0 \quad \hat{M} = \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix} \]

\[ \text{double concave} \quad \text{double convex} \]
\[ \text{lens} \quad \text{lens} \]
\[ \Rightarrow \text{defocusing} \quad \Rightarrow \text{focusing} \]

\[ q_1 = \frac{A_q g_0 + B_q}{C_q d_0 + D_q} = -\frac{q_0}{f + 1} \]
\[ \frac{1}{q_1} = \frac{-q_0}{f + 1} = \frac{1}{-f} + \frac{i \lambda_s}{2 \pi w_0^2} \quad \text{for} \quad q_0 = -i \frac{\pi w_0^2}{\lambda_s} \]

Be careful: Gaussian optics describes the evolution of the beam’s width and
phase curvature only!
\[ \Rightarrow \text{Changes of amplitude and reflection are not included!} \]

2.8.4 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).

2.8.4.1 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 \( \rightarrow z_2 - z_1 = d \)
\[ - \] because \( R(z) = z + \frac{z^2}{z} \rightarrow \text{radius of wave front <0 for z <0} \)
\[ - \] from Chapter 1: beam hits concave mirror \( \rightarrow \text{radius} \quad R_i(i=1,2) < 0. \)
\[ - \] beam hits convex mirror \( \rightarrow \text{radius} \quad R_i(i=1,2) > 0. \)

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:

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

\[ R_1 = z_1 + \frac{z_0}{z_1}, \quad -R_2 = z_2 + \frac{z_0}{z_2} \]

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 \( z_2 = z_1 + d \) 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!

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

The denominator \( (R_1 + R_2 + 2d)^2 \) is always positive we need to fulfill

\[ -d(R_2 + d)(R_1 + R_2 + d) > 0 \]

If we introduce the so-called resonator parameters

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

We can re-express the stability condition as

\[ -d(R_2 + d)(R_1 + R_2 + d) = d g_1 g_2 R_1 R_2 \frac{(1 - g_1 g_2) R_1 R_2}{d} = g_1 g_2 (1 - g_1 g_2) (R_1 R_2)^2 > 0. \]

This inequality is fulfilled for

\[ 0 < g_1 g_2 < 1 \quad \text{or} \quad 0 < \left(1 + \frac{d}{R_1}\right) \left(1 + \frac{d}{R_2}\right) < 1 \]

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](image)

Resonator stability diagram. A spherical-mirror resonator is stable if the parameters \( g_1 = 1 + d/R_1 \) and \( g_2 = 1 + d/R_2 \) lie in the unshaded regions bounded by the lines \( g_1 = 0 \) and \( g_2 = 0 \), and the hyperbola \( g_1 g_2 = 1/g_1 \). \( R \) 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 \( g_1 = g_2 \).

Examples for a stable and an unstable resonator:

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

![Stable resonator](image)

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

![Unstable resonator](image)
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 → radius $R_i (i=1,2) > 0$.
- beam hits convex mirror → radius $R_i (i=1,2) < 0$.
- $z_{1,2}$ is the distance of mirror ‘1‘, ‘2‘ to the focus!
- $d$ is the distance between the two mirrors → $z_2 + z_1 = d$

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:

$$R_i = R(z_i), \ R'_i = R(z'_i)$$

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

$$g_i = \left(1 - \frac{d}{R_i}\right), \quad g'_i = \left(1 - \frac{d}{R'_i}\right)$$

the same stability condition

$$g_1g_2(1 - g_1g_2) \left(R_1R_2\right)^2 > 0, \quad 0 < g_1g_2 < 1.$$ 

2.8.4.2 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:

$$v_0(x,y,z) = A \cdot \frac{w_0}{w(z)} \exp \left[-\frac{x^2 + y^2}{w^2(z)} \right] \exp \left[i \frac{k}{2} \frac{x^2 + y^2}{R(z)} \right] \exp[i \phi(z)].$$

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

$$u_{lm}(x,y,z) = A_{lm} \cdot \frac{w_0}{w(z)} \frac{\sqrt{2}}{\sqrt{w(z)}} H_l \frac{\sqrt{2}}{\sqrt{w(z)}} x \exp \left[i \frac{k}{2} \frac{x^2 + y^2}{R(z)} \right] \exp[i k z] \exp[i (l + m + 1) \phi(z)].$$

The functions $G_i$ are given by the so-called Hermite polynomials:

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

Several low-order Hermite-Gaussian functions: (a) $G_0(u)$; (b) $G_1(u)$; (c) $G_2(u)$; and (d) $G_3(u)$.
2.9 Dispersion of pulses in homogeneous isotropic media

2.9.1 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 $T_e$ of $10^{-13}s (100fs) \leq T_e \leq 10^{-10}s (100ps)$.

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

$$f(t) = \exp(-\frac{4t^2}{T_e^2})$$

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

$\Rightarrow$ spectral width: $4 \cdot 10^{10}s^{-1} \leq \omega_s \leq 4 \cdot 10^{13}s^{-1}$

- center frequency of visible light: $\omega_0 = 2\pi \nu \approx 4 \cdot 10^{15}s^{-1}$

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

Hence, we have the following order of magnitudes:

$$\omega_s << \omega_0 \Rightarrow \omega - \omega_0 = \bar{\omega} << \omega_0$$

In this situation it can be helpful to replace the complicated frequency dependence (dispersion relation) of the wave vector $k(\omega)$ or the wave number $k(\omega)$ by a Taylor expansion at the central frequency $\omega = \omega_0$.

In most cases, a parabolic (or cubic) approximation of the frequency dependence in the dispersion relation will be sufficient:

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

The following terminology for the individual expansion coefficients is commonly used in the literature. It associates the physics, which is inherited in the dispersion relation, with the three parameters of the Taylor expansion.

**The three expansion coefficients and their physical significance**

**A) Phase velocity** $v_p$

$$k(\omega_0) = k_0 \Rightarrow k_0 = \frac{\omega_0}{\omega_0} = c$$

$\Rightarrow$ velocity of the phase front for the light at the central frequency $\omega = \omega_0$

**B) Group velocity or group index** $v_g$

$$\frac{\partial k}{\partial \omega_{\omega_0}} = \frac{1}{v_g}$$

$\Rightarrow$ group velocity is the velocity of the center of the pulse (see below)

$$k(\omega) = \frac{\omega}{c} n(\omega) \Rightarrow \frac{1}{v_g} = \frac{\partial k}{\partial \omega_{\omega_0}} = \frac{1}{c} \left[ n(\omega_0) + \omega_0 \frac{\partial n}{\partial \omega_{\omega_0}} \right]$$
\[ \nu_g = \frac{c}{n_n^2} \left( \frac{n_n}{\rho_0} + \frac{\partial n}{\partial \omega} \frac{\partial \omega}{\partial \rho_0} \right) = \frac{c}{n_n^2} n_n \frac{\partial n}{\partial \omega} \]

\[ n_n = n_n(\omega_0) + \frac{n_n}{\rho_0} \frac{\partial n}{\partial \rho_0} \quad \rightarrow \text{group index} \]

normal dispersion: \( \partial n / \partial \omega > 0 \cap n_g > n, \nu_g < \nu_p \)
anomalous dispersion: \( \partial n / \partial \omega < 0 \cap n_g < n, \nu_g > \nu_p \)

**C) Group velocity dispersion (GVD) or simply dispersion** \( D \)

\[ \frac{\partial^2 k}{\partial \omega^2} = D \]

\( \rightarrow \) GVD changes pulse shape upon propagation (see below)

\[ D = D_0 = \frac{\partial^2 k}{\partial \omega^2} = \frac{\partial}{\partial \omega} \left( \frac{1}{\nu_g} \right) \]

\[ D = \frac{\partial}{\partial \omega} \left( \frac{1}{\nu_g} \right) = -\frac{1}{\nu_g^2} \frac{\partial \nu_g}{\partial \omega} \]

\[ D = D_0 > 0 \cap \frac{\partial \nu_g}{\partial \omega} < 0 \]

\[ D = D_0 < 0 \cap \frac{\partial \nu_g}{\partial \omega} > 0 \]

Alternatively in telecommunication one often uses

\[ D_\lambda = \frac{\partial}{\partial \lambda} \left( \frac{1}{\nu_g} \right) = -\frac{2\pi}{\lambda^2} c D_0 \]

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

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

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

**dispersion relation:** \( k^2(\omega) = \frac{\omega^2}{c^2} \varepsilon(\omega) \)

Hence, we need to consider the propagation of the Fourier spectrum (Fourier transform in space and time):

\[ U(\alpha, \beta, \omega; z) = U_0(\alpha, \beta, \omega_0) \exp \left[ i \gamma(\alpha, \beta, \omega) \right] \]

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

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

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

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

\[ U(\alpha, \beta, \omega; z) \approx U_0(\alpha, \beta, \omega) \exp \left[ i k(\omega) z \right] \exp \left[ -i \frac{\alpha^2 + \beta^2}{2k_0(\omega)} \frac{z}{2} \right] \]

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

\[ H_p(\alpha, \beta, \omega; z) = \exp \left[ i k(\omega) \right] \exp \left[ -i \frac{\alpha^2 + \beta^2}{2k_0(\omega)} \frac{z}{2} \right] \]

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

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

Moreover, in the second term \( \exp[-i(\alpha^2 + \beta^2)z / (2k(\omega))t] \) of the transfer function (which is already small due to paraxiality) we can approximate the frequency dependence of the wave number by \( k(\omega) \approx k(\omega_0) = k_0 \). This approximation is sufficiently accurate to describe the diffraction of pulsed beams which are not too short. But this approximation would break down for \( T_0 \lesssim 15 \) fs since for such short pulses the frequency spectrum would become very wide. By introducing this approximation, we obtain the so-called parabolic approximation:

\[ H_p(\alpha, \beta, \omega; z) \approx \exp \left[ i k_0 z \right] \exp \left[ -i \frac{\alpha^2 + \beta^2}{2k_0} \frac{z}{2} \right] \]

with \( \omega = \omega - \omega_0 \)
Based on the last line of the above equation we can introduce a new variant of the propagation function, where the frequency argument is replaced by the frequency difference $\omega$ from the center frequency $\bar{\omega}$.

The new transfer function $H_{\alpha\beta}(\alpha, \beta, \bar{\omega}; z)$ is the propagation function for the slowly varying envelope $v(x, y, z, t)$:

$$u(x, y, z, t) = \exp\left[i(k_z z - \omega_0 t)\right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_0(\alpha, \beta, \omega) H_{\alpha\beta}(\alpha, \beta, \bar{\omega}; z) \times \exp\left[i(\alpha x + \beta y - \omega_0 t)\right] d\alpha d\beta d\omega$$

Illustration of the slowly varying envelope in the spectral domain

$$v(x, y, z, t) = v(x, y, z, t) \exp\left[i(k_z z - \omega_0 t)\right]$$

In order to complete the formalism, we also need to define the initial spectrum of the slowly varying envelope $u(\omega(t)) = u(t) \exp(-i\omega_0 t)$.

Thus, the propagation of the slowly varying envelope is given by:

$$v(x, y, z, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V_0(\alpha, \beta, \bar{\omega}) H_{\alpha\beta}(\alpha, \beta, \bar{\omega}; z) \exp\left[i(\alpha x + \beta y - \bar{\omega} t)\right] d\alpha d\beta d\bar{\omega}$$

Co-moving reference frame

The next step is to introduce a co-moving reference frame with $\bar{H}_{\alpha\beta}(\alpha, \beta, \bar{\omega}; z)$, the propagation function for the slowly varying envelope in the co-moving frame of the pulse:

$$u(\tau) = u(t) \exp(-i\tau)$$

This frame is called co-moving because $\bar{H}_{\alpha\beta}(\alpha, \bar{\omega}; z)$ 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:

$$\tilde{v}(x, y, z, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V_0(\alpha, \beta, \bar{\omega}) \bar{H}_{\alpha\beta}(\alpha, \beta, \bar{\omega}; z) \exp\left[i(\alpha x + \beta y - \bar{\omega} t)\right] d\alpha d\beta d\bar{\omega}$$

The optical field $u$ reads in the co-moving frame as:

$$\tilde{u}(x, y, z, t) = \tilde{v}(x, y, z, t) \exp\left[i(k_z z - \omega_0 t)\right] = \tilde{v}(x, y, z, t) \exp\left[i\left(k_z z - \frac{\omega_0}{v_g} z - \omega_0 t\right)\right]$$

Propagation equation in real space

Finally, let us derive the propagation equation for the slowly varying envelope in the co-moving frame. We start from the transfer function:
\begin{align}
\hat{V}(\alpha,\beta,\omega_0, z) &= V_o(\alpha,\beta,\omega_0) \exp \left[ i \frac{z}{2} \left( D\omega^2 - \frac{\alpha^2 + \beta^2}{k_0} \right) \right]
\end{align}

Then we take the spatial derivative of the transfer function long the propagation direction \( z \)
\begin{align}
\frac{1}{i} \frac{\partial \hat{V}(\alpha,\beta,\omega_0, z)}{\partial z} &= - \frac{1}{2} \left( D\omega^2 - \frac{\alpha^2 + \beta^2}{k_0} \right) \hat{V}(\alpha,\beta,\omega_0, z)
\end{align}

As before in the case of monochromatic beams, we use Fourier back-transformation to get the differential equation in the time-position domain
\begin{align}
\frac{1}{i} \frac{\partial \hat{v}(x,y,z,t)}{\partial z} - \frac{D}{2} \frac{\partial^2}{\partial t^2} \hat{v}(x,y,z,t) + \frac{1}{2k_0} \Delta^2 \hat{v}(x,y,z,t) = 0
\end{align}

This is the scalar paraxial equation for propagation of so-called pulsed beams.

**Comment: Extension to inhomogeneous media**

By using the slowly varying envelope approximation, it is possible to generalize the scalar paraxial equation also for inhomogeneous media, when a weak index contrast is assumed.
\begin{align}
\frac{1}{i} \frac{\partial \hat{v}(x,y,z,t)}{\partial z} - \frac{D}{2} \frac{\partial^2}{\partial t^2} \hat{v}(x,y,z,t) + \frac{1}{2k_0} \Delta^2 \hat{v}(x,y,z,t) + \frac{\Delta k_0(r) - \Delta k_0}{2k_0} \hat{v}(x,y,z,t) = 0
\end{align}

with \( \Delta k_0 \approx \langle \Delta k_0 \rangle \)

For \( D = 0 \) the equation would be reduced to simple diffraction, as in the beam propagation scheme which was derived earlier.

### 2.9.2 Infinite transverse extension - pulse propagation

Diffraction plays no role for sufficiently small propagation lengths \( z \ll L_0 \). For broad beams, the diffraction length \( L_0 \) can be rather large and we can assume \( \alpha = \beta \approx 0 \), corresponding to the assumption that we have a single plane wave propagating in \( z \)-direction.

**Description in frequency domain**

1. initial condition: \( u_0(t) = v_o(t) \exp(-i\omega_0 t) \)
2. initial spectrum: \( \hat{V}_o(\omega) = U_v(\omega) \)
3. propagation of the spectrum: \( \hat{V}(\omega, z) = \hat{V}_o(\omega) \exp \left[ i \frac{z D}{2} \right] \)
4. back-transformation to \( \tau \) leads to the following evolution of the slowly varying envelope in the co-moving frame:
\begin{align}
\hat{v}(z, \tau) &= \int_{-\infty}^{\infty} \hat{V}_o(\omega) \exp \left[ i \frac{z D}{2} \right] \exp \left[ -i \frac{\omega \tau}{2} \right] d\omega
\end{align}

**Description in time domain**

A) In time domain it is possible to describe pulse propagation by means of a response function:

FT-1 of \( \hat{V}(\omega, z) = \exp \left[ i \frac{z D}{2} \right] \rightarrow \hat{v}(\tau, z) = \sqrt{\frac{2}{\pi Dz}} \exp \left[ -i \frac{\tau^2}{2Dz} \right] \)

and the evolution is described by
\begin{align}
\hat{v}(z, \tau) &= \int_{-\infty}^{\infty} \hat{h}_\tau(\tau - \tau'; z) \hat{v}_0(\tau') d\tau'
\end{align}

B) The evolution equation for slowly varying envelope in the co-moving frame reads
\begin{align}
\frac{1}{i} \frac{\partial \hat{v}(z, \tau)}{\partial z} - \frac{D}{2} \frac{\partial^2}{\partial \tau^2} \hat{v}(z, \tau) = 0
\end{align}

**Analogy of diffraction and dispersion**

<table>
<thead>
<tr>
<th>DIFFRACTION</th>
<th>DISPERSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta k_0(r) \approx \langle \Delta k_0 \rangle ) &amp; ( \Delta )</td>
<td></td>
</tr>
<tr>
<td>( (x,y) \approx \tau ) &amp; ( \nabla \approx \frac{\partial}{\partial \tau} )</td>
<td></td>
</tr>
<tr>
<td>( \frac{1}{k_0} \approx -D ) but ( D \ll 0 ) can vary &amp; ( -D )</td>
<td></td>
</tr>
</tbody>
</table>

In the following we will study two typical examples of pulse propagation.

#### 2.9.3 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
2. Initial pulse spectrum

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

\( V_0(\omega) \) describes Gaussian pulse \( z_0 = \left(\frac{k}{2} v_0^2\right) \rightarrow \]

Hence anomalous GVD is equivalent to 'normal' diffraction.

Dispersion length: \( L_0 = 2|z_0| \)

3. Evolution of the amplitude

\[ \tilde{v}(z, \tau) = \tilde{A}(z) \frac{T_0}{T(z)} \exp\left(-\frac{\tau^2}{T(z)^2}\right) \exp\left[\frac{\omega T_0^2}{2D} \left(\frac{z}{z_0}\right)^2\right] \exp[\Delta \phi(z)] \]

with

\[ \tilde{A}(z) = \tilde{A}_0 \sqrt{1 + \left(\frac{z}{z_0}\right)^2}, \quad T(z) = T_0 \sqrt{1 + \left(\frac{z}{z_0}\right)^2} \]

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

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

Remember: The phase of a Gaussian beam \( \Phi(x, y, z) \) has the following shape:

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

For monochromatic fields the temporal dynamics of the phase is:

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

\( \rightarrow \) arbitrary time dependence of phase

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

The chirp of a pulse describes the variation of the temporal frequency of the electric field in the pulse.

