Fundamentals of Modern Optics

Winter Term 2015/2016

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

Table of content

This script originates from the lecture series “Theoretische Optik” given by Falk Lederer at the FSU Jena for many years between 1990 and 2012. Later the script was adapted by Stefan Skupin and Thomas Pertsch for the international education program of the Abbe School of Photonics.

0. Introduction

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

> UNDOUBTEDLY, NEARLY ALL WHO READ THIS BOOK HAVE, at one time or another, pondered the question, "What is light?" To answer, "Light is that which permits vision," begs the question, for such an answer provides us with no understanding of the nature of light. It says no more than "light is light."

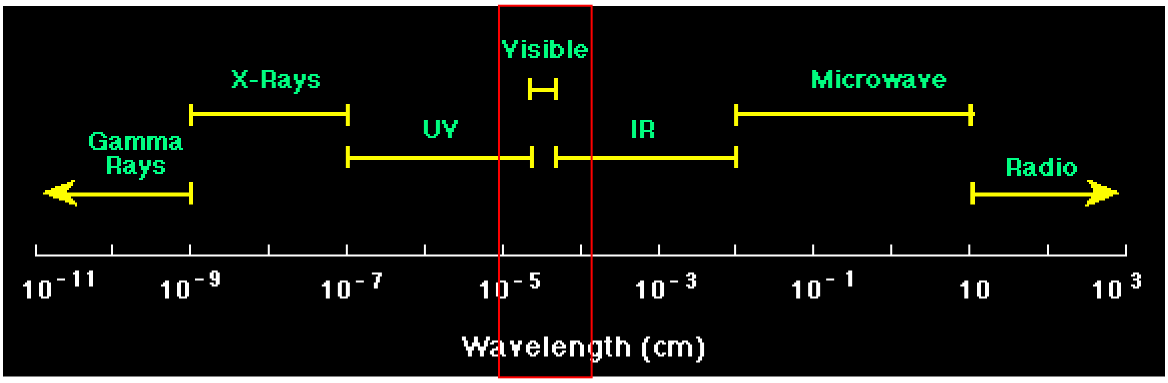
D.J. Lovell, Optical Anecdotes

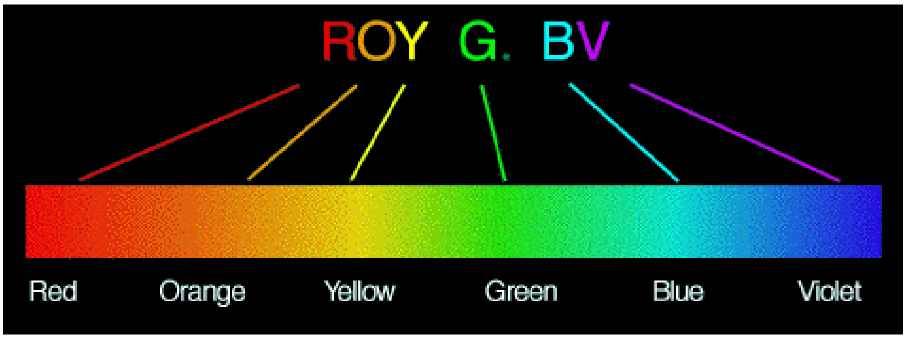
- Light is the origin and requirement for life → photosynthesis
- 90% of information we get is visual

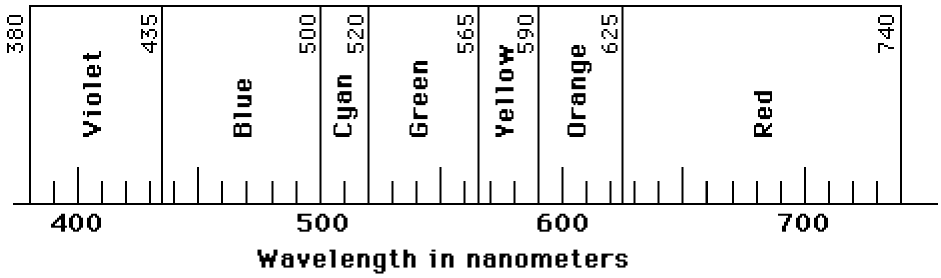
A) What is light?

- electromagnetic wave propagating with the speed of $c = 3\times10^{8} m/s$
- wave = evolution of
  - amplitude and phase → complex description
  - polarization → vectorial field description
  - coherence → statistical description

Spectrum of Electromagnetic Radiation				
Region	**Wavelength [nm]**	**Wavelength [m] (nm=10^{-9}m)**	**Frequency [Hz] (THz=10^{12}Hz)**	**Energy [eV]**
Radio	> 10^{8}	> 10^{-1}	< 3×10^{9}	< 10^{-5}
Microwave	10^{8} - 10^{5}	10^{-1} – 10^{-4}	3×10^{9} - 3×10^{12}	10^{-5} - 0.01
Infrared	10^{5} - 700	10^{-4} - 7×10^{-7}	3×10^{12} - 4.3×10^{14}	0.01 - 2
Visible	700 - 400	7×10^{-7} - 4×10^{-7}	4.3×10^{14} - 7.5×10^{14}	2 - 3
Ultraviolet	400 - 1	4×10^{-7} - 10^{-9}	7.5×10^{14} - 3×10^{17}	3 - 10^{3}
X-Rays	1 - 0.01	10^{-9} - 10^{-11}	3×10^{17} - 3×10^{19}	10^{3} - 10^{5}
Gamma Rays	< 0.01	< 10^{-11}	> 3×10^{19}	> 10^{5}







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

PHYSICAL REVIEW VOLUME 112, NUMBER 6 DECEMBER 15, 1958

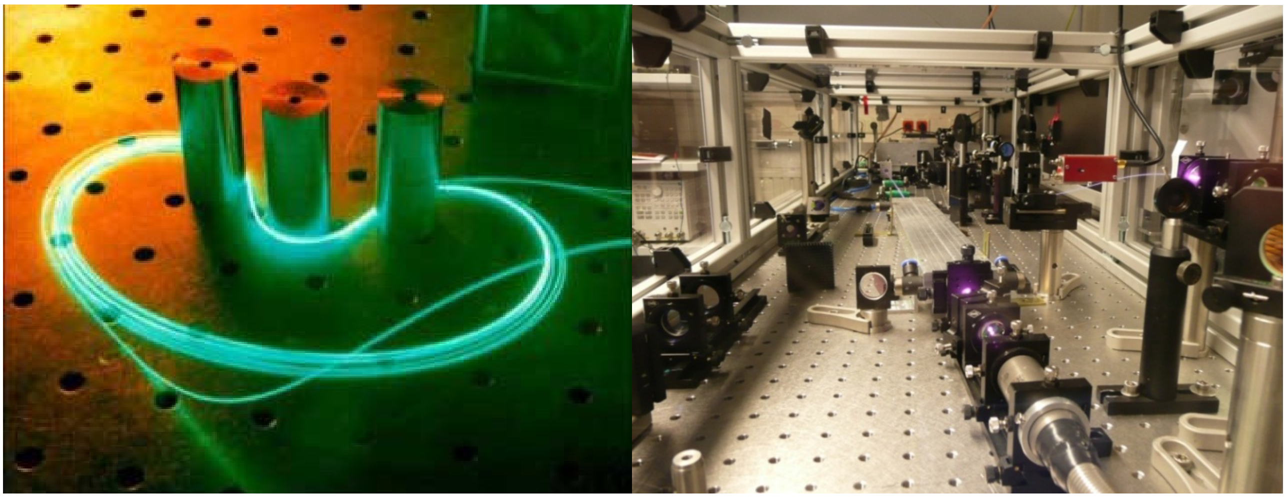
Infrared and Optical Masers

A. L. Schawlow and C. H. Townes*
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received August 26, 1958)

The extension of maser techniques to the infrared and optical region is considered. It is shown that by using a resonant cavity of centimeter dimensions, having many resonant modes, maser oscillation at these wavelengths can be achieved by pumping with reasonable amounts of incoherent light. For wavelengths much shorter than those of the ultraviolet region, maser-type amplification appears to be quite impractical. Although use of a multimode cavity is suggested, a single mode may be selected by making only the end walls highly reflecting, and defining a suitably small angular aperture. Then extremely monochromatic and coherent light is produced. The design principles are illustrated by reference to a system using potassium vapor.

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

- laser → artificial light source with new and unmatched properties (e.g. coherent, directed, focused, monochromatic)
- applications of laser: fiber-communication, DVD, surgery, microscopy, material processing, ...



Fiber laser: Limpert, Tünnermann, IAP Jena, ~10kW CW (world record)

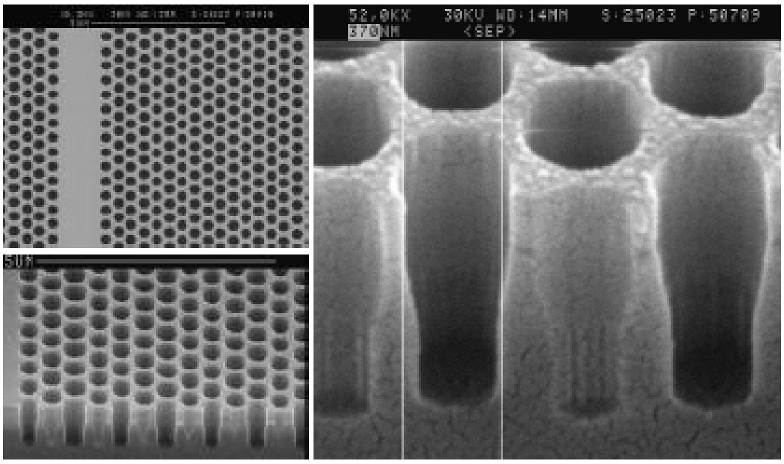
C) Propagation of light through matter

- light-matter interaction (G: Licht-Materie-Wechselwirkung)

effect	dispersion	diffraction	absorption	scattering
	↓	↓	↓	↓
governed by	frequency spectrum	spatial frequency	center of frequency spectrum	wavelength

- 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)

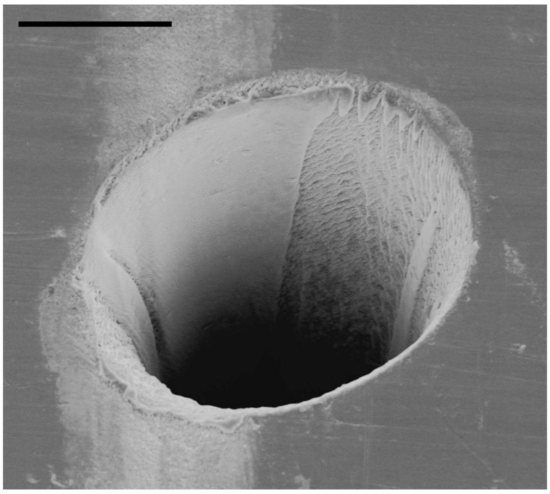
- design media with desired properties: glasses, polymers, semiconductors, compounded media (effective media, photonic crystals, meta-materials)



Two-dimensional photonic crystal membrane.

D) Light can modify matter

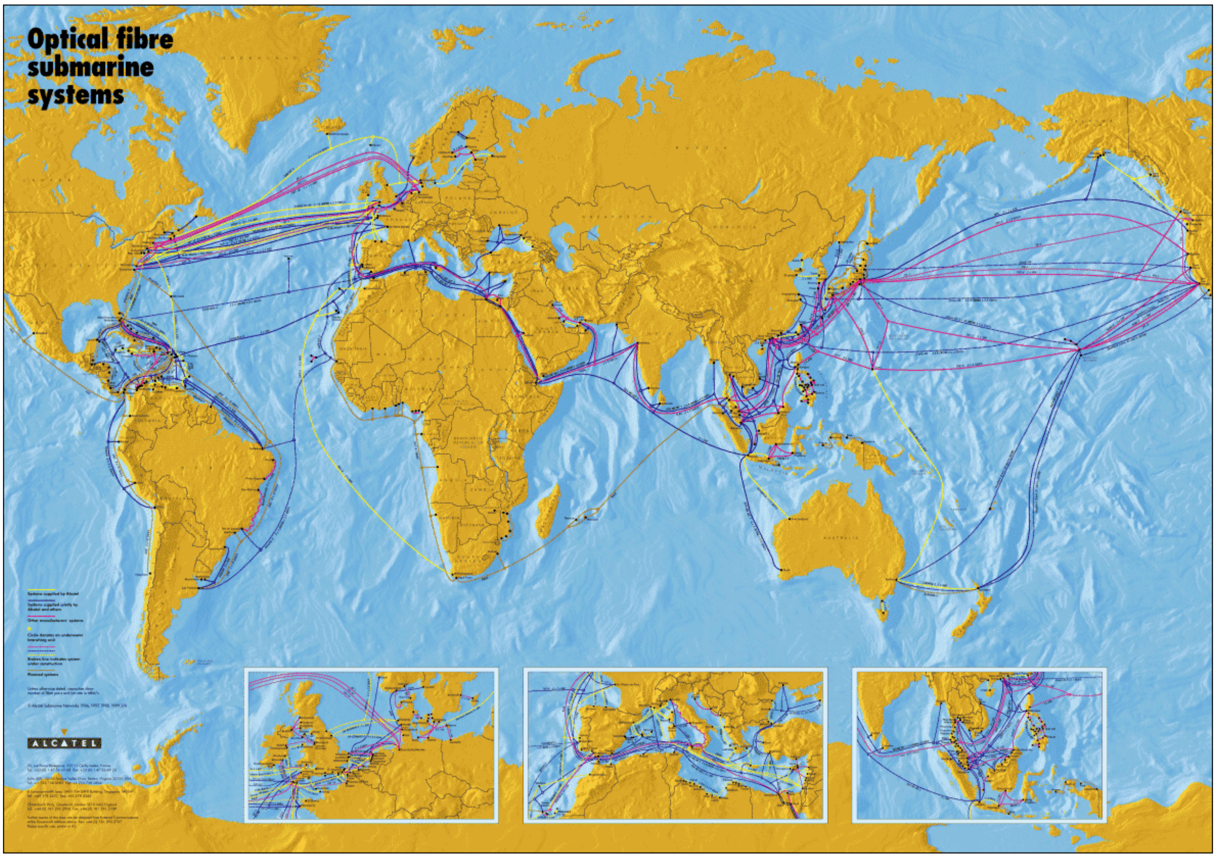
- light induces physical, chemical and biological processes
- used for lithography, material processing, or modification of biological objects (bio-photonics)



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

E) Optical telecommunication

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



1000 m telecommunication fiber is installed every second.

F) Optics in medicine and life sciences

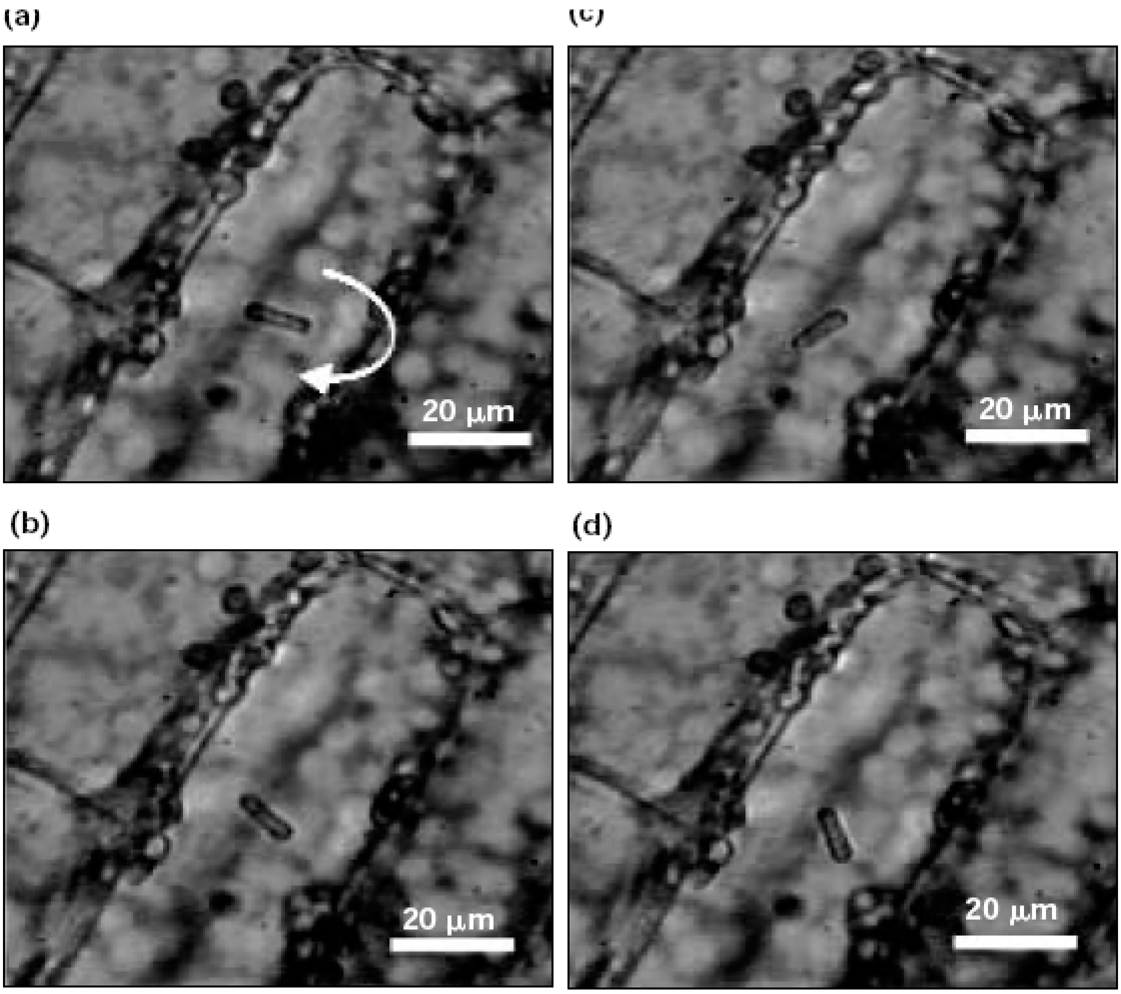
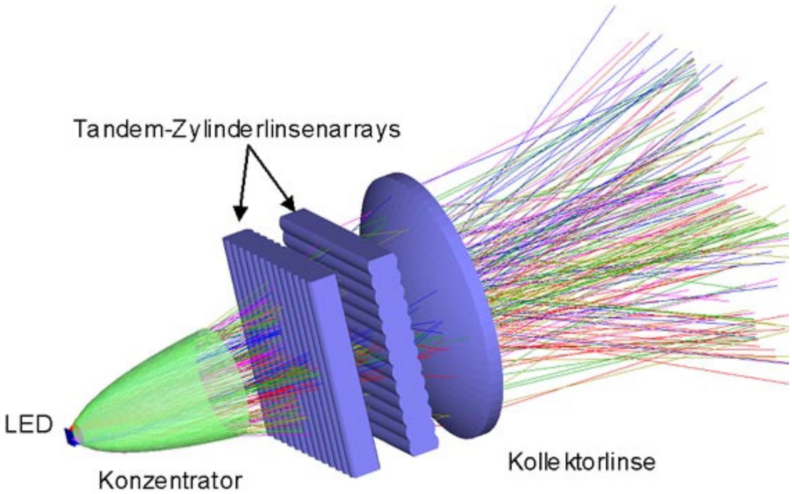


Figure 1. Rotation of an intracellular object inside *Elodea densa* plant cell using the rotating line tweezers. The rod shaped structure was trapped using 25 mW power and rotated at a speed of 4 Hz. The direction of rotation is shown by arrow (a). Clockwise rotation by angles of 45° (b); 145° (c); and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

G) Light sensors and light sources

- new light sources to reduce energy consumption

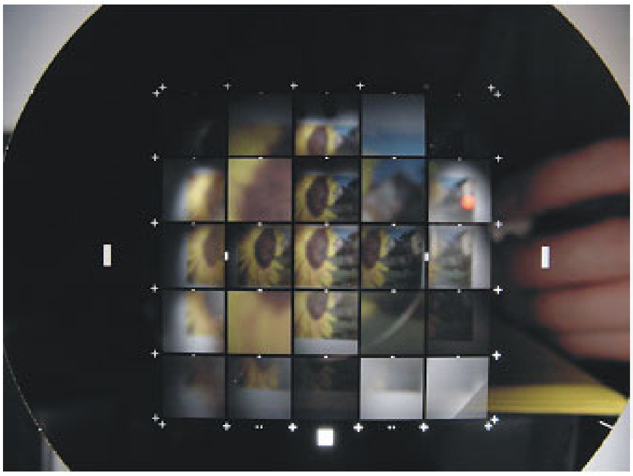
- new projection techniques

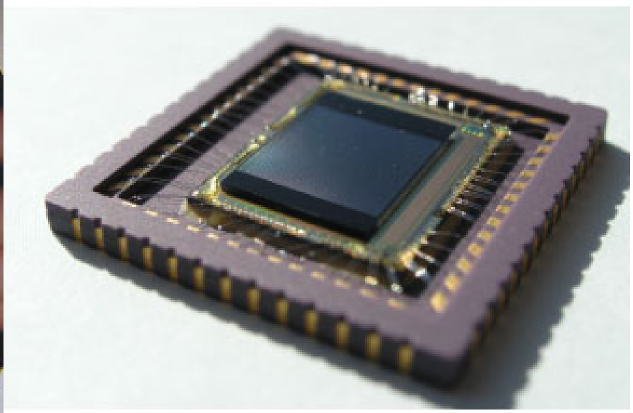


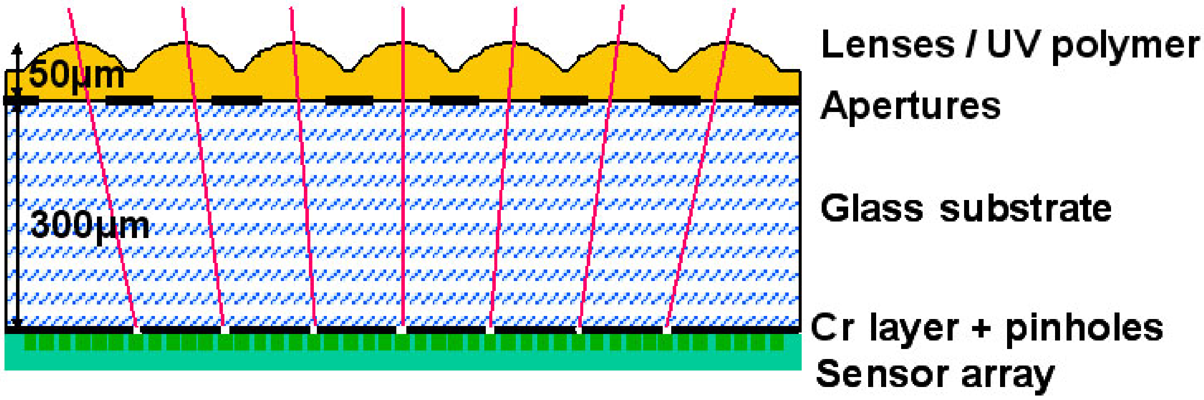
Deutscher Zukunftspreis 2008 - IOF Jena + OSRAM

H) Micro- and nano-optics

- ultra small camera







Insect inspired camera system develop at Fraunhofer Institute IOF Jena

I) Relativistic optics

Figure 3. Two relativistic lasers: (a) Helios circa 1980 at LANL was the first relativistic laser with $a_0 \sim 1$ at a millihertz repetition rate. [Courtesy of LANL.] (b) The λ^3 laser at the University of Michigan has an $a_0 \sim 1$ at a kilohertz repetition rate.