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

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

integration leads to:

\[ -\frac{\partial \Phi(\tau)}{\partial \tau} = \omega(\tau) = \omega_0 + \frac{2C}{T_0} \frac{\tau}{T_0} \]

\[ -\Phi(\tau) = \omega_0 \tau + C \frac{\tau^2}{T_0^2} \]

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

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

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

Complete phase:

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

\( \rightarrow \)

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

with

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

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

Complete field:
\[
\begin{equation*}
\begin{align*}
&u(z, \tau) = A_0 \sqrt{\frac{T_0}{T(z)}} \exp\left[ -\frac{\tau^2}{T(z)^2} \right] \exp\left[ -4C(z) \frac{\tau^2}{T_0^2} \right] \exp\left[ i\varphi(z) \right] \exp\left[ i\left(k_0z - \omega_0\tau\right)\right]
\end{align*}
\end{equation*}
\]

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

- important parameter \( \rightarrow \) dispersion parameter \( z_0 = -\frac{T_0^2}{2D} \)
- \( z << |z_0| \): no effect
- \( z = |z_0| \): similar to beam diffraction
- \( z \gg |z_0| \): asymptotic dependence

\[
T(z) \approx T_0 \left| \frac{z}{z_0} \right| \rightarrow T(z) / z = T_0 / |z_0| = \frac{2|D|}{T_0} = \text{const.}
\]

\[
T(z) = \frac{2|D|}{T_0} z
\]

Gaussian pulse spreading as a function of distance \( z \). For large distances, the width increases at a rate \( 2|D|/T_0 \), which is inversely proportional to the initial width \( T_0 \).

\( D \leq 0 \) is only important if initial pulse is chirped, since otherwise the same quadratic dependence is observed, independent from the sign of \( D \).

**2.9.4 Example 2: Chirped Gaussian pulse**

Important because of:
- short pulse lasers \( \rightarrow \) chirped pulses
- Chirp is introduced on purpose, for subsequent pulse compression

- analogy to curved phase \( \rightarrow \) focusing
- chirped pulse amplification (CPA) \( \rightarrow \) Petawatt lasers

### 1. Input pulse shape

\[
\nu_0(t) = A_0 \exp\left[ -\frac{\tau^2(1 + iC_0)}{T_0^2} \right] \quad C_0 \text{ - initial chirp}
\]

### 2. Input pulse spectrum

\[
\nu_s(\omega) = A_0 \exp\left[ -\frac{\omega^2 T_0^2 (1 - iC_0)}{4(1 + C_0^2)} \right]
\]

\( \rightarrow \) spectral width: \( \delta \omega_s = \frac{4(1 + C_0^2)}{T_0^2} \)

\( \rightarrow \) spectral width of chirped pulse is larger than that of unchirped pulse

\( (\delta \omega_s = 4 / T_0^2) \) only for transform limited pulses

Aim: calculation of pulse width and chirp in dependence on \( z \) for given initial conditions

Gaussian beam \( \rightarrow q \)-parameter \( \rightarrow \) similar to Gaussian pulse

Use analogy: however it is limited to homogeneous space

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

Remember beams:

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

\[
k \rightarrow \frac{1}{D}, \quad w^2(z) \rightarrow T^2(z), \quad \frac{1}{R(z)} \rightarrow \frac{2DC(z)}{T_0^2}
\]

\[
\frac{1}{q(z)} = \frac{2DC(z)}{T_0^2} - \frac{2D}{T^2(z)}\quad (*)
\]

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

hence at \( z = 0 \):

\[
\frac{1}{q(0)} = \frac{2D}{T_0^2} [C_0 - 1] \quad \text{with} \quad C_0 = C(0)
\]

Idea:
a) \( q(z) = q(0) + z \) and \( \frac{1}{q(0)} = \frac{2D}{T_0^2} \left[ C_0 - 1 \right] \) insert into \( \frac{1}{q(z)} \)

b) \( \frac{1}{q(z)} = \frac{2D}{T_0^2} \left[ C(z) - 1 + \frac{T_0^2}{T^4(z)} \right] \)

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

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

1) **Determination of q parameter at input**

\[ q(0) = \frac{T_0^2}{2D} \left( C_0 + 1 \right) \]

2) **Evolution of q parameter**

\[ q(d) = q(0) + d = \frac{T_0^2}{2D} \left( C_0 + 1 \right) + d = \frac{1}{2D \left( 1 + C_0^2 \right)} \left[ 2Dd \left( 1 + C_0^2 \right) + C_0 T_0^2 + i T_0^2 \right] \]

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

\[ \frac{1}{q(d)} = \frac{2D}{T_0^2} \left[ C(d) - 1 + \frac{T_0^2}{T^4(d)} \right] \]

\[ q(d) = \frac{T_0^2 T^4(d)}{2D \left[ C^2(d) T^4(d) + T_0^4 \right]} \left[ C(d) T^2(d) + i T_0^2 \right] \]

4) **Set two equations equal**

\[ \frac{2Dd \left( 1 + C_0^2 \right) + C_0 T_0^2 + i T_0^2}{2D \left( 1 + C_0^2 \right)} = \frac{C_0 T_0^2 T^4(d)}{2D \left( C^2(d) T^4(d) + T_0^4 \right)} \]

a) real part

\[ \frac{2Dd \left( 1 + C_0^2 \right) + C_0 T_0^2}{\left( 1 + C_0^2 \right)} = \frac{C(d) T_0^2 T^4(d)}{C^2(d) T^4(d) + T_0^4} \]

b) imaginary part

\[ \frac{T_0^2 T^4(d)}{\left( 1 + C_0^2 \right)} = \frac{C^2(d) T^4(d) + T_0^4}{T_0^4} \]

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: \( C_0, T_0 \) & in the focus: \( C(d) = 0 \)

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

a) real part must be zero

\[ 2Dd \left( 1 + C_0^2 \right) + C_0 T_0^2 = 0 \]

\[ \rightarrow d = \frac{1}{2} \frac{T_0^2 C_0}{D \left( 1 + C_0^2 \right)} = \frac{1}{2} \frac{\text{sgn}(D) \left( 1 + C_0^2 \right)}{D} T_0^2 \]

b) \[ T^2(d) = \frac{T_0^4}{\left( 1 + C_0^2 \right)} \]

**Resulting properties**

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

2) The possible compression increases with initial chirp.

**Physical interpretation**

If e.g. \( C_0 < 0 \) and \( D > 0 \) → \( \partial v_g / \partial \omega < 0 \) → 'red' is faster than 'blue'

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. At this propagation distance the pulse has no remaining chirp.

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.

\[ C(z) = -\frac{z}{z_0 \left( 1 + \frac{z}{z_0} \right)} \quad z_0 = \frac{T_0^2}{2D} \]
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 $(x,y) = 0$ for $x > a$ (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
   → not important, generally plane wave (no diffraction)

ii) multiply field distribution of illuminating wave by transfer function $u_c(x,y,z) = t(x,y)u(x,y,z)$

iii) propagation of modified field distribution behind the aperture through homogeneous space

$$u(x,y,z) = \int_{-\infty}^{\infty} H(\alpha,\beta;z-z_a)U_c(\alpha,\beta;z_a)\exp\left\{i(\alpha x + \beta y)\right\}d\alpha d\beta$$

or

$$u(x,y,z) = \int_{-\infty}^{\infty} h(x-x',y-y',z-z_a)u_c(x',y',z_a)dx'dy'$$

with $h = \frac{1}{(2\pi)^2}FT^{-1}[H]$.

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

$$H(\alpha,\beta;z_a) = \exp\left\{\gamma(\alpha,\beta;z_a)\right\}$$

where $\gamma = k^2 - \alpha^2 - \beta^2$.

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:

$$u(x,y,z) = \int_{-\infty}^{\infty} U_c(\alpha,\beta;z_a)H(\alpha,\beta;z_a)\exp\left\{i(\alpha x + \beta y)\right\}d\alpha d\beta$$

where $U_c(\alpha,\beta) = FT[u_c(x,y)]$

**Derivation of the response function**

We start from the Weyl-representation of a spherical wave:

$$\frac{1}{r}\exp(ikr) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left\{i(\alpha x + \beta y + \gamma z)\right\}d\alpha d\beta$$

Now we can compute the response function $h$, which we did not do in the previous chapter, where we computed only $h$ (Fresnel approximation). The following trick shows that

$$-2\pi \frac{\partial}{\partial z} \left\{\frac{1}{r}\exp(ikr)\right\} = \int_{-\infty}^{\infty} \exp\left\{i(\alpha x + \beta y + \gamma z)\right\}d\alpha d\beta = FT^{-1}[H] = (2\pi)^2 h$$

and therefore

$$h(x,y,z) = -\frac{\partial}{\partial z} \left[\frac{1}{r}\exp(ikr)\right]$$

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’:

$$u(x,y,z) = \int_{-\infty}^{\infty} h(x-x',y-y',z-z_a)u_c(x',y',z_a)dx'dy'$$

3.2.2 Fresnel approximation (paraxial approximation)

From the previous chapter we know that we can apply the Fresnel approximation if $\alpha^2 + \beta^2 << k^2$ which is valid for a limited angular spectrum, and therefore a large size of the structures inside the aperture. Then

$$H_f(\alpha,\beta;z_a) = \exp(ikz_a)\exp\left[-\frac{i z_a}{2k}(\alpha^2 + \beta^2)\right]$$
\[ h_p(x, y, z_n) = -\frac{i}{\lambda z_n} \exp(i k z_n) \exp \left[ \frac{k}{2z_n} \left( x^2 + y^2 \right) \right] \]

### 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 \( N_r \)

\[ N_r \leq 0.1 \quad \text{with} \quad N_r = \frac{a}{\lambda z_n} \]

where \( a \) is the (largest) size of the aperture (like the "beam width"). Obviously, a larger aperture needs a larger distance \( z_B \) to fulfill \( N_r \approx 0.1 \).

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:

\[
\begin{align*}
H_p(x, y, z_n) &= \int_{-\infty}^{\infty} H_p(\alpha, \beta; z_n) U_p(\alpha, \beta) \exp \left[ i (\alpha x + \beta y) \right] d\alpha d\beta \\
&= \int_{-\infty}^{\infty} h_p(x - x', y - y'; z_n) u_p(x', y') dx'dy'
\end{align*}
\]

In this situation it is easier to treat the beam propagation in position space, because

\[ u_p(x, y) = t(x,y) u_p(x, y), \quad \text{and} \quad t(x,y) = 0 \quad \text{for} \quad |x|,|y| > a \quad \text{(aperture)} \]

\[ \rightarrow u_p(x, y) = 0 \quad \text{for} \quad |x|,|y| > a \]

This means that we need to integrate only over the aperture in the above integral:

\[
\begin{align*}
H_p(x, y, z_n) &= -\frac{i}{\lambda z_n} \exp(i k z_n) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_p(x', y') \exp \left[ i \frac{k}{2z_n} \left( (x - x')^2 + (y - y')^2 \right) \right] dx'dy'
\end{align*}
\]

Now, let us have a closer look at the exponential expression:

\[
\begin{align*}
\exp \left[ i \frac{k}{2z_n} \left( (x - x')^2 + (y - y')^2 \right) \right]
&= \exp \left[ i \frac{k}{2z_n} \left[ x^2 - 2 \alpha x' + \alpha^2 + y^2 - 2 \beta y' + \beta^2 \right] \right]
\end{align*}
\]

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

\[ \rightarrow \frac{k}{2z_n} \left[ x^2 + y^2 \right] < \frac{ka^2}{z_n} = 2\pi N_r \]

\[ \rightarrow \text{for} \quad N_r \leq 0.1 \rightarrow \exp \left[ i \frac{k}{2z_n} \left[ x^2 + y^2 \right] \right] \approx 1 \]

This means that we can neglect the quadratic phase term in \( x', y' \) and we get for the far field:

\[
\begin{align*}
H_p(x, y, z_n) &= -\frac{i}{\lambda z_n} \exp(i k z_n) \exp \left[ i \frac{k}{2z_n} (x^2 + y^2) \right] \\
&= \frac{i}{\lambda z_n} \exp(i k z_n) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_p(k \frac{x}{z_n}, k \frac{y}{z_n}) \exp \left[ i \frac{k}{2z_n} \left( \alpha^2 + \beta^2 \right) \right] dx'dy'
\end{align*}
\]

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

\[ I_p(x, y, z_n) = \frac{1}{\lambda z_n^2} \left| U_p(k \frac{x}{z_n}, k \frac{y}{z_n}) \right|^2 \]

### 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 \( U_p(kx / z_n, ky / z_n) \) 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) $\alpha^2 + \beta^2 < k^2 \Rightarrow$ smallest features $\Delta x, \Delta y \gg \lambda \Rightarrow$ narrow angular spectrum (paraxiality)

B) $N_p = \frac{\Delta x}{z_a} < 1 \Rightarrow$ largest feature $a$ determines $z_a > k \Rightarrow$ far field

Example: $\Delta x, \Delta y = 10 \lambda, \quad a = 100 \lambda, \quad \lambda = 1 \mu m \Rightarrow z_a > 10^4 \lambda = 1 \text{ cm}$

### 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 $kz_a > 1 \rightarrow z_a > \lambda$.

$$N_p \lesssim 0.1 \quad \text{with} \quad N_p = \frac{a}{\lambda z_a} = \frac{1}{\pi z_a}$$

$$u_{za}(x, y, z_a) = -\frac{1}{\lambda \sqrt{x^2 + y^2 + z_a^2}} \frac{z_a}{\sqrt{x^2 + y^2 + z_a^2}} \times U_{\lambda} \left( \frac{kx}{\sqrt{x^2 + y^2 + z_a^2}}, \frac{ky}{\sqrt{x^2 + y^2 + z_a^2}} \right) \exp \left( \frac{i k \sqrt{x^2 + y^2 + z_a^2}}{z_a} \right)$$

### 3.3 Fraunhofer diffraction at plane masks (paraxial)

Let us now plug things together and investigate diffraction patterns induced by plane masks in (paraxial) Fraunhofer approximation.

For inclined incidence of the excitation the field behind the mask is given by:

$$u_{\lambda}(x, y, z_a) = u_\lambda(x, y, z_a) \lambda (x, y) = A \exp \left[ i \left( k \lambda x + k \lambda y + k \lambda z_a \right) \right] \lambda(x, y)$$

Form the previous chapter we know that the diffraction pattern in the far field in paraxial Fraunhofer approximation is given as:

$$I(x, y, z_a) \sim \left| u_\lambda(x, y, z_a) \right|^2 \sim \frac{1}{(\lambda z_a)} \left| U_\lambda \left( k \frac{x}{z_a}, k \frac{y}{z_a} \right) \right|^2$$

Hence, the diffraction pattern is proportional to the spectrum of the field behind the mask at

$$\alpha = k \frac{x}{z_a}, \beta = k \frac{y}{z_a}.$$ 

This spectrum is calculated by the Fourier transform of the field as:

$$U_\lambda \left( k \frac{x}{z_a}, k \frac{y}{z_a} \right) = \frac{A}{(2\pi)^3} \int \int t(x', y') \exp \left[ -i \left( k \frac{x}{z_a} - k_i \right) x' - i \left( k \frac{y}{z_a} - k_j \right) y' \right] dx' dy'$$

$$= A \exp \left( i k z_a \right) T \left( k \frac{x}{z_a} - k, k \frac{y}{z_a} - k \right)$$

Hence, the intensity distribution of the diffraction pattern is given as:

$$I(x, y, z_a) \sim \frac{1}{(\lambda z_a)} \left| T \left( k \frac{x}{z_a} - k, k \frac{y}{z_a} - k \right) \right|^2$$

This is the absolute square of the Fourier transform of the aperture function. In paraxial approximation an inclination of the illuminating plane wave just shifts the pattern transversely.

### Examples

A) Rectangular aperture illuminated by normal plane wave

$$t(x, y) = \begin{cases} 1 & \text{for } |x|, |y| \leq b \\ 0 & \text{elsewhere} \end{cases}$$

$$I(x, y, z_a) \sim \text{sinc}^2 \left( \frac{ka x}{z_a} \right) \text{sinc}^2 \left( \frac{kb y}{z_a} \right)$$
Fraunhofer diffraction pattern from a rectangular aperture. The central lobe of the pattern has half-angular widths $\theta_x = \frac{\lambda}{D_x}$ and $\theta_y = \frac{\lambda}{D_y}$.

**B) Circular aperture (pinhole) illuminated by normal plane wave**

$$t(x, y) = \begin{cases} 1 & \text{for } x^2 + y^2 \leq a^2 \\ 0 & \text{elsewhere} \end{cases}$$

$$I(x, y, z_a) \sim \left[ \frac{J_1\left(\frac{\rho}{z_a}\sqrt{x^2 + y^2}\right)}{\frac{\rho}{z_a}\sqrt{x^2 + y^2}} \right]^2 \rightarrow \text{Airy disk}$$

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):

$$\frac{ka}{z_a} = 1.22\pi \rightarrow \frac{\rho}{z_a} = 0.61\frac{\lambda}{a} \quad \text{with } \rho^2 = x^2 + y^2$$

So-called angle of aperture: $\Theta = \frac{2\rho}{z_a} = \frac{1.22\lambda}{a}$

**C) One-dimensional periodic structure (grating) 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 periodic slit aperture with: period $b$ and a size of each slit $2a$:

$$t_i(x) = \begin{cases} 1 & \text{for } |x| \leq a \\ 0 & \text{elsewhere} \end{cases}$$

Then, we can express the mask function $t$ as:

$$t(x) = \sum_{n=0}^{\infty} t_i(x - nb)$$

The Fourier transform of the mask is given as

$$T\left( \frac{k x}{z_a} \right) = \sum_{n=0}^{\infty} \int t_i(x') \exp\left(-i k \frac{x}{z_a} x'\right) dx'$$

With the new variable $x' - nb = X'$ we can simplify further:

$$T\left( \frac{k x}{z_a} \right) = \sum_{n=0}^{\infty} \int t_i(X') \exp\left(-i k \frac{x}{z_a} X'\right) \exp\left(-i k \frac{x}{z_a} nb\right) dX'$$

We see that the Fourier transform $T_i$ of the elementary slit $t_i$ 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 by using the following formula for the definition of the sinc-function:

$$\sum_{n=0}^{\infty} \exp(-i \delta n) = \frac{\sin(N \delta)}{\sin(\delta)}$$

Thus we finally write:

$$T\left( \frac{k x}{z_a} \right) = T_i\left( \frac{k x}{z_a} \right) \sum_{n=0}^{\infty} \exp\left(-i k \frac{x}{z_a} nb\right)$$

For the particular case of a simple grating of slit apertures with $t_i(x) = 1$ we have
We find three important parameters for the diffraction pattern of a grating:

- **Global width of diffraction pattern** → first zero of slit function $T_s$
  \[ k \frac{x}{z_a} = \pi \Rightarrow x_s = \frac{\lambda z_a}{2a} \]
  The width of the total far-field diffraction pattern $x_N$ (largest length scale in the pattern) is determined by the size $a$ of the individual slit (smallest length scale in the mask).