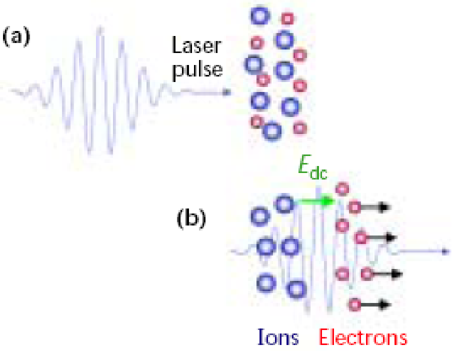
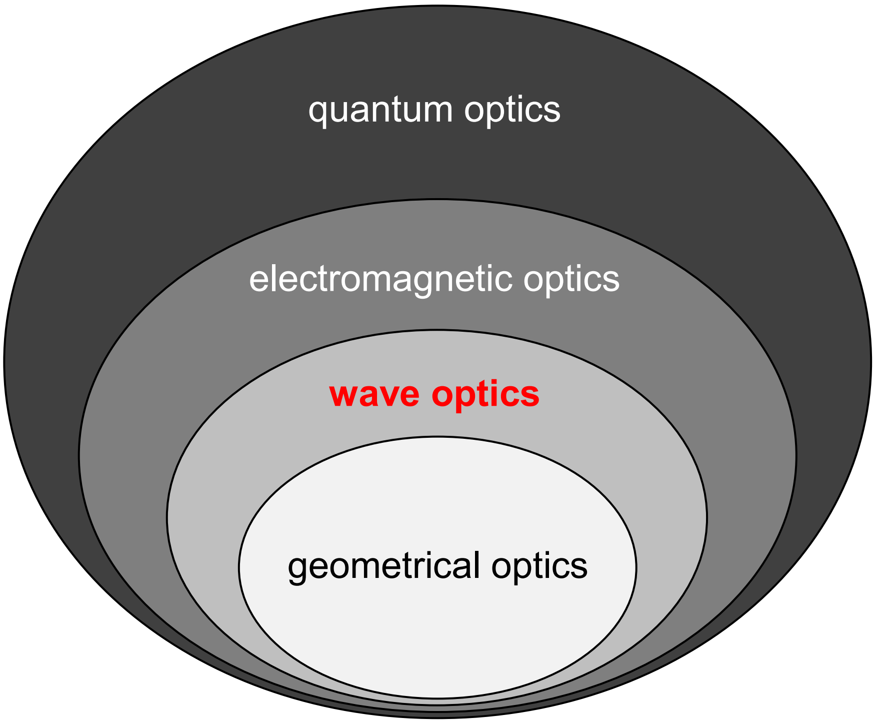


Figure 4. Relativistic rectification in plasma: (a) a high-intensity pulse before it enters the plasma and (b) the $\vec{v} \times \vec{B}$ that pushes the first plasma electrons. The electrons drag the heavy ions behind them like a horse pulling a cart. The electrostatic field that is generated is almost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

field is created within a plasma, the medium cannot be broken down further and can support these extremely high fields. This effect, predicted by Tajima and Dawson in 1989,[9] had to await the development of relativistic intensity lasers to be demonstrated.[14] Electron energies as high as 200 MeV (Ref. 15) have been demonstrated by use of a 50-TW laser focused on a gas jet. Because the electrons drag the ions, a beam of ions is also observed concomitant with the electron beam. The generation of protons with energy as high as 56 MeV has been reported by LLNL.[16]

J) Schematic of optics



- geometrical optics
  - λ << size of objects $\rightarrow$ daily experience
  - optical instruments, optical imaging
  - intensity, direction, ~~coherence, phase, polarization, photons~~
    G: Intensität, Richtung, ~~Kohärenz, Phase, Polarisation, Photon~~

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

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

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

- in this lecture
  - electromagnetic optics and wave optics

- no quantum optics → subject of advanced lectures

K) Literature

- Fundamental

1. **Saleh, Teich, 'Fundamenals of Photonics', Wiley (1992)**
   in German: "Grundlagen der Photonik" Wiley (2008)
2. **Hecht, 'Optic', Addison-Wesley (2001)**
   in German: "Optik", Oldenbourg (2005)
3. Mansuripur, 'Classical Optics and its Applications', Cambridge (2002)
4. Menzel, 'Photonics', Springer (2000)
5. Lipson, Lipson, Tannhäuser, 'Optik'; Springer (1997)
6. Born, Wolf, 'Principles of Optics', Pergamon
7. Sommerfeld, 'Optik'

- Advanced

1. W. Silvast, 'Laser Fundamentals',
2. Agrawal, 'Fiber-Optic Communication Systems', Wiley
3. Band, 'Light and Matter', Wiley, 2006
4. Karthe, Müller, 'Integrierte Optik', Teubner
5. Diels, Rudolph, 'Ultrashort Laser Pulse Phenomena', Academic
6. Yariv, 'Optical Electronics in modern Communications', Oxford
7. Snyder, Love, 'Optical Waveguide Theory', Chapman&Hall
8. Römer, 'Theoretical Optics', Wiley,2005.

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

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

1.1 Introduction

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

What is covered in this chapter?

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

1.2 Postulates

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

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

$$n = \frac{c}{c_n}$$ c_n – speed of light in the medium

C) optical path length ~ delay

i) homogeneous media

$$\boxed{nl}$$

ii) inhomogeneous media

$$\boxed{\int_A^B n(\mathbf{r})ds}$$

D) Fermat's principle

$$\boxed{\delta\int_A^B n(\mathbf{r})ds = 0}$$

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

1.3 Simple rules for propagation of light

A) Homogeneous media

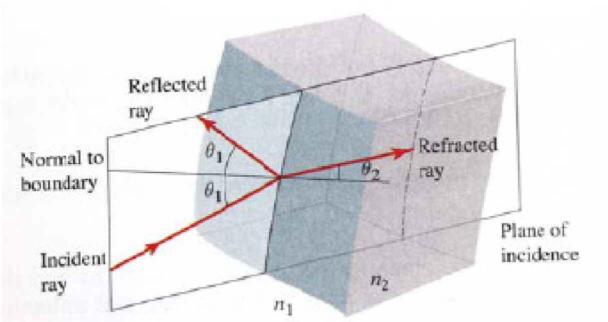
- n = const. → minimum delay = minimum distance
- Rays of light propagate on straight lines.

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 → reflected ray plus refracted ray
- The reflected ray obeys b).
- The refracted ray lies in the plane of incidence.



- The angle of refraction θ_2 depends on the angle of incidence θ_1 and is given by Snell's law:

$$\boxed{n_1 \sin\theta_1 = n_2 \sin\theta_2}$$

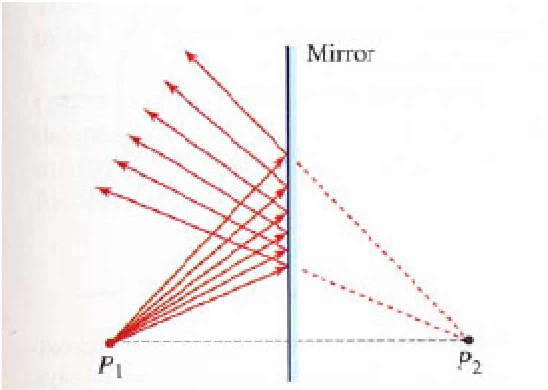
- no information about amplitude ratio.

1.4 Simple optical components

A) Mirror

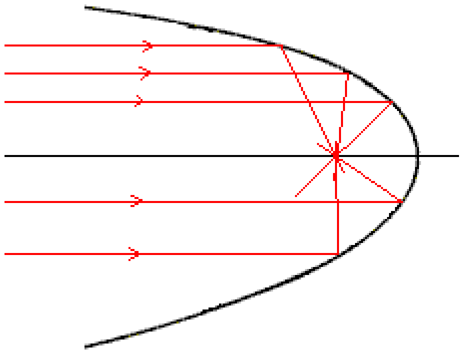
i) Planar mirror

- Rays originating from P_1 are reflected and seem to originate from P_2.



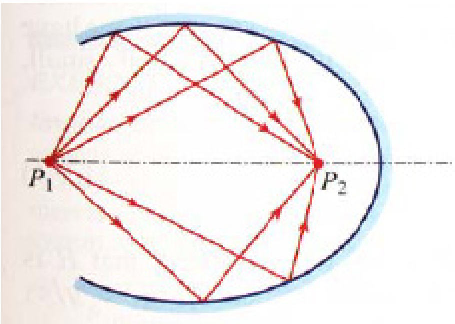
ii) Parabolic mirror

- Parallel rays converge in the focal point (focal length *f*).
- Applications: Telescope, collimator



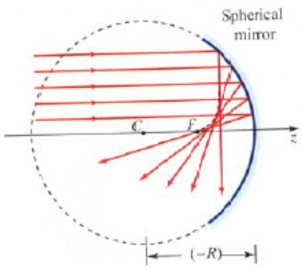
iii) Elliptic mirror

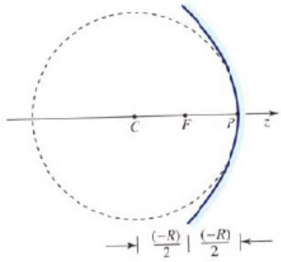
- Rays originating from focal point P_1 converge in the second focal point P_2



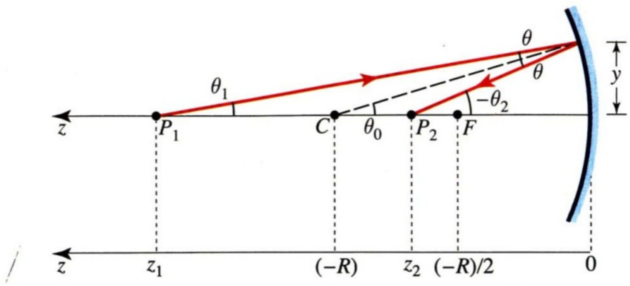
iv) Spherical mirror

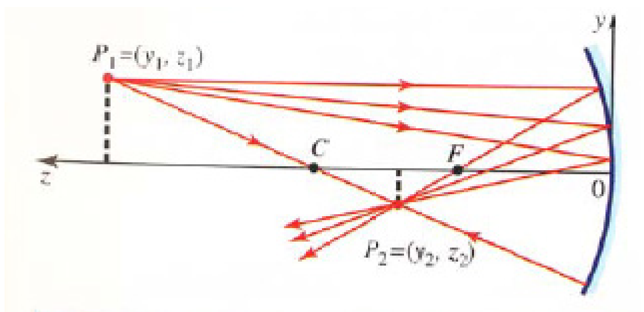
- Neither imaging like elliptical mirror nor focusing like parabolic mirror
- parallel rays cross the optical axis at different points
- connecting line of intersections of rays → *caustic*





- parallel, paraxial rays converge to the focal point $f = (-R)/2$
- 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 P1 are reflected and converge in point P2





$$\boxed{\frac{1}{z_1} + \frac{1}{z_2} \approx \frac{2}{(-R)}}$$ (imaging formula)

paraxial imaging: imaging formula and magnification

$\boxed{m = -z2 / z1}$ (proof given in exercises)

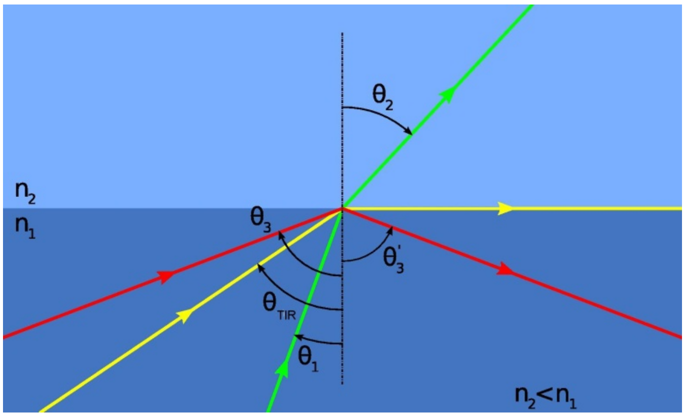
B) Planar interface

Snell's law: $n_1 \sin\theta_1 = n_2 \sin\theta_2$

for paraxial rays: $n_1\theta_1 = n_2\theta_2$

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

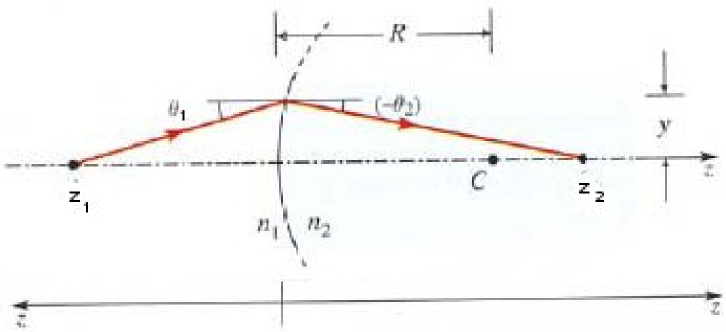
$$\theta_2 = \frac{\pi}{2} \rightarrow \boxed{\sin\theta_1 = \sin\theta_{\mathrm{TIR}} = \frac{n_2}{n_1}}$$

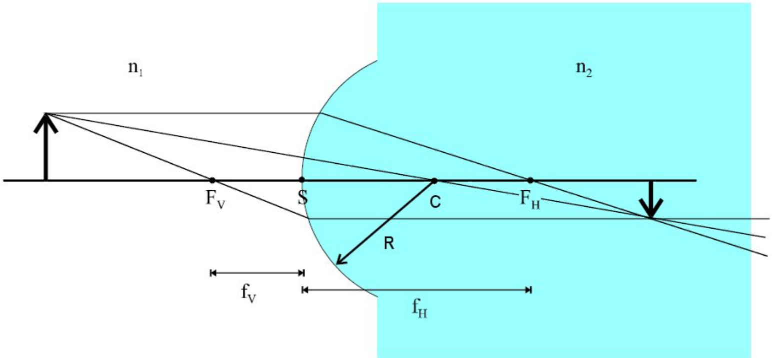


C) Spherical interface (paraxial)

– paraxial imaging

$$\theta_2 \approx \frac{n_1}{n_2}\theta_1 - \frac{n_2-n_1}{n_2}\frac{y}{R} \;(*)$$





$$\frac{n_1}{z_1} + \frac{n_2}{z_2} \approx \frac{n_2-n_1}{R} \text{ (imaging formula)}$$

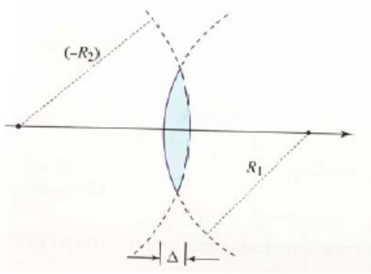
$$m = -\frac{n_1}{n_2}\frac{z_2}{z_1} \text{ (magnification)}$$

(Proof: exercise)

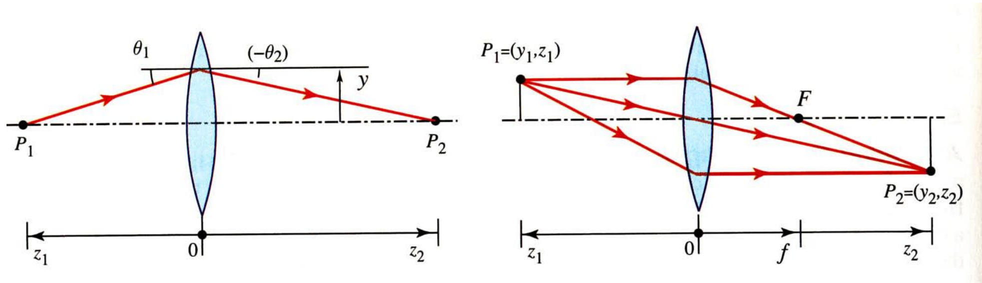
– if paraxiality is violated → aberration

– 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, Δ) apply (*) two times and assume y=const (Δ small)



$$\theta_2 \approx \theta_1 - \frac{y}{f} \text{ with focal length: } \frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$$\frac{1}{z_1} + \frac{1}{z_2} \approx \frac{1}{f} \text{ (imaging formula)} \qquad m = -\frac{z_2}{z_1} \text{ (magnification)}$$

(compare to spherical mirror)

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

– $n(\mathbf{r})$ - continuous function, fabricated by, e.g., doping

– curved trajectories → graded-index layer can act as, e.g., a lens

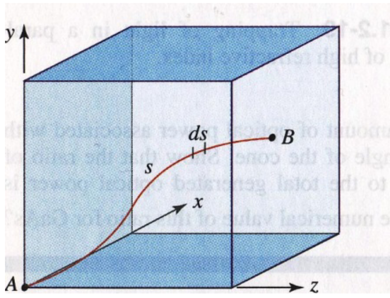
1.5.1 Ray equation

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

$$\boxed{\delta \int_A^B n(\mathbf{r}) ds = 0}$$

computation:

$$L = \int_A^B n\left[\mathbf{r}(s)\right] ds$$



variation of the path: $\mathbf{r}(s)+\delta\mathbf{r}(s)$

$$\delta L=\int_A^B \delta n ds+\int_A^B n\delta ds$$

$$\delta n=\operatorname{grad} n\cdot\delta\mathbf{r}$$

$$\begin{aligned}\delta ds&=\sqrt{\left(d\mathbf{r}+d\delta\mathbf{r}\right)^2}-\sqrt{\left(d\mathbf{r}\right)^2}\\&=\sqrt{\left(d\mathbf{r}\right)^2+2d\mathbf{r}\cdot d\delta\mathbf{r}+\left(d\delta\mathbf{r}\right)^2}-\sqrt{\left(d\mathbf{r}\right)^2}\\&\approx ds\sqrt{1+2\frac{d\mathbf{r}}{ds}\cdot\frac{d\delta\mathbf{r}}{ds}}-ds\\&\approx ds\left(1+\frac{d\mathbf{r}}{ds}\cdot\frac{d\delta\mathbf{r}}{ds}\right)-ds\\&=ds\frac{d\mathbf{r}}{ds}\cdot\frac{d\delta\mathbf{r}}{ds}\end{aligned}$$

$$\begin{aligned}\delta L&=\int_A^B\left(\operatorname{grad} n\cdot\delta\mathbf{r}+n\frac{d\mathbf{r}}{ds}\cdot\frac{d\delta\mathbf{r}}{ds}\right)ds\\&=\int_A^B\left(\operatorname{grad} n-\frac{d}{ds}\left(n\frac{d\mathbf{r}}{ds}\right)\right)\cdot\delta\mathbf{r}ds\end{aligned}$$

integration by parts and A,B fix

$\delta L=0$ for arbitrary variation

$$\boxed{\operatorname{grad} n=\frac{d}{ds}\left(n\frac{d\mathbf{r}}{ds}\right)}$$ ray equation

Possible solutions:

A) trajectory

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

– solve for $x(z)$, $y(z)$

– paraxial rays → ($ds\approx dz$)

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

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

B) homogeneous media

– straight lines

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

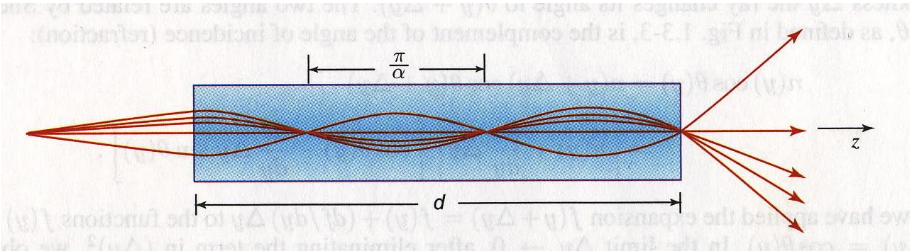
paraxial → $\frac{dy}{dz}\ll 1$ and $dz\approx ds$

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

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

for $n(y)-n_0 \ll 1$: $\frac{d^2y}{dz^2}=-\alpha^2y$ →

$$\boxed{\begin{aligned}y(z)&=y_0\cos\alpha z+\frac{\theta_0}{\alpha}\sin\alpha z\\ \theta(z)&=\frac{dy}{dz}=-y_0\alpha\sin\alpha z+\theta_0\cos\alpha z\end{aligned}}$$



1.5.2 The eikonal equation

– bridge between geometrical optics and wave

– eikonal S(**r**) = constant → planes perpendicular to rays

– from S(**r**) we can determine direction of rays ~ grad S(**r**) (like potential)

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

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

– geometrical optics: Fermat's or eikonal equation

$$S(\mathbf{r}_B)-S(\mathbf{r}_A)=\int_A^B\left|\operatorname{grad}S(\mathbf{r})\right|ds=\int_A^B n(\mathbf{r})ds$$

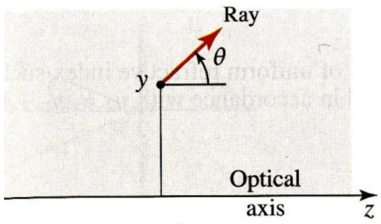
eikonal → optical path length ~ phase of the wave

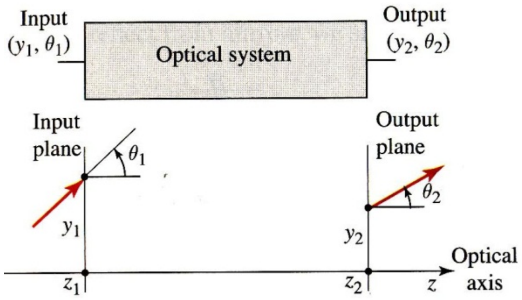
1.6 Matrix optics

- 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 (θ) → two algebraic equation → 2 x 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:

$$\begin{aligned} y_2 &= Ay_1 + B\theta_1 \\ \theta_2 &= Cy_1 + D\theta_1 \end{aligned} \rightarrow \begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} \rightarrow \mathbf{M} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

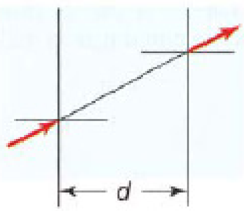
A=0: same θ_1 → same y_2 → focusing

D=0: same y_1 → same θ_2 → collimation

1.6.2 Matrices of optical elements

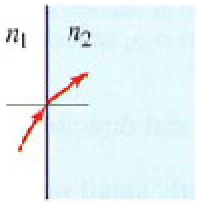
A) free space

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



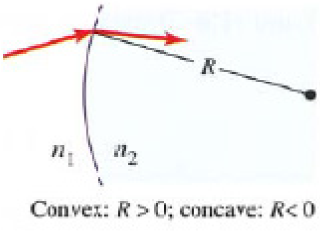
B) refraction on planar interface

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & n_1/n_2 \end{bmatrix}$$



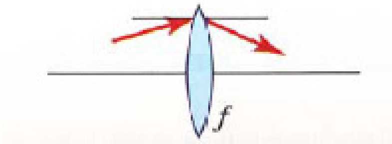
C) refraction on spherical interface

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ -(n_2 - n_1)/n_2 R & n_1/n_2 \end{bmatrix}$$



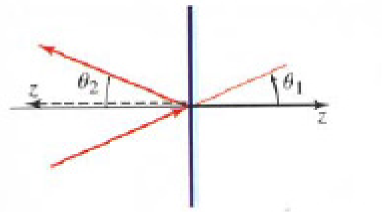
D) thin lens

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix}$$



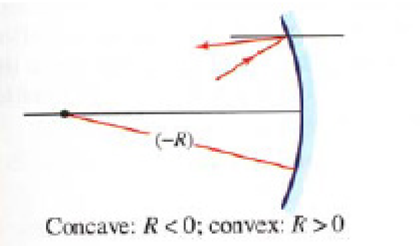
E) reflection on planar mirror

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



F) reflection on spherical mirror (compare to lens)

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 2/R & 1 \end{bmatrix}$$



1.6.3 Cascaded elements

$$\begin{bmatrix} y_{N+1} \\ \theta_{N+1} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} \rightarrow \mathbf{M} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \qquad M = M_N \ldots M_2 M$$

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 their 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 the influence of many point charges on the electromagnetic field in a homogeneously distributed response of the solid state on the excitation by 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 electrodynamic theory. Here we will not redo this derivation. We will rather start directly from the averaged Maxwell's equations.

$$\boxed{\begin{array}{ll} \mathrm{rot}\,\mathbf{E}(\mathbf{r},t) = -\dfrac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t} & \mathrm{div}\,\mathbf{D}(\mathbf{r},t) = \rho_{\mathrm{ext}}(\mathbf{r},t) \\ \mathrm{rot}\,\mathbf{H}(\mathbf{r},t) = \mathbf{j}_{\mathrm{makr}}(\mathbf{r},t) + \dfrac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t} & \mathrm{div}\,\mathbf{B}(\mathbf{r},t) = 0 \end{array}}$$

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

$$\boxed{\begin{array}{l} \mathbf{D}(\mathbf{r},t) = \varepsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t) \\ \mathbf{H}(\mathbf{r},t) = \dfrac{1}{\mu_0}\left[\mathbf{B}(\mathbf{r},t) - \mathbf{M}(\mathbf{r},t)\right] \end{array}}$$

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

  electric constant and magnetic constant are connected by the speed of light in vacuum c as

$$\varepsilon_0 = \frac{1}{\mu_0 c^2}$$

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

- The effect of the medium gives rise to polarization $\mathbf{P}(\mathbf{r},t) = f\left[\mathbf{E}\right]$ and magnetization $\mathbf{M}(\mathbf{r},t) = f\left[\mathbf{B}\right]$. → In order to solve Maxwell's equations we need material models which describe these quantities.
- In optics at visible wavelength, we generally deal with non-magnetizable media. Hence we can assume $\mathbf{M}(\mathbf{r},t) = 0$. Exceptions to this general property are metamaterials which might possess some artificial effective magnetization properties resulting in $\mathbf{M}(\mathbf{r},t) \neq 0$.