- **Position of local maxima of diffraction pattern** → maxima of grid function
  \[
  \max \left( \frac{\sin^2 \left( \frac{N \frac{x}{z_a} + b}{\lambda z_a} \right)}{\sin^2 \left( \frac{\lambda z_a}{2z_a} \right)} \right) \Rightarrow \frac{k x_n b}{\lambda z_a} = n \pi \Rightarrow x_p = \frac{n \lambda z_a}{b}
  \]
  These are the so-called diffraction orders, which are determined exclusively by the grating period.

- **Width of local maxima** → zero-points of grid function

\[
N \frac{k x_n b}{2 z_a} = \pi \Rightarrow x_p = \frac{\lambda z_a}{N b}
\]

The width of a maximum in the far-field diffraction pattern $x_N$ (smallest length scale in the pattern) is determined by $N * b$ which is 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 and vice versa.

### 3.4 Remarks on Fresnel diffraction

**Fresnel number** $N_F = \frac{a^2}{\lambda z_a}$

**Fresnel diffraction length from Gaussian beams** $z_0 = \frac{\pi a^2}{\lambda}$

- $N_F \gtrsim 10$ (large, $\lambda z_a$ small, $z_0 < 1/30 \ z_0$) → shadow
- $N_F \lesssim 0.1$ (small, $z_0 > 3z_0$) → Fraunhofer → FT of aperture
- $10 \gtrsim N_F \gtrsim 0.1$ (1/30 $z_0 < z_a < 3z_0$) → 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 $N_F = 10, 1, 0.5, 0.1$. The shaded area represents the geometrical shadow of the slit. The dashed lines at $|x| = (\lambda / D) x_D$ 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 $N_F = a^2 / D^2 = 0.5$.  

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 $N_F = 10, 1, 0.5, 0.1$. The shaded area represents the geometrical shadow of the slit. The dashed lines at $|x| = (\lambda / D) x_D$ 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 \( N_f = \frac{a^2}{\lambda d} = 0.5 \).

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. While detailed designs of high resolution optical systems have to consider non-paraxial effects, usually the paraxial approximation is sufficient to obtain a principle understanding of optical systems. Hence we use the paraxial approximation here.

Many imaging systems exploit the appearance of the Fourier transform of the original object in the so-called Fourier plane of the system in order to manipulate the angular spectrum of the object in this plane. Accordingly this field of optical science is called Fourier optics. 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 in Fourier plane.
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 \( \rightarrow \) calculated in Fourier domain
• interaction with lens or filter \( \rightarrow \) calculated 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

\[
\frac{\frac{i}{\lambda f} \exp (i k f) \exp \left[ \frac{i k}{2 f} (x^2 + y^2) \right]}{\text{spherical wave}} t_w(x, y) = -\frac{\frac{i}{\lambda f} \exp (i k f)}{\text{plane wave}}
\]

And therefore the response function for a thin lens is given as (see chapter 2.8.3):
$$t_e(x,y) = \exp \left[-\frac{1}{2} k f (x^2 + y^2)\right]$$

By Fourier transforming the response function we find consequently the transfer function in the Fourier domain as

$$T_e(\alpha,\beta) = -\frac{i \lambda f}{2\pi} \exp \left[\frac{i f}{2k} (\alpha^2 + \beta^2)\right]$$

### 4.1.2 Optical imaging using the 2f-setup

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 recipe for light propagation.

**A) Spectrum in object plane**

$$U_0(\alpha,\beta) = \text{FT}[u_0(x,y)]$$

**B) Propagation from object to lens**

_lens positioned at distance f_

$$U_1(\alpha,\beta; f) = H_0(\alpha,\beta; f) U_0(\alpha,\beta)$$

$$U_1(\alpha,\beta; f) = \exp(\pm i k f) \exp \left[-\frac{i}{2k} (\alpha^2 + \beta^2) f\right] U_0(\alpha,\beta)$$

**C) Interaction with lens**

_multiplication in position space or convolution in Fourier domain_

$$u_1(x,y,f) = t_e(x,y) u_0(x,y,f)$$

**D) Propagation from lens to image plane**

$$U_2(\alpha,\beta; 2f) = H_f(\alpha,\beta; f) U_1(\alpha,\beta; f)$$

$$U_2(\alpha,\beta; 2f) = -\frac{i \lambda f}{2\pi} \exp(2i k f) \int_{-\infty}^{\infty} \exp \left[\frac{i f}{2k} ((\alpha - \alpha')^2 + (\beta - \beta')^2)\right] \cdot \exp \left[-\frac{i}{2k} (\alpha^2 + \beta^2) f\right] U_0(\alpha',\beta') d\alpha' d\beta'$$

Quadratic terms $\left[-\frac{i}{2k} (\alpha^2 + \beta^2) f\right]$ and $\left[-\frac{i}{2k} (\alpha' + \beta') f\right]$ in the exponent cancel with quadratic terms from $\frac{i f}{2k} ((\alpha - \alpha')^2 + (\beta - \beta')^2)$ and only the mixed terms remain.

$$U_2(\alpha,\beta; 2f) = -\frac{i \lambda f}{2\pi} \exp(2i k f) \int_{-\infty}^{\infty} U_0(\alpha',\beta') \exp \left[-\frac{i f}{k} (\alpha + \beta')\right] d\alpha' d\beta'$$

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**

$$u(x,y,2f) = \text{FT}^{-1}[U(\alpha,\beta; 2f)]$$

$$u(x,y,2f) = -\frac{i \lambda f}{2\pi} \exp(2i k f) \int_{-\infty}^{\infty} u_0 \left(-\frac{f}{k}, -\frac{f}{k}\right) \exp \left[i(\alpha x + \beta y)\right] d\alpha d\beta$$

With the coordinate transformation
we get:

\[ u(x, y, 2f) = -\frac{1}{\lambda f} \exp(2\pi k f) \int \int u_0(x', y') \exp \left[ -\frac{i}{f} (xx' + yy') \right] dx' dy' \]

The image in the second focal plane is the Fourier transform of the optical field in the object plane. This is similar to the far field in Fraunhofer approximation, but for \( z_0 \leftrightarrow f \). This finding allows us to perform an optical Fourier transform over shorter distances. And in the Fourier plane it is possible to manipulate the spectrum.

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:

We know that the image in Fourier plane is the FT of the optical field in object plane.

\[ u(x, y, 2f) = AU_0 \left( \frac{k}{f}, \frac{k}{f}, \frac{k}{f}, \frac{k}{f} \right) \]

We have to compute the imaging with the second lens after manipulation. Our final goal is the transmission function \( H_k(\alpha, \beta; 4f) \) of the complete imaging system:

\[ u(-x, -y, 4f) = \int \int H_k(\alpha, \beta; 4f) U_0(\alpha, \beta) \exp \left[ -\frac{i}{f} (\alpha x + \beta y) \right] d\alpha d\beta \]

Note: We will see in the following calculation that the second lens does a Fourier transform \( \sim \exp \left[ -\frac{i}{f} (\alpha x + \beta y) \right] \). 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 constraints of the system (e.g. a limited aperture) and optical filtering (which we can design).

**A)** Field behind transmission mask

\[ u_i(x, y, 2f) = u(x, y, 2f) p(x, y) \sim AU_0 \left( \frac{k}{f}, \frac{k}{f}, \frac{k}{f}, \frac{k}{f} \right) p(x, y) \]

**B)** Second lens \( \Rightarrow \) Fourier back-transform of field distribution

\[ u(x, y, 4f) = -\frac{1}{\lambda f} \exp(2\pi k f) U_0 \left( \frac{k}{f}, \frac{k}{f}, \frac{k}{f}, 2f \right) \]

Now we can make the link to the initial spectrum in the object plane \( U_0 \):
We introduce the coordinate transform \( \begin{align*}
\alpha &= \frac{k}{f} x', \\
\beta &= \frac{k}{f} y'
\end{align*} \)

\[
h_\alpha(x,y) \sim \int_{-\infty}^{\infty} P(x',y') \exp \left\{ \frac{-k}{f} (x' + y y') \right\} dx'dy'
\]

The response-function is proportional to the Fourier transform of the transmission mask.

### 4.2.2 Examples of aperture functions

**Example 1: The ideal image (infinite aperture)**

Be careful, we use paraxial approximation \( \rightarrow \) limited angular spectrum

\[
p = 1 \rightarrow P - \delta(x-x') \delta(y-y') \rightarrow u(-x,-y,4f) - u_0(x,y) \rightarrow \text{mirrored original}
\]

**Example 2: Finite aperture**

\[
p(x,y) = \begin{cases} 
1 & \text{for } x^2 + y^2 \leq (D/2)^2 \\
0 & \text{elsewhere}
\end{cases}
\]

The 4f system performs a Fourier transform followed by an inverse Fourier transform, so that the image is a perfect replica of the object (perfect only within the Fresnel approximation).

The response function is given as the Fourier transform of the transfer function:

\[
\begin{align*}
\mathcal{H}_\alpha(\alpha,\beta;4f) &= \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \mathcal{H}_\alpha(\alpha,\beta;4f) \exp \left\{ i \left( \alpha x + \beta y \right) \right\} d\alpha d\beta \\
\end{align*}
\]

From above we have \( \mathcal{H}_\alpha(\alpha,\beta;4f) \sim p\left( \frac{f}{k} \alpha, \frac{f}{k} \beta \right) \)

\[
h_\alpha(x,y) \sim \int_{-\infty}^{\infty} p\left( \frac{f}{k} \alpha, \frac{f}{k} \beta \right) \exp \left\{ i \left( \alpha x + \beta y \right) \right\} d\alpha d\beta
\]
Spatial filtering. The transparencies in the object and Fourier planes have complex amplitude transmittances \( f(x, y) \) and \( p(x, y) \). A plane wave traveling in the \( z \) direction is modulated by the object transparency, Fourier transformed by the first lens, multiplied by the transmittance of the mask in the Fourier plane and inverse Fourier transformed by the second lens. As a result, the complex amplitude in the image plane \( g(x, y) \) is a filtered version of the original field \( f(x, y) \) in the object plane. The system has a transfer function \( H(v_x, v_y) = p(v_x, v_y) \).

Transmission function:

\[
H_\Delta(\alpha, \beta; 4f) \sim \begin{cases} 
1 & \text{for } \left(\frac{\alpha}{\lambda}/\Delta f\right)^2 + \left(\frac{\beta}{\lambda}/\Delta f\right)^2 \leq \left(\frac{D}{2}\right)^2 \\
0 & \text{elsewhere}
\end{cases}
\]

- finite aperture truncates large angular frequencies (low pass)
- determines optical resolution

**4.2.3 Optical resolution**

A finite aperture acts as a low pass filter for angular frequencies.

\[
\left(\frac{k \alpha}{\lambda}\right)^2 + \left(\frac{k \beta}{\lambda}\right)^2 \leq \left(\frac{D}{2}\right)^2 \rightarrow \left(\alpha\right)^2 + \left(\beta\right)^2 \leq \left(\frac{k}{f} D / 2\right)^2
\]

With \( \rho^2 = \alpha^2 + \beta^2 \) we can define an upper limit for the angular frequencies \( \rho_{\text{max}} \) which are transmitted (bandwidth of the system)

\[
\rho_{\text{max}}^2 = \left(\frac{k D}{f} / 2\right)^2 \Rightarrow \rho_{\text{max}} = \frac{2\pi n D}{\lambda f}
\]

Translated to position space, the smallest transmitted structural information is given by:

\[
\Delta x_{\text{min}} = \frac{2\pi}{\rho_{\text{max}}} = \frac{2\lambda f}{nD}
\]

A more precise definition of the optical resolution can be derived the following way:

\[
u(-x, -y, 4f) \sim \iint \frac{1}{f} \frac{1}{f} \left(\frac{k}{f} (x' - x), \frac{k}{f} (y' - y)\right) u_0(x', y') dx' dy'
\]

One point of the object \( x_0, y_0 \) gives an Airy disk (pixel) in the image:

\[
\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi}} \left[ \frac{\lambda D}{f} \sqrt{(x_0 - x)^2 + (y_0 - y)^2}\right]^{-\frac{1}{2}}
\]

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:

\[
\frac{kD}{2f} \Delta x_{\text{min}} = 1.22\pi
\]

Hence we find:

\[
\Delta x_{\text{min}} = \frac{1.22\lambda f}{nD}
\]
Further examples for 4f filtering

Examples of object, mask, and filtered image for three spatial filters: (a) low-pass filter; (b) high-pass filter; (c) vertical pass filter. Black means the transmittance is zero and white means the transmittance is unity.

5. The polarization of electromagnetic waves

5.1 Introduction

We are interested in the temporal evolution of the electric field vector $\mathbf{E}_z(r,t)$. 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 $\mathbf{E}(\mathbf{r},t) = \mathbf{E} \exp \left[ i \left( \mathbf{k} \cdot \mathbf{r} - \omega t \right) \right]$.

If we assume propagation in $z$ direction ($\mathbf{k}$-vector points in $z$-direction), $\nabla \mathbf{E}(\mathbf{r},t) = 0$ 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

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

Because the field is transversal we have two free complex field components

$$\mathbf{E} = \begin{pmatrix} E_x \exp(i\phi_x) \\ E_y \exp(i\phi_y) \\ 0 \end{pmatrix} \quad \text{with } E_{x,y} \text{ and } \phi_{x,y} \text{ being real}$$

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 $\rightarrow \delta = \phi_y - \phi_x$

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 \( \rightarrow E_x = E_y = E, \delta = \pm \pi / 2 \)

\( \delta = + \pi / 2 \rightarrow \) counterclockwise rotation

\( \delta = - \pi / 2 \rightarrow \) 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 \( kz = 0 \) with \( \delta = \pm \pi / 2 \).

\[
E \begin{pmatrix}
\cos(\omega t) \\
\cos(\omega t)
\end{pmatrix} + E \begin{pmatrix}
\sin(\omega t) \\
-\sin(\omega t)
\end{pmatrix} = 2E \begin{pmatrix}
\cos(\omega t) \\
0
\end{pmatrix}
\]
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 $\rightarrow$ ???

In the following we will write $E \rightarrow E$, because we assume monochromatic light and the frequency $\omega$ is just a parameter.

Before (isotropic)

$$P(r, \omega) = e_0 \chi(\omega) \ E(r, \omega)$$
$$D(r, \omega) = e_0 \varepsilon(\omega) \ E(r, \omega)$$

In the following we will write $E \rightarrow \mathbf{E}$, because we assume monochromatic light and the frequency $\omega$ is just a parameter.

Now (anisotropic)

$$P(r, \omega) = e_0 \sum_{\beta=1}^{l=3} \chi_{\beta}(\omega) E_{\beta}(r, \omega)$$

The linear susceptibility tensor has $3 \times 3 = 9$ 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 $\chi_{\beta}$ depend on the structure of crystal. However, we do not need to know the microscopic structure because of the different length scales involved (optics: $5 \cdot 10^{-7}$ m; crystal: $5 \cdot 10^{-10}$ m), but the field is influenced by the symmetries of the crystal (see next section).

In complete analogy we find for the $D$ field:

$$D(r, \omega) = e_0 \delta(\omega) E(r, \omega)$$

As for the polarization we find:
- $\mathbf{D} \neq \mathbf{E}$

We introduce the following notation:
- $\chi = (\chi_{\beta}) \rightarrow$ susceptibility tensor
- $\varepsilon = (\varepsilon_{\beta}) \rightarrow$ dielectric tensor
- $\sigma = (\varepsilon)^{-1} = (\sigma_{\beta}) \rightarrow$ inverse dielectric tensor

The following properties of the dielectric and inverse dielectric tensor are important:
- $\sigma_{\beta}, \varepsilon_{\beta}$ 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 $\varepsilon_{\beta} = \varepsilon_{\beta}, \sigma_{\beta} = \sigma_{\beta}$.
- 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 $\sigma_{\beta}$, it means that we are looking for directions where $\mathbf{D} = \mathbf{E}$, i.e., our principal axes:

$$\sigma_{\beta} E_{\beta} = \sum_{\beta=1}^{l=3} \sigma_{\beta} D_{\beta} = \lambda D_{\beta}$$

This is a so-called eigenvalue problem, with eigenvalues $\lambda$. If we want to solve for the eigenvalues we get

$$\det[\sigma_{\beta} - \lambda I_{\beta}] = 0, \ \text{with } I_{\beta} = \delta_{\beta}$$

This leads to a third order equation in $\lambda$, hence we expect three solutions (roots) $\lambda^{(n)}$. The corresponding eigenvectors can be computed from

$$\sum_{\beta=1}^{l=3} \sigma_{\beta} D^{(n)} = \lambda^{(n)} D^{(n)}.$$ 

The eigenvectors are orthogonal:

$$D^{(b)} D^{(a)} = 0 \ \text{for } \lambda^{(a)} \neq \lambda^{(b)}.$$
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:

\[
\varepsilon_i = \varepsilon_i \delta_{ij}, \quad \sigma_j = \sigma_j \delta_{ij} = \frac{1}{\varepsilon_i} \delta_{ij} \\
\left( \varepsilon_j \right)_j = \varepsilon_i (\omega) 0 0 \\
             0  e_2 (\omega) 0 \\
             0  0  e_3 (\omega) \\
\]

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,...)

\[
\varepsilon_i (\omega) = \varepsilon_i (\omega) = \varepsilon_i (\omega) \Rightarrow D_i = \varepsilon_i (\omega) E_i
\]

Cubic crystals behave like gas, amorphous solids, liquids, and have no anisotropy.

B) uniaxial
- two crystallographic equivalent directions
- trigonal (quartz, lithium niobate), tetragonal, hexagonal

\[
\varepsilon_i (\omega) = \varepsilon_i (\omega) \neq \varepsilon_i (\omega)
\]

6.3 The index ellipsoid

The index ellipsoid offers a simple geometrical interpretation of the inverse dielectric tensor \( \varepsilon = [\varepsilon_j]^{-1} \). The defining equation for the index ellipsoid is

\[
\sum_{i=1}^{3} \sigma_i x_i x_j = 1
\]

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:

\[
\sum_{i=1}^{3} \sigma_i D_i D_j = \varepsilon_0 \sum_{i=1}^{3} E_i D_i = 2 \omega e_k
\]

In the principal coordinate system the defining equation of the index ellipsoid reads:

\[
\frac{\sigma_1 x_1^2 + \sigma_2 x_2^2 + \sigma_3 x_3^2}{\varepsilon_1 + \varepsilon_2 + \varepsilon_3} = 1
\]

This equation can be interpreted as the defining equation of an ellipsoid having semi-principal axes of length \( \sqrt{\varepsilon_i} \). From our discussion of the normal modes in isotropic media we know that \( \sqrt{\varepsilon_i} \) corresponded to the refractive index of the normal modes. We will show in the following discussion that also for anisotropic media, there will be special cases where the \( \sqrt{\varepsilon_i} \) determine the phase velocity of normal modes. Hence the elements of the dielectric tensor which define the semi-principal axes of the epsilon ellipsoid can be related to refractive indexes

\[
n_i = \sqrt{\varepsilon_i}
\]

This is the reason why the ellipsoid, which represents graphically the epsilon tensor, is called index ellipsoid.

C) biaxial
- no crystallographic equivalent directions
- orthorhombic, monoclinic, triclinic
Graphical representation of the epsilon tensor of an anisotropic crystal by the so-called index ellipsoid.

Summary:
- anisotropic media $\rightarrow$ tensor instead of scalar $\Rightarrow$ 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 $n_x = n_z$

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 $\Rightarrow$ most simple solution $\sim \exp\left[i\left(k(\omega)\cdot r - \omega t\right]\right]$
- 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

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

with the dispersion relation

$$k^2(\omega) = \frac{\omega^2}{c^2} \varepsilon(\omega)$$

with $\varepsilon(\omega) > 0$ and real as well as $k \cdot E = k \cdot D = 0$. 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 ($k \rightarrow k_z$). Then, the fields are arbitrary in the x,y-plane

$$D_x, D_y \neq 0$$

and

$$D_z = \varepsilon_x \varepsilon_z E_z$$

In general we have $E \neq D$, but here $k \cdot D = 0 \Rightarrow k \cdot E = 0$, and the two polarization directions ($x = 1, y = 2$) are decoupled:

$$D_x, \varepsilon_1 \cap D_x \exp[i(\omega_1 z - \omega_1 t)] = D_1 \exp[i\varphi_1(z)]\exp(-i\omega_1 t)$$

with $k_1^2 = \frac{\omega_1^2}{c^2} \varepsilon_1(\omega)$

$$D_y, \varepsilon_2 \cap D_y \exp[i(\omega_2 z - \omega_2 t)] = D_2 \exp[i\varphi_2(z)]\exp(-i\omega_2 t)$$

with $k_2^2 = \frac{\omega_2^2}{c^2} \varepsilon_2(\omega)$

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 linear polarization in the direction of a principal axis (x or y) only the phase changes during propagation, thus we found our normal modes:
For light propagation in principle direction we find two perpendicular linearly polarized normal modes with $E \parallel D$. 

**Remark on the indices in the index ellipsoid**

The indices $n_i$ in the index ellipsoid are connected to the indices $n_a$ and $n_b$ of the normal modes propagating along the principal axis. However please be careful about the direction correspondence. For example, the two normal modes propagating along the $z$ direction have phase velocities determined by the indices $n_1 = n_i$ and $n_2 = n_i$ determined by the direction of their electric field rather than by the direction of their propagation.

### 6.4.2 Normal modes for arbitrary propagation direction

#### 6.4.2.1 Geometrical construction

Before we will do the mathematical derivation and actually calculate normal modes and dispersion relation, let us preview the results visualized in the index ellipsoid. Actually, it is possible to construct the normal modes geometrically. We start from the normal modes which we have determined for propagation in the principal directions of the crystal and try to generalize to arbitrary propagation directions:

- For a specific crystal and a given frequency $\omega$ we take the $e_i$ in the principal axis system and construct the index ellipsoid.
- We then fix the propagation direction of the normal mode which we would like to look at $k / k = u$.
- We draw a plane through the origin of index ellipsoid which is perpendicular to $k$.

- 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 $u = k / k$.
- The directions of the principal axes of the index ellipse are the polarization directions of the normal modes $D_a$ and $D_b$.
- The electric field vectors of the normal modes $E_a$ and $E_b$ follow from 

$$E_i = \frac{D_i}{\epsilon_i k_i}.$$ 

Thus, $D^{(a,b)} \parallel E^{(a,b)}$, and $E^{(a,b)}$ are not perpendicular to $k$.

- This has a direct consequence on the pointing vector:

$$\langle S \rangle = \frac{1}{2} \Re \{ E \times H^* \}$$

hence $k$ is not parallel to $\langle S \rangle$ because $\langle S \rangle \perp E$.

- If the index ellipse is a circle, the direction of this particular $k$-vector defines the optical axis of the crystal.
6.4.2.2 Mathematical derivation of dispersion relation

Let us now derive mathematically the dispersion relation for normal modes of the form

\[ E(x, t) = E \exp \left( i \left[ k(\omega) \cdot r - \omega t \right] \right) \]

\[ D(x, t) = D \exp \left( i \left[ k(\omega) \cdot r - \omega t \right] \right) \]

In the isotropic case we found the dispersion relation

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

where the absolute value of the 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 \( \exp \left( i \left[ k(\omega) \cdot r - \omega t \right] \right) \), but the wavenumber depends on the direction \( u \) of propagation

\[ k = k(\omega, u) \]

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 \( k \):

\[ \mathbf{k} = \left( \begin{array}{c} k_1 \\ k_2 \\ k_3 \end{array} \right) = k \left( \begin{array}{c} u_1 \\ u_2 \\ u_3 \end{array} \right) \quad \text{with} \quad u_1^2 + u_2^2 + u_3^2 = 1 \]

Our aim is to derive \( \omega = \omega(k, u_1, u_2, u_3) \) 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:

\[ \mathbf{k} \cdot \mathbf{D} = 0 \quad \mathbf{k} \times \mathbf{E} = \omega \mu_0 \mathbf{H} \]

\[ \mathbf{k} \cdot \mathbf{H} = 0 \quad \mathbf{k} \times \mathbf{H} = -\omega \mathbf{D} \]

Now we follow the usual derivation of the wave equation:

\[ -\left[ \mathbf{k} \times (\mathbf{k} \times \mathbf{E}) \right] = \frac{\omega^2}{c^2} \varepsilon_0 \mathbf{D} \Rightarrow -\mathbf{k}(\mathbf{k} \cdot \mathbf{E}) + k^2 \mathbf{E} = \frac{\omega^2}{c^2} \varepsilon_0 \mathbf{D} \]

- Here \( \mathbf{k} \cdot \mathbf{E} \) does not vanish as in the isotropic case!
- In the principal coordinate system and with \( D_i = \varepsilon_i E_i \), we find

\[ -k_i \sum_j k_j E_j + k^2 E_i = \frac{\omega^2}{c^2} \varepsilon_i E_i \]

\[ \left( \frac{\omega^2}{c^2} \varepsilon_i - k^2 \right) E_i = -k_i \sum_j k_j E_j \]

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:

\[ \frac{\omega^2}{c^2} \varepsilon_i - k^2 \]

\[ k_i \]

\[ \frac{\omega^2}{c^2} \varepsilon_1 - k^2 \]

\[ k_i \]

\[ \frac{\omega^2}{c^2} \varepsilon_2 - k^2 \]

\[ k_i \]

\[ \frac{\omega^2}{c^2} \varepsilon_3 - k^2 \]

\[ k_i \]

\[ \left( \begin{array}{ccc} E_1 \\ E_2 \\ E_3 \end{array} \right) = \left( \begin{array}{ccc} 0 \\ 0 \end{array} \right) \]

The general way to solve this problem is using \( \det \left[ \begin{array}{ccc} k_i \end{array} \right] = 0 \), which gives the dispersion relation \( \omega = \omega(k) \) for given \( k_i / k \).

However, there is an easy way to show some general properties of the dispersion relation. We start from the following trick:

\[ \left( \frac{\omega^2}{c^2} \varepsilon_i - k^2 \right) E_i = -k_i \sum_j k_j E_j \Rightarrow E_i = -\frac{k_i}{\left( \frac{\omega^2}{c^2} \varepsilon_i - k^2 \right)} \sum_j k_j E_j \]

Now we multiply this equation by \( k_i \), perform a summation over the index 'i', and rename \( i \leftrightarrow j \) on the l.h.s:

\[ \sum_j k_i E_j = -\sum_j \left( \frac{\omega^2}{c^2} \varepsilon_i - k^2 \right) k_i E_j \]

Because \( \text{div} \mathbf{E} = \sum_j k_i E_j \neq 0 \) we can divide and get the (implicit) dispersion relation:

\[ \sum_j \frac{k_i^2}{k^2 - \frac{\omega^2}{c^2} \varepsilon_i} = 1 \]

With \( \left[ \begin{array}{ccc} k_1 \\ k_2 \\ k_3 \end{array} \right] = k(\omega) \left( \begin{array}{c} u_1 \\ u_2 \\ u_3 \end{array} \right) = \frac{\omega}{c} n(\omega) \left( \begin{array}{c} u_1 \\ u_2 \\ u_3 \end{array} \right) \) we can write

\[ \sum_j \frac{k_i^2 u_j^2}{k^2 - \frac{\omega^2}{c^2} \varepsilon_i} = 1 \rightarrow \sum_j \frac{u_j^2}{1 - \frac{\varepsilon_i}{n^2}} = 1 \]

\[ \sum_j \frac{u_j^2}{n^2 - \varepsilon_i} = 1 \]

final form of DR

Discussion of results

For given \( \varepsilon_i(\omega) \) and direction \( (u_1, u_2) \) we can compute the refractive index \( n(\omega, u_1, u_2) \) seen by the normal mode. Because \( u_1^2 + u_2^2 + u_3^2 = 1 \), it is sufficient to fix two components \( (u_1, u_2) \) of \( u \) to determine the direction.
A more explicit form of the dispersion relation can be obtained by multiplying with denominators:

$$u_i^2(n^2 - e_1) (n^2 - e_3)n^2 + u_i^2(n^2 - e_1)(n^2 - e_3)n^2 + u_i^2(n^2 - e_1)(n^2 - e_3)n^2 = (n^2 - e_1) (n^2 - e_3) (n^2 - e_1)$$

The resulting equation is quadratic in $n^2$ since the $n^4$-terms cancel. Hence, we get two (positive) solutions $n_a, n_b$ and therefore $k_a = n_a(\omega / c)$ and $k_b = n_b(\omega / c)$ for the two orthogonally polarized normal modes $\mathbf{D}^a$ and $\mathbf{D}^b$.

In particular, for the propagation in direction of the principal axis ($u_1 = 1$ and $u_1 = u_2 = 0$, see 6.4.1) we find:

$$\begin{align*}
(n^2 - e_1)(n^2 - e_3)n^2 &= (n^2 - e_1)(n^2 - e_3)(n^2 - e_1) \\
(n^2 - e_1)(n^2 - e_3)e_3 &= 0 \\
n_a^2 &= e_1, \quad n_b^2 = e_3
\end{align*}$$

Finally, we can derive some properties of the fields of the normal modes, i.e. the eigenfunctions.

We start from the eigenvalue equation, which we had derived above

$$\left(\frac{\omega^2}{c^2} - k^2\right)E_i = -k \sum_j k_j E_j.$$ 

For cases where the first factor of the l.h.s. is unequal zero (propagation directions not parallel to the principal axes) we can divide by this term

$$E_i = -\frac{k_i}{\left(\frac{\omega^2}{c^2} - k^2\right)} \sum_j k_j E_j.$$ 

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 derive a relation of the individual field components from the first part of the equation

$$\begin{align*}
E_1 &: E_2 : E_3 = \frac{k_1}{\omega^2 e_1 - k^2} : \frac{k_2}{\omega^2 e_2 - k^2} : \frac{k_3}{\omega^2 e_3 - k^2} \\
&\quad \Leftrightarrow E_i : E_i = \frac{\omega^2}{c^2} e_i - k^2
\end{align*}$$

and with $D_i = \varepsilon_{ijk} E_j$

$$\begin{align*}
D_1 &: D_2 : D_3 = \frac{\varepsilon_{i1k} k_i}{\omega^2 e_1 - k^2} : \frac{\varepsilon_{i2k} k_i}{\omega^2 e_2 - k^2} : \frac{\varepsilon_{i3k} k_i}{\omega^2 e_3 - k^2}
\end{align*}$$

Please be aware that this relation can only be applied for propagation directions not parallel to the principal axes.

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: $\mathbf{E}^a \cdot \mathbf{E}^b \neq 0$)

$$\mathbf{D}^a \cdot \mathbf{D}^b = \sum \frac{k_i k_j \varepsilon_{ijk}^2}{k_i^2 - \frac{\omega^2}{c^2} e_i} = \frac{c^2}{\omega^2} \frac{k_b}{k_a} \sum k_i^2 \sum \frac{\varepsilon_{ijk}^2}{k_i^2 - \frac{\omega^2}{c^2} e_i}$$

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:

$$1 = \sum \frac{k_i^2 \varepsilon_{ijk}^2}{k_i^2 - \frac{\omega^2}{c^2} e_i} = 1 + \frac{\omega^2}{c^2} \sum \frac{\varepsilon_{ijk}^2}{k_i^2 - \frac{\omega^2}{c^2} e_i}$$

6.4.3 Normal surfaces of normal modes

In addition to the index ellipsoid, which is a graphical representation of the material properties of crystals from which the properties of the normal modes can be interpreted as shown above, we can derive a direct graphical representation of the dispersion relation of normal modes in crystals. This graphical representation of the dispersion relation is called normal surfaces:

If we plot the refractive indices (wave number or norm of the k-vector divided by $k_0$) of the normal modes in the $k_i$-space (normal surfaces), we get a centro-symmetric, two layer surface.
Normal surfaces as the graphical representation of the dispersion relation of normal modes in crystals.

**isotrop:** sphere

**uniaxial:** 2 points with \( n_a = n_b \) in the poles → connecting line defines the optical axis (for \( e_i = e_2 = e_{ex} \), \( e_1 = e_a \), the z-axis is the optical axis)

**biaxial:** 4 points with \( n_a = n_b \) → connecting lines define two optical axes

How to read the figure:
- fix propagation direction \((u, u_2)\) → intersection with surfaces
- distances from origin to intersections with surfaces correspond to refractive indices of normal modes
- definition of optical axis → \( n_a = n_b \)

**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 \( n_a, n_b \) (refractive indices of the normal modes)
- optical axis \( \sim \) 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 \( n_a, n_b \)
- optical axis connects points with \( n_a = n_b \)

**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 investigate the special and 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 on plane. 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 rotationally symmetric around the z-axis, and we have \( e_i = e_2 = e_{ex}, e_1 = e_a \), which we call ordinary and extraordinary refractive indices.

Then, we expect two normal modes:

- **A)** ordinary wave → \( n_a \) independent of propagation direction
- **B)** extraordinary wave → \( n_b \) depends on propagation direction

The z-axis is, according to definition, the optical axis with \( n_a = n_b \).

- The ordinary wave \( \mathbf{D}^{(o)} \) 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 \( \mathbf{D}^{(o)} \).

Let us now derive the dispersion relation: From above we know the implicit form

$$
\sum \left[ \frac{u^2}{n^2 - e_i} \right] = \frac{1}{n^2}
$$

For uniaxial crystals this leads to

$$
\frac{u_1^2}{n^2 - e_{ex}} + \frac{u_2^2}{n^2 - e_{ex}} = \frac{1}{n^2}
$$

$$
n^2 \left[ n^2 - e_{ex} \right] \left( u_1^2 + u_2^2 \right) + n^2 \left( n^2 - e_{ex} \right) u_1^2 = \left[ n^2 - e_{ex} \right] \left( n^2 - e_{ex} \right)^2
$$

**A)** ordinary wave: independent of direction

\( n_a^2 = e_{ex} \) \( \rightarrow \) \( k_a^2 = \frac{\omega^2}{c^2} n_a^2 = k_{ex}^2 e_{ex} \)
B) extraordinary wave (derivation is your exercise): dependent on direction

\[ \left( \frac{u_1^2 + u_3^2}{e_a} \right) + \frac{u_2^2}{e_{ae}} = \frac{1}{n_a^2}, \quad k_a^2 = \frac{\omega^2}{c^2} n_a^2 (u_1, u_2, u_3) \]

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 \( k_2, k_3 \) or \( y, z \) plane (\( u_1 = 0 \)).

\[ \text{Normal surfaces for a uniaxial crystal.} \]

We have with

\[ k_i^2 = k_0^2 n_i^2 u_i \]

A) ordinary wave

\[ k_a^2 = k_1^2 + k_2^2 + k_3^2 = k_0^2 e_{ae} \]

B) extraordinary wave

\[ \frac{1}{e_a} \left( k_1^2 + k_2^2 \right) + \frac{1}{e_{ae}} k_3^2 = 1 \]

What about the fields? We know from before that

\[ D_1 : D_2 : D_3 = \frac{e_{ae} k_1}{\varepsilon_0 \varepsilon_{ae} - k^2} : \frac{e_{ae} k_2}{\varepsilon_0 \varepsilon_{ae} - k^2} : \frac{e_{ae} k_3}{\varepsilon_0 \varepsilon_a - k^2} \]

For the extraordinary wave all denominators are finite, and in particular \( k_i = 0 \) implies \( D^{(e)} = 0 \), hence \( D^{(e)} \) is polarized in the \( y \)-\( z \) plane. Then, \( D^{(or)} \perp D^{(e)} \) implies that \( D^{(or)} \) is polarized in x-direction.