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 Maxwell's equations.

- free charge density (G: Dichte freier Ladungsträger)

  $\rho_{\mathrm{ext}}(\mathbf{r},t)$ [As/m³]

- macroscopic current density (G: makroskopische Stromdichte) consisting oft wo contributions

  $\mathbf{j}_{\mathrm{makr}}(\mathbf{r},t) = \mathbf{j}_{\mathrm{cond}}(\mathbf{r},t) + \mathbf{j}_{\mathrm{conv}}(\mathbf{r},t)$ [A/m²]

  - *conductive current density (G: Konduktionsstromdichte)*

    $\mathbf{j}_{\mathrm{cond}}(\mathbf{r},t) = f\left[\mathbf{E}\right]$

  - *convective* current density (G: Konvektionsstromdichte)

    $\mathbf{j}_{\mathrm{conv}}(\mathbf{r},t) = \rho_{\mathrm{ext}}(\mathbf{r},t)\mathbf{v}(\mathbf{r},t)$

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

$$\frac{\partial \rho_{\mathrm{ext}}(\mathbf{r},t)}{\partial t} \approx 0 \rightarrow \mathbf{j}_{\mathrm{conv}}(\mathbf{r},t) = 0$$

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

$$\boxed{\begin{array}{ll} \mathrm{rot}\,\mathbf{E}(\mathbf{r},t) = -\mu_0 \dfrac{\partial \mathbf{H}(\mathbf{r},t)}{\partial t} & \varepsilon_0\,\mathrm{div}\,\mathbf{E}(\mathbf{r},t) = -\mathrm{div}\,\mathbf{P}(\mathbf{r},t) \\ \mathrm{rot}\,\mathbf{H}(\mathbf{r},t) = \mathbf{j}(\mathbf{r},t) + \dfrac{\partial \mathbf{P}(\mathbf{r},t)}{\partial t} + \varepsilon_0 \dfrac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t} & \mathrm{div}\,\mathbf{H}(\mathbf{r},t) = 0 \end{array}}$$

  - In optics, the medium (or more precisely the mathematical material model) determines the dependence of the induced polarization on the electric field $\mathbf{P}(\mathbf{E})$ and the dependence of the induced (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 $\mathbf{P}$ and current density $\mathbf{j}$ are zero (most simple material model). Hence we can solve Maxwell's equations directly.

Remarks:

- Even though the Maxwell's equations, in the way they have been written above are derived to describe light, i.e. electromagnetic fields at optical frequencies, they are simultaneously valid for other frequency ranges as well. Furthermore, since Maxwell's equations are linear equations as long as the material does not introduce any nonlinearity the principle of superposition holds. Hence we can decompose the comprehensive electromagnetic fields into components of different frequency ranges. In turn this means that we do not have to take care of any slow electromagnetic phenomena, e.g. electrostatics or radio wave, in our formulation of Maxwell's equations since they can be split off from our optical problem and can be treated separately.
- We can define a bound charge density (G: Dichte gebundener Ladungsträger) as the source of spatially changing polarization $\mathbf{P}$.

$$\rho_{\mathrm{b}}(\mathbf{r},t) = -\mathbf{div}\,\mathbf{P}(\mathbf{r},t)$$

- Analogously we can define a bound current density (G: Stromdichte gebundener Ladungsträger) as the source of the temporal variation of the polarization $\mathbf{P}$.

$$\mathbf{j}_{\mathrm{b}}(\mathbf{r},t) = \frac{\partial \mathbf{P}(\mathbf{r},t)}{\partial t}$$

- This essentially means that we can describe the same physics in two different ways since currents are in principle moving charges (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 way.
- 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 $\sin(x)$.
- Hence we use the following convention in this lecture to distinguish between the two types of fields.

real physical field: $\mathbf{E}_{\mathrm{r}}(\mathbf{r},t)$

complex mathematical representation: $\mathbf{E}(\mathbf{r},t)$

– Here we define their relation as

$$\mathbf{E}_{\mathbf{r}}(\mathbf{r},t)=\tfrac{1}{2}\left[\mathbf{E}(\mathbf{r},t)+\mathbf{E}^{*}(\mathbf{r},t)\right]=\mathrm{Re}\left[\mathbf{E}(\mathbf{r},t)\right]$$

However, this relation can be defined differently in different textbooks.

– This means in general: For calculations we use the complex fields $[\mathbf{E}(\mathbf{r},t)]$ 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 in the description of the material response or the field dynamics. In this case you would have to go back to real quantities before you compute these multiplication. This becomes relevant for, e.g., calculation of the Poynting vector, as can be seen in a 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 (stationary fields)

– harmonic dependence on temporal coordinate

– $\sim\exp(-\mathbf{i}\omega t)$ → phase is fixed → coherent, infinite wave train e.g.:

$$\mathbf{E}(\mathbf{r},t)=\mathbf{E}(\mathbf{r})\exp(-\mathbf{i}\omega t)$$

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

B) polychromatic light (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)

$$\boxed{\begin{aligned}\mathbf{E}(\mathbf{r},t)&=\int_{-\infty}^{\infty}\bar{\mathbf{E}}(\mathbf{r},\omega)\exp(-\mathbf{i}\omega t)d\omega\\ \bar{\mathbf{E}}(\mathbf{r},\omega)&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{E}(\mathbf{r},t)\exp(\mathbf{i}\omega t)dt\end{aligned}}$$

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

In order to solve Maxwell's equations more easily we would like to introduce a Fourier decompositions of the fields in the Maxwell's equations.

For this purpose, we need to find out how a time derivative of a dynamic variable can be calculated in Fourier space.

A simple rule for the transformation a time derivative into Fourier space ca be obtained using integration by parts:

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

Thus, a time derivative in real space transforms into a simple product with $-i\omega$ in Fourier space.

→ rule: $\boxed{\frac{\partial}{\partial t}\xrightarrow{\text{FT}}-i\omega}$

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

$$\boxed{\begin{array}{ll}\operatorname{rot}\overline{\mathbf{E}}(\mathbf{r},\omega) = i\omega\mu_0\overline{\mathbf{H}}(\mathbf{r},\omega) & \varepsilon_0\operatorname{div}\overline{\mathbf{E}}(\mathbf{r},\omega) = -\operatorname{div}\overline{\mathbf{P}}(\mathbf{r},\omega) \\ \operatorname{rot}\overline{\mathbf{H}}(\mathbf{r},\omega) = \overline{\mathbf{j}}(\mathbf{r},\omega) - i\omega\overline{\mathbf{P}}(\mathbf{r},\omega) - i\omega\varepsilon_0\overline{\mathbf{E}}(\mathbf{r},\omega) & \operatorname{div}\mathbf{H}(\mathbf{r},\omega) = 0\end{array}}$$

2.1.4 From Maxwell's equations to the wave equation

Maxwell's equations provide the basis to derive all possible mathematical solutions of electromagnetic problems. However very often we are interested just in the radiation fields which can be described more easily by an adapted equation, which is the so-called wave equation. From Maxwell's equations it is straight forward to derive the wave equation by using the two curl equations.

A) Time domain derivation

We start from applying the curl operator (rot) a second time on $\operatorname{rot}\mathbf{E}(\mathbf{r},t)=\ldots$ and substitute $\operatorname{rot}\mathbf{H}$ with the other Maxwell equation.

$$\mathbf{rotrot}\;\;\mathbf{E}(\mathbf{r},t) = -\mu_0\,\mathbf{rot}\;\frac{\partial\mathbf{H}(\mathbf{r},t)}{\partial t} = -\mu_0\frac{\partial}{\partial t}\left[\mathbf{j}(\mathbf{r},t)+\frac{\partial\mathbf{P}(\mathbf{r},t)}{\partial t}+\varepsilon_0\frac{\partial\mathbf{E}(\mathbf{r},t)}{\partial t}\right]$$

We find the wave equation for the electric field as

$$\boxed{\mathbf{rotrot}\;\;\mathbf{E}(\mathbf{r},t)+\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{E}(\mathbf{r},t) = -\mu_0\frac{\partial\mathbf{j}(\mathbf{r},t)}{\partial t}-\mu_0\frac{\partial^2\mathbf{P}(\mathbf{r},t)}{\partial t^2}}$$

The blue terms require knowledge of the material model. Additionally, we have to make sure that all other Maxwell's equations are fulfilled. This holds in particular for the divergence of the electric field:

$$\boxed{\mathbf{div}\left[\varepsilon_0\mathbf{E}(\mathbf{r},t)+\mathbf{P}(\mathbf{r},t)\right]=0}$$

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

$$\boxed{\frac{\partial \mathbf{H}(\mathbf{r},t)}{\partial t} = -\frac{1}{\mu_0}\mathbf{rot}\ \mathbf{E}(\mathbf{r},t)}$$

Remarks:

- An analog procedure is possible also for the magnetic field $\mathbf{H}$, i.e., we can derive a wave equation for the magnetic field as well.
- Normally, the wave equation for the electric field $\mathbf{E}$ is more convenient, because the material model defines $\mathbf{P}(\mathbf{E})$.
- However, for inhomogeneous media $\mathbf{H}$ can sometimes be the better choice for the numerical solution of the partial differential equation since it forms a hermitian operator.
- analog procedure possible for $\mathbf{H}$ instead of $\mathbf{E}$
- generally, wave equation for $\mathbf{E}$ is more convenient, because $\mathbf{P}(\mathbf{E})$ given
- for inhomogeneous media $\mathbf{H}$ can sometimes be better choice

B) Frequency domain derivation

We can do the same procedure to derive the wave equation also directly in the Fourier domain and find

$$\boxed{\mathbf{rotrot}\ \bar{\mathbf{E}}(\mathbf{r},\omega) - \frac{\omega^2}{c^2}\bar{\mathbf{E}}(\mathbf{r},\omega) = \mathtt{i}\omega\mu_0\bar{\mathbf{j}}(\mathbf{r},\omega) + \mu_0\omega^2\bar{\mathbf{P}}(\mathbf{r},\omega)}$$

and

$$\boxed{\mathbf{div}\left[\varepsilon_0\bar{\mathbf{E}}(\mathbf{r},\omega) + \bar{\mathbf{P}}(\mathbf{r},\omega)\right] = 0}$$

- magnetic field:

$$\boxed{\bar{\mathbf{H}}(\mathbf{r},\omega) = -\frac{\mathtt{i}}{\omega\mu_0}\mathbf{rot}\ \bar{\mathbf{E}}(\mathbf{r},\omega)}$$

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

$$\mathbf{E}(\mathbf{r},t) = \int_{-\infty}^{\infty} \bar{\mathbf{E}}(\mathbf{r},\omega)\exp(-\mathtt{i}\omega t)d\omega$$

2.1.5 Decoupling of the vectorial wave equation

So far we have seen that for the general problem of electromagnetic waves all 3 vectorial 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.

$$\mathbf{rot}\ \mathbf{rot}\ \overline{\mathbf{E}} = \mathbf{grad}\ \mathbf{div}\ \overline{\mathbf{E}} - \Delta\overline{\mathbf{E}} = \begin{bmatrix} \frac{\partial}{\partial x}\left(\frac{\partial \overline{E}_x}{\partial x} + \frac{\partial \overline{E}_z}{\partial z}\right) \\ 0 \\ \frac{\partial}{\partial z}\left(\frac{\partial \overline{E}_x}{\partial x} + \frac{\partial \overline{E}_z}{\partial z}\right) \end{bmatrix} - \begin{bmatrix} \Delta^{(2)}\overline{E}_x \\ \Delta^{(2)}\overline{E}_y \\ \Delta^{(2)}\overline{E}_z \end{bmatrix}$$

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

- decomposition of electric field

$$\overline{\mathbf{E}} = \overline{\mathbf{E}}_\perp + \overline{\mathbf{E}}_\shortparallel \quad \rightarrow \quad \overline{\mathbf{E}}_\perp = \begin{pmatrix} 0 \\ \overline{E}_y \\ 0 \end{pmatrix}, \qquad \overline{\mathbf{E}}_\shortparallel = \begin{pmatrix} \overline{E}_x \\ 0 \\ \overline{E}_z \end{pmatrix}$$

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

Hence we obtain two wave equations for the $\overline{\mathbf{E}}_\perp$ and $\overline{\mathbf{E}}_\shortparallel$ fields.

- gives two decoupled wave equations

$$\boxed{\begin{aligned} \Delta^{(2)}\overline{\mathbf{E}}_\perp(\mathbf{r},\omega) + \frac{\omega^2}{c^2}\overline{\mathbf{E}}_\perp(\mathbf{r},\omega) &= -\mathbf{i}\omega\mu_0\overline{\mathbf{j}}_\perp(\mathbf{r},\omega) - \mu_0\omega^2\overline{\mathbf{P}}_\perp(\mathbf{r},\omega) \\ \Delta^{(2)}\overline{\mathbf{E}}_\shortparallel(\mathbf{r},\omega) + \frac{\omega^2}{c^2}\overline{\mathbf{E}}_\shortparallel(\mathbf{r},\omega) - \mathbf{grad}^{(2)}\ \mathbf{div}^{(2)}\ \overline{\mathbf{E}}_\shortparallel &= -\mathbf{i}\omega\mu_0\overline{\mathbf{j}}_\shortparallel(\mathbf{r},\omega) - \mu_0\omega^2\overline{\mathbf{P}}_\shortparallel(\mathbf{r},\omega) \end{aligned}}$$

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 $\bar{\mathbf{E}}_\perp$ and $\bar{\mathbf{E}}_{\|}$ can be treated separately
- propagation of $\bar{\mathbf{E}}_\perp$ is described by scalar equation
- similarly the other field components can be described by a scalar equation for $\bar{\mathbf{H}}_\perp$
- alternative notations:

  $\perp$ → s → TE (transversal electric)

  $\|$ → p → 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 $\mathbf{P}(\mathbf{E})$ and $\mathbf{j}(\mathbf{E})$ are linear (no $\mathbf{E}^2$ etc.), we can use linear response theory.

For the polarization $\mathbf{P}(\mathbf{E})$ (for $\mathbf{j}(\mathbf{E})$ very similar):

- description in both time and frequency domain possible

- **In time domain:** we introduce the response function (G: Responsfunktion)
  $\mathbf{E}(\mathbf{r},t)$ → medium (response function) → $\mathbf{P}(\mathbf{r},t)$

$$P_i(\mathbf{r},t)=\varepsilon_0\sum_j\int_{-\infty}^{t}R_{ij}(\mathbf{r},t-t')E_j(\mathbf{r},t')dt'$$

with $\hat{\mathrm{R}}$ being a 2nd rank tensor
$i=x,y,z$ and summing over $j=x,y,z$

- **In frequency domain:** we introduce the *susceptibility (G: Suszeptibilität)*
  $\overline{\mathbf{E}}(\mathbf{r},\omega)$ → medium (susceptibility) → $\overline{\mathbf{P}}(\mathbf{r},\omega)$

$$\overline{P}_i(\mathbf{r},\omega)=\varepsilon_0\sum_j\chi_{ij}(\mathbf{r},\omega)\overline{E}_j(\mathbf{r},\omega)$$

- response function and susceptibility are linked via Fourier transform (convolution theorem)

$$R_{ij}(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\chi_{ij}(\omega)\exp(-\mathrm{i}\omega t)d\omega$$

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

$$\boxed{\begin{gathered}\mathbf{rotrot}\ \overline{\mathbf{E}}(\mathbf{r},\omega)-\frac{\omega^2}{c^2}\overline{\mathbf{E}}(\mathbf{r},\omega)=\mu_0\omega^2\overline{\mathbf{P}}(\mathbf{r},\omega)\\ \text{or}\\ \Delta\overline{\mathbf{E}}(\mathbf{r},\omega)+\frac{\omega^2}{c^2}\overline{\mathbf{E}}(\mathbf{r},\omega)-\mathbf{graddiv}\overline{\mathbf{E}}(\mathbf{r},\omega)=-\mu_0\omega^2\overline{\mathbf{P}}(\mathbf{r},\omega)\end{gathered}}$$

- and for auxiliary fields

$$\boxed{\overline{\mathbf{D}}(\mathbf{r},\omega)=\varepsilon_0\overline{\mathbf{E}}(\mathbf{r},\omega)+\overline{\mathbf{P}}(\mathbf{r},\omega)}$$

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

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

frequency domain		time domain description
$\bar{\mathbf{P}}(\mathbf{r},\omega) = \varepsilon_0 \chi \ \bar{\mathbf{E}}(\mathbf{r},\omega)$	$\leftrightarrow$	$\mathbf{P}(\mathbf{r},t) = \varepsilon_0 \chi \mathbf{E}(\mathbf{r},t)$ (unphysical!)
$\bar{\mathbf{D}}(\mathbf{r},\omega) = \varepsilon_0 \varepsilon \ \bar{\mathbf{E}}(\mathbf{r},\omega)$	$\leftrightarrow$	$\mathbf{D}(\mathbf{r},t) = \varepsilon_0 \varepsilon \mathbf{E}(\mathbf{r},t)$ with $\varepsilon = 1 + \chi$

Maxwell: $\mathbf{div}\,\bar{\mathbf{D}} = 0 \ \to \ \mathbf{div}\,\bar{\mathbf{E}}(\mathbf{r},\omega) = 0$ for $\varepsilon(\omega) \neq 0$

$$\to \boxed{\Delta\ \bar{\mathbf{E}}(\mathbf{r},\omega) + \frac{\omega^2}{c^2}\varepsilon\ \bar{\mathbf{E}}(\mathbf{r},\omega) = 0} \to \boxed{\Delta\ \mathbf{E}(\mathbf{r},t) - \frac{\varepsilon}{c^2}\frac{\partial^2}{\partial t^2}\ \mathbf{E}(\mathbf{r},t) = 0}$$

- 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 → $\chi(\omega)$

$$\bar{\mathbf{P}}(\mathbf{r},\omega) = \varepsilon_0 \chi(\omega)\ \bar{\mathbf{E}}(\mathbf{r},\omega)$$
$$\bar{\mathbf{D}}(\mathbf{r},\omega) = \varepsilon_0 \varepsilon(\omega)\ \bar{\mathbf{E}}(\mathbf{r},\omega)$$

$$\mathbf{div}\ \bar{\mathbf{D}}(\mathbf{r},\omega) = 0 \curvearrowright \mathbf{div}\ \bar{\mathbf{E}}(\mathbf{r},\omega) = 0 \text{ for } \varepsilon(\omega) \neq 0$$

$$\to \boxed{\Delta\ \bar{\mathbf{E}}(\mathbf{r},\omega) + \frac{\omega^2}{c^2}\varepsilon\left(\omega\right)\ \bar{\mathbf{E}}(\mathbf{r},\omega) = 0}$$ Helmholtz equation

- This description is sufficient for many materials.

C) linear, **inhomogeneous**, isotropic, dispersive media → $\chi(\mathbf{r},\omega)$

$$\bar{\mathbf{P}}(\mathbf{r},\omega) = \varepsilon_0 \chi(\mathbf{r},\omega)\ \bar{\mathbf{E}}(\mathbf{r},\omega),$$
$$\bar{\mathbf{D}}(\mathbf{r},\omega) = \varepsilon_0 \varepsilon(\mathbf{r},\omega)\ \bar{\mathbf{E}}(\mathbf{r},\omega).$$

$$\operatorname{div}\bar{\mathbf{D}}(\mathbf{r},\omega) = 0$$
$$\operatorname{div}\bar{\mathbf{D}}(\mathbf{r},\omega) = \varepsilon_0\varepsilon(\mathbf{r},\omega)\mathbf{div}\ \bar{\mathbf{E}}(\mathbf{r},\omega) + \varepsilon_0\bar{\mathbf{E}}(\mathbf{r},\omega)\bullet\mathbf{grad}\ \varepsilon(\mathbf{r},\omega) = 0,$$

$$\to \operatorname{div}\bar{\mathbf{E}}(\mathbf{r},\omega) = -\frac{\mathbf{grad}\ \varepsilon(\mathbf{r},\omega)}{\varepsilon(\mathbf{r},\omega)}\bullet\bar{\mathbf{E}}(\mathbf{r},\omega).$$

$$\boxed{\Delta\ \bar{\mathbf{E}}(\mathbf{r},\omega) + \frac{\omega^2}{c^2}\varepsilon\left(\mathbf{r},\omega\right)\ \bar{\mathbf{E}}(\mathbf{r},\omega) = -\mathbf{grad}\left\{\frac{\mathbf{grad}\ \varepsilon(\mathbf{r},\omega)}{\varepsilon(\mathbf{r},\omega)}\bullet\bar{\mathbf{E}}(\mathbf{r},\omega)\right\}}$$

- All field components couple.

D) linear, homogeneous, **anisotropic**, dispersive media $\to \chi_{ij}(\omega)$

$$\bar{P}_i(\mathbf{r},\omega) = \varepsilon_0 \sum_j \chi_{ij}(\omega)\ \bar{E}_j(\mathbf{r},\omega)$$
$$\bar{D}_i(\mathbf{r},\omega) = \varepsilon_0 \sum_j \varepsilon_{ij}(\omega)\ \bar{E}_j(\mathbf{r},\omega).$$

→ see chapter on crystal optics

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

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

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

I) Interaction of light with bound electrons and the lattice

The contributions of bound electrons and lattice vibrations in dielectrics and semiconductors give rise to the polarization $\mathbf{P}$. The lattice vibrations (phonons) are the ionic part of the material model. 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 ω).