In summary, we find for the polarizations of the fields:

A) ordinary: \( D \perp k, D \parallel E \)

B) extraordinary: \( D \perp k \) and in the plane \( k \)-optical axis

\[ D \parallel \mathbf{E} \]

If we introduce an angle \( \Theta \), as in the figures below, to describe the propagation direction, a simple computation of \( n_a^2(\Theta) \) for the extraordinary wave is possible (exercise):

\[ n_a^2(\Theta) = \frac{\varepsilon_{ae}^2 \varepsilon_a}{\varepsilon_{ae}^2 \sin^2 \Theta + \varepsilon_a \cos^2 \Theta} \]

The following classification for uniaxial crystals is commonly used

\( \varepsilon_{ae} > \varepsilon_a \rightarrow \text{negative uniaxial} \)

\( \varepsilon_{ae} < \varepsilon_a \rightarrow \text{positive uniaxial} \)
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 subsequent interfaces)
- system of layers (arbitrary number of subsequent interfaces)

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 in inhomogeneous space \( \rightarrow \) “guided” waves (propagation of confined light beams without 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 homogeneous space \( \rightarrow \) no interaction of normal modes 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 at the inhomogeneities 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

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

In general, for isotropic media all field components are coupled due to the rotrot 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 \( \rightarrow \partial / \partial y = 0 \)

\[
\text{rot rot } \mathbf{E} = \text{grad div } \mathbf{E} - \Delta \mathbf{E} = \begin{pmatrix} \frac{\partial}{\partial x} \left( \frac{\partial E_x}{\partial x} + \frac{\partial E_z}{\partial z} \right) \\ 0 \\ \frac{\partial}{\partial z} \left( \frac{\partial E_x}{\partial x} + \frac{\partial E_z}{\partial z} \right) \end{pmatrix} - \begin{pmatrix} \Delta^{(1)} E_x \\ \Delta^{(2)} E_x \\ \Delta^{(2)} E_z \end{pmatrix}
\]

Then, we can split the electric field as \( \mathbf{E} = \mathbf{E}_\perp + \mathbf{E}_\parallel \) with

\[
\mathbf{E}_\perp = \begin{pmatrix} \frac{\partial E_x}{\partial x} \\ 0 \\ \frac{\partial E_z}{\partial z} \end{pmatrix}, \quad \mathbf{E}_\parallel = \begin{pmatrix} E_x \\ 0 \\ E_z \end{pmatrix}, \quad V^{(2)} = \begin{pmatrix} \frac{\partial}{\partial x} \\ 0 \\ \frac{\partial}{\partial z} \end{pmatrix}, \quad \Lambda^{(2)} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}
\]

\( \mathbf{E}_\perp \) is polarized perpendicular to the plane of propagation, \( \mathbf{E}_\parallel \) is polarized parallel to this plane.

Common notations are:
- perpendicular: \( \perp \rightarrow s \rightarrow \text{TE} \) (transversal electric)
- parallel: \( \parallel \rightarrow p \rightarrow \text{TM} \) (transversal magnetic)

Both components are decoupled and can be treated independently:

\[
\Lambda^{(3)} \mathbf{E}_\perp + \frac{\omega^2}{c^2} \mathbf{E}_\perp(r, \omega) = -\mathbf{j}(r, \omega) + \mathbf{P}_\perp(r, \omega) - \mu\omega^2 \mathbf{P}_\parallel(r, \omega)
\]

\[
\Lambda^{(2)} \mathbf{E}_\parallel + \frac{\omega^2}{c^2} \mathbf{E}_\parallel(r, \omega) - \text{grad}^{(2)} \text{div}^{(2)} \mathbf{E}_\parallel = -\mathbf{j}(r, \omega) + \mathbf{P}_\perp(r, \omega) - \mu\omega^2 \mathbf{P}_\parallel(r, \omega)
\]

From

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

we can conclude that the corresponding magnetic fields are
\[ \mathbf{E}_{\text{TE}} = \begin{pmatrix} 0 \\ E_y \\ 0 \end{pmatrix}, \quad \mathbf{H}_{\text{TE}} = \begin{pmatrix} H_x \\ 0 \\ H_z \end{pmatrix} \]

\[ \mathbf{E}_{\text{TM}} = \begin{pmatrix} E_x \\ 0 \\ E_z \end{pmatrix}, \quad \mathbf{H}_{\text{TM}} = \begin{pmatrix} 0 \\ H_y \\ 0 \end{pmatrix} \]

7.1.3 Interfaces and symmetries

Up to now we treated plane waves of the form

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

- homogeneous space implies: \( \exp(\mathbf{i} \mathbf{k} \cdot \mathbf{r}) \)
- monochromaticity leads to: \( \exp(-\mathbf{i} \omega t) \)

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

Now, we will break the homogeneity in x-direction by considering an interface in y-z-plane which is infinite in y and z.

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

7.1.4 Transition conditions

From Maxwell's equations follow transition conditions for the field components. Here we will use that \( E_i, H_i \) (transverse components) are continuous at an interface between two media. This implies for the:

A) Continuity of fields
   - TE: \( E = E_i \) and \( H = H_i \) continuous
   - TM: \( E = E_i \) and \( H = H_i \) continuous

B) Continuity of wave vectors
   homogeneous in z-direction \( \rightarrow \) phase \( e^{\mathbf{i} \mathbf{k} \cdot \mathbf{r}} \rightarrow k \), 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 \( \varepsilon(\omega) \).

- 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, \( E(x,z;\omega) \), \( H(x,z;\omega) \)) and in the following we will often omit \( \omega \) in the notation.

TE-polarization

We have to solve the wave equation (no \( y \)-dependence because of translational invariance):

\[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + \frac{\omega^2}{c^2} \varepsilon_i E_i(x,z) = 0 \]

We use the ansatz from above:

\[ E_{\text{TE}}(x,z) = E_{\text{TE}}(x) \exp(i k_z z) \text{ and } H_{\text{TE}}(x,z) = H_{\text{TE}}(x) \exp(i k_z z) \]

\[ \left[ \frac{\partial^2}{\partial x^2} + \frac{\omega^2}{c^2} \varepsilon_i - k_z^2 \right] E_{\text{TE}}(x) = 0 \]

with: \( H_{\text{TE}}(x,z) = -\frac{i}{\omega \varepsilon_i} \text{rot} E_{\text{TE}}(x,z) \)
Now let us extract the equations for transversal fields $E_y = E_y$, $H_z$:

$$\left[ \frac{\partial^2}{\partial x^2} + k_0^2 (k_z, \omega) \right] E(x) = 0 \text{ with } k_0^2 (k_z, \omega) = \frac{\omega^2}{c^2} \varepsilon_f (\omega) - k_x^2$$

$$H_z(x) = -i \frac{\partial}{\omega \mu_0 \partial x} E(x)$$

This makes sense since the wave equation for the y-component of the electric field is a second order differential equation. Hence we need to specify the field and its first derivative as initial condition at $x = 0$ to determine a unique solution.

**TM-polarization**

analog for transversal components $H_y = H_y$, $E_z$:

$$\left[ \frac{\partial^2}{\partial x^2} + k_0^2 (k_z, \omega) \right] H(x) = 0$$

$$E_z(x) = -i \frac{\partial}{\omega \varepsilon_0 \varepsilon_f \partial x} H(x)$$

Again, we succeed describing everything in transversal components.

Now 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 $x = 0$
- calculate the fields at $x = d$
- 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

$E, H \rightarrow F$  \hspace{1cm} \text{generalized field 1}$

$i \omega \mu_0 H_y$, $-i \omega \varepsilon_0 E_z \rightarrow G$  \hspace{1cm} \text{generalized field 2}$

and write down the problem to solve:

$$\left[ \frac{\partial^2}{\partial x^2} + k_0^2 (k_z, \omega) \right] F(x) = 0$$

$$G(x) = \alpha_{\text{TE}} \frac{\partial}{\partial x} F(x)$$

with $\alpha_{\text{TE}} = 1$, $\alpha_{\text{TM}} = \frac{1}{\varepsilon_f}$

We know the general solution of this system (harmonic oscillator equation):

$$F(x) = C_1 \exp(ik_{\text{Tx}} x) + C_2 \exp(-ik_{\text{Tx}} x)$$

$$G(x) = \alpha_{\text{TE}} \frac{\partial}{\partial x} F(x) = i \alpha_{\text{TE}} k_{\text{Tx}} \left[ C_1 \exp(ik_{\text{Tx}} x) - C_2 \exp(-ik_{\text{Tx}} x) \right]$$

We have as initial conditions $F(0), G(0)$ given:

$$F(0) = C_1 + C_2$$

$$G(0) = i \alpha_{\text{TE}} k_{\text{Tx}} \left[ C_1 - C_2 \right]$$

from which we can compute the constants $C_1, C_2$:

$$C_1 = \frac{1}{2} \left[ F(0) - \frac{i}{\alpha_{\text{TE}} k_{\text{Tx}}} G(0) \right]$$

$$C_2 = \frac{1}{2} \left[ F(0) + \frac{i}{\alpha_{\text{TE}} k_{\text{Tx}}} G(0) \right]$$

The final solution of the initial value problem is therefore:

$$F(x) = \cos(k_{\text{Tx}} x) F(0) + \frac{1}{\alpha_{\text{TE}} k_{\text{Tx}}} \sin(k_{\text{Tx}} x) G(0)$$

$$G(x) = -\alpha_{\text{TE}} k_{\text{Tx}} \sin(k_{\text{Tx}} x) F(0) + \cos(k_{\text{Tx}} x) G(0)$$

By resubstituting we have the electromagnetic field in the layer $0 \leq x \leq d$.

### 7.2.2 The fields in a system of layers

In the previous subchapter we have seen how to compute the electromagnetic field in a single dielectric layer, dependent on the transverse field components $E_y$, $H_z$ (TE) and $H_y$, $E_z$ (TM) at $x = 0$. We can generalize our results to systems of dielectric layers, which are used in many optical devices:

- Bragg mirrors
- chirped mirrors for dispersion compensation
- interferometer
- multi-layer waveguides
- Bragg waveguide
- metallic interfaces and layers

We can even go further and “discretize” an arbitrary inhomogeneous (in one dimension) refractive index distribution. This is important for so-called 'GRIN' - Graded-Index-Profiles.
From above, we know the fields in one layer:

\[
F(x) = \cos(k_{tx}x)F(0) + \frac{1}{\alpha_t k_{tx}}\sin(k_{tx}x)G(0)
\]

\[
G(x) = -\alpha_t k_{tx}\sin(k_{tx}x)F(0) + \cos(k_{tx}x)G(0)
\]

We can write this formally in matrix notation as

\[
\begin{bmatrix}
F(x) \\
G(x)
\end{bmatrix} = \hat{m}(x) \begin{bmatrix}
F(0) \\
G(0)
\end{bmatrix}
\]

where the 2x2-matrix \( \hat{m} \) describes propagation of the fields:

\[
\hat{m}(x) = \begin{bmatrix}
\cos(k_{tx}x) & \frac{1}{k_{tx}\alpha_t}\sin(k_{tx}x) \\
-k_{tx}\alpha_t\sin(k_{tx}x) & \cos(k_{tx}x)
\end{bmatrix}
\]

- To compute the field at the end of the layer we set \( x = d \).
- We assume no absorption in the layer \( ||\hat{m}(x)|| = 1 \).
- A system of layers is characterized by \( c_i, d_i \).

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 since it just requires matrix multiplications:

**A) Two layers**

\[
\begin{bmatrix}
F \\ G
\end{bmatrix}_{d_1 + d_2} = \hat{m}_1(d_1) \begin{bmatrix}
F \\ G
\end{bmatrix}_{d_1} = \hat{m}_2(d_2) \hat{m}_1(d_1) \begin{bmatrix}
F \\ G
\end{bmatrix}_{d_1}
\]

**B) N layers**

\[
\begin{bmatrix}
F \\ G
\end{bmatrix}_{d_1 + d_2 + \ldots + d_N} = \prod_{i=1}^{N} \hat{m}_i(d_i) \begin{bmatrix}
F \\ G
\end{bmatrix}_{d_1} = \hat{M} \begin{bmatrix}
F \\ G
\end{bmatrix}_{d_1}
\]

with \( \hat{M} = \prod_{i=1}^{N} \hat{m}_i(d_i) \)

All matrices \( \hat{m}_i \) have the same form, but different \( \alpha'_i, d_i, k_{tx}'_i = \sqrt{\frac{\epsilon_0^2 c_0^2}{\epsilon_i} - k^2} \).

**Summary of matrix method**

- \( F(0) \) and \( G(0) \) given \( F, H \) for TE, \( E_z, H \) for TM
- \( k, \alpha'_i, c_i, d_i \) 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 \( \rightarrow \) 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 field on the other side. We have seen that after splitting in the TE(TM)-polarizations, continuous (transversal) field components are sufficient to describe the whole field. What we will do now is to link those field components with the fields, which are accessible in an experimental configuration, 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 the incident \( (k_z) \), reflected \( (k_z') \) and transmitted \( (k_z'') \) fields:

\[
\begin{bmatrix}
k_x'' \\ k_x'
\end{bmatrix} = \begin{bmatrix}
\frac{k_{sx}}{} \\ 0
\end{bmatrix}, \quad \begin{bmatrix}
k_x' \\ 0
\end{bmatrix}, \quad \begin{bmatrix}
k_x'' \\ k_x'
\end{bmatrix}
\]

with

\[
k_{sx} = \sqrt{\frac{\epsilon_0^2 c_0^2}{\epsilon_s} - k^2} = \sqrt{k^2(\omega) - k^2}, \quad k_x' = \sqrt{\frac{\epsilon_0^2 c_0^2}{\epsilon_c} - k^2} = \sqrt{k^2(\omega) - k^2},
\]

where \( \epsilon_s(\omega) \) and \( \epsilon_c(\omega) \) are the dielectric functions of the substrate and cladding and \( k_z \) is the tangential component of the wave vector which is continuous throughout the layer system.
As we have seen before, the $k_z$ component of the wave vector is conserved and $k_{\pm}$ determines the direction of the wave (forward or backward). The total length of the wave vector in each layer is given by the dispersion relation for dispersive, isotropic, homogeneous media. As a consequence, the $k_x$ component changes its value in each layer.

**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. $k_x \sin \varphi_x = k_x \sin \varphi_x \cap \varphi_z = \varphi_n$ (reflection)
2. $k_x \sin \varphi_x = k_x \sin \varphi_x \cap n_x \sin \varphi_z = n_x \sin \varphi_x$ (Snellius)

Let us now rewrite the fields in order to solve the reflection transmission problem:

**A) Field in substrate**

The fields in the substrate can be expressed based on the complex amplitudes of the incident $F_z$ and reflected field $F_n$ as:

$$F_z(x,z) = \exp(i k_z z) \left[ F_x \exp(i k_{ax} x) + F_n \exp(-i k_{ax} x) \right]$$

$$G_z(x,z) = i \alpha \alpha k_{ax} \exp(i k_z z) \left[ F_x \exp(i k_{ax} x) - F_n \exp(-i k_{ax} x) \right]$$

**B) Field in layer system**

The fields inside the layer system can be expressed as

$$F_i(x,z) = \exp(i k_z z) F(x)$$

$$G_i(x,z) = \exp(i k_z z) G(x)$$

where the amplitudes $F(x)$ and $G(x)$ are given by matrix method as

$$\begin{pmatrix} F \\ G \end{pmatrix}_x = M(x) \begin{pmatrix} F \\ G \end{pmatrix}_0$$

**C) Field in cladding**

The fields in the cladding can be expressed based on the complex amplitude of the transmitted field $F_T$ as:

$$F_z(x,z) = \exp(i k_{ax} z) F_x \exp\left[ i k_{ax} \left( x - D \right) \right]$$

$$G_z(x,z) = i \alpha \alpha k_{ax} \exp(i k_{ax} z) F_x \exp\left[ i k_{ax} \left( x - D \right) \right]$$

Note that in the cladding we consider a forward (transmitted) wave only.

**Reflection transmission problem**

We want to compute $F_R$ and $F_T$ for given $F_I$, $k_z$ ($\sin \varphi_z$), $k_x$, $d$. We know that $F$ and $G$ are continuous at the interfaces, in particular at $x = 0$ and $x = D$.

We have:

$$\begin{pmatrix} F \\ G \end{pmatrix}_D = M(D) \begin{pmatrix} F \\ G \end{pmatrix}_0$$

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:

$$\begin{pmatrix} F_T \\ 0 \end{pmatrix} = \begin{pmatrix} M(I) & M(D) \\ 0 & M(D) \end{pmatrix} \begin{pmatrix} F_I \\ F_R \end{pmatrix}$$

We consider $F_T$ as known, and $F_R$ and $F_I$ as unknown and get:

$$F_R = \frac{(\alpha_k M_{22} - \alpha_k M_{12}) M_1 - \alpha_k \alpha_k M_{11} M_{22}}{(\alpha_k M_{22} + \alpha_k M_{12}) + \alpha_k \alpha_k M_{11}} F_T$$

$$F_I = \frac{2 \alpha_k M_{12} (M_{11} M_{22} - M_{12} M_{21})}{\alpha_k M_{22} + \alpha_k M_{12} + i(M_{21} - \alpha_k \alpha_k M_{11})} F_T$$

Those are the general formulas for reflected and transmitted amplitudes. Please remember that the matrix elements depend on the polarization direction $\rightarrow M^{TE} \neq M^{TM}$.

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

\[ F = E = E_y, \quad \alpha_{TE} = 1 \]

1) reflected field

\[ E_{rE}^{TE} = R_{TE}^{TE} E_i^{TE} \]

with the reflection coefficient

\[
R_{TE}^{TE} = \frac{\left( k_{ex} M_{22}^{TE} - k_{cx} M_{11}^{TE} \right) - i \left( M_{21}^{TE} + k_{ex} k_{cx} M_{12}^{TE} \right)}{\left( k_{ex} M_{22}^{TE} + k_{cx} M_{11}^{TE} \right) + i \left( M_{21}^{TE} - k_{ex} k_{cx} M_{12}^{TE} \right)}
\]

2) transmitted field

\[ E_{tE}^{TE} = T_{TE}^{TE} E_i^{TE} \]

with the transmission coefficient

\[
T_{TE}^{TE} = \frac{2k_{ex}}{k_{ex} M_{22}^{TE} + k_{cx} M_{11}^{TE}} = \frac{2k_{ex}}{N_{TE}^{TE}}.
\]

We get complex coefficients for reflection and transmission, which determine the amplitude and phase of the reflected and transmitted light.

B) TM-polarization

\[ F = H = H_y, \quad \alpha_{TM} = \frac{1}{\varepsilon} \]

In the case of TM polarization we have the problem that an analog calculation to TE would lead to \( H_{rT}/H_i \), i.e., relations between the magnetic field. However, we want \( E_{rT}/E_i \). Therefore, we have to convert the \( H \)-field to the \( E \)-field:

\[
\frac{E_x}{E_y} = \frac{-\sin \phi}{k} = \frac{-k_{ex}}{k}
\]

\[
\bigcap E_{TM} = \frac{k}{k_{ex}} E_x,
\]

With Maxwell we can link \( E \) to \( H \):

\[
E = -\frac{1}{\omega \varepsilon_0 c} (k \times H) \quad \Rightarrow \quad E_x = \frac{1}{\omega \varepsilon_0 c} k_x H_y \quad \Rightarrow \quad E_{TM} = -\frac{k}{k_{ex}} H_y = -\frac{1}{\varepsilon_0 c} \sqrt{\varepsilon_x/\varepsilon_{TM}}
\]

relevant for transmission only

Hence we find the following for TM polarization:

\[ E_{rT}^{TM} = R_{TM}^{TM} E_i^{TM} \]

with the reflection coefficient

\[
R_{TM}^{TM} = \frac{\left( \varepsilon_{ex} k_{ex} M_{22}^{TM} - \varepsilon_{cx} k_{cx} M_{11}^{TM} \right) - i \left( \varepsilon_{ex} k_{ex} M_{21}^{TM} + k_{ex} k_{cx} M_{12}^{TM} \right)}{\left( \varepsilon_{ex} k_{ex} M_{22}^{TM} + \varepsilon_{cx} k_{cx} M_{11}^{TM} \right) + i \left( \varepsilon_{ex} k_{ex} M_{21}^{TM} - k_{ex} k_{cx} M_{12}^{TM} \right)}
\]

\[ E_{tT}^{TM} = T_{TM}^{TM} E_i^{TM} \]

with the transmission coefficient

\[
T_{TM}^{TM} = \frac{2\sqrt{\varepsilon_x/\varepsilon_{TM}}}{\varepsilon_{ex} k_{ex} M_{22}^{TM} + \varepsilon_{cx} k_{cx} M_{11}^{TM}} = \frac{2\sqrt{\varepsilon_x/\varepsilon_{TM}}}{N_{TM}^{TM}}.
\]

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

\[
R_{TE}^{TE} = \frac{\left( k_{ex} M_{22}^{TE} - k_{cx} M_{11}^{TE} \right) - i \left( M_{21}^{TE} + k_{ex} k_{cx} M_{12}^{TE} \right)}{\left( k_{ex} M_{22}^{TE} + k_{cx} M_{11}^{TE} \right) + i \left( M_{21}^{TE} - k_{ex} k_{cx} M_{12}^{TE} \right)}
\]

\[
T_{TE}^{TE} = \frac{2k_{ex}}{N_{TE}^{TE}}
\]

\[
R_{TM}^{TM} = \frac{\left( \varepsilon_{ex} k_{ex} M_{22}^{TM} - \varepsilon_{cx} k_{cx} M_{11}^{TM} \right) - i \left( \varepsilon_{ex} k_{ex} M_{21}^{TM} + k_{ex} k_{cx} M_{12}^{TM} \right)}{\left( \varepsilon_{ex} k_{ex} M_{22}^{TM} + \varepsilon_{cx} k_{cx} M_{11}^{TM} \right) + i \left( \varepsilon_{ex} k_{ex} M_{21}^{TM} - k_{ex} k_{cx} M_{12}^{TM} \right)}
\]

\[
T_{TM}^{TM} = \frac{2\sqrt{\varepsilon_x/\varepsilon_{TM}}}{N_{TM}^{TM}}
\]
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:

\[ \langle S \rangle e_x = \frac{1}{2} \text{Re}(E \times H^*) e_x \]

With

\[ H^* = \frac{1}{\omega \mu_0} (k^* \times E^*) \]

we find

\[ \langle S \rangle e_x = \frac{1}{2 \omega \mu_0} \text{Re}(k^* e_x) |E|^2 = \frac{1}{2 \omega \mu_0} \text{Re}(k_x) |E|^2. \]

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

\[ k_{ss} = \sqrt{\frac{\omega^2}{c^2} \varepsilon_s - k_s^2} = \sqrt{k_s^2(\omega) - k_s^2} \]

is supposed to be real-valued, because we have our incident wave coming from there. The total energy flux from the substrate to the layer system is given as

\[ \langle S \rangle e_x = \frac{1}{2 \omega \mu_0} k_{ss} |E_s|^2 - k_{ss} |E_n|^2 \]

In contrast, in the cladding

\[ k_{cs} = \sqrt{\frac{\omega^2}{c^2} \varepsilon_c - k_s^2} = \sqrt{k_s^2(\omega) - k_s^2} \]

may be complex-valued. The energy flux from the layer system into the cladding is

\[ \langle S \rangle e_x = \frac{1}{2 \omega \mu_0} \text{Re}(k_{cs}) |E_c|^2. \]

Because we have energy conservation:

\[ \langle S \rangle e_x = \langle S \rangle e_x \cap \]

\[ |E|^2 = |E_s|^2 + \frac{\text{Re}(k_{cs})}{k_{cs}} |E_c|^2. \]

Now we 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 \( \rho_{TE, TM} \) and transmissivities \( \tau_{TE, TM} \). We know:

\[ E_n = E_n^{TE} + E_n^{TM}, \quad E_s = E_s^{TE} + E_s^{TM} \]

\[ |E_s|^2 = |E_s^{TE}|^2 + |E_s^{TM}|^2 + \frac{\text{Re}(k_{cs})}{k_{cs}} |E_c^{TE}|^2 + |E_c^{TM}|^2 \]

\[ = \left[ |R_{TE}^s|^2 + \frac{\text{Re}(k_{cs})}{k_{cs}} |I_{TE}^c|^2 \right] |E_s^{TE}|^2 + \left[ |R_{TM}^s|^2 + \frac{\text{Re}(k_{cs})}{k_{cs}} |I_{TM}^c|^2 \right] |E_s^{TM}|^2. \]

Here, we just substituted the reflected and transmitted field amplitudes by incident amplitudes times Fresnel coefficients. Now, we decompose the incident field as follows:

\[ E_{ci} = |E_{ci}| e^{i \delta} \]

and

\[ E_{c1} = |E_{c1}| e^{i \delta} \]

Then, we can divide by the (arbitrary) amplitude \( |E_s|^2 \) and write

\[ 1 = \left( |R_{TE}^s|^2 + \frac{\text{Re}(k_{cs})}{k_{cs}} |I_{TE}^c|^2 \right) \cos^2 \delta + \left( |R_{TM}^s|^2 + \frac{\text{Re}(k_{cs})}{k_{cs}} |I_{TM}^c|^2 \right) \sin^2 \delta \]

\[ 1 = \left( |R_{TE}^s|^2 \cos^2 \delta + |R_{TM}^s|^2 \sin^2 \delta \right) + \frac{\text{Re}(k_{cs})}{k_{cs}} \left( |I_{TE}^c|^2 \cos^2 \delta + |I_{TM}^c|^2 \sin^2 \delta \right) \]

The red and blue terms can be identified as

\[ 1 = \rho + \tau \]

The global reflectivity and transmissivity are therefore given as

\[ \rho = \rho_{TE} \cos^2 \delta + \rho_{TM} \sin^2 \delta \]

\[ \tau = \tau_{TE} \cos^2 \delta + \tau_{TM} \sin^2 \delta \]

with the reflectivities

\[ \rho_{TE, TM} = |R_{TE, TM}^s|^2, \quad \tau_{TE, TM} = \frac{\text{Re}(k_{cs})}{k_{cs}} |I_{TE, TM}^c|^2 \]

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):

\[
\begin{pmatrix}
  k_x \\
  k_y \\
  k_z
\end{pmatrix} = \begin{pmatrix}
  k_{\text{ex}} \\
  k_{\text{ey}} \\
  k_{\text{ez}}
\end{pmatrix},
\begin{pmatrix}
  -k_{\text{ex}} \\
  k_{\text{ey}} \\
  k_{\text{ez}}
\end{pmatrix},
\begin{pmatrix}
  k_{\text{ex}} \\
  -k_{\text{ey}} \\
  k_{\text{ez}}
\end{pmatrix},
\begin{pmatrix}
  k_{\text{ex}} \\
  k_{\text{ey}} \\
  -k_{\text{ez}}
\end{pmatrix},
\begin{pmatrix}
  k_{\text{ex}} \\
  -k_{\text{ey}} \\
  -k_{\text{ez}}
\end{pmatrix},
\begin{pmatrix}
  k_{\text{ex}} \\
  k_{\text{ey}} \\
  -k_{\text{ez}}
\end{pmatrix},
\begin{pmatrix}
  -k_{\text{ex}} \\
  k_{\text{ey}} \\
  -k_{\text{ez}}
\end{pmatrix},
\begin{pmatrix}
  -k_{\text{ex}} \\
  -k_{\text{ey}} \\
  k_{\text{ez}}
\end{pmatrix}
\]

The continuous component of the wave vector, expressed in terms of the angle of incidence, is

\[
k_\perp = \frac{\omega}{c} \sqrt{\varepsilon_\perp} \sin \phi_\perp = \frac{\omega}{c} n_\perp \sin \phi_\perp.
\]

Then, the discontinuous component is given as

\[
k_\parallel = \frac{\omega^2}{c^2} \varepsilon_\parallel - k_\perp^2 = \frac{\omega^2}{c^2} \varepsilon_\parallel - \frac{\omega^2}{c^2} \varepsilon_\perp \sin^2 \phi_\perp = \frac{\omega}{c} \sqrt{n_\parallel^2 - n_\perp^2 \sin^2 \phi_\perp} \cos \phi_\perp,
\]

As above, we can assume that \( k_\parallel \) is always real, because otherwise we have no incident wave. \( k_\parallel \) is real for \( n_\parallel > n_\perp \sin \phi_\perp \) (total internal reflection).

The matrix for a single interface is the unit matrix

\[
\hat{\mathbf{M}} = \hat{\mathbf{m}}(d = 0) = \begin{pmatrix}
  1 & 0 \\
  0 & 1
\end{pmatrix}
\]

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

\[
R_{\text{TE}} = \frac{(k_{\text{ex}} M_{22} - k_{\text{ey}} M_{11}) - \mathbf{i} (M_{21} + k_{\text{ex}} k_{\text{ey}} M_{12})}{(k_{\text{ex}} M_{22} - k_{\text{ey}} M_{11}) + \mathbf{i} (M_{21} - k_{\text{ex}} k_{\text{ey}} M_{12})}, \quad T_{\text{TE}} = \frac{2k_{\text{ex}}}{N_{\text{TE}}}
\]

\[
R_{\text{TM}} = \frac{(k_{\text{ex}} - k_{\text{ey}}) - n_\parallel \cos \phi_\perp - \sqrt{n_\parallel^2 - n_\perp^2 \sin^2 \phi_\perp}}{(k_{\text{ex}} + k_{\text{ey}}) + n_\parallel \cos \phi_\perp + \sqrt{n_\parallel^2 - n_\perp^2 \sin^2 \phi_\perp}} = \frac{2n_\parallel \cos \phi_\perp}{n_\parallel \cos \phi_\perp + n_\perp \cos \phi_\perp}
\]

\[
T_{\text{TM}} = \frac{2k_{\text{ex}}}{(k_{\text{ex}} + k_{\text{ey}})^2} = \frac{2n_\parallel \cos \phi_\perp}{n_\parallel \cos \phi_\perp + n_\perp \cos \phi_\perp}
\]

B) TM-Polarisation

\[
R_{\text{TM}} = \frac{(k_{\text{ex}} M_{22} - k_{\text{ey}} M_{11}) - \mathbf{i} (M_{21} + k_{\text{ex}} k_{\text{ey}} M_{12})}{(k_{\text{ex}} M_{22} - k_{\text{ey}} M_{11}) + \mathbf{i} (M_{21} - k_{\text{ex}} k_{\text{ey}} M_{12})}, \quad T_{\text{TM}} = \frac{2k_{\text{ex}}}{N_{\text{TM}}}
\]

with: \( \hat{\mathbf{M}}(d = 0) = \begin{pmatrix}
  1 & 0 \\
  0 & 1
\end{pmatrix} \)
Remark
It may seem that we have a problem for $\varphi_0 = 0$. For $\varphi_0 = 0$, TE and TM polarization should be equivalent, because the fields are always polarized parallel to the interface. However, formally we have $R_{TM} = -R_{TE}$, $T_{TE} = T_{TM}$. The “strange” behavior of the coefficient of reflection can be explained by the following figures:

### 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.

$$\rho_{TE} = \frac{k_{sx} - k_{cx}}{k_{sx} + k_{cx}}$$

$$\rho_{TM} = \frac{k_{sx} e_s - k_{cx} e_s}{k_{sx} e_s + k_{cx} e_s}$$

With $k_{cx} = \frac{\omega}{c} \sqrt{n_c^2 - n_s^2 \sin^2 \varphi_1}$ we can compute the smallest angle of incidence with $\rho_{TE, TM} = 1$:

$$k_{cx} = 0 \Leftrightarrow n_s = n_s \sin \varphi_{\text{tot}}$$

For angles of incidence larger than this limit angle, $\varphi_1 > \varphi_{\text{tot}}$, we have

$$k_{cx} = \frac{i \omega}{c} \sqrt{n_s^2 \sin^2 \varphi_1 - n_c^2} = \frac{i \mu_c}{k_{sx}} = \frac{i}{i^2} \frac{k_{sx} - \mu_c}{k_{sx} + \mu_c} \Leftrightarrow \text{imaginary}$$

$$\rightarrow \Re (k_{cx}) = 0 \rightarrow \text{TIR}$$

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):

$$\rho_{TE} = \frac{k_{sx} - i \mu_c}{k_{sx} + i \mu_c} = 1$$

$$\tau_{TE} = \frac{4 k_{sx} \Re (k_{cx})}{k_{sx} + k_{cx}} = 0.$$  

### A) TE-polarization

$$R_{TE} = 1 \cdot \exp (i \Theta_{\text{m}}) = \frac{k_{sx} - i \mu_c}{k_{sx} + i \mu_c} = \frac{Z}{Z'} = \frac{\exp (i \alpha)}{\exp (-i \alpha)} = \exp (2i \alpha)$$

$$\tan \alpha = \tan \frac{\Theta_{\text{m}}}{2} = \frac{\mu_c}{k_{sx}} = \frac{\sqrt{n_s^2 \sin^2 \varphi_1 - n_c^2}}{n_s \cos \varphi_1}.$$  

### B) TM-polarization

$$R_{TM} = 1 \cdot \exp (i \Theta_{\text{m}}) = \frac{k_{sx} e_s - i \mu_c e_s}{k_{sx} e_s + i \mu_c e_s} = \frac{Z}{Z'} = \frac{\exp (i \alpha)}{\exp (-i \alpha)} = \exp (2i \alpha)$$

$$\tan \alpha = \tan \frac{\Theta_{\text{m}}}{2} = \frac{-i \mu_c e_s}{k_{sx} e_s} = \frac{\varepsilon_s}{\varepsilon_c} \tan \frac{\Theta_{\text{m}}}{2}.$$  

In conclusion, we have seen that the phase shifts of the reflected light at TIR is different for TE and TM polarization, and because $\varepsilon_s > \varepsilon_c$

$$|\Theta_{\text{m}}| > |\Theta_{\text{m}}|.$$  

As a consequence, incident linearly polarized light gets generally elliptically polarized after TIR $\Leftrightarrow$ Fresnel prism

Remark
For metals in visible range (below the plasma frequency) we have always TIR, because: $\Re (\varepsilon_0) < 0 \rightarrow k_{cx} = \frac{\omega}{c} \sqrt{n_c^2 - n_s^2 \sin^2 \varphi_1}$ always imaginary.

In the case of TIR the modulus of the coefficient of reflection is one, but the coefficient itself is complex $\rightarrow$ nontrivial phase shift for reflected light:

A) TE-polarization

$$R_{TE} = 1 \cdot \exp (i \Theta_{\text{m}}) = \frac{k_{sx} - i \mu_c}{k_{sx} + i \mu_c} = \frac{Z}{Z'} = \frac{\exp (i \alpha)}{\exp (-i \alpha)} = \exp (2i \alpha)$$

$$\tan \alpha = \tan \frac{\Theta_{\text{m}}}{2} = \frac{-i \mu_c e_s}{k_{sx} e_s} = \frac{\varepsilon_s}{\varepsilon_c} \tan \frac{\Theta_{\text{m}}}{2}.$$  

B) TM-polarization

$$R_{TM} = 1 \cdot \exp (i \Theta_{\text{m}}) = \frac{k_{sx} e_s - i \mu_c e_s}{k_{sx} e_s + i \mu_c e_s} = \frac{Z}{Z'} = \frac{\exp (i \alpha)}{\exp (-i \alpha)} = \exp (2i \alpha)$$

$$\tan \alpha = \tan \frac{\Theta_{\text{m}}}{2} = \frac{-i \mu_c e_s}{k_{sx} e_s} = \frac{\varepsilon_s}{\varepsilon_c} \tan \frac{\Theta_{\text{m}}}{2}.$$  

In conclusion, we have seen that the phase shifts of the reflected light at TIR is different for TE and TM polarization, and because $\varepsilon_s > \varepsilon_c$

$$|\Theta_{\text{m}}| > |\Theta_{\text{m}}|.$$  

As a consequence, incident linearly polarized light gets generally elliptically polarized after TIR $\Leftrightarrow$ Fresnel prism

**Remark**

The field in the cladding is evanescent $\sim \exp (ik_{sx} x) = \exp (-\mu_c x)$.

$\Rightarrow$ The averaged energy flux in the cladding normal to the interface vanishes.

$$\langle S \rangle = \frac{1}{20 \mu_0} \Re \langle k | E |^2 \rangle = \frac{1}{20 \mu_0} \Re \langle k_{sx} | E |^2 \rangle = 0.$$  

### 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 $\varphi_B$ we find $R_{TM} = 0$:

$$\rho_{TM} = \frac{k_{sx} e_s - k_{cx} e_s}{k_{sx} e_s + k_{cx} e_s} = 0.$$
With the last two lines we can write the final result for the Brewster angle:
\[
\tan \phi_b = \frac{\varepsilon_s}{\varepsilon_c}.
\]

The Brewster angle exists only for TM polarization, but for any \( n_s \leq n_c \).

There is a simple physical interpretation, why there is no reflection at the interface for the Brewster angle.

\[
\tan \phi_b = \frac{\sin \phi_s}{\cos \phi_s} = \frac{n_s}{n_c}.
\]

At the same time the angle of the transmitted light is always
\[
n_s \sin \phi_s = n_c \cos \phi_s = n_c \sin \left( \frac{\pi}{2} - \phi_s \right).
\]

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 \( \varepsilon_s > \varepsilon_c \).