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 $\mathbf{j}$. 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 bound electrons / lattice vibrations and free electrons separately. Later we will join the results in a generalized material model which holds for many common optical materials..

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 $\mathbf{E}(\mathbf{r},t)$ gives rise to a displacement $\mathbf{s}(\mathbf{r},t)$ of charged particles from their equilibrium positions. In the easiest approach this can be modeled by a driven harmonic oscillator:

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

- resonance frequency (electronic transition) → ω_0
- damping → g
- charge → q
- mass → m

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

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

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

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

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

$$\frac{\partial^2}{\partial t^2}\mathbf{P}(\mathbf{r},t) + g\frac{\partial}{\partial t}\mathbf{P}(\mathbf{r},t) + \omega_0^2\mathbf{P}(\mathbf{r},\mathbf{t}) = \frac{q^2 N}{m}\mathbf{E}(\mathbf{r},t) = \varepsilon_0 f\mathbf{E}(\mathbf{r},t)$$

with oscillator strength $f = \frac{1}{\varepsilon_0}\frac{e^2 N}{m}$, for q=-e (electrons)

This equation is easy to solve in Fourier domain:

$$-\omega^2\bar{\mathbf{P}}(\mathbf{r},\omega) - \mathbf{i}g\omega\bar{\mathbf{P}}(\mathbf{r},\omega) + \omega_0^2\bar{\mathbf{P}}(\mathbf{r},\omega) = \varepsilon_0 f\ \bar{\mathbf{E}}(\mathbf{r},\omega)$$

$$\rightarrow \boxed{\bar{\mathbf{P}}(\mathbf{r},\omega) = \frac{\varepsilon_0 f}{\left(\omega_0^2 - \omega^2\right) - \mathbf{i}g\omega}\bar{\mathbf{E}}(\mathbf{r},\omega)}$$

with $\bar{\mathbf{P}}(\mathbf{r},\omega) = \varepsilon_0\chi(\omega)\bar{\mathbf{E}}(\mathbf{r},\omega)$ → $\chi(\omega) = \frac{f}{\left(\omega_0^2 - \omega^2\right) - \mathbf{i}g\omega}$

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

$$\bar{\mathbf{P}}(\mathbf{r},\omega) = \varepsilon_0\sum_j\left\{\frac{f_j}{\left(\omega_{0j}^2 - \omega^2\right) - \mathbf{i}g_j\omega}\right\}\bar{\mathbf{E}}(\mathbf{r},\omega) = \varepsilon_0\chi(\omega)\bar{\mathbf{E}}(\mathbf{r},\omega)$$

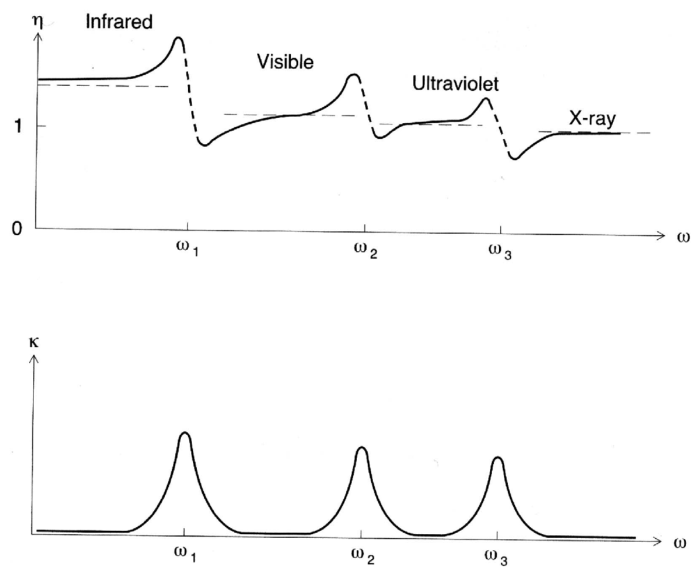
$$\boxed{\chi(\omega) = \sum_j\left\{\frac{f_j}{\left(\omega_{0j}^2 - \omega^2\right) - \mathbf{i}g_j\omega}\right\}}$$

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

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

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

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



2.2.3 Conductive current and conductivity

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

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

The resulting induced current density is given by

$$\mathbf{j}(\mathbf{r},t)=-Ne\frac{\partial}{\partial t}\mathbf{s}(\mathbf{r},t)$$

and the governing dynamic equation reads as

$$\frac{\partial}{\partial \mathbf{t}}\mathbf{j}(\mathbf{r},t)+g\mathbf{j}(\mathbf{r},t)=\frac{e^2N}{m}\mathbf{E}(\mathbf{r},t)=\varepsilon_0\omega_{\mathbf{p}}^2\mathbf{E}(\mathbf{r},t)$$

with plasma frequency $\omega_{\mathbf{p}}^2=f=\frac{1}{\varepsilon_0}\frac{e^2N}{m}$

Again we solve this equation in Fourier domain:

$$-\mathbf{i}\omega\bar{\mathbf{j}}(\mathbf{r},\omega)+g\bar{\mathbf{j}}(\mathbf{r},\omega)=\varepsilon_0\omega_{\mathbf{p}}^2\ \bar{\mathbf{E}}(\mathbf{r},\omega)$$

$$\rightarrow \boxed{\bar{\mathbf{j}}(\mathbf{r},\omega)=\frac{\varepsilon_0\ \omega_{\mathbf{p}}^2}{g-\mathbf{i}\omega}\bar{\mathbf{E}}(\mathbf{r},\omega)=\sigma(\omega)\bar{\mathbf{E}}(\mathbf{r},\omega).}$$

Here we introduced the complex frequency dependent conductivity

$$\boxed{\sigma(\omega)=\frac{\varepsilon_0\ \omega_{\mathbf{p}}^2}{g-\mathbf{i}\omega}=-\mathbf{i}\frac{\varepsilon_0\ \omega\omega_{\mathbf{p}}^2}{-\omega^2-\mathbf{i}g\omega}.}$$

Remarks on plasma frequency

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

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

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):

$$\frac{\partial}{\partial t}\mathbf{j}+g\mathbf{j}=\varepsilon_0\omega_{\mathbf{p}}^2\mathbf{E}(\mathbf{r},t)$$

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

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

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

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

We can substitute the divergence of the current density and find:

$$-\frac{\partial^2}{\partial t^2}\rho-g\frac{\partial}{\partial t}\rho=\omega_{\mathbf{p}}^2\rho$$

$$\boxed{\frac{\partial^2}{\partial t^2}\rho+g\frac{\partial}{\partial t}\rho+\omega_{\mathrm{p}}^2=0}\ \rightarrow \text{harmonic oscillator equation}$$

Hence, the plasma frequency $\omega_{\mathbf{p}}$ 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 $\bar{\mathbf{j}}(\mathbf{r},\omega)$ and $\bar{\mathbf{P}}(\mathbf{r},\omega)$ into the wave equation (in Fourier domain)

$$\mathbf{rot}\mathbf{rot}\ \bar{\mathbf{E}}(\mathbf{r},\omega)-\frac{\omega^2}{c^2}\bar{\mathbf{E}}(\mathbf{r},\omega)=\mu_0\omega^2\bar{\mathbf{P}}(\mathbf{r},\omega)+\mathbf{i}\omega\mu_0\bar{\mathbf{j}}(\mathbf{r},\omega)$$

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

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

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

$$\boxed{\textbf{rotrot }\overline{\mathbf{E}}(\mathbf{r},\omega)=\frac{\omega^2}{c^2}\varepsilon(\omega)\overline{\mathbf{E}}(\mathbf{r},\omega)}$$

Here, we introduced the generalized complex dielectric function

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

So, in general we have

$$\boxed{\varepsilon(\omega)=1+\sum_j\left\{\frac{f_j}{\left(\omega_{0j}^2-\omega^2\right)-\mathbf{i}g_j\omega}\right\}+\frac{\omega_{\mathbf{p}}^2}{-\omega^2-\mathbf{i}g\omega},}$$

because (from before)

$$\chi(\omega)=\sum_j\left\{\frac{f_j}{\left(\omega_{0j}^2-\omega^2\right)-\mathbf{i}g_j\omega}\right\},\qquad \sigma(\omega)=-\mathbf{i}\frac{\varepsilon_0\ \omega\omega_{\mathbf{p}}^2}{-\omega^2-\mathbf{i}g\omega}.$$

$\varepsilon(\omega)$ 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:

$$\varepsilon(\omega)=1+\sum_j\left\{\frac{f_j}{\left(\omega_{0j}^2-\omega^2\right)-\mathbf{i}g_j\omega}\right\}+\frac{f}{\left(\omega_0^2-\omega^2\right)-\mathbf{i}g\omega}\quad\text{with }\omega_0\ll\omega_{0j}\text{ and }\omega\sim\omega_0$$

$$\rightarrow\ \varepsilon(\omega)=\varepsilon_\infty+\frac{f}{\left(\omega_0^2-\omega^2\right)-\mathbf{i}g\omega}$$

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 ε_∞.

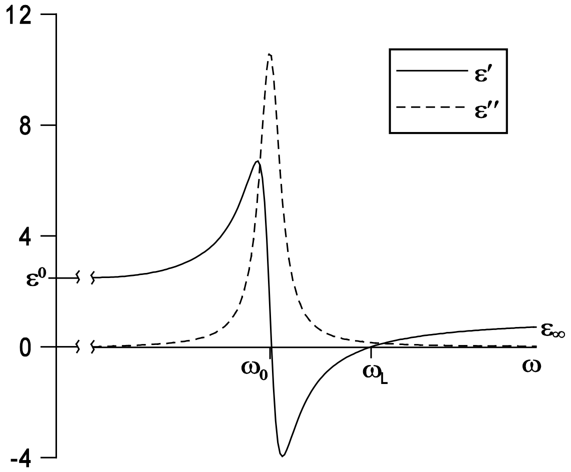
Let us study the real and the imaginary part of the resulting $\varepsilon(\omega)$ separately:

ε_∞ → vacuum and electronic transitions

→ $\varepsilon(\omega) = \Re\varepsilon(\omega) + \mathbf{i}\ \Im\varepsilon(\omega) = \varepsilon'(\omega) + \mathbf{i}\ \varepsilon''(\omega)$

$$\boxed{\varepsilon'(\omega) = \varepsilon_\infty + \frac{f\ \left(\omega_0^2 - \omega^2\right)}{\left(\omega_0^2 - \omega^2\right)^2 + g^2\omega^2},}$$

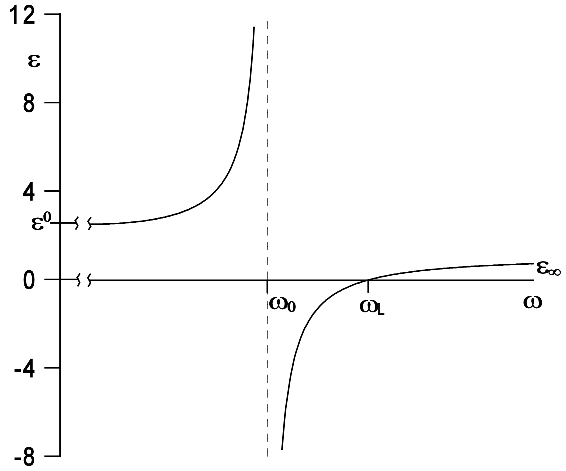
$$\boxed{\varepsilon''(\omega) = \frac{gf\ \ \omega}{\left(\omega_0^2 - \omega^2\right)^2 + g^2\omega^2}.}$$ → Lorentz curve



properties:

- resonance frequency: ω_0
- width of resonance peak: g
- static dielectric constant in the limit $\omega \to 0$: $\varepsilon^0 = \varepsilon_\infty + \frac{f}{\omega_0^2}$
- so called longitudinal frequency ω_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 $g \to 0$



– relation between resonance frequency ω_0 and longitudinal frequency ω_L (Lyddane-Sachs-Teller relation)

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

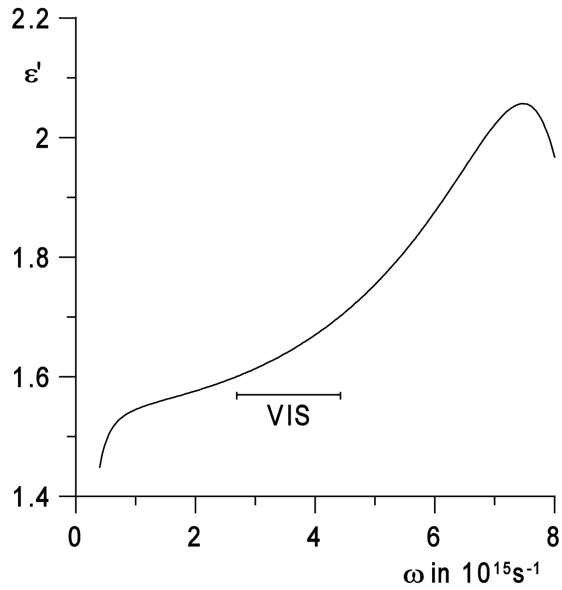
$$\rightarrow \boxed{\omega_L = \omega_0\sqrt{\frac{\varepsilon^0}{\varepsilon_\infty}}.}$$

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).

$$\boxed{\varepsilon(\omega) = \varepsilon_\infty + \frac{f_p}{\left(\omega_{0p}^2 - \omega^2\right) - ig_p\omega} + \frac{f_e}{\left(\omega_{0e}^2 - \omega^2\right) - ig_e\omega},} \text{ with } \omega_{0p} \ll \omega \ll \omega_{0e}$$

ε_∞ → contribution of vacuum and other (far away) resonances



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

$$\boxed{\varepsilon'(\omega)-1=\sum_j \frac{\bar{f}_j\omega_{0j}^2}{\left(\omega_{0j}^2-\omega^2\right)},}$$

- with j being the number of resonances taken into account
- describes many media very well (dispersion of absorption is neglected)
- oscillator strengths and resonance frequencies are often fit parameters to match experimental data

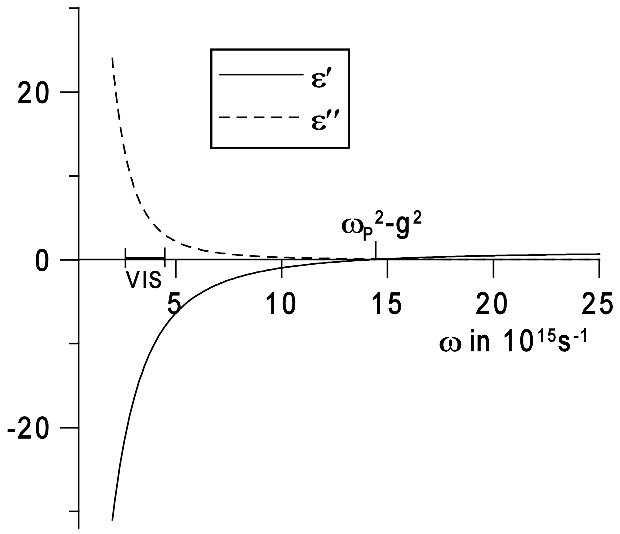
C) Metals in the visible spectral range

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

$$\boxed{\varepsilon(\omega)=1-\frac{\omega_p^2}{\omega^2+\mathbf{i}g\omega}} \quad \text{with } \omega_p >> \omega$$

$$\varepsilon'(\omega)=1-\frac{\omega_p^2}{\omega^2+g^2}, \qquad \varepsilon''(\omega)=\frac{g\omega_p^2}{\omega\left(\omega^2+g^2\right)}.$$

Metals show a large negative real part of the dielectric function $\varepsilon'(\omega)$ which gives rise to decay of the fields. Eventually this results in reflection of light at metallic surfaces.



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:

$$\bar{\mathbf{D}}(\mathbf{r},\omega)=\varepsilon_0\varepsilon(\omega)\bar{\mathbf{E}}(\mathbf{r},\omega)$$
$$\bar{\mathbf{P}}(\mathbf{r},\omega)=\varepsilon_0\chi(\omega)\bar{\mathbf{E}}(\mathbf{r},\omega).$$

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

$$R(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\chi(\omega)\exp\left(-\mathbf{i}\omega t\right)d\omega \qquad \chi(\omega)=\int_{-\infty}^{\infty}R(t)\exp\left(\mathbf{i}\omega t\right)dt$$

To prove this, we can use the convolution theorem

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

$$=\varepsilon_0\int_{-\infty}^{\infty}\chi(\omega)\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{E}(\mathbf{r},t')\exp(i\omega t')\,dt'\exp(-i\omega t)\,d\omega$$

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

$$\mathbf{P}(\mathbf{r},t)=\varepsilon_0\int_{-\infty}^{\infty}\underbrace{\frac{1}{2\pi}\int_{-\infty}^{\infty}\chi(\omega)\exp(-i\omega(t-t'))\,d\omega}_{R(t-t')}\mathbf{E}(\mathbf{r},t')\,dt'$$

$$=\varepsilon_0\int_{-\infty}^{\infty}R(t-t')\mathbf{E}(\mathbf{r},t')\,dt'$$

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

$$\mathbf{E}(\mathbf{r},t'')=\mathbf{e}\delta(t''-t_0)\ \rightarrow\ \mathbf{P}(\mathbf{r},t)=\varepsilon_0 R(t-t_0)\mathbf{e}\ \rightarrow \text{Green's function}$$

Examples

A) instantaneous media (unphysical simplification)

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

$$R(t)=\chi\delta(t)\qquad\rightarrow\qquad \mathbf{P}(\mathbf{r},t)=\varepsilon_0\chi\mathbf{E}(\mathbf{r},t)\qquad\text{(unphysical!)}$$

B) dielectrics

$$R_P(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\chi(\omega)\exp(-i\omega t)\,d\omega=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f}{\omega_0^2-\omega^2-ig\omega}\exp(-i\omega t)\,d\omega,$$

– Using the residual theorem we find:

$$R(t)=\begin{cases}\dfrac{f}{\Omega}\exp\left(-\dfrac{g}{2}t\right)\sin\Omega t & t\geq 0\\ 0 & t<0\end{cases}\qquad\text{with}\qquad \Omega=\sqrt{\omega_0^2-\frac{g^2}{4}}$$

$$\boxed{\mathbf{P}(\mathbf{r},t)=\frac{f}{\Omega}\int_{-\infty}^{t}\exp\left[-\frac{g}{2}(t-t')\right]\sin[\Omega(t-t')]\mathbf{E}(\mathbf{r},t')\,dt'}$$

C) metals

$$R_j(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\sigma(\omega)\exp(-i\omega t)\,d\omega=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\varepsilon_0\omega_p^2}{g-i\omega}\exp(-i\omega t)\,d\omega,$$

– Using again the residual theorem we find:

$$R(t)=\begin{cases}\frac{\omega_p^2}{g}\exp(-gt) & t\geq 0\\ 0 & t<0\end{cases}$$

$$\boxed{\mathbf{j}(\mathbf{r},t)=\varepsilon_0\omega_p^2\int_{-\infty}^{t}\exp[-g(t-t')]\mathbf{E}(\mathbf{r},t')dt'}$$

2.3 The Poynting vector and energy balance

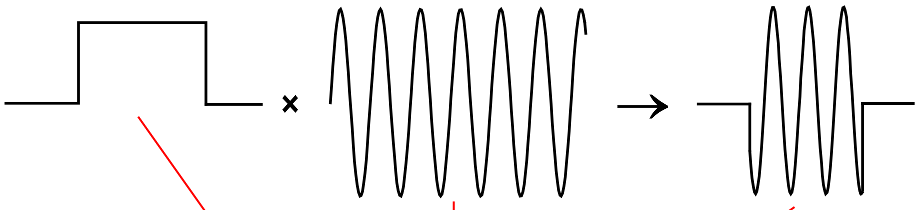
2.3.1 Time averaged Poynting vector

The energy flux of the electromagnetic field is given by the Poynting vector $\mathbf{S}$. In practice, we always measure the energy flux through a surface (detector), $\mathbf{S}\cdot\mathbf{n}$, where $\mathbf{n}$ is the normal vector of surface. To be more precise, the Poynting vector $\mathbf{S}(\mathbf{r},t)=\mathbf{E}_r(\mathbf{r},t)\times\mathbf{H}_r(\mathbf{r},t)$ 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: $T_0=2\pi/\omega_0\leq 10^{-14}$ s
- pulse duration: T_p in general $T_p>>T_0$
- duration of measurement: T_m in general $T_m>>T_0$

Hence, in general the detector does not recognize the fast oscillations of the optical field $\sim e^{-i\omega_0 t}$ (optical cycles) and only delivers a time averaged value. For the situation described above, the electro-magnetic fields factorize in slowly varying envelopes and fast carrier oscillations:



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

For such pulses, the momentary Poynting vector reads:

$$\begin{aligned}\mathbf{S}(\mathbf{r},t) &= \mathbf{E}_{\mathbf{r}}(\mathbf{r},t)\times\mathbf{H}_{\mathbf{r}}(\mathbf{r},t)\\ &= \frac{1}{4}\left[\tilde{\mathbf{E}}(\mathbf{r},t)\times\tilde{\mathbf{H}}^{*}(\mathbf{r},t)+\tilde{\mathbf{E}}^{*}(\mathbf{r},t)\times\tilde{\mathbf{H}}(\mathbf{r},t)\right]\\ &\quad+\frac{1}{4}\left[\tilde{\mathbf{E}}(\mathbf{r},t)\times\tilde{\mathbf{H}}(\mathbf{r},t)\exp\left(-2\mathbf{i}\omega_0 t\right)+\tilde{\mathbf{E}}^{*}(\mathbf{r},t)\times\tilde{\mathbf{H}}^{*}(\mathbf{r},t)\exp\left(2\mathbf{i}\omega_0 t\right)\right]\\ &= \underbrace{\frac{1}{2}\Re\left[\tilde{\mathbf{E}}(\mathbf{r},t)\times\tilde{\mathbf{H}}^{*}(\mathbf{r},t)\right]}_{\text{slow}}+\underbrace{\frac{1}{2}\Re\left[\tilde{\mathbf{E}}(\mathbf{r},t)\times\tilde{\mathbf{H}}(\mathbf{r},t)\right]\cos\left(2\omega_0 t\right)}_{\text{fast}}\\ &\quad+\underbrace{\frac{1}{2}\Im\left[\tilde{\mathbf{E}}^{*}(\mathbf{r},t)\times\tilde{\mathbf{H}}^{*}(\mathbf{r},t)\right]\sin\left(2\omega_0 t\right)}_{\text{fast}}.\end{aligned}$$

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 $\sim\cos\left(2\omega_0 t\right),\sim\sin\left(2\omega_0 t\right)$ changing over time scales of the optical cycle T_0. Now, a measurement of the Poynting vector over a time interval T_{m} leads to a time average of $\mathbf{S}(\mathbf{r},t)$.

$$\left\langle\mathbf{S}(\mathbf{r},t)\right\rangle=\frac{1}{T_{\mathbf{m}}}\int_{t-T_{\mathrm{m}}/2}^{t+T_{\mathrm{m}}/2}\mathbf{S}(\mathbf{r},t')dt'$$

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

$$\left\langle\mathbf{S}(\mathbf{r},t)\right\rangle=\frac{1}{2}\frac{1}{T_{\mathbf{m}}}\int_{t-T_{\mathrm{m}}/2}^{t+T_{\mathrm{m}}/2}\Re\left[\tilde{\mathbf{E}}(\mathbf{r},t')\times\tilde{\mathbf{H}}^{*}(\mathbf{r},t')\right]dt'$$

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

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

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

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

$I=\left|\left\langle\mathbf{S}(\mathbf{r},t)\right\rangle\right|$ → 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 $\mathbf{H}_\mathrm{r}$ resp. $\mathbf{E}_\mathrm{r}$ (note that we use real fields):

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

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

Next, we subtract the two equations and get

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

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

$$\mathbf{H}_\mathrm{r}\cdot\mathbf{rot}\mathbf{E}_\mathrm{r}-\mathbf{E}_\mathrm{r}\cdot\mathbf{rot}\mathbf{H}_\mathrm{r}=\mathbf{div}\left(\mathbf{E}_\mathrm{r}\times\mathbf{H}_\mathrm{r}\right)$$

Finally, with the substitution $\mathbf{E}_\mathrm{r}\cdot\partial\mathbf{E}_\mathrm{r}/\partial t=\tfrac{1}{2}\partial\mathbf{E}_\mathrm{r}^2/\partial t$ we find Poynting's theorem

$$\boxed{\frac{1}{2}\varepsilon_0\frac{\partial}{\partial t}\mathbf{E}_\mathrm{r}{}^2+\frac{1}{2}\mu_0\frac{\partial}{\partial t}\mathbf{H}_\mathrm{r}{}^2+\mathbf{div}\left(\mathbf{E}_\mathrm{r}\times\mathbf{H}_\mathrm{r}\right)=-\mathbf{E}_\mathrm{r}\cdot\left(\mathbf{j}_\mathrm{r}+\frac{\partial}{\partial t}\mathbf{P}_\mathrm{r}\right)}\ (*)$$

This equation has the general form of a balance equation. Here it represents the energy balance. Apart from the appearance of the divergence of the Poynting vector (energy flux), we can identify the vacuum energy density $u=\tfrac{1}{2}\varepsilon_0\mathbf{E}_\mathrm{r}{}^2+\tfrac{1}{2}\mu_0\mathbf{H}_\mathrm{r}{}^2$. The right-hand-side of the Poynting's theorem contains the so-called source terms.

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

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

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

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

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

$$\left\langle\frac{1}{2}\varepsilon_0\frac{\partial}{\partial t}\mathbf{E}_\mathbf{r}^2(\mathbf{r},t)+\frac{1}{2}\mu_0\frac{\partial}{\partial t}\mathbf{H}_\mathbf{r}^2(\mathbf{r},t)+\mathbf{div}\left[\mathbf{E}_\mathbf{r}(\mathbf{r},t)\times\mathbf{H}_\mathbf{r}(\mathbf{r},t)\right]\right\rangle=\frac{1}{2}\mathbf{div}\left\{\Re\left[\mathbf{E}(\mathbf{r})\times\mathbf{H}^*(\mathbf{r})\right]\right\}$$

$$=\mathbf{div}\left\langle\mathbf{S}(\mathbf{r},t)\right\rangle.$$

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

$$-\left\langle \left[\mathbf{j}_{\mathrm{r}}(\mathbf{r},t) + \frac{\partial}{\partial t}\mathbf{P}_{\mathrm{r}}(\mathbf{r},t) \right] \mathbf{E}_{\mathrm{r}}(\mathbf{r},t) \right\rangle$$

$$= -\frac{1}{4}\left\langle \left[\sigma(\omega_0)\mathbf{E}(\mathbf{r})e^{-i\omega_0 t} - \mathbf{i}\omega_0\varepsilon_0\chi(\omega_0)\mathbf{E}(\mathbf{r})e^{-i\omega_0 t} c.c. \right] \left[\mathbf{E}(\mathbf{r})e^{-i\omega_0 t} + c.c. \right] \right\rangle$$

Now we use our generalized dielectric function:

$$= -\frac{1}{4}\left\langle \left[-\mathbf{i}\omega_0\varepsilon_0 \left(\chi(\omega_0) + \mathbf{i}\frac{\sigma(\omega_0)}{\omega_0\varepsilon_0} \right) \mathbf{E}(\mathbf{r})\exp(-\mathbf{i}\omega_0 t) + c.c. \right] \left[\mathbf{E}(\mathbf{r})\exp(-\mathbf{i}\omega_0 t) + c.c. \right] \right\rangle$$

$$= \frac{1}{4}\mathbf{i}\omega_0\varepsilon_0 \left[\varepsilon(\omega_0) - 1 \right] \mathbf{E}(\mathbf{r})\mathbf{E}(\mathbf{r})^* + c.c.$$

Again, all fast oscillating terms $\sim \exp(\pm 2\mathbf{i}\omega_0 t)$ cancel due to the time average. Finally, splitting $\varepsilon(\omega_0)$ into real and imaginary part yields

$$= \frac{1}{4}\mathbf{i}\omega_0\varepsilon_0 \left[\varepsilon'(\omega_0) - 1 + \mathbf{i}\varepsilon''(\omega_0) \right] \mathbf{E}(\mathbf{r})\mathbf{E}(\mathbf{r})^* + c.c. = -\frac{1}{2}\omega_0\varepsilon_0\varepsilon''(\omega_0)\mathbf{E}(\mathbf{r})\mathbf{E}^*(\mathbf{r}).$$

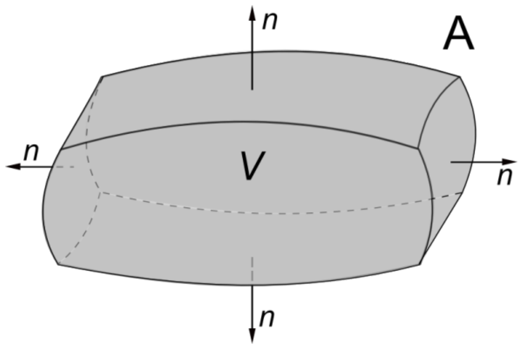
Hence, the divergence of the time averaged Poynting vector is related to the imaginary part of the generalized dielectric function:

$$\rightarrow \boxed{\mathbf{div}\langle \mathbf{S} \rangle = -\frac{1}{2}\omega_0\varepsilon_0\varepsilon''(\omega_0)\mathbf{E}(\mathbf{r})\mathbf{E}^*(\mathbf{r}).}$$

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

$$\int_V \mathbf{div}\langle \mathbf{S} \rangle dV = \int_A \langle \mathbf{S} \rangle \cdot \mathbf{n} dA$$



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 to derive the solution of the wave equation

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

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. These normal modes are the stationary solutions of the wave equation. Hence they usually correspond to waves which are infinitely extended in space and time. However as we will see later, by taking into account the principle of superposition we can construct the general solutions from these normal modes, by superimposing multiple normal modes to construct also the transient waves.

We start from the wave equation in Fourier domain, which reads as

$$\mathbf{rotrot}\ \bar{\mathbf{E}}(\mathbf{r},\omega) = \frac{\omega^2}{c^2}\varepsilon(\omega)\bar{\mathbf{E}}(\mathbf{r},\omega)$$

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

$$\varepsilon_0\left[1+\chi(\omega)\right]\mathbf{div}\ \bar{\mathbf{E}}(\mathbf{r},\omega) = 0$$

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

for $1+\chi(\omega) \neq 0$ → $\mathbf{div}\ \bar{\mathbf{E}}(\mathbf{r},\omega) = 0$ (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:

$$\bar{\mathbf{E}}(\mathbf{r},\omega) = \bar{\mathbf{E}}(\omega)\exp\left(\mathbf{i}\mathbf{k}\mathbf{r}\right),$$

$\mathbf{k}$ = unknown complex wave-vector

The corresponding stationary field in time domain is given by:

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

→ monochromatic plane wave → 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 these waves are transverse

$\mathbf{k}\perp\bar{\mathbf{E}}(\omega)$ → transverse wave

Here transverse means that the electric field and the wave-vector are oriented perpendicular to each other.

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

- planes of constant phase $\mathbf{k}'\mathbf{r}=\mathrm{const}$
- planes of constant amplitude $\mathbf{k}''\mathbf{r}=\mathrm{const}$

In the following we will call the solutions

A) homogeneous waves → if these planes are identical

B) evanescent waves → if these planes are perpendicular

C) inhomogeneous waves → otherwise

We will see that in dielectrics $\left(\sigma\left(\omega\right)=0\right)$ we can find a second, exotic type of wave solutions: At $\omega=\omega_{\mathrm{L}}\rightarrow\varepsilon(\omega_{\mathrm{L}})=0$, so-called longitudinal waves $\mathbf{k}\,\|\,\bar{\mathbf{E}}(\omega)$ appear.

2.4.1 Transverse waves

Let us have a closer look at the transverse nature of the fields first. As pointed out above, for $\omega\neq\omega_{\mathrm{L}}$ the electric field becomes free of divergence:

$$\varepsilon_0\varepsilon(\omega)\,\mathrm{div}\,\mathbf{E}(\mathbf{r},\omega)=0 \;\rightarrow\; \mathrm{div}\,\bar{\mathbf{E}}(\mathbf{r},\omega)=0$$

Then, the wave equation reduces to the Helmholtz equation:

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

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

$$\bar{\mathbf{E}}(\mathbf{r},\omega)=\bar{\mathbf{E}}(\omega)\exp\left(\mathrm{i}\mathbf{kr}\right)$$

Immediately we see that the wave is transversal:

$$0=\mathbf{div}\bar{\mathbf{E}}(\omega)=\mathrm{i}\mathbf{k}\cdot\bar{\mathbf{E}}(\mathbf{r},\omega)\rightarrow\boxed{\mathbf{k}\perp\bar{\mathbf{E}}(\omega).}$$

Hence, we have to solve

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

which leads to the following dispersion relation

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

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 transverse plane waves are solutions to Maxwell's equations in homogeneous, isotropic media, only if the dispersion relation $\boxed{k(\omega)}$ is fulfilled.

In general, $\mathbf{k}=\mathbf{k}'+i\mathbf{k}''$ is complex. Alternatively if $\mathbf{k}'\parallel\mathbf{k}''$, it is sometimes useful to introduce the complex refractive index $\hat{n}$:

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

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.

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

Hence, the coincidence of the complex refractive index $\hat{n}(\omega)$ with $\sqrt{\varepsilon(\omega)}$ holds only for homogeneous media since $\hat{n}(\omega)$ cannot not be a local property while $\sqrt{\varepsilon(\omega)}$ could.

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

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

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

2.4.2 Longitudinal waves

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

$$\rightarrow\ \mathbf{rotrot}\ \bar{\mathbf{E}}(\mathbf{r},\omega_L)=0$$

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

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

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

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

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

$$\bar{\mathbf{E}}(\mathbf{r},\omega)=\bar{\mathbf{E}}(\omega)\exp(\mathrm{i}\mathbf{kr})=\bar{\mathbf{E}}_\perp(\omega)\exp(\mathrm{i}\mathbf{kr})+\bar{\mathbf{E}}_{\parallel}(\omega)\exp(\mathrm{i}\mathbf{kr})$$

with $\bar{\mathbf{E}}_\perp(\omega)\perp\mathbf{k}$ and $\bar{\mathbf{E}}_{\parallel}(\omega)\parallel\mathbf{k}$

This decomposed field is inserted into the wave equation:

$$\mathbf{k}\times\left[\mathbf{k}\times\left(\bar{\mathbf{E}}_\perp+\bar{\mathbf{E}}_{\parallel}\right)\right]\exp(\mathrm{i}\mathbf{kr})=0$$

$$\mathbf{k}\times\left[\mathbf{k}\times\bar{\mathbf{E}}_\perp\right]\exp(\mathrm{i}\mathbf{kr})+\mathbf{k}\times\underbrace{\left[\mathbf{k}\times\bar{\mathbf{E}}_{\parallel}\right]}_{=0}\exp(\mathrm{i}\mathbf{kr})=0$$

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

$$k^2\bar{\mathbf{E}}_\perp=0$$

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

$$\rightarrow\ \boxed{\bar{\mathbf{E}}(\mathbf{r},\omega_L)=\bar{\mathbf{E}}_{\parallel}(\omega_L)\exp(\mathrm{i}\mathbf{kr})}$$

2.4.3 Plane wave solutions in different frequency regimes

The dispersion relation $\mathbf{k}^2=k^2=k_x^2+k_y^2+k_z^2=\left(\omega^2/c^2\right)\varepsilon(\omega)$ for plane wave solutions dictates the (complex) wavenumber k only. Thus, different solutions for the complex wave vector $\mathbf{k}=\mathbf{k}'+\mathrm{i}\mathbf{k}''$ 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 favorable regime for optics. We have transparency, and the frequency of light is far from resonances of the medium. The dispersion relation gives

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

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

A.1) Real valued wave-vector $\mathbf{k}''=0$

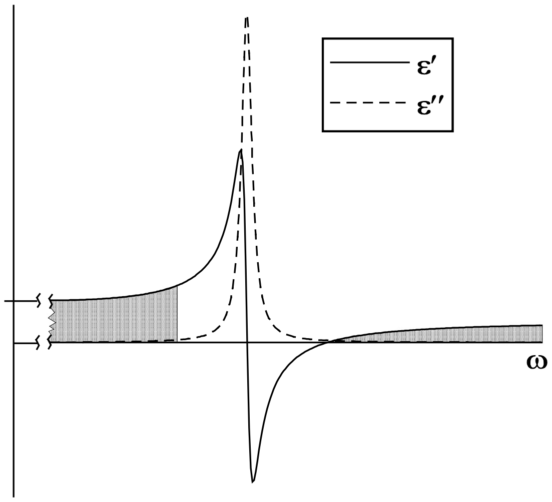
- In this case the wave vector is real and we find the dispersion relation

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

- Because $\mathbf{k}''=0$ 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

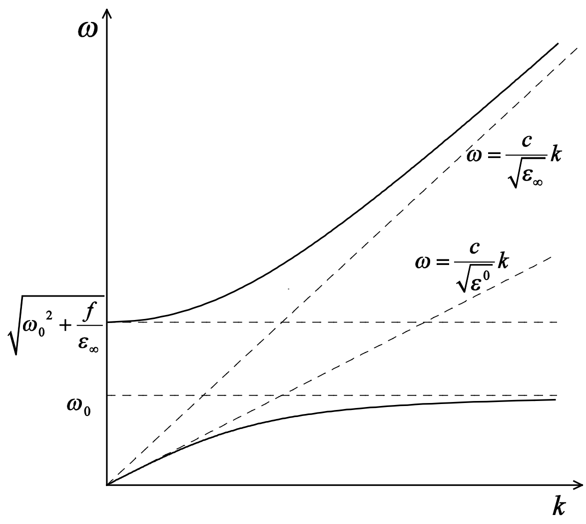
– for lattice vibrations (phonons)



– Now the imaginary part of $\varepsilon(\omega)$ is neglected, which mathematically corresponds to an undamped resonance

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

– We can invert the dispersion relation $k(\omega)=\frac{\omega}{c}\sqrt{\varepsilon(\omega)}$ → $\omega(k)$:



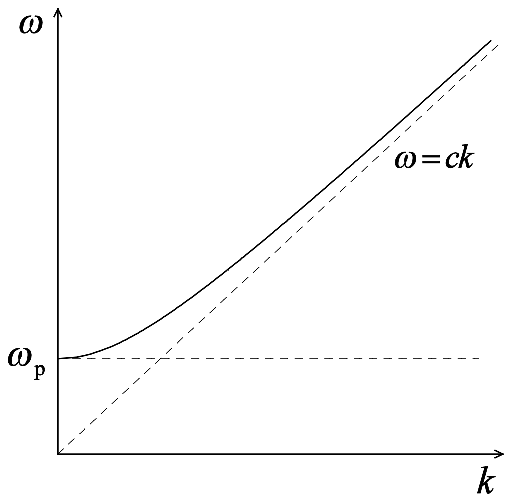
Example 2: free electrons

– for plasma and metal

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

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

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



A.2) Complex valued wave-vector $\mathbf{k}' \perp \mathbf{k}''$

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

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

- This means that

→ $\boxed{\mathbf{k}''^2 \neq 0 \text{ and } \mathbf{k}'^2 > k^2}$

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

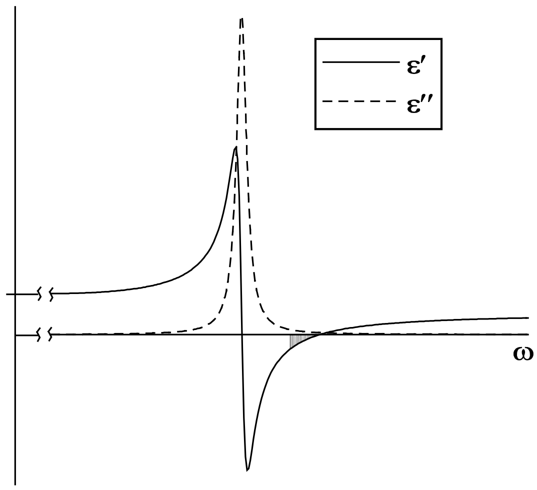
$$\bar{\mathbf{E}}(\mathbf{r},\omega)=\mathbf{E}(\omega)\exp\left\{\mathrm{i}\left[\mathbf{k}'(\omega)\mathbf{r}\right]\right\}\exp\left(-\mathbf{k}''(\omega)\mathbf{r}\right)$$

- The planes defined by the equation $\mathbf{k}''(\omega)\mathbf{r}=\mathbf{const.}$ are the so-called planes of constant amplitude, those defined by $\mathbf{k}'(\omega)\mathbf{r}=\mathbf{const.}$ are the planes of constant phase. Because of $\mathbf{k}' \perp \mathbf{k}''$ these planes are perpendicular to each other.
- The factor $\exp\left(-\mathbf{k}''(\omega)\mathbf{r}\right)$ 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 ($\omega_0<\omega<\omega_L$) or below the plasma frequency ($\omega<\omega_p$) in metals. Then the dispersion relation gives

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



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

B.1) $\mathbf{k}'=0$

→ $\mathbf{k}''^2=\frac{\omega^2}{c^2}\left|\varepsilon'(\omega)\right|$ → $\bar{\mathbf{E}}(\mathbf{r},\omega)\sim\exp\left(-\mathbf{k}''\mathbf{r}\right)$ → strong damping

B.2) $\mathbf{k}'\cdot\mathbf{k}''=0$

→ $\mathbf{k}'\perp\mathbf{k}''$ → evanescent waves

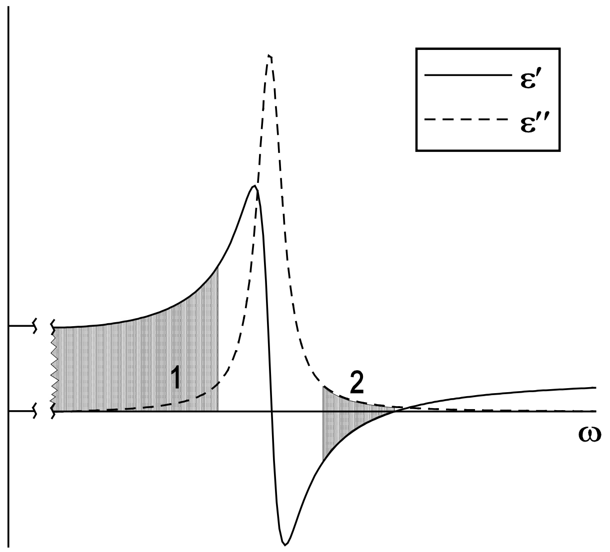
$$k^2=\mathbf{k}'^2-\mathbf{k}''^2=-\frac{\omega^2}{c^2}\left|\varepsilon'(\omega)\right|$$

$$\boxed{\mathbf{k}''^2=\frac{\omega^2}{c^2}\left|\varepsilon'(\omega)\right|+\mathbf{k}'^2.}$$

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 $\mathbf{k}'^2$. In particular, case B.1) ($\mathbf{k}'=0$) is included. Hence, we can conclude that for $\varepsilon(\omega)=\varepsilon'(\omega)<0$ we find only evanescent waves!

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

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



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

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

Since $\mathbf{k}'$ and $\mathbf{k}''$ are almost parallel:

$$\rightarrow\ \left|\mathbf{k}'\right| = \frac{\omega}{c}n(\omega), \qquad \left|\mathbf{k}''\right| = \frac{\omega}{c}\kappa(\omega)$$

The dispersion relation in terms of the complex refractive index gives

$$\mathbf{k}^2 = k^2 = \frac{\omega^2}{c^2}\varepsilon(\omega) = \frac{\omega^2}{c^2}\left[n(\omega)+\mathbf{i}\kappa(\omega)\right]^2$$

Here we have

$$\varepsilon(\omega) = \varepsilon'(\omega)+\mathbf{i}\varepsilon''(\omega) = n^2(\omega)-\kappa^2(\omega)+2\mathbf{i}n(\omega)\kappa(\omega),$$

and therefore ➔ $\begin{aligned}\varepsilon'(\omega) &= n^2(\omega)-\kappa^2(\omega)\\ \varepsilon''(\omega) &= 2n(\omega)\kappa(\omega)\end{aligned}$

$$\boxed{n^2(\omega) = \frac{\varepsilon'}{2}\left[\operatorname{sgn}(\varepsilon')\sqrt{1+(\varepsilon''/\varepsilon')^2}+1\right],}$$

$$\boxed{\kappa^2(\omega) = \frac{\varepsilon'}{2}\left[\operatorname{sgn}(\varepsilon')\sqrt{1+(\varepsilon''/\varepsilon')^2}-1\right].}$$

Two important limiting cases of quasi-homogeneous plane waves:

C.1) $\varepsilon',\varepsilon''>0,\ \varepsilon''\ll\varepsilon',$ ***(dielectric media)***

$$\boxed{n(\omega)\approx\sqrt{\varepsilon'(\omega)}, \qquad \kappa(\omega)\approx\frac{1}{2}\frac{\varepsilon''(\omega)}{\sqrt{\varepsilon'(\omega)}}}$$

In this regime propagation dominates ($n(\omega) \gg \kappa(\omega)$), and we have weak absorption:

$$\mathbf{k'}^2 - \mathbf{k''}^2 = \frac{\omega^2}{c^2}\varepsilon'(\omega), \qquad 2\mathbf{k'}\cdot\mathbf{k''} = \frac{\omega^2}{c^2}\varepsilon''(\omega).$$

$$|\mathbf{k'}| = \frac{\omega}{c}n(\omega) \approx \frac{\omega}{c}\sqrt{\varepsilon'(\omega)}, \qquad |\mathbf{k''}| = \frac{\omega}{c}\kappa(\omega) \approx \frac{1}{2}\frac{\omega}{c}\frac{\varepsilon''(\omega)}{\sqrt{\varepsilon'(\omega)}}$$

- $\mathbf{k'}\cdot\mathbf{k''} \approx |\mathbf{k'}||\mathbf{k''}|$,
- $\mathbf{k'}$ and $\mathbf{k''}$ almost parallel → homogeneous waves

→ in homogeneous, isotropic media, next to resonances, we find damped, homogeneous plane waves, $\mathbf{k'} \,||\, \mathbf{k''} \,||\, \mathbf{e}_\mathrm{k}$ with $\mathbf{e}_\mathrm{k}$ being the unit vector along $\mathbf{k}$

$$\boxed{\bar{\mathbf{E}}(\mathbf{r},\omega) = \mathbf{E}(\omega)\exp\left(i\mathbf{kr}\right) = \mathbf{E}(\omega)\exp\left\{i\left[\frac{\omega}{c}n(\omega)\left(\mathbf{e}_\mathrm{k}\mathbf{r}\right)\right]\right\}\exp\left[-\frac{\omega}{c}\kappa(\omega)\left(\mathbf{e}_\mathrm{k}\mathbf{r}\right)\right].}$$

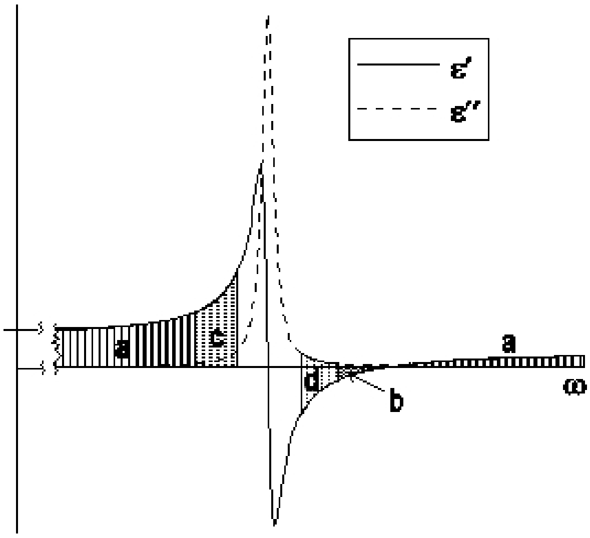
C.2) $\varepsilon' < 0,\ \varepsilon'' > 0,\quad \varepsilon'' << |\varepsilon'|$, ***(metals and dielectric media in so-called Reststrahl domain)***

$$\boxed{n(\omega) \approx \frac{1}{2}\frac{\varepsilon''(\omega)}{\sqrt{|\varepsilon'(\omega)|}}, \qquad \kappa(\omega) \approx \sqrt{|\varepsilon'(\omega)|},}$$

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

Here we summarize our previous findings about the properties of normal modes for different parameters of the material which they exist in. We do this at the example of a dielectric media for frequencies close to one resonance. Here we can identify different frequency intervals in which we find the typical behavior of the material's reponse.