**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:
\[ E_z(x, z) = E_z \exp\left[i(k_{sx}x + k_{sz}z)\right] \rightarrow E_z(x, z) = E_z \exp\left[i(az + \gamma_xz)\right] \]

\[ E_x(x, z) = E_x \exp\left[i(\alpha z - \gamma_xz)\right] \exp\left[i\Theta(\alpha)\right] \]

with
\[ \alpha = \frac{\omega}{c} n_s \sin \varphi_{10} \]
\[ \gamma_x = \sqrt{\frac{\omega^2}{c^2} n_s^2 - \alpha^2} \]

The reflected plane wave gets a phase shift, which depends on the angle of incidence (here characterized by the transverse wave number \( \alpha \)). We want to treat beams, which we can write as a superposition of plane waves (Fourier amplitude \( e_\alpha(\alpha) \)):
\[ E_z(x, z) = \int d\alpha e_\alpha(\alpha) \exp\left[i(\alpha z + \gamma_xz)\right] \]

We assume a mean angle of incidence \( \varphi_{10} \):
\[ \alpha_0 = \pm n_s \sin \varphi_{10} \quad \text{mean angular frequency} \]
\[ \alpha = \alpha_0 + \varepsilon \]

In the Fourier integral, we have to integrate over angular frequencies with non-zero amplitudes \( e_\alpha(\alpha_0 + \varepsilon) \neq 0 \) for \( -\Delta \leq \varepsilon \leq \Delta \) only
\[ E_z(x, z) = \int_{-\Delta}^{\Delta} d\varepsilon e_\alpha(\alpha_0 + \varepsilon) \exp\left[i(\varepsilon z + \gamma_xz)\right] \exp\left[i\Theta(\alpha_0 + \varepsilon)\right] \]

\[ E_x(x, z) = \int_{-\Delta}^{\Delta} d\varepsilon e_\alpha(\alpha_0 + \varepsilon) \exp\left[i(\varepsilon z - \gamma_xz)\right] \exp\left[i\Theta(\alpha_0 + \varepsilon)\right] \]

Let us make the following assumptions:
- small divergence of beam (narrow spectrum, \( \Delta \ll \frac{\omega}{c} n_s \))
- all Fourier component undergo TIR \( \Theta(\alpha) > \Theta_{\text{tot}} \)

Then, it is justified to expand the phase angle into a Taylor series up to first order:
\[ \Theta(\alpha_0 + \varepsilon) = \Theta(\alpha_0) + \frac{\partial \Theta}{\partial \alpha} \mid_{\alpha_0} \varepsilon = \Theta(\alpha_0) + \Theta'\varepsilon \]

Then, the reflected beam at the interface at \( x = 0 \) is given as:
\[ E_x(0, z) = \exp\left[i\left[\alpha_0z + \Theta(\alpha_0)\right]\right] \int_{-\Delta}^{\Delta} d\varepsilon e_\alpha(\alpha_0 + \varepsilon) \exp\left[i(z + \Theta)\varepsilon\right] \]

We can identify the remaining integral as a shifted version of the incident beam profile at \( x = 0 \)
\[ E_x(0, z) = \exp\left[i\left[\Theta(\alpha_0) - \alpha_0\Theta'\right]\right] E_x(z + \Theta'). \]

Thus, the reflected beam appears shifted by \( d = -\Theta' \) (Goos-Hänchen Shift).

\[ x_{\text{ext}} = \frac{-d}{\Theta'} \tan \varphi_{10} \]

\[ \Theta' = \frac{\partial \Theta}{\partial \alpha} \mid_{\alpha_0} = -2 \frac{\varepsilon}{1 + \varepsilon^2} \frac{k_{sx}^2}{\mu_{sx}^2} = -2 \frac{\varepsilon}{1 + \varepsilon^2} \frac{k_{sx}^2}{\mu_{sx}^2} = -2 \frac{\varepsilon}{1 + \varepsilon^2} \frac{\mu_{sx}^2}{k_{sx}^2} = -2 \frac{\alpha}{\gamma_s} \]

\[ x_{\text{ext}} = \frac{1}{\mu_{sx}} \frac{1}{\sqrt{\alpha_0^2 - \gamma_s^2}} \quad \text{and} \quad \tan \varphi_{10} = \frac{\gamma_s}{\mu_{sx}} = \frac{\alpha}{\gamma_s} \]

\[ \Rightarrow x_{\text{ext}} \quad \text{depth of penetration} \]

### 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 periodic 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 building 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 \( x > 0, (\varepsilon_i, d_i), (\varepsilon_2, d_2) \)
- TE-polarization only
- monochromatic light

At the interface between substrate and Bragg-mirror \( (x = 0) \) we have incident and reflected electric field:

\[
E_i + E_r = E_0 \quad \text{and} \quad i k_x (E_i - E_r) = \frac{\partial E}{\partial x} = E'_{y}
\]

or

\[
E_i = \frac{E_0}{2} + \frac{i E'_y}{2k_x} \quad \text{and} \quad E_r = \frac{E_0}{2} + \frac{i E'_y}{2k_x}
\]

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).

As a particular example we will investigate an infinite system consisting of just two periodically repeated layers. The two layers should consist of homogeneous material with \( \varepsilon_1 \) and \( \varepsilon_2 \) having a thickness of \( d_i \) and \( d_2 \), respectively. Hence we have:

\[
\varepsilon(x) = \varepsilon(x + \Lambda) \quad \text{with the period} \quad \Lambda = d_1 + d_2
\]

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:

\[
E(x, z; \omega) = \exp \left\{ i \left[ k_x (k_x, \omega) x + k_z z \right] \right\} E_x (x)
\]

with \( E_x (x + \Lambda) = E_x (x) \) 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

\[
k_x, (k_x, \omega), \quad \text{in complete analogy to the DR for plane waves} \quad k_x^2 = \frac{\omega^2}{c^2} \varepsilon(\omega) - k_z^2
\]

homogeneous media. In order to make the difference to the homogeneous case more obvious, we change the notation for the Bloch vector: \( k \rightarrow K \).

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 \( N + 1 \)):

\[
\begin{pmatrix} E \\ E' \end{pmatrix}_{[N+1]} = \exp (i K \lambda) \begin{pmatrix} E \\ E' \end{pmatrix}_{[N]}
\]

On the other hand, we know from our matrix method:

\[
\begin{pmatrix} E \\ E' \end{pmatrix}_{[N+1]} = \hat{M} \begin{pmatrix} E \\ E' \end{pmatrix}_{[N]} + \hat{M}
\]

with \( \hat{M} = \hat{m} (d_1) \hat{m} (d_2) \rightarrow M_\mu = \sum m_\mu m_\mu^* \)

If the Bloch wave is a solution to our problem, we can set the two expressions equal:

\[
\left\{ \hat{M} - \exp (i K \lambda) \hat{I} \right\} \begin{pmatrix} E \\ E' \end{pmatrix}_{[N]} = 0
\]

and with \( \mu = \exp (i K \lambda) \) we have to solve the following eigenvalue problem:

\[
\left\{ \hat{M} - \mu (K) \hat{I} \right\} \begin{pmatrix} E \\ E' \end{pmatrix}_{[N]} = 0
\]

This eigenvalue problem determines the Bloch vector \( K \) and will finally give our dispersion relation.

As usual, we use the solvability condition \( \det \{ \hat{M} - \mu \hat{I} \} = 0 \) to compute the dispersion relation expressed in \( \mu = \exp (i K \lambda) \). Hence we still need to compute \( K \) afterwards.

\[
\mu = \exp (i K \lambda) \frac{(M_{11} + M_{22})}{2} + \sqrt{\frac{(M_{11} + M_{22})^2}{4} - 1}
\]

Note that we used \( \det \{ \hat{M} \} = 1 \), which explains why the off-diagonal elements of the matrix do not appear in the formula. Moreover, because of \( \det \{ \hat{M} \} = 1 \) we have \( \mu, \mu_\mu = 1 \).

The corresponding eigenvectors (field and its derivative at \( x = N \Lambda \)) can be computed from

\[
\left\{ \hat{M} - \exp (i K \lambda) \hat{I} \right\} \begin{pmatrix} E \\ E' \end{pmatrix}_{[N]} = 0
\]
From the first row the following condition can be derived:
\[
(M_{11} - \mu)E + M_{12}E' = 0
\]

Since the investigated system is linear and invariant for the phase, the absolute amplitude and phase of the \( E \)-component of the eigenvector can be chosen arbitrarily. Here we take \( E = 1 \) and get for the full eigenvector
\[
E = \begin{pmatrix} 1 \\ (\mu - M_{11}) / M_{12} \end{pmatrix} \cdot E_{\alpha}. 
\]

If field values of the Bloch mode, i.e. the function
\[
E_{\alpha}(x + \Lambda) = E_{\alpha}(x), 
\]
inside the layers are desired, they can be computed by using the matrix formalism and the above eigenvector \( (E, E')_{\alpha} \).

**Physical properties of infinite multilayer systems**

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
\[
\rho = \frac{E_0}{E_1} \text{ with } E_0 = \frac{E_0}{2} + i \frac{E_0'}{2k_{\alpha}} \text{ and } E_1 = \frac{E_1}{2} - i \frac{E_1'}{2k_{\alpha}} \text{ from before}
\]
\[
\rho = \frac{k_{\alpha}E_0 + i E_0'}{k_{\alpha}E_0 - i E_0'}
\]

With our knowledge of the eigenvector from above we can compute the reflectivity:
\[
E'_{\alpha} = \frac{\mu - M_{11}}{M_{12}} E_0 \rightarrow \rho = \frac{k_{\alpha}E_0 + i E_0'}{k_{\alpha}E_0 - i E_0'}^2 = \frac{k_{\alpha}^2 + \frac{1}{k_{\alpha}^2} - \frac{M_{11}}{M_{12}}}{k_{\alpha}^2 - \frac{1}{k_{\alpha}^2} - \frac{M_{11}}{M_{12}}}
\]

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
\[
\frac{(M_{11} + M_{22})}{2} \geq 1
\]

with [for our example \( (e_1, e_2, d_1, d_2) \)]

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
\[
\frac{(M_{11} + M_{22})}{2} < 1
\]

We can compute the explicit dispersion relation:
\[
\mu = \exp(\pm i K \Lambda) = \frac{(M_{11} + M_{22})}{2} \pm \sqrt{\left(\frac{(M_{11} + M_{22})}{2}\right)^2 - 1}
\]

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 \( k^2 = \epsilon(\omega) - k_z^2 \).

**Interpretation**

- For the case of total internal reflection (\( \mu \) real, \( \frac{(M_{11} + M_{22})}{2} \geq 1 \)) the Bloch vector \( K \) is complex, \( \mu = \exp(\pm i K \Lambda) = \exp(\pm i \Re(K \Lambda)) \exp(-\Im(K \Lambda)) \).

Hence \( \Re[K(k_z, \omega) \Lambda] = n \pi \) and
\[
\Im[K(k_z, \omega) \Lambda] = -\ln \left( -1 \right)^n \left( \frac{(M_{11} + M_{22})}{2} \pm \sqrt{\left(\frac{(M_{11} + M_{22})}{2}\right)^2 - 1} \right)
\]

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 \( n = 1, 2, \ldots \). These band gaps are interesting for Bragg mirrors and Bragg waveguides. The band gaps correspond to “forbidden” frequency ranges, where no propagating solution exists.

- The limits of the **bands** are given by
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 between the angle of incidence and frequency of light determining the reflection properties of multilayer systems. Therefore let us have a look at the simpler case of normal incidence \( k_z = 0 \).

In a graphical representation of the dispersion relation for \( k_z = 0 \) it is common to use the following dimensionless quantities

\[
\frac{\omega}{c G} \text{ and } \frac{K}{G} \text{ with the scaling constant } G = \frac{2 \pi}{\Lambda}.
\]

Examples for normal incidence

\[ n_1 = 1.4, \quad d_1 = 0.5 \Lambda \]

\[ n_2 = 3.4, \quad d_2 = 0.5 \Lambda \]

It is common to use the reduced bandstructure - Brillouin zone, where the information for all possible Bloch vectors is mapped onto the Bloch vectors in the following interval \(-0.5 \leq (k / G) \leq 0.5\).

Because of \( e^{ik_z} \) we need only \(-\pi \leq K \Lambda \leq \pi \to \left| \frac{K}{G} \right| \leq 0.5 \) to describe the dispersion relation.

Inside the band gap, we find damped solutions:

\[ n_1 = 1.4, \quad d_1 = 0.5 \Lambda \]

\[ n_2 = 3.4, \quad d_2 = 0.5 \Lambda \]

Let us quantify the damping. In our example \( (n_1, n_2, d_1, d_2) \) we have

\[
\left( \frac{M_{11} + M_{22}}{2} \right) = \cos \left( \frac{\omega}{c} n_i d_i \right) \cos \left( \frac{\omega}{c} n_i d_z \right) - \frac{1}{2} \left( \frac{n_2}{n_1} + \frac{n_1}{n_2} \right) \sin \left( \frac{\omega}{c} n_i d_i \right) \sin \left( \frac{\omega}{c} n_i d_z \right)
\]

In the middle of the first band gap (optimum configuration for high reflection) we have

\[
\frac{\omega}{c} n_i d_i = \frac{\omega}{c} n_i d_z = \frac{\pi}{2}, \quad \text{with } \omega_i \text{ being the Bragg frequency, and}
\]

\[
\left( \frac{M_{11} + M_{22}}{2} \right) = -\frac{1}{2} \left( \frac{n_2}{n_1} + \frac{n_1}{n_2} \right) < -1.
\]

If we plug this (for \( n = 1 \)) in our...
expression for $\Delta K$ and assume a small index contrast $|n_2 - n_1| \ll (n_2 + n_1)$ we find

$$\Delta K_{\text{max}} \approx 2 \frac{n_2 - n_1}{n_2 + n_1}$$

(do derivation as an exercise)

$\triangleright$ Damping is proportional to index contrast of the subsequent layers $|n_2 - n_1|

The spectral width of the gap $\left| \frac{M_{11} + M_{22}}{2} \right| \geq 1$ is then

$$\Delta \omega_{\text{gap}} \approx \frac{2 \omega_0}{\pi} \Delta K_{\text{max}}$$

(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, the so-called Fabry-Perot-resonator. To construct a Fabry-Perot-resonator, one can start from highly reflecting periodic multi-layer system (Bragg reflector). If one changes just a single layer somewhere in the middle of the otherwise periodic layer system, a so-called cavity is formed, and we are interested in the forward and backward propagating fields of the entire layer system. While the single layer which is distinguished from the periodic stack forms the cavity, 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. This task could be achieved employing the matrix method which we developed in the previous sections. However we will take a different approach to achieve deeper physical insight into the cavity's behavior.

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 $R_0$, $T_0$, $R_D$, and $T_D$.

Using the known coefficients of reflection and transmission of the two mirrors, we can eliminate $E_\pm$ and $E_T$ by connecting the field amplitudes:

A) At the lower mirror inside the cavity:

$$T_0 E_x + R_0 E_r(0) = E_r(0)$$

B) At the upper mirror outside the cavity:

$$E_x = T_0 E_r(D)$$

And with $E_r(D) = E_r(0) \exp(\pm ik_x D) E_x(0) = \frac{E_x}{T_0} \exp(-\pm ik_x D)$

C) At the upper mirror inside the cavity:

$$E_r(D) = R_0 E_x(D) = \frac{R_0}{T_0} E_x$$

and with $E_r(0) = E_x(D) \exp(\pm ik_x D)$

$$E_r(0) = \frac{R_0}{T_0} E_x \exp(\pm ik_x D)$$

D) we substitute $E_x(0)$ and $E_r(0)$ in A)

$$T_0 E_x + R_0 E_r(0) = E_r(0)$$

$$T_0 E_x + R_0 \frac{R_0}{T_0} E_x \exp(\pm ik_x D) = \frac{E_x}{T_0} \exp(-\pm ik_x D)$$

$$\Rightarrow E_x = \frac{1}{T_D/T_0} \left[ \exp(-\pm ik_x D) - R_0 R_0 \exp(\pm ik_x D) \right] E_x.$$
Thus, the coefficient of transmission for the whole FP-resonator expressed by the coefficients of the mirrors \(R, T, D\), the cavity properties and the angle of incidence \(D, k_{\text{eff}} = \sqrt{\frac{\omega^2}{c^2} - k_x^2}\) reads:

\[
T_{\text{FP}} = \frac{E_x}{E_i} = \frac{T_R T_T \exp(\pm k_{\text{eff}} D)}{1 - R_R R_T \exp(\pm 2k_{\text{eff}} D)}
\]

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 \(\phi_R, \phi_T\). Obviously, only the phase shifts induced by the coefficients of reflection \(R, T\) are important for the transmissivity of the FP resonator \(\tau = |T_{\text{FP}}|^2\).

For given \(|R_0|, |R_1|, |T_0|, |T_1|\) and \(\phi_R, \phi_T\), the general transmissivity of a lossless Fabry-Perot resonator reads:

\[
|T_{\text{FP}}|^2 = |T|^2 = \frac{|T_0|^2 |T_1|^2}{1 + |R_0|^2 |R_1|^2 - 2|R_0||R_1|^\cos(2k_{\text{eff}} D + \phi_R + \phi_T)}
\]

\[
\tau = \frac{k_{\text{eff}}}{k_{\text{eff}}} |T|^2
\]

Here we introduced the phase-shift \(\delta\), which the field acquires in one round-trip in the cavity.

**Discussion**

Depending on whether the two mirrors have identical properties we distinguish between symmetric and asymmetric FP-resonators.

**a) asymmetric FP-resonator**

\[
\tau = \frac{k_{\text{eff}}}{k_{\text{eff}}} \frac{|T_0|^2 |T_1|^2}{1 + |R_0|^2 |R_1|^2 - 2|R_0||R_1|\cos(2k_{\text{eff}} D + \phi_R + \phi_T)}
\]

Because we assume no losses we can use energy conservation at each mirror to eliminate \(T_{0,1}\):

\[
|T_0|^2 |T_1|^2 = \frac{k_{\text{eff}}}{k_{\text{eff}}} (1 - |R_0|^2) (1 - |R_1|^2)
\]

\[
\tau = \frac{k_{\text{eff}}}{k_{\text{eff}}} (1 - |R_0|^2) (1 - |R_1|^2)
\]

Note: \(\tau\) and \(\rho\) for a lossless mirror are the same for both sides of the mirror. For lossy mirrors only \(\tau\) is the same, \(\rho\) is then side-dependent.