There are the following frequency ranges, which give rise to the typical properties of normal modes:

a) Frequency far below or far above resonance where $\varepsilon'(\omega) > 0$, $\varepsilon''(\omega) \approx 0$
Typical normal modes:
- undamped homogeneous waves
- evanescent waves

b) Frequency above resonance where $\varepsilon'(\omega) < 0$, $\varepsilon''(\omega) \approx 0$
Typical normal modes:
- evanescent waves

c) Frequency close to and below resonance where $\varepsilon'(\omega) > 0$, $\varepsilon''(\omega) > 0$
Typical normal modes:
- weakly damped quasi-homogeneous waves

d) Frequency close to and above resonance where $\varepsilon'(\omega) > 0$, $\varepsilon''(\omega) > 0$
Typical normal modes:
- strongly damped quasi-homogeneous waves

Optical systems work mainly in regime a) since here you find light propagating undamped over long distances through space. Furthermore one sometimes exploits also regime b) when one would like to have reflection of light at surfaces (e.g. at a metallic mirror).

2.4.4 Time averaged Poynting vector of plane waves

$$\left\langle \mathbf{S}(\mathbf{r},t) \right\rangle = \frac{1}{2}\frac{1}{T_\mathrm{m}}\int_{t-T_\mathrm{m}/2}^{t+T_\mathrm{m}/2} \Re\left[\tilde{\mathbf{E}}(\mathbf{r},t') \times \tilde{\mathbf{H}}^*(\mathbf{r},t')\right] dt',$$

For plane waves we find:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}\exp\left(\mathrm{i}\mathbf{k}\mathbf{r} - \mathrm{i}\omega t\right) = \mathbf{E}\exp\left(\mathrm{i}\mathbf{k}'\mathbf{r} - \mathbf{k}''\mathbf{r} - \mathrm{i}\omega t\right)$$

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

assuming a stationary case $\mathbf{E}(t)=\bar{\mathbf{E}}(\omega)\exp(-\mathrm{i}\omega t)$

$$\rightarrow \boxed{\langle \mathbf{S}(\mathbf{r},t)\rangle = \frac{1}{2}\frac{\mathbf{k}'}{\omega\mu_0}\exp\left[-2\mathbf{k}''\bullet\mathbf{r}\right]|\mathbf{E}|^2 = \frac{1}{2}n\sqrt{\frac{\varepsilon_0}{\mu_0}}\mathbf{e}_{\mathbf{k}'}\exp\left[-2\frac{\omega}{c}\kappa\left(\mathbf{e}_{\mathbf{k}''}\bullet\mathbf{r}\right)\right]|\mathbf{E}|^2}$$

with $\mathbf{e}_{k'}$ being the unit vector along $\mathbf{k}'$ and $\mathbf{e}_{k''}$ being the unit vector along $\mathbf{k}''$.

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:

$$\mathbf{P}_{\mathrm{r}}(\mathbf{r},t)=\varepsilon_0\int_{-\infty}^{t}R(t-t')\mathbf{E}_{\mathrm{r}}(\mathbf{r},t')dt' \quad\leftrightarrow\quad \mathbf{P}_{\mathrm{r}}(\mathbf{r},t)=\varepsilon_0\int_{0}^{\infty}R(\tau)\mathbf{E}_{\mathrm{r}}(\mathbf{r},t-\tau)d\tau$$

Reality of the response function implies:

$$R(\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\omega\,\chi(\omega)\mathrm{e}^{-i\omega\tau}=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\omega\,\chi^*(\omega)\mathrm{e}^{i\omega\tau} \quad\rightarrow\quad \chi(\omega)=\chi^*(-\omega)$$

Causality of the response function implies:

$$R(\tau)=\theta(\tau)y(\tau)\ \text{ with }\ \theta(\tau)=\begin{cases}1 \text{ for } \tau>0\\ \frac{1}{2}\text{for } \tau=0\\ 0 \text{ for } \tau<0\end{cases} \rightarrow \text{Heaviside distribution}$$

In the following, we will make use of the Fourier transform of Heaviside distribution:

$$\boxed{2\pi\bar{\theta}(\omega)=\int_{-\infty}^{\infty}dt\theta(t)\mathrm{e}^{i\omega t}=\mathrm{P}\frac{i}{\omega}+\pi\delta(\omega)}$$ → defined as integral only

In Fourier space, the Heaviside distribution consists of the Dirac delta distribution

$$\int_{-\infty}^{\infty}d\omega\delta(\omega-\omega_0)f(\omega)=f(\omega_0)$$ → Dirac delta distribution

and the expression $\mathrm{P}(i/\omega)$ involving a Cauchy principal value:

$$\mathrm{P}\int_{-\infty}^{\infty}d\omega\frac{i}{\omega}f(\omega)=\lim_{\alpha\to 0}\left[\int_{-\infty}^{-\alpha}d\omega\frac{i}{\omega}f(\omega)+\int_{\alpha}^{\infty}d\omega\frac{i}{\omega}f(\omega)\right]$$ → 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:

$$\rightarrow \begin{aligned}\chi(\omega)&=\int_{-\infty}^{\infty}d\tau R(\tau)\mathrm{e}^{i\omega\tau}=\int_{-\infty}^{\infty}d\tau\theta(\tau)y(\tau)\mathrm{e}^{i\omega\tau}\\&=\int_{-\infty}^{\infty}d\bar{\omega}\;\bar{\theta}(\omega-\bar{\omega})\bar{y}(\bar{\omega})\end{aligned}\qquad \bar{\theta}(\omega)=\frac{1}{2\pi}\mathrm{P}\frac{i}{\omega}+\frac{1}{2}\delta(\omega)$$

$$\boxed{\chi(\omega)=\frac{1}{2\pi}\mathrm{P}\int_{-\infty}^{\infty}d\bar{\omega}\frac{i\,\bar{y}(\bar{\omega})}{\omega-\bar{\omega}}+\frac{\bar{y}(\omega)}{2}\qquad(*)}$$

In order to derive the Kramers-Kronig relation we can use a small trick (this trick saves us using complex integration in the derivation). Because of the Heaviside function, we can choose the function $y(\tau)$ for $\tau < 0$ arbitrarily without altering the susceptibility! In particular, we can choose:

a) $y(-\tau)=y(\tau)$ even function

b) $y(-\tau)=-y(\tau)$ odd function

a) $y(-\tau)=y(\tau)$

In this case $y(-\tau)=y(\tau)$ is a real valued and even function. We can exploit this property and show that

$$\rightarrow\bar{y}(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\tau\,y(\tau)\mathrm{e}^{i\omega\tau}=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\tau\,y(\tau)\mathrm{e}^{-i\omega\tau}=\bar{y}^*(\omega)$$ is real as well

Hence, we can conclude from equation (*) above that

$$\rightarrow\ \chi^*(\omega) = -\frac{1}{2\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{i\,\bar{y}(\varpi)}{\omega-\varpi} + \frac{\bar{y}(\omega)}{2}$$

Here $\mathrm{P}\int$ 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:

$$\chi(\omega)+\chi^*(\omega) = \frac{1}{2\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{i\,\bar{y}(\varpi)}{\omega-\varpi} + \frac{\bar{y}(\omega)}{2} - \frac{1}{2\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{i\,\bar{y}(\varpi)}{\omega-\varpi} + \frac{\bar{y}(\omega)}{2} = \bar{y}(\omega)$$

$$\chi(\omega)-\chi^*(\omega) = \ldots = \frac{1}{\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{i\,\bar{y}(\varpi)}{\omega-\varpi}$$

Plugging the last two equations together we find the first Kramers-Kronig relation:

$$\rightarrow\ \boxed{\Im\chi(\omega) = -\frac{1}{\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{\Re\,\chi(\varpi)}{\varpi-\omega}} \qquad \text{1. K-K relation}$$

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

b) $y(-\tau) = -y(\tau)$

The second K-K relation can be found by a similar procedure when we assume that $y(-\tau) = -y(\tau)$ is a real odd function. We can show that in this case

$$\rightarrow\ \bar{y}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\tau\, y(\tau)\mathrm{e}^{i\omega\tau} = -\frac{1}{2\pi}\int_{-\infty}^{\infty} d\tau\, y(\tau)\mathrm{e}^{-i\omega\tau} = -\bar{y}^*(\omega) \text{ is purely imaginary}$$

With equation (*) we then find that

$$\rightarrow\ \chi^*(\omega) = \frac{1}{2\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{i\,\bar{y}(\varpi)}{\omega-\varpi} - \frac{\bar{y}(\omega)}{2} \quad \text{(see (*)) and}$$

Again we can then compute real and imaginary part of the susceptibility

$$\chi(\omega)-\chi^*(\omega) = \frac{1}{2\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{i\,\bar{y}(\varpi)}{\omega-\varpi} + \frac{\bar{y}(\omega)}{2} - \frac{1}{2\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{i\,\bar{y}(\varpi)}{\omega-\varpi} + \frac{\bar{y}(\omega)}{2} = \bar{y}(\omega)$$

$$\chi(\omega)+\chi^*(\omega) = \ldots = \frac{1}{\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{i\,\bar{y}(\varpi)}{\omega-\varpi}$$

and finally obtain

$$\rightarrow\ \boxed{\Re\chi(\omega) = \frac{1}{\pi}\mathrm{P}\int_{-\infty}^{\infty} d\varpi\, \frac{\Im\,\chi(\varpi)}{\varpi-\omega}} \qquad \text{2. K-K relation}$$

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 ω:

K-K relation for ε:

- $\chi(\omega)=\chi^*(-\omega)\rightarrow \chi'(\omega)=\chi'(-\omega) \quad \chi''(\omega)=-\chi''(-\omega)$ and

$$\chi(\omega)=\varepsilon(\omega)-1=\left[\varepsilon'(\omega)-1\right]+i\varepsilon''(\omega)$$

$$\boxed{\begin{aligned}\varepsilon'(\omega)-1&=\frac{2}{\pi}\mathrm{P}\int_0^\infty\frac{\bar{\omega}\varepsilon''(\bar{\omega})}{\bar{\omega}^2-\omega^2}d\bar{\omega},\\ \varepsilon''(\omega)&=-\frac{2}{\pi}\omega\mathrm{P}\int_0^\infty\frac{\left[\varepsilon'(\bar{\omega})-1\right]}{\bar{\omega}^2-\omega^2}d\bar{\omega}.\end{aligned}}$$

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

Example:

$$\varepsilon''(\omega)\sim\delta(\omega-\omega_0) \rightarrow \varepsilon'(\omega)-1\sim\frac{\omega_0}{\omega_0^2-\omega^2}$$ → 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:

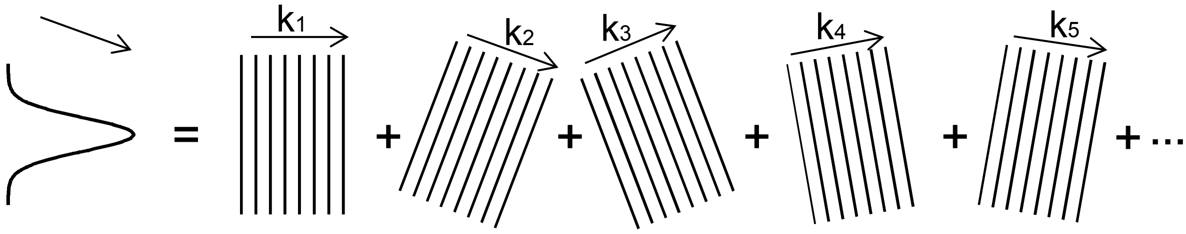
- Pulses with finite spatial width (i.e. pulsed beams) are superposition of normal modes (in frequency- and spatial frequency domain).
- Spatio-temporally localized optical excitations delocalize during propagation because of different phase evolution for different frequencies and spatial frequencies (different propagation directions of normal modes).

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

A) beam → finite transverse width → diffraction

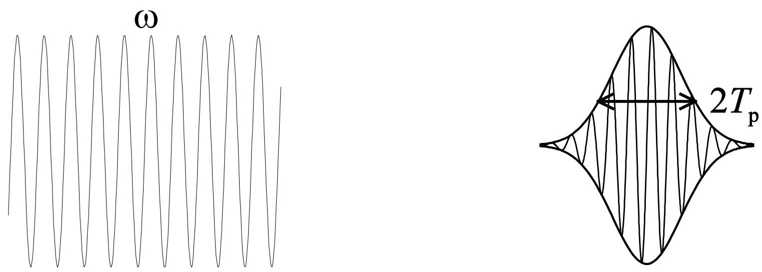
plane wave (normal mode) *beam*

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



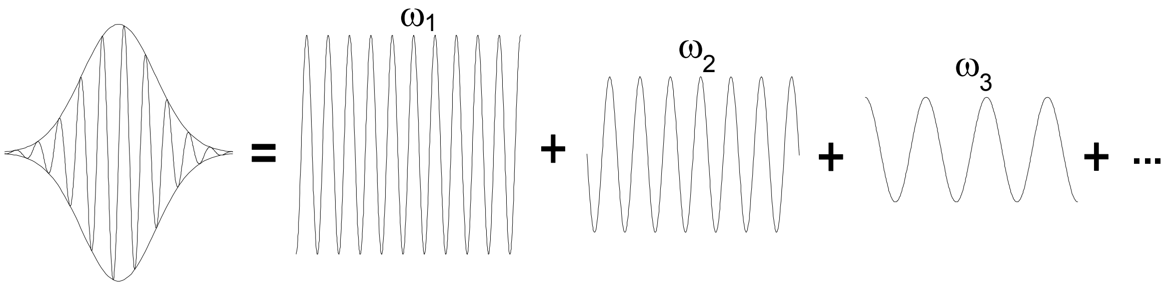
$$\boxed{\mathbf{E}(\mathbf{r},t)=\int_{-\infty}^{\infty}\bar{\mathbf{E}}(\mathbf{k})\exp\left[\mathrm{i}\left(\mathbf{kr}-\omega\, t\right)\right]d^3k}$$

B) pulse → finite duration → dispersion



stationary wave (normal mode) *pulse*

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



$$\mathbf{E}(\mathbf{r},t)=\int_{-\infty}^{\infty}\bar{\mathbf{E}}(\omega)\exp\left[\mathrm{i}\left(\mathbf{k}\cdot\mathbf{r}-\omega\, t\right)\right]d\omega.$$

C) pulsed beams → finite transverse width and finite duration → diffraction and dispersion

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

$$\mathbf{E}(\mathbf{r},t)=\int_{-\infty}^{\infty}\bar{\mathbf{E}}(\mathbf{k},\omega)\exp\left[\mathrm{i}\left(\mathbf{k}\cdot\mathbf{r}-\omega t\right)\right]d^3k\,d\omega$$

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's equations) 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 $\omega=\mathrm{const}$. For technical convenience and because it is sufficient for many important problems, we will make the following assumptions and approximations:

- $\varepsilon(\omega)=\varepsilon'(\omega)>0,$ → optical transparent regime → normal modes are stationary homogeneous and evanescent plane waves
- scalar approximation

$$\bar{\mathbf{E}}(\mathbf{r},\omega)\rightarrow\bar{E}_{\mathrm{y}}(\mathbf{r},\omega)\mathbf{e}_{\mathrm{y}}\rightarrow\bar{E}_{\mathrm{y}}(\mathbf{r},\omega)\rightarrow u(\mathbf{r},\omega).$$

  – exact for one-dimensional beams and linear polarization
  – approximation in two-dimensional case

In homogeneous isotropic media we have to solve the Helmholtz equation

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

In scalar approximation and for fixed frequency ω it reads

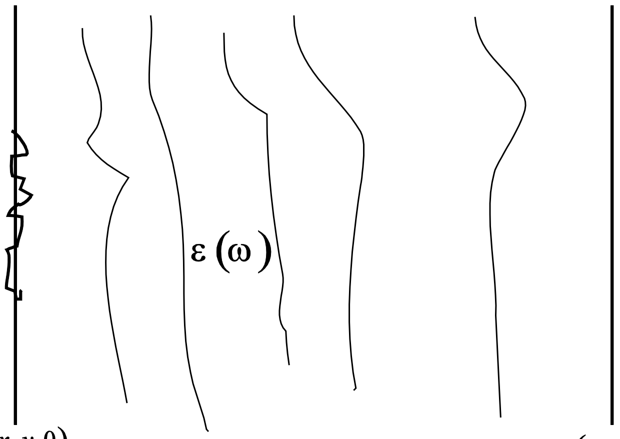
$$\Delta u(\mathbf{r},\omega)+\frac{\omega^2}{c^2}\varepsilon(\omega)u(\mathbf{r},\omega)=0,$$
$$\Delta u(\mathbf{r},\omega)+k^2(\omega)u(\mathbf{r},\omega)=0.$$

scalar Helmholtz equation

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 ω.

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

$$\Delta u(\mathbf{r},\omega)+k^2(\omega)u(\mathbf{r},\omega)=0$$

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

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

which can be interpreted as a superposition of normal modes with different propagation directions and wavenumbers $k(\omega)$ (here the absolute value of the wave-vector $\mathbf{k}$). 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:

$$\mathbf{k}^2=k^2=k_{\mathbf{x}}^2+k_{\mathbf{y}}^2+k_{\mathbf{z}}^2=\frac{\omega^2}{c^2}\varepsilon(\omega)$$

$\rightarrow$ only two components of $\mathbf{k}$ are independent, e.g., $k_{\mathbf{x}},k_{\mathbf{y}}$.

Our naming convention is in the following: $\boxed{k_{\mathbf{x}}=\alpha,\quad k_{\mathbf{y}}=\beta,\quad k_{\mathbf{z}}=\gamma.}$

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 the transverse coordinates x and y, when z is the direction of propagation:

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

In analogy to the frequency ω we call α, β spatial frequencies.

Now we plug this expression into the scalar Helmholtz equation

$$\Delta u(\mathbf{r})+k^2(\omega)u(\mathbf{r})=0$$

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

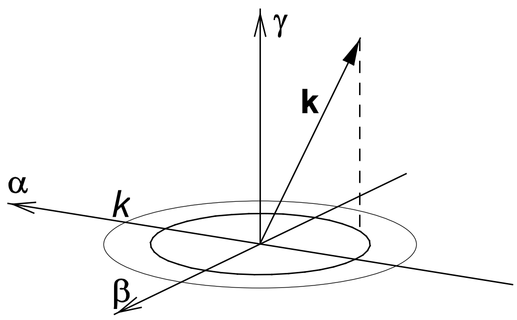
$$\boxed{U(\alpha,\beta;z)=U_1(\alpha,\beta)\exp\left[\mathbf{i}\gamma(\alpha,\beta)z\right]+U_2(\alpha,\beta)\exp\left[-\mathbf{i}\gamma(\alpha,\beta)z\right],}$$

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

We can identify two types of solutions:

A) Homogeneous waves

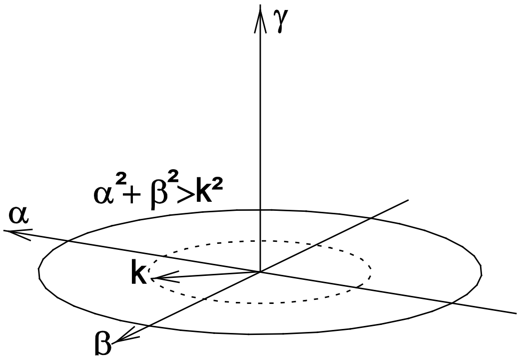
$\gamma^2\geq 0$, → $\alpha^2+\beta^2\leq k^2$, i.e., **k real** → homogeneous waves



B) Evanescent waves

$\gamma^2<0$, → $\alpha^2+\beta^2>k^2$, i.e., **k** complex, because $\gamma=k_z$ imaginary. Then, we have $\mathbf{k}=\mathbf{k}'+\mathbf{i}\mathbf{k}''$, with $\mathbf{k}'=\alpha\mathbf{e}_x+\beta\mathbf{e}_y$ and $\mathbf{k}''=\gamma\mathbf{e}_z$.

→ $\mathbf{k}'\perp\mathbf{k}''$ → evanescent waves



We see immediately that in the half-space $z>0$ the solution ~ $\exp(-\mathbf{i}\gamma z)$ grows exponentially. Because this does not make sense, this component of the solution must vanish $U_2(\alpha,\beta)=0$. In fact, we will see later that $U_2(\alpha,\beta)$ corresponds to backward running waves, i.e., light propagating in the opposite direction. We therefore find the solution:

$$\boxed{U(\alpha,\beta;z)=U_1(\alpha,\beta)\exp\left[\mathbf{i}\gamma(\alpha,\beta)z\right]}$$

Furthermore the following boundary condition holds

$$U_1(\alpha,\beta)U(\alpha,\beta;0)=U_0(\alpha,\beta)$$

which determines uniquely the entire solution for $z\neq 0$ as

$$\boxed{\begin{aligned}U(\alpha,\beta;z)&=U(\alpha,\beta;0)\exp\left[\mathbf{i}\gamma(\alpha,\beta)z\right]\\&\doteq U_0(\alpha,\beta)\exp\left[\mathbf{i}\gamma(\alpha,\beta)z\right]\end{aligned}}$$

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

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

$$\boxed{u(\mathbf{r})=\iint_{-\infty}^{\infty}U_0(\alpha,\beta)\exp\left[\mathbf{i}\gamma\left(\alpha,\beta\right)z\right]\exp\left[\mathbf{i}\left(\alpha x+\beta y\right)\right]d\alpha d\beta.}$$

For homogeneous waves (real γ) 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 of the different normal modes when these modes are traveling in propagation direction. The individual phase shifts are determined according to different spatial frequencies α,β of the normal modes.