For discussing the effect of the phase shift \(\delta\) we rewrite

\[
\cos \delta = \cos^2 \frac{\delta}{2} - \sin^2 \frac{\delta}{2} = 1 - 2\sin^2 \frac{\delta}{2}
\]

Plugging everything in we get

\[
\tau = \frac{(1 - |R_0|^2)(1 - |R_1|^2)}{1 + |R_0|^2 |R_1|^2 - 2|R_0||R_1|^\cos(2k_{\text{eff}} D + \phi_R + \phi_T)}
\]

\[
= \frac{(1 - |R_0|^2)(1 - |R_1|^2)}{(1 - |R_0|^2)(1 - |R_1|^2) + 4|R_0||R_1|\sin^2 \frac{\delta}{2}}^{-1}
\]

and with

\[
\rho = |R_0|^2, \rho = |R_1|^2
\]

\[
\tau = \frac{(1 - \sqrt{\rho_0 \rho_1})^2}{(1 - \rho_0)(1 - \rho_1) + 4\rho_0 \rho_1 \sin^2 \frac{\delta}{2}}^{-1}
\]

**b) symmetric FP-resonator (Airy-formula for transmissivity)**

\[
\rho_m = |R_0| = |R_1| = \rho_0 = \rho, \quad \phi = \phi_0 = \phi_R
\]

\[
\tau = \frac{(1 - \rho_m)^2}{(1 - \rho_m)^2 + 4\rho_m \sin^2 \frac{\delta}{2}}^{-1}
\]

\[
= \frac{1 + F \sin^2 \frac{\delta}{2}}{2}\right)^{-1}
\]

with \(F = \frac{4\rho_m}{(1 - \rho_m)^2}\) and \(\delta = k_{\text{eff}} D + \phi\)

The 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 → $R^\text{TM}$ resp. $\rho^\text{TM}$
- Resonances of the cavity with $\tau_{\text{max}} = 1$ occur for $\delta / 2 = k_D D_{\text{max}} + \varphi = m\pi$ with
  \[
  k_{\text{fs}} = \frac{2\pi}{\lambda} \sqrt{n_x^2 - n_s^2 \sin^2 \varphi_1}
  \]
  \[
  \sqrt{D_{\text{max}}} = \frac{m\pi - \varphi}{k_{\text{fs}}} = \frac{\lambda}{2} \sqrt{n_x^2 - n_s^2 \sin^2 \varphi_1}
  \]
  \[
  \sqrt{\frac{\lambda}{2}} \ \text{cavity}
  \]
  where $\varphi_1$ is the angle of incidence in the substrate
- transmission properties of a given resonator depend on $\varphi_2$ and $\lambda$.
- minimum transmission is given as $\tau_{\text{min}} = \frac{1}{1 + F}$
- it is favorable to have large $F$, e.g.:
  \[
  100 = F = \frac{4\rho_m}{(1 - \rho_m)}
  \]
  \[
  \rho_m = 1 - \tau_m \cdot \frac{4(1 - \tau_m)}{\tau_m^2} = \frac{4}{\tau_m} \approx 100 \ \text{and} \ \rho_m = 0.8.
  \]
- it is favorable to have large $F$, e.g.:
- there are always some range of incident angles $\varphi_2$, they 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.

To calculate the FWHM $\varepsilon$ we can start from:

\[
\left[1 + F \sin^2 \left(\frac{m\pi \pm \varepsilon}{2}\right)\right]^{-1} = \frac{1}{2}
\]

For narrow resonances (small line width $\varepsilon$) we can write

\[
\left[1 + F \left(\frac{\varphi}{2}\right)^2\right]^{-1} \approx \frac{1}{2} \Rightarrow F \left(\frac{\varepsilon}{2}\right)^2 = 1 \cap \varepsilon = \frac{2}{\sqrt{F}}
\]

\[
\Phi = \frac{\Delta}{\varepsilon} = \frac{\pi}{2\sqrt{F}} = \frac{\pi}{\sqrt{1 - \rho_m}}
\]

- The line width $\varepsilon$ (FWHM) is inversely proportional to the finesse $\Phi$.
- The Airy-formula can be expressed in terms of the finesse

\[
\tau = \left[1 + \left(\frac{2\Phi}{\pi}\right) \sin \frac{\varphi}{2}\right]^{-1}
\]

- 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:

\[
kD + \varphi = m\pi
\]

reduced transmission by factor $1/2$

\[
\rightarrow kD + \varphi \pm \frac{\Delta k}{2}D = m\pi \pm \frac{\varepsilon}{2} = m\pi \pm \frac{\pi}{2\Phi}
\]

\[
\cap |\Delta k| = \frac{2\pi}{\lambda} n_x |\Delta \lambda| = \frac{\pi}{\Phi D}
\]

\[
|\Delta \lambda| = \frac{\lambda}{n_x D \Phi} \cap \phi D = \frac{\lambda}{n_x D \Phi} - \frac{\varepsilon}{D}
\]

example: $\lambda = 5 \times 10^{-7}$m, $\Phi = 30$, $n_x D = 4 \times 10^{-7}$m $\cap \Delta \lambda = 2 \times 10^{-12}$m

- The field amplitudes (here forward field) inside the cavity are given as:

\[
E_x = T_x E_x(D)
\]

Because $|f|^2 = 1/(1 - \rho) - \Phi$ these intra-cavity fields can be very high $\cap$ important for nonlinear effects

- Lifetime of photons in cavity: Via the "uncertainty relation":

$\Delta \omega T_c = \text{const.} \approx 1$ it is possible to define a lifetime of photons inside the cavity:
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, e.g. where the waveguide is a fiber.

7.4 Guided waves in layer systems

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 $\Delta \omega = \frac{\pi}{c} \frac{\Delta n}{\Phi D}$, we calculated $E_{\text{RT}}$. Inside each layer we have plane waves $E(x, z, t) = \exp\left[\pm ik_{z}z + ik_{x}x - \omega t\right]$. 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:

$$E_{\text{sc}}(x, z) = E_{\text{sc}} \exp\left[\pm ik_{z}z + ik_{x}x - \omega t\right]$$

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?

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:

$$k_{z}\text{max} \leq \left|k_{z}\right| \leq k_{\text{sc}}\text{max}$$

The field structure in substrate and cladding is given as:

- oscillating solution (standing wave) in the core (layers, fiber core, film)
- $A\sin(k_{z}x) + B\cos(k_{z}x)$

If we go back to our usual reflection transmission problem, we have here reflected and transmitted (evanescent) field for zero incident field $I_{0} \rightarrow 0$. 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 $E_{c}, E_{sc} \neq 0$ for $E_{i} \rightarrow 0$, and this can be exploited. The coefficients for reflection and transmission are

$$T_{\text{TE,TM}} = \begin{pmatrix} F_{z} \\ E_{z} \end{pmatrix}, R_{\text{TE,TM}} = \begin{pmatrix} E_{z} \\ F_{z} \end{pmatrix} \quad \text{with } E_{i} \rightarrow 0$$

and we find $R, T \rightarrow \infty$ in the case of guided waves. In this sense, guided waves are resonances of the system. Let us compare to a driven harmonic oscillator

$$x = \frac{F}{\omega^{2} - \omega_{0}^{2}}$$
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 the expressions for $R, T$. 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:

$R = \frac{F_x}{F_z} = \left(\frac{\alpha_x k_x M_{22} - \alpha_x k_x M_{11}}{\alpha_x k_x M_{22} + \alpha_x k_x M_{11}}\right) - \frac{i}{2} \left(M_{21} + \alpha_x k_x \alpha_x k_x M_{12}\right)$

The pole is then given as:

$\left(\frac{\alpha_x k_x M_{22} + \alpha_x k_x M_{11}}{\alpha_x k_x M_{22} - \alpha_x k_x M_{11}}\right) - \frac{i}{2} \left(M_{21} - \alpha_x k_x \alpha_x k_x M_{12}\right) = 0$

With (because we have evanescent waves in substrate and cladding)

$k_x = i \mu_x = i \sqrt{k_x^2 - \frac{\omega^2}{c^2} \varepsilon_x(\omega)}$, \quad $k_{cx} = i \mu_{cx} = i \sqrt{k_{cx}^2 - \frac{\omega^2}{c^2} \varepsilon_{cx}(\omega)}$

We can write the general dispersion relation in an arbitrary layer system

$M_{11}^{TE,TM} + \alpha_x k_x M_{12}^{TE,TM} + \frac{1}{\alpha_x k_x} M_{21}^{TE,TM} + \frac{\varepsilon_x}{\alpha_x k_x} M_{22}^{TE,TM} = 0$

Here, as usual, we have $\alpha_{TE} = 1$, $\alpha_{TM} = 1/\varepsilon$.

In addition to the material dispersion in the layers

$k_x^2(\omega) = \frac{\omega^2}{c^2} \varepsilon_x(\omega) = k_{ax}^2 + k_{cx}^2$

we get the waveguide dispersion relation $k_x(\omega, geometry)$

The waveguide dispersion relation gives a discrete set of solutions, so-called waveguide modes. For given $\varepsilon_s$, $d_s$, $\omega$ 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:

$F(x) = F \exp(\mu x), \quad G(x) = \frac{\partial}{\partial x} \left[F \exp(\mu x)\right]$

Hence, for given $F(0)$ we get $G(0) = \alpha \mu F(0) \quad (F(0) \rightarrow \text{free parameter})$

$\begin{bmatrix} F \\ G \end{bmatrix} = M(\alpha \mu) \begin{bmatrix} F \\ G \end{bmatrix} = \tilde{M}(\alpha \mu) F(0)$.

### 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 is that on both sides of the interface we have evanescent waves: $k_x^2 > \frac{\omega^2}{c^2} \varepsilon_{ax}$, because

$$\mu_{ax} = \sqrt{k_x^2 - \frac{\omega^2}{c^2} \varepsilon_{ax}} > 0$$

The general dispersion relation we derived before reads

$M_{11}^{TE,TM} + \alpha_x \mu_x M_{12}^{TE,TM} + \frac{1}{\alpha_x k_x} M_{21}^{TE,TM} + \frac{\varepsilon_x}{\alpha_x k_x} M_{22}^{TE,TM} = 0$
and with the matrix for a single interface: 
\[ \mathbf{M} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \]

we get the dispersion relation 
\[ 1 + \frac{\alpha_1 M_1}{\alpha_2 M_2} = 0 \]

A) TE-polarization (\( \alpha = 1 \))
\[ \mu_a + \mu \rightarrow \text{no solution because } \mu, \mu > 0 \]

B) TM-polarization (\( \alpha = 1 / \varepsilon \))
\[ \frac{\varepsilon_a + \varepsilon}{\varepsilon_a - \varepsilon} = 0 \]

with \( \mu, \mu > 0 \) and \( \varepsilon \cdot \varepsilon < 0 \),

\( \rightarrow \) on of the media has to have negative \( \varepsilon \) (dielectric near resonance or metal)

In dielectrics \( \omega_0 < \omega < \omega_s \) 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
\[ k \cdot \varepsilon = 0, \]

Let us now compute the explicit dispersion relation for surface polaritons. W.o.i.g. we assume \( \varepsilon_\omega < 0, \) and because \( \varepsilon_\omega \) is near a resonance it will show a much stronger \( \omega \) dependence than \( \varepsilon_c \).

\[ (\mu_\omega \varepsilon_\omega)^2 = (\mu_\omega \varepsilon_c)^2 \]
\[ \varepsilon_\omega^2 \left( k_\omega^2 - \frac{\omega^2}{c^2} \varepsilon_\omega^2 \right) = \varepsilon_c^2 \left( k_c^2 - \frac{\omega^2}{c^2} \varepsilon_c^2 \right) \]
\[ k_\omega(\omega) = \frac{\omega}{c} \sqrt{\varepsilon_\omega(\omega) \varepsilon_c} \]

There is a second condition for existence of surface polaritons: \( \varepsilon_c + \varepsilon_\omega(\omega) < 0 \)

**Conclusion:**
\[ k_\omega(\omega) = \frac{\omega}{c} \sqrt{\varepsilon_\omega(\omega) \varepsilon_c} \]

TM polarization \( \varepsilon_\omega(\omega) < 0, \) \( \varepsilon_c > 0 \)
\[ \left| \varepsilon_\omega(\omega) \right| > \varepsilon_c \]

Surface polaritons may have very small effective wavelengths in z-direction:
\[ \varepsilon_\omega = 1, \left| \varepsilon_\omega(\omega) \right| = 1 \rightarrow k_\omega(\omega) = \frac{2 \pi}{\lambda} \sqrt{\varepsilon_\omega(\omega) \varepsilon_c} = \frac{2 \pi}{\lambda_{\text{eff}}} \rightarrow \lambda_{\text{eff}} \ll \lambda \]

**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
\[ \frac{\omega^2}{c^2} \varepsilon_\omega(\omega) > k_\omega^2(\omega). \]
Such film waveguides are the basis of integrated optics. Typical parameters are:

\[ d \approx \text{a few wavelengths} \]
\[ \Delta \varepsilon \approx 10^{-3} - 10^{-4} \]

Fabrication of film waveguides can be achieved by coating, diffusion or ion implantation.

The matrix of a single layer (film) is given as:

\[ \mathbf{M}^{TE, TM} = \mathbf{m}^{TE, TM}(d) = \begin{pmatrix} \cos(k_{ex}d) & -\frac{1}{\alpha_{e}k_{ex}} \sin(k_{ex}d) \\ -\frac{1}{\alpha_{e}k_{ex}} \sin(k_{ex}d) & \cos(k_{ex}d) \end{pmatrix} \]

From this matrix we can compute the dispersion relation for guided modes:

\[ M_{11} + \alpha_{e} \mu_{a} M_{12} + \frac{1}{\alpha_{e} \mu_{a}} M_{21} + \frac{\alpha_{e} \mu_{s}}{\alpha_{e} \mu_{a}} M_{22} \geq 0 \]

\[ \cos(k_{ex}d) + \frac{\alpha_{e} \mu_{s}}{\alpha_{e} k_{ex}} \sin(k_{ex}d) - \frac{\alpha_{e} k_{ex}}{\alpha_{e} \mu_{a}} \sin(k_{ex}d) + \frac{\alpha_{e} \mu_{s}}{\alpha_{e} \mu_{a}} \cos(k_{ex}d) = 0 \]

\[ \frac{\sin(k_{ex}d)}{\cos(k_{ex}d)} = \tan(k_{ex}d) = \frac{1 + \frac{\alpha_{e} \mu_{s}}{\alpha_{e} k_{ex}}}{\frac{\alpha_{e} \mu_{a}}{\alpha_{e} \mu_{s}} - \frac{\alpha_{e} \mu_{s}}{\alpha_{e} k_{ex}}} = \frac{\alpha_{e} k_{ex}}{\alpha_{e} \mu_{s} - \alpha_{e} k_{ex} + \alpha_{e} \mu_{s}} \]

\[ \tan(k_{ex}d) = \frac{\alpha_{e} k_{ex} \left( \mu_{e} + \mu_{s} \right)}{\alpha_{e} \mu_{s} - \mu_{s} \mu_{a}} \]

Here: TE-Polarisation (\( \alpha = 1 \))

\[ \tan(k_{ex}d) = \frac{k_{ex} \left( \mu_{e} + \mu_{s} \right)}{k_{ex}^2 - \mu_{s} \mu_{a}} \]

This waveguide dispersion relation is an implicit equation for \( k_{ex} \). For given frequency \( \omega \) and thickness \( d \) we get several solutions with index \( k_{ex} \).

Here is an example for fixed frequency \( \omega \), the effective index \( n_{eff} = k_{ex} / \left( \frac{\omega}{c} \right) \) 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 $\rightarrow$ no guiding

$$ k_s^2 = \frac{\omega^2}{c^2} \varepsilon_s $$

We can plug this cut-off condition in the DR:

$$ \tan \left( k_s d \right) = \frac{k_s \left( \mu_s + \mu_a \right)}{k_s - \mu_a \mu_s}, $$

$$ \tan \left( \frac{\omega}{c} \sqrt{\varepsilon_s - \varepsilon_c} d \right) = \frac{\sqrt{\varepsilon_s - \varepsilon_c} \sqrt{\varepsilon_s - \varepsilon_e}}{\varepsilon_s - \varepsilon_c} = \frac{\varepsilon_s - \varepsilon_e}{\varepsilon_s - \varepsilon_c} $$

$$ (od)_{\omega}^{TE} = \frac{c}{\sqrt{\varepsilon_s - \varepsilon_e}} \left\{ \arctan \left( \frac{\varepsilon_s - \varepsilon_e}{\varepsilon_s - \varepsilon_c} \right) \right\} $$

$$ \rightarrow (od)_{\omega}^{TE} = \frac{c}{\sqrt{\varepsilon_s - \varepsilon_e}} \left\{ \arctan \left( a + \frac{\pi}{2} \right) \right\} $$

$$ \varepsilon_s \approx \varepsilon_e \quad a \rightarrow 0 $$

$$ \varepsilon_s \approx \varepsilon_c \quad a \rightarrow \infty $$

with parameter of asymmetry $a$: $\varepsilon_s \approx \varepsilon_c$

$\rightarrow$ we can define a cut-off frequency for $k_s(\omega)$ when we keep $d$ fix

$\rightarrow$ we can define a cut-off thickness for $k_s(d)$ when we keep $\omega$ fix

In a symmetric waveguide the fundamental mode $(\nu = 0)$ has cut-off $= 0$!

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):

$$ E(x,z) = \sum \alpha_v E_v(x) \exp(i k_v x) $$

$$ \rightarrow E(x,0) = \sum \alpha_v E_v(x) \| E_v(x) $$

with: $P_v = \int_{-\infty}^{\infty} |E_v(x)|^2 dx$

$$ a_v = \frac{k_{x_v}}{2 \omega \mu_0 P_v} \int_{-\infty}^{\infty} E_{in}(x) E_v(x) dx $$

$\rightarrow$ 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

$$k_z > \frac{\omega}{c} \sqrt{\varepsilon_{ex}}$$

but dispersion relation for waves in bulk media dictates

$$k_z = \sqrt{\frac{\omega^2}{c^2} \varepsilon_{ex} - k_{\perp}^2} < \frac{\omega}{c} \sqrt{\varepsilon_{ex}}$$

We got a problem!

There are two solutions:

i) coupling by prism
we bring a medium with $\varepsilon_p > \varepsilon_e$ (prism) near the waveguide

$$k_z < \frac{\omega}{c} \sqrt{\varepsilon_p} \quad \text{and} \quad k_z < \frac{\omega}{c} \sqrt{\varepsilon_p}$$

$\Rightarrow$ light can couple to the waveguide via optical tunneling: ATR ('attenuated total reflection')

ii) coupling by grating
grating on waveguide (modulated thickness of layer d):

$$d(z) = d + \zeta(z)$$

$$\zeta(z) = A \sin (gz) \quad \text{mit} \quad g = \frac{2\pi}{P} \quad \text{period}$$

coupling works for m'th diffraction order:

$$k_{m} = k_{z} + mg$$

$$= \frac{\omega}{c} n_i \sin \phi + mg.$$