The initial spatial frequency spectrum or angular spectrum at $z=0$ forms the initial condition of the initial value problem and follows from $u_0(x,y)=u(x,y,0)$ by Fourier transform:

$$U_0(\alpha,\beta)=\left(\frac{1}{2\pi}\right)^2\iint_{-\infty}^{\infty}u_0(x,y)\exp\left[-\mathbf{i}\left(\alpha x+\beta y\right)\right]dxdy,$$

As mentioned above the wave-vector components α,β 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 pane wave. For example $\alpha/k=\cos\theta_x$ 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 $\exp\left[\mathbf{i}\gamma\left(\alpha,\beta\right)z\right]$
4. new spectrum: $U(\alpha,\beta;z)=U_0(\alpha,\beta)\exp\left[\mathbf{i}\gamma\left(\alpha,\beta\right)z\right]$
5. new field distribution: $u(x,y,z)$ by Fourier back transform

This scheme allows for two interpretations:

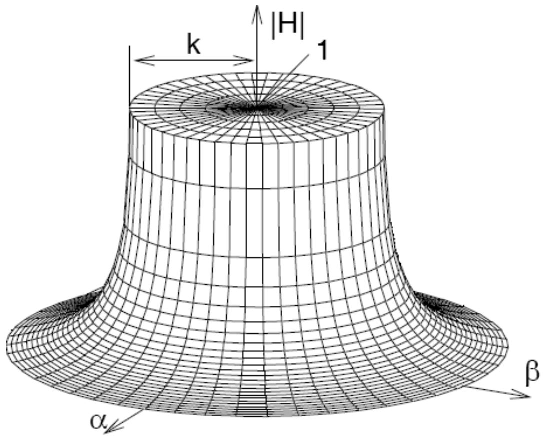
1) The resulting field distribution is the Fourier transform of the propagated spectrum

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

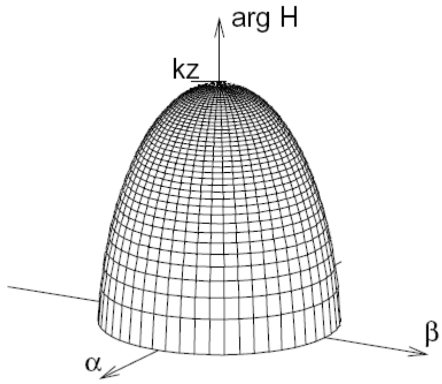
2) The resulting field distribution is a superposition, i.e. interference, of homogeneous and evanescent plane waves ('plane-wave spectrum') which obey the dispersion relation

$$\boxed{u(\mathbf{r})=\underbrace{\iint_{-\infty}^{\infty}d\alpha d\beta}_{\text{interference of eigenstates to form the field pattern after propagation}}\ \underbrace{U_0\left(\alpha,\beta\right)}_{\text{amplitude of the excited eigenstates}}\ \underbrace{\exp\left\{\mathbf{i}\gamma\left(\alpha,\beta\right)z\right\}}_{\text{phase factor which is accumulated by the eigenstates during propagation}}\ \underbrace{\exp\left\{\mathbf{i}\left[\alpha x+\beta y\right]\right\}}_{\text{shape of eigenstates (plan waves)}}}$$

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



amplitude



phase

Obviously, $H(\alpha,\beta;z)=\exp\left[\mathbf{i}\gamma\left(\alpha,\beta\right)z\right]$ acts differently on homogeneous and evanescent waves:

A) Homogeneous waves for $\alpha^2+\beta^2\le k^2$

$$\curvearrowright \quad \left|\exp\left[\mathbf{i}\gamma\left(\alpha,\beta\right)z\right]\right|=1,\ \arg\left(\exp\left[\mathbf{i}\gamma\left(\alpha,\beta\right)z\right]\right)\neq 0$$

→Upon propagation the homogeneous waves are multiplied by the phase factor

$$\exp\left[\mathbf{i}\sqrt{k^2-\alpha^2-\beta^2}z\right]$$

→Each homogeneous wave keeps its amplitude.

B) Evanescent waves for $\alpha^2+\beta^2>k^2$

$$\curvearrowright \quad \left|\exp\left[\mathbf{i}\gamma\left(\alpha,\beta\right)z\right]\right|=\exp\left[-\sqrt{\alpha^2+\beta^2-k^2}z\right],\ \arg\left(\exp\left[\mathbf{i}\gamma\left(\alpha,\beta\right)z\right]\right)=0$$

→Upon propagation the evanescent waves are multiplied by an amplitude factor <1

$$\exp\left[-\sqrt{\alpha^2+\beta^2-k^2}z\right]<1$$

This means that their contribution gets damped with increasing propagation distance z.

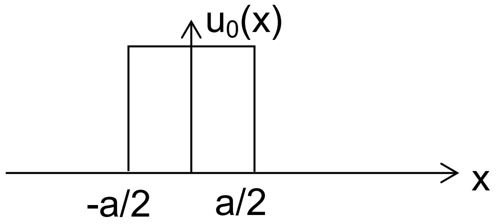
→Each evanescent wave keeps its phase.

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.

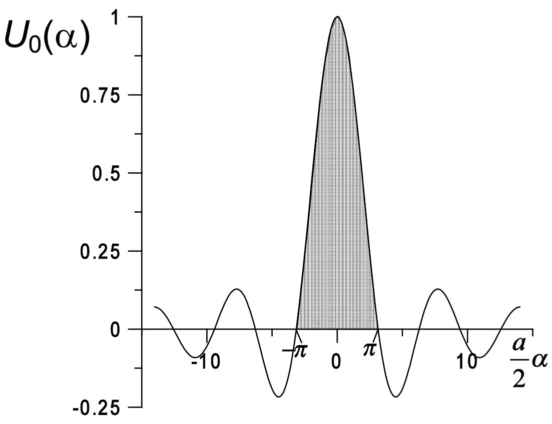
Example: Slit

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

$$u_0(x)=\begin{cases}1 & \text{for } \; |x|\le \dfrac{a}{2}.\\ 0 & \text{otherwise}\end{cases}$$



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



- All spatial frequencies ($-\infty \rightarrow \infty$) are excited.
- Important spectral information is contained in the interval $|\alpha|=2\pi/a$.
  → Largest important spectral frequency for a structure with width a is $\alpha=2\pi/2$.
- Evanescent waves appear for $|\alpha|>k$.
- To represent the relevant information by homogeneous waves the following condition must be fulfilled: $\dfrac{2\pi}{a}<k=\dfrac{2\pi}{\lambda}n \;\rightarrow\; \boxed{a>\dfrac{\lambda}{n}}$

General result

We have seen in the example above that evanescent waves appear for structures < wavelength in the initial condition. Information about these small structures gets lost for $z>>\lambda$.

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

$$u_0(x,y) \xrightarrow{\mathbf{FT}^{-1}} U_0(\alpha,\beta) \quad \rightarrow \quad U(\alpha,\beta;z) = H(\alpha,\beta;z)U_0(\alpha,\beta) \quad \xrightarrow{\mathbf{FT}} \quad u(x,y,z)$$

with the transfer function $H(\alpha,\beta;z) = \exp\left[\mathbf{i}\gamma(\alpha,\beta)z\right]$

Remark: diffraction free beams

With our understanding of diffraction it is straight forward 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

$$U(\alpha,\beta;z) = U_0(\alpha,\beta)\exp\left[\mathbf{i}\gamma(\alpha,\beta)z\right] \equiv U_0(\alpha,\beta)\exp\left[\mathbf{i}Cz\right]$$

$$\rightarrow \; u(x,y,z) = \exp\left[\mathbf{i}Cz\right]u_0(x,y)$$

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

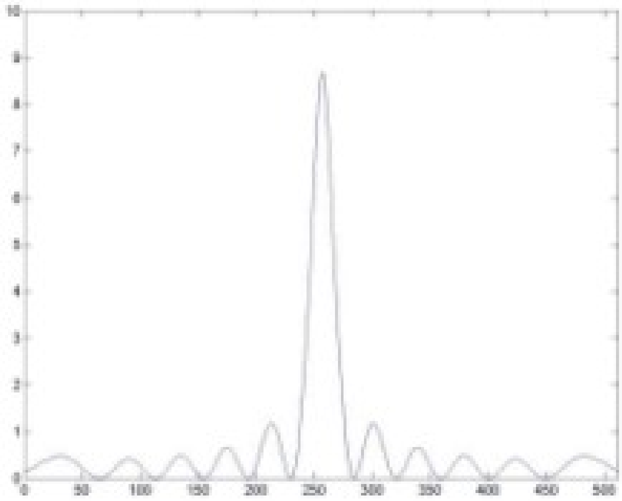
$$U_0(\alpha,\beta) \neq 0 \text{ only for } \gamma(\alpha,\beta) = \sqrt{k^2 - \alpha^2 - \beta^2} = C$$

It is straightforward to see that the excited spatial frequencies must lie on a circular ring in the (α,β) plane.

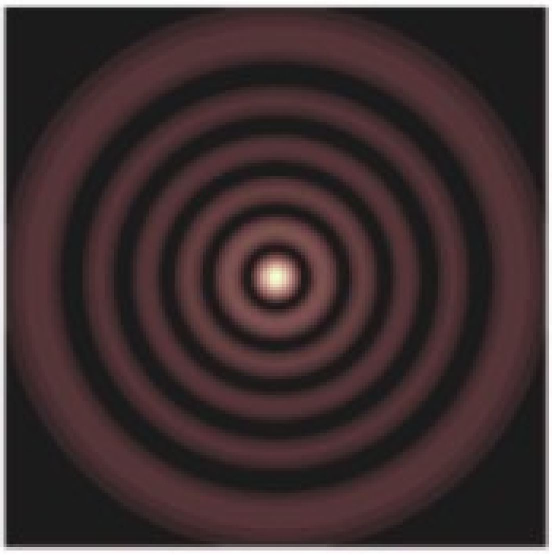
$$\alpha^2 + \beta^2 = \rho_0^2$$

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

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



Bessel-beam (profile)



Bessel-beam

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

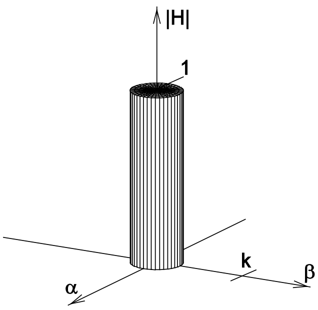
$$\boxed{U_0(\alpha,\beta)\neq 0 \qquad \text{for} \qquad \alpha^2+\beta^2 \ll k^2}$$

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:

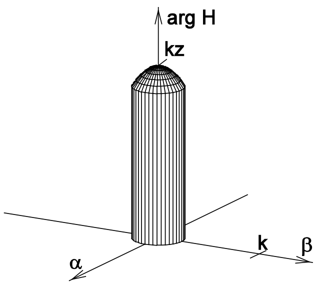
$$\gamma(\alpha,\beta)=\sqrt{k^2-\alpha^2-\beta^2}\approx k\left(1-\frac{\alpha^2+\beta^2}{2k^2}\right)=k-\frac{\alpha^2+\beta^2}{2k}$$

The resulting expression for the transfer function in Fresnel approximation reads:

$$\boxed{H=\exp\left(\mathbf{i}\gamma(\alpha,\beta)z\right)\approx\exp\left(\mathbf{i}kz\right)\exp\left(-\mathbf{i}\frac{\alpha^2+\beta^2}{2k}z\right)=H_{\mathrm{F}}(\alpha,\beta;z)}$$



Amplitude



Phase

We can see that this $H_{\mathbf{F}}(\alpha,\beta;z)$ is always real valued. Hence it does not account for the physics of evanescent waves. However, we must remember that the derivation of $H_{\mathbf{F}}(\alpha,\beta;z)$ as an approximation of $H(\alpha,\beta;z)$ required the assumption that the spatial frequency spectrum is narrow (paraxial waves). Thus, already at the beginning we had excluded the excitation of evanescent waves 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:

$$|\Delta x|,|\Delta y| > 10\lambda / n \gg \lambda / n$$

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. This includes the condition that the phase of the beam should not have a strong inclination to the principal propagation direction.

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:

$$\boxed{U_{\mathbf{F}}(\alpha,\beta;z) = H_{\mathbf{F}}(\alpha,\beta;z)U_0(\alpha,\beta)}$$

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 $H_{\mathbf{F}}(\alpha,\beta;z)$.

Description in real space

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

$$\begin{aligned} u_{\mathbf{F}}(x,y,z) &= \iint_{-\infty}^{\infty} U_{\mathbf{F}}(\alpha,\beta;z)\exp\left[\mathbf{i}\left(\alpha x+\beta y\right)\right]d\alpha d\beta \\ &= \iint_{-\infty}^{\infty} H_{\mathbf{F}}(\alpha,\beta;z)U_0(\alpha,\beta)\exp\left[\mathbf{i}\left(\alpha x+\beta y\right)\right]d\alpha d\beta \\ &= \iint_{-\infty}^{\infty} h_{\mathbf{F}}(x-x',y-y';z)u_0(x',y')dx'dy' \end{aligned}$$

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

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

This Fourier integral can be solved and we find:

$$\boxed{h_{\mathrm{F}}(x,y;z)=\exp\left(\mathbf{i}kz\right)\left\{-\frac{\mathbf{i}k}{2\pi z}\exp\left[\mathbf{i}\frac{k}{2z}\left(x^2+y^2\right)\right]\right\}=-\frac{\mathbf{i}k}{2\pi z}\exp\left\{\mathbf{i}kz\left[1+\frac{\left(x^2+y^2\right)}{2z^2}\right]\right\},}$$

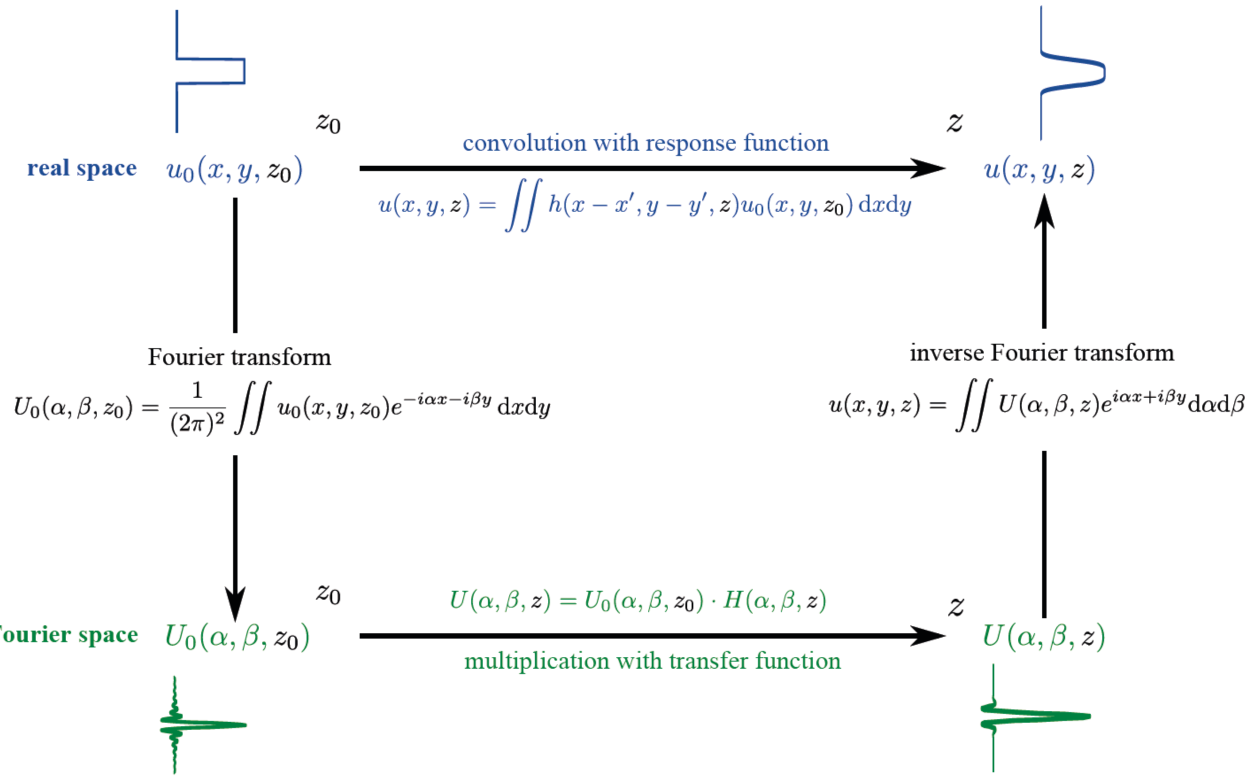
The response function corresponds to a spherical wave in paraxial approximation. Similar to Huygens principle, where from each point in the object plane a spherical wave is emitted towards the image plane, here paraxial approximations of spherical waves are emitted.

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

$$\boxed{u_{\mathrm{F}}(x,y,z)=-\frac{\mathbf{i}k}{2\pi z}\exp\left(\mathbf{i}kz\right)\iint_{-\infty}^{\infty}u_0(x',y')\exp\left\{\mathbf{i}\frac{k}{2z}\left[\left(x-x'\right)^2+\left(y-y'\right)^2\right]\right\}dx'dy'.}$$

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:

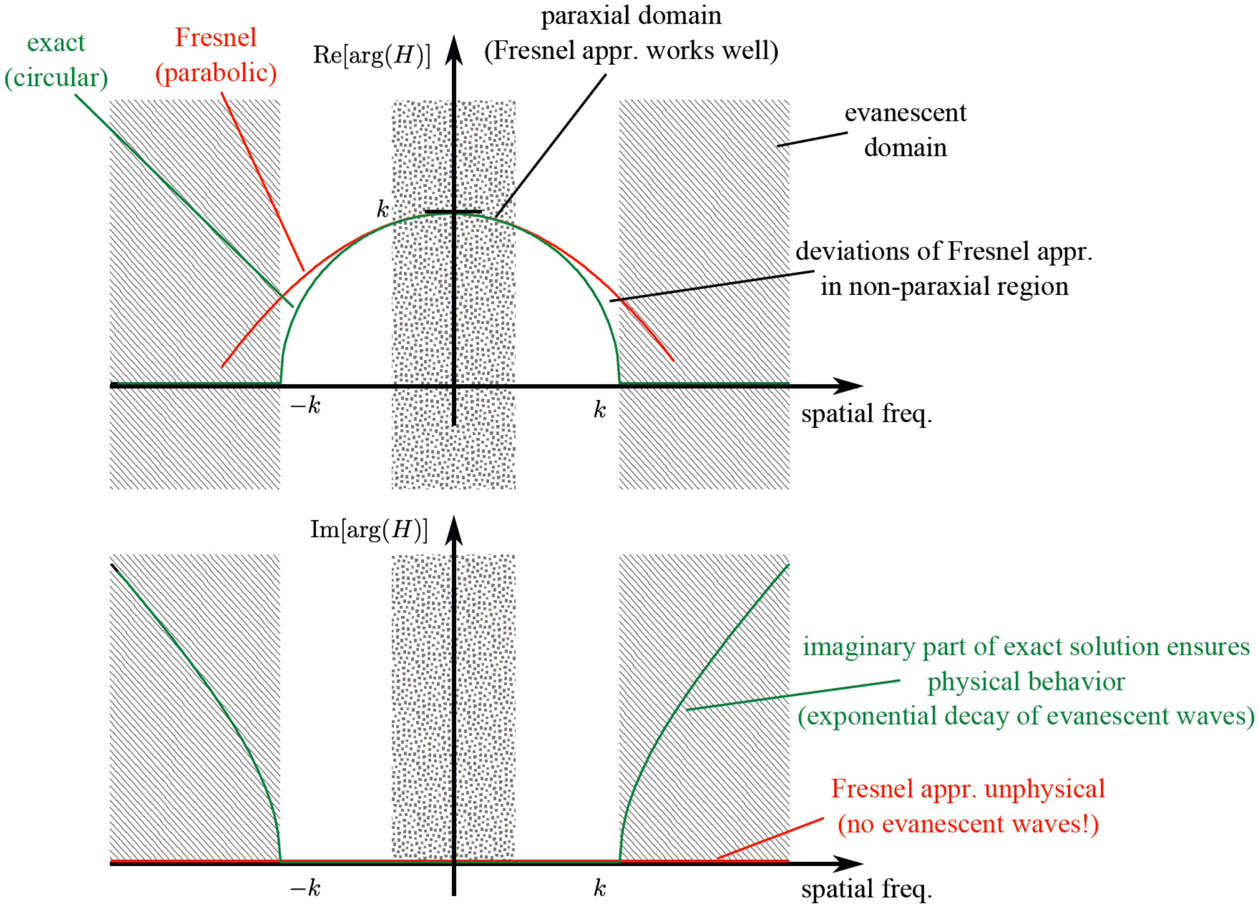
$$h(x,y;z)=\frac{1}{(2\pi)^2}\iint_{-\infty}^{\infty}H(\alpha,\beta;z)\exp\left[\mathrm{i}\left(\alpha x+\beta y\right)\right]d\alpha d\beta$$

Transfer functions for homogeneous space

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

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

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



Remark on the validity of the scalar approximation

In the previous description of the propagation of arbitrary beams we have used the scalar approximation of the vectorial fields. It is interesting to see, to what extend this approximation stays in correspondence to the conditions which where necessary to derive the Fresnel approximation.

$$\bar{\mathbf{E}}(\mathbf{r},\omega)=\iint\hat{\bar{\mathbf{E}}}(\alpha,\beta,\omega)e^{i(\alpha x+\beta y+\gamma z)}d\alpha d\beta$$

$$\boxed{\mathbf{div}\bar{\mathbf{E}}(\mathbf{r},\omega)=0\rightarrow\alpha\hat{\bar{E}}_x+\beta\hat{\bar{E}}_y+\gamma\hat{\bar{E}}_z=0}$$

A) One-dimensional beams

- translational invariance in y-direction: $\beta = 0$
- and linear polarization in y-direction: $\hat{\bar{E}}_y = U$, $\hat{\bar{E}}_x = 0$

→scalar approximation is exact since divergence condition is strictly fulfilled

B) Two-dimensional beams

- Finite beam which is localized in the x,y-plane: $\alpha, \beta \neq 0$
- and linear polarization, w.l.o.g. in y-direction: $\hat{\bar{E}}_x = 0$, $\hat{\bar{E}}_y = U$

→divergence condition: $\beta \hat{\bar{E}}_y + \gamma \hat{\bar{E}}_z = 0$

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

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

$$U_F(\alpha,\beta;z) = H_F(\alpha,\beta;z)U_0(\alpha,\beta)$$

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

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

$$\rightarrow\ U_F(\alpha,\beta;z) = \exp(\mathbf{i}kz)V(\alpha,\beta;z)\ \rightarrow\ V(\alpha,\beta;z) = \exp\left(-\mathbf{i}\frac{\alpha^2+\beta^2}{2k}z\right)V_0(\alpha,\beta).$$

Differentiation of V with respect to z gives:

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

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

$$\mathbf{i}\frac{\partial}{\partial z}\iint_{-\infty}^{\infty} V(\alpha,\beta;z)\exp\left[\mathbf{i}(\alpha x+\beta y)\right]d\alpha d\beta$$

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

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

$$\rightarrow \boxed{\mathbf{i}\frac{\partial}{\partial z}v(x,y,z)+\frac{1}{2k}\Delta^{(2)}v(x,y,z)=0}$$ paraxial wave equation

The slowly varying envelope $v(x,y,z)$ (Fourier transform of the slowly varying spectrum) relates to the scalar field as $u_{\mathbf{F}}(x,y,z)=v(x,y,z)\exp(\mathbf{i}kz)$.

Extension of the wave equation to weakly inhomogeneous media

(slowly varying envelope approximation - SVEA)

There is an alternative, more general way to derive the paraxial wave equation, the so-called slowly varying envelope approximation. This approximation even allows us to treat wave propagation in inhomogeneous media. We will include inhomogeneous media in this derivation even though the current chapter of this lecture is devoted to inhomogeneous media.

We start from the scalar Helmholtz equation. However, we should mention that the extrapolation of our previous discussion towards inhomogeneous media requires another approximation. This approximation assumes small spatial fluctuations of $\varepsilon(\mathbf{r},\omega)$. This is equivalent to having a weak index contrast between different spatial positions.

$$\Delta u(x,y,z)+k^2(\mathbf{r},\omega)u(x,y,z)=0 \text{ with } k^2(\mathbf{r},\omega)=\frac{\omega^2}{c^2}\varepsilon(\mathbf{r},\omega)$$

We make the ansatz

$$u(x,y,z)=v(x,y,z)\exp\left(\mathbf{i}\tilde{k}z\right) \text{ with } \tilde{k}=\langle k\rangle$$

where $\tilde{k}$ is the spatially averaged wavenumber. This is the mean wave number in the volume, in which wave propagation is considered. Hence it is the average of the spatially varying material properties of the medium.

With the SVEA condition

$$\left|\tilde{k}v\right|\gg\left|\partial v/\partial z\right|$$

we can simplify the scalar Helmholtz equation as follows:

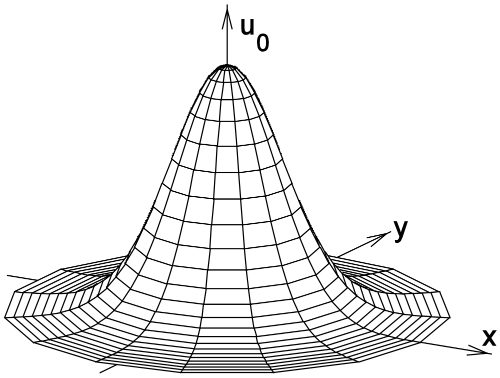
$$\underbrace{\frac{\partial^2}{\partial z^2}v(x,y,z)}_{\approx 0}+2\mathbf{i}\tilde{k}\frac{\partial}{\partial z}v(x,y,z)+\Delta^{(2)}v(x,y,z)+\left[k^2(\mathbf{r},\omega)-\tilde{k}^2\right]v(x,y,z)=0,$$

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

This is the paraxial wave equation for inhomogeneous media having a weak index contrast.

2.8 Propagation of Gaussian beams

The propagation of Gaussian beams is an important special case. Optical beams of Gaussian shape are import because many beams in reality have a shape which is at least close to the shape of a Gauss function. In some cases they even have exactly the Gaussian shape as e.g. the transversal fundamental modes of many lasers. The second reason why Gaussian beams are important, is the fact that in paraxial approximation it is possible to compute the evolution of Gaussian beams analytically.



Fundamental Gaussian beam in focus

The general form of a Gaussian beam at one plane ($z=\mathrm{const}$) is elliptic, with curved phase

$$u_0(x,y)=v_0(x,y)=A_0\exp\left[-\left(\frac{x^2}{w_\mathbf{x}^2}+\frac{y^2}{w_\mathbf{y}^2}\right)\right]\exp\left[\mathbf{i}\varphi(x,y)\right].$$

Here, we will restrict ourselves to rotational symmetry ($w_\mathbf{x}^2=w_\mathbf{y}^2=w_0^2$) and (initially) 'flat' phase $\varphi(x,y)=0$. Later we will see that this corresponds to the focus of a Gaussian beam. The Gaussian beam in such a focal plane with flat phase is characterized by amplitude A and width w_0, which is defined as

$$u_0(x^2+y^2=w_0^2)=A_0\exp(-1)=A_0/e.$$

In practice, the so-called 'full width at half maximum' (FWHM) of the intensity is often used instead of w_0. The FWHM of the intensity is connected to w_0 of the field by

$$\left|u_0\left(x^2+y^2\right)\right|^2=\exp\left(-\frac{w_{\mathrm{FWHM}}^2}{2\mathrm{w}_0^2}\right)\doteq\frac{1}{2}$$

$$-\frac{w_{\mathrm{FWHM}}^2}{2\mathrm{w}_0^2}=-\ln 2 \;\rightarrow\; \boxed{w_{\mathrm{FWHM}}^2=2\ln 2\mathrm{w}_0^2\approx 1.386\mathrm{w}_0^2}$$

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 $z=0$:

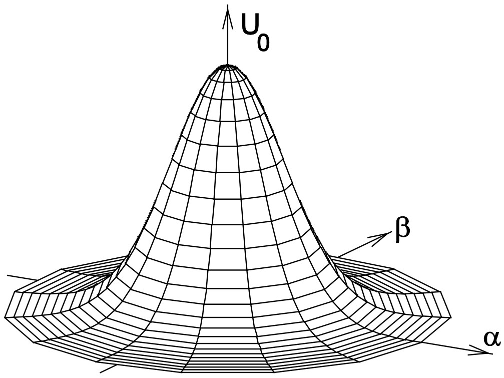
$$u_0(x,y)=v_0(x,y)=A_0\exp\left(-\frac{x^2+y^2}{w_0^2}\right).$$

2) Angular spectrum at $z=0$:

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

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

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 $w_s\cdot w_0=2$.



Angular spectrum in the focal plane

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 $k^2\gg\left(\alpha^2+\beta^2\right)$

→ $k^2\gg 16/w_0^2$

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

→ paraxial approximation works for $\boxed{w_0\gtrsim 10\frac{\lambda}{n}=10\lambda_n}$

3) Propagation of the angular spectrum:

$$V(\alpha,\beta;z)=U_0(\alpha,\beta)\mathrm{H}_\mathrm{F}(\alpha,\beta;\mathrm{z})$$

$$V(\alpha,\beta;z)=U_0(\alpha,\beta)\exp\left(-\mathrm{i}\frac{\alpha^2+\beta^2}{2k}z\right)$$
$$=\frac{A_0}{4\pi}w_0^2\exp\left[-w_0^2\frac{\alpha^2+\beta^2}{4}\right]\exp\left(-\mathrm{i}\frac{\alpha^2+\beta^2}{2k}z\right).$$

4) Fourier back-transformation to position space, when still using the slowly varying envelope v

$$v(x,y,z)=\frac{A_0}{4\pi}w_0^2\iint_{-\infty}^{\infty}\exp\left[-\left(\alpha^2+\beta^2\right)\left(\frac{w_0^2}{4}+\frac{\mathrm{i}}{2k}z\right)\right]\exp\left[\mathrm{i}\left(\alpha x+\beta y\right)\right]d\alpha d\beta$$
$$=A_0\frac{1}{1+\frac{2\mathrm{i}z}{kw_0^2}}\exp\left[-\frac{x^2+y^2}{w_0^2\left(1+2\mathrm{i}z/(kw_0^2)\right)}\right]$$
$$=A_0\frac{1}{1+\mathrm{i}\frac{z}{z_0}}\exp\left[-\frac{x^2+y^2}{w_0^2\left(1+\mathrm{i}z/z_0\right)}\right].$$

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

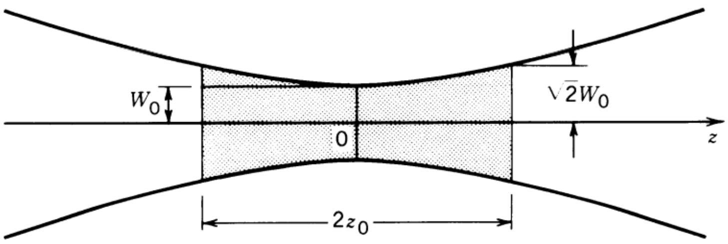
$$\boxed{z_0=\frac{kw_0^2}{2}=\frac{\pi}{\lambda_n}w_0^2.}$$

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" $L_B=2z_0$, a measure for the "focus depth" of the Gaussian beam. E.g.: $w_0\gtrsim 10\lambda_n$ → $L_B\gtrsim 600\lambda_n$.



The depth of focus of a Gaussian beam.

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

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

For practical use, we can write this expression in terms of z-dependent amplitude, width, and shape of the phase front:

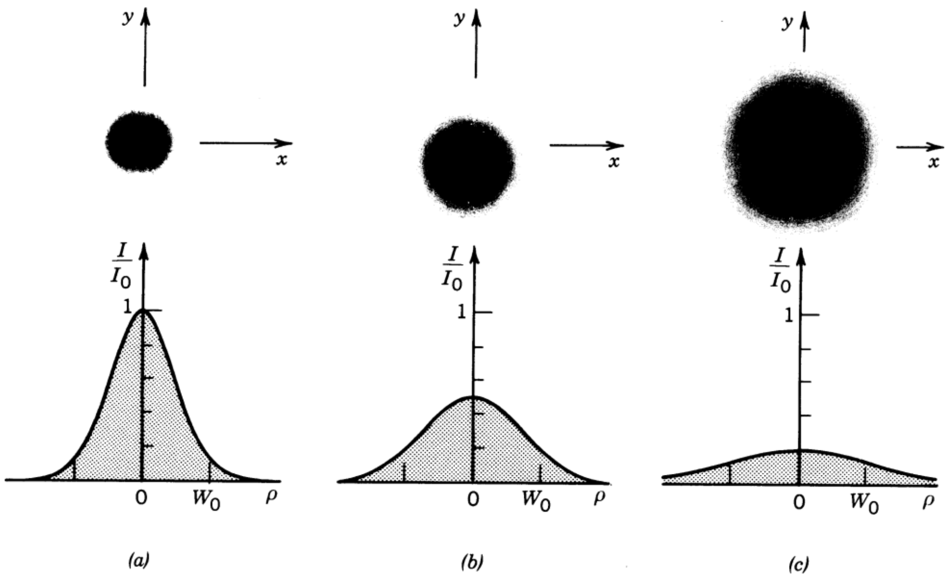
$$v(x,y,z)=A_0\frac{1-\mathrm{i}\frac{z}{z_0}}{1+(z/z_0)^2}\exp\left\{-\frac{x^2+y^2}{w_0^2\left[1+(z/z_0)^2\right]}\right\}\exp\left\{\mathrm{i}\frac{(x^2+y^2)z/z_0}{w_0^2\left[1+(z/z_0)^2\right]}\right\}$$

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

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

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

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



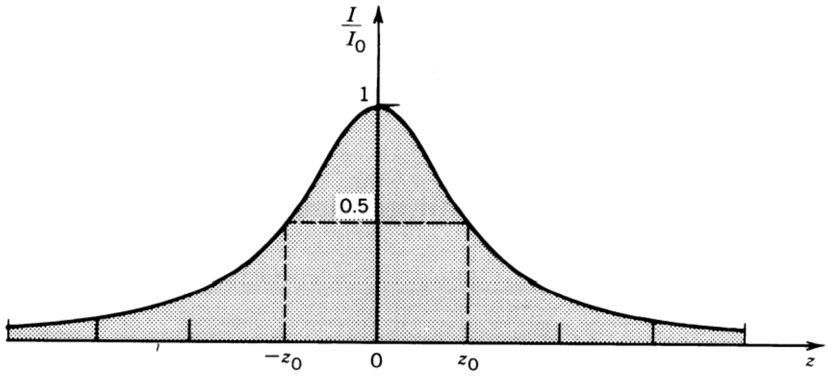
The normalized beam intensity I/I_0 as a function of the radial distance ρ at different axial distances: (a) $z=0$; (b) $z=z_0$; (c) $z=2z_0$.

Discussion

The amplitude evolves along z like:

$$\boxed{A(z) = A_0 \frac{1}{\sqrt{1+\left(\frac{z}{z_0}\right)^2}} = A_0 \frac{1}{\sqrt{1+\left(\frac{2z}{l_B}\right)^2}},}$$

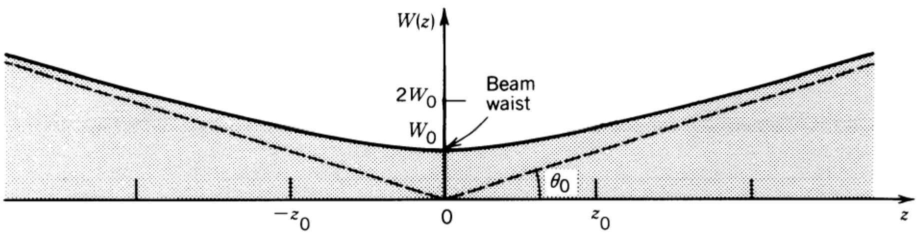
Hence, we get for the Intensity profile $I \sim A^2$:



The normalized beam intensity I / I_0 at points on the beam axis ($\rho = \sqrt{x^2+y^2} = 0$) as a function of the propagation distance z.

The beam width evolves along z like:

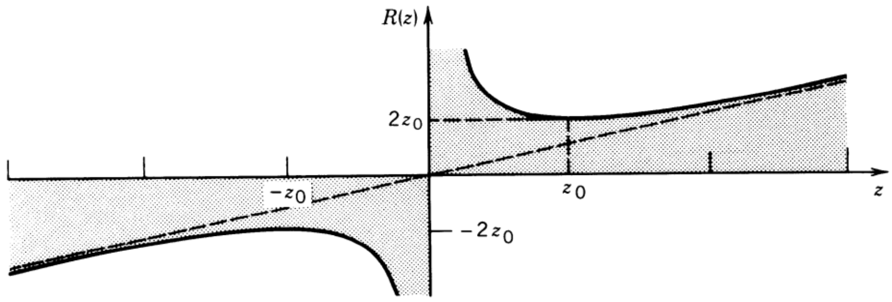
$$\boxed{w(z) = w_0 \sqrt{1+\left(\frac{z}{z_0}\right)^2} = w_0 \sqrt{1+\left(\frac{2z}{l_B}\right)^2}}$$



The beam radius $W(z)$ has its minimum value W_0 at the waist ($z=0$), reaches $\sqrt{2}W_0$ at $z = \pm z_0$, and increases linearly with z for large z.

The radius of the phase curvature is given by

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

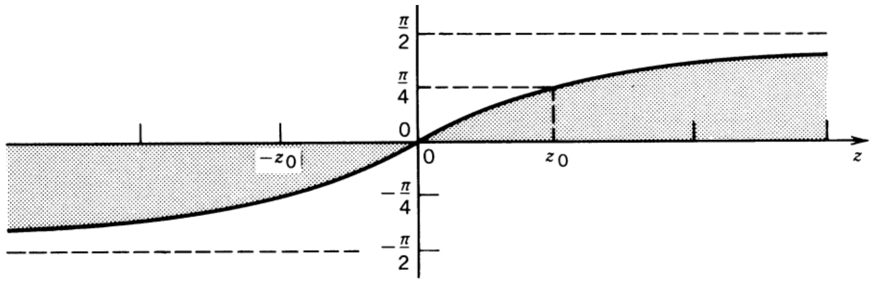


The radius of curvature $R(z)$ of the wavefronts of a Gaussian beam. The dashed line is the radius of curvature of a spherical wave.

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

$$\boxed{\tan\varphi = -\frac{z}{z_0} = -\frac{2z}{L_B}}$$

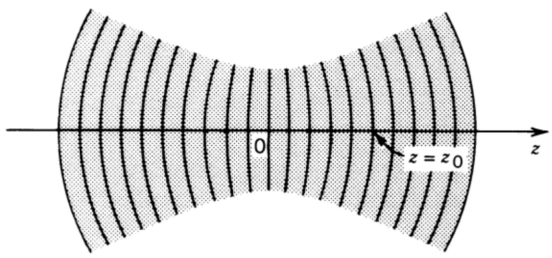


The so-called Gouy phase is the retardation $\varphi(z)$ of the phase of a Gaussian beam relative to a uniform plane wave at the points of the beam axis.

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 wave fronts (planes of constant phase) of a Gaussian beam are given by

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



Wavefronts of a Gaussian beam.

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

$$q(z) = z - \mathbf{i} z_0 \qquad \text{q-parameter}$$

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

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

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

$$\boxed{\frac{1}{q(z)} = \frac{1}{R(z)} + \mathbf{i}\frac{\lambda_{\mathrm{n}}}{\pi w^2(z)}} \quad \text{because } z_0 = \frac{k w_0^2}{2} = \frac{\pi}{\lambda_{\mathrm{n}}} w_0^2$$

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

Example: propagation in homogeneous space by $z = d$

A) initial conditions: $\dfrac{1}{q(0)} = \dfrac{1}{R(0)} + \mathbf{i}\dfrac{\lambda_{\mathrm{n}}}{\pi w^2(0)}$

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

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

$$\frac{1}{q(d)} = \frac{1}{q(0) + d} \doteq \frac{1}{R(d)} + \mathbf{i}\frac{\lambda_{\mathrm{n}}}{\pi w^2(d)}$$

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_0, w_0 (i.e. q_0) → 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:

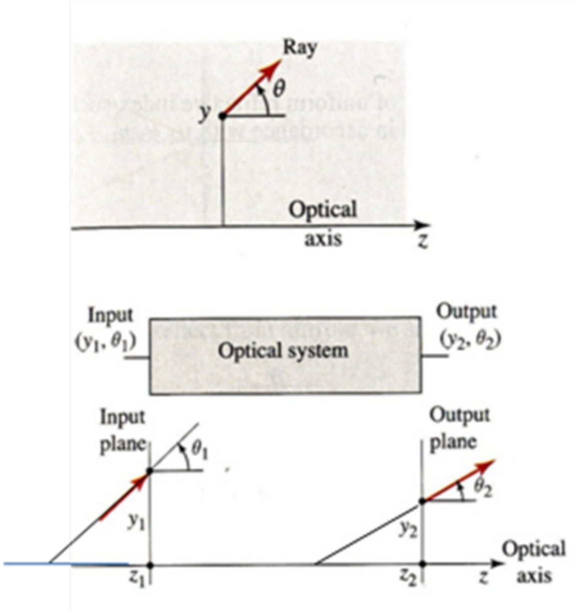
$$\hat{\mathbf{M}} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

B) propagation through multiple elements:

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

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

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



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". Using the small angle approximation ($\tan(\Theta_1) \approx \Theta_1 = y_1/z_1$) we can define this distance as:

$$z_1 = \frac{y_1}{\Theta_1} \rightarrow z_2 = \frac{y_2}{\Theta_2} = \frac{Ay_1 + B\Theta_1}{Cy_1 + D\Theta_1} = \frac{A\frac{y_1}{\Theta_1} + B}{C\frac{y_1}{\Theta_1} + D} = \frac{Az_1 + B}{Cz_1 + D}$$

The distances z_1, z_2 are connected by matrix elements, but not by normal matrix vector multiplication.

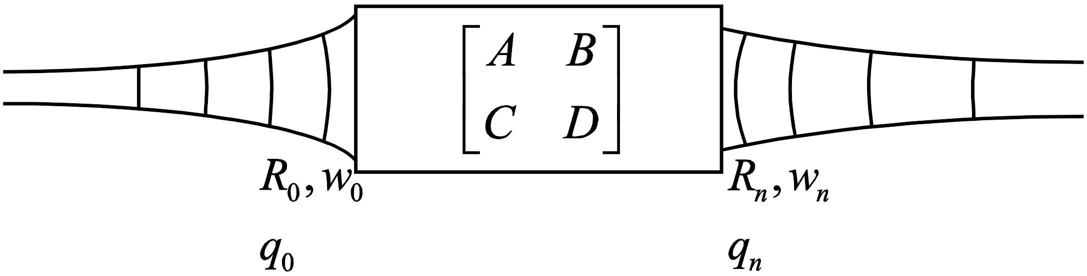
It turns out that we can pass to Gaussian optics by replacing z by the complex beam parameter q. The propagation of q-parameters through an optical element is given by:

$$\boxed{q_1 = \frac{A_1 q_0 + B_1}{C_1 q_0 + D_1}}$$

→ propagation through N elements:

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

with the matrix $\hat{\mathbf{M}} = \hat{\mathbf{M}}_{\mathbf{N}} \hat{\mathbf{M}}_{\mathbf{N}-1} .. \hat{\mathbf{M}}_1 = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$.



→ 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) $q(d) = q(0) + d$

$$\hat{\mathbf{M}} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \rightarrow A_1 = 1,\ B_1 = d,\ C_1 = 0,\ D_1 = 1$$

$$q_1 = \frac{A_1 q_0 + B_1}{C_1 q_0 + D_1} = \frac{q_0 + d}{0 + 1} = q_0 + d$$

ii) thin lens with focal length f

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

– no change of the width

– but change of phase curvature R_f: $\times \exp\left[\mathbf{i}\frac{k}{2}\frac{\left(x^2+y^2\right)}{R_f}\right]$

How can we see that?

Trick:

We start from the focus which is produced by the lens with $z_0 = z_f = \dfrac{\pi w_f^2}{\lambda_n}$

and w_f is the focal width. Hence the q-parameter is

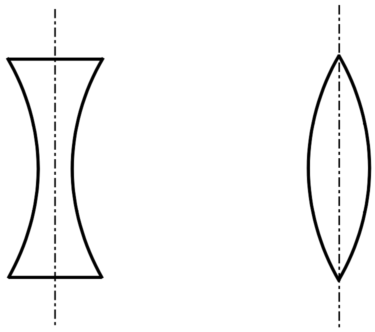
$$\frac{1}{q_0} = i\frac{\lambda_n}{\pi w_f^2}$$

The radius of curvature evolves as:

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

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

$f < 0$ $\qquad$ $f > 0$ $\qquad$ $\hat{\mathbf{M}}_{\text{thin lens}} = \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix}$



double concave lens → defocusing $\qquad$ *double convex lens → focusing*

$$q_1 = \frac{A_1 q_0 + B_1}{C_1 q_0 + D_1} = \frac{q_0}{-q_0/f + 1}$$

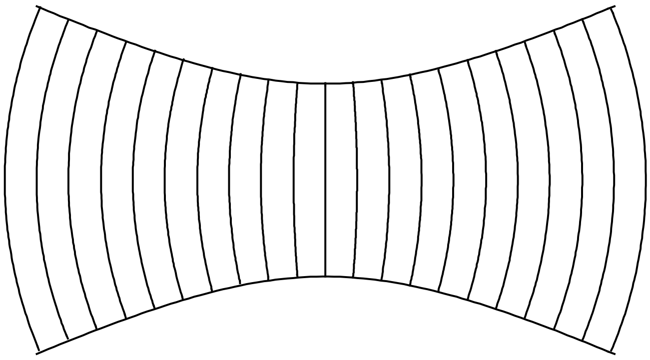
$$\frac{1}{q_1} = \frac{-q_0/f + 1}{q_0} = \frac{1}{-f} + \frac{\mathbf{i}\lambda_n}{\pi w_0^2} \qquad \text{for} \quad q_0 = -\mathbf{i}\frac{\pi w_0^2}{\lambda_n}$$

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



Wave fronts, i.e. planes of constant phase, of a Gaussian beam.

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

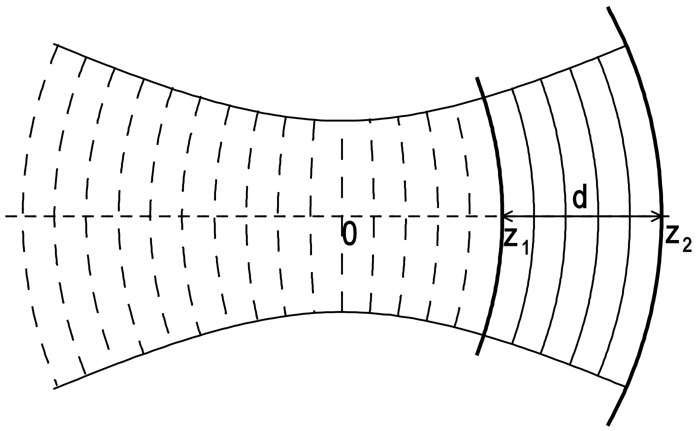
In paraxial approximation and Gaussian beams this condition is easily fulfilled: The radii of mirror and wave front have to be identical!

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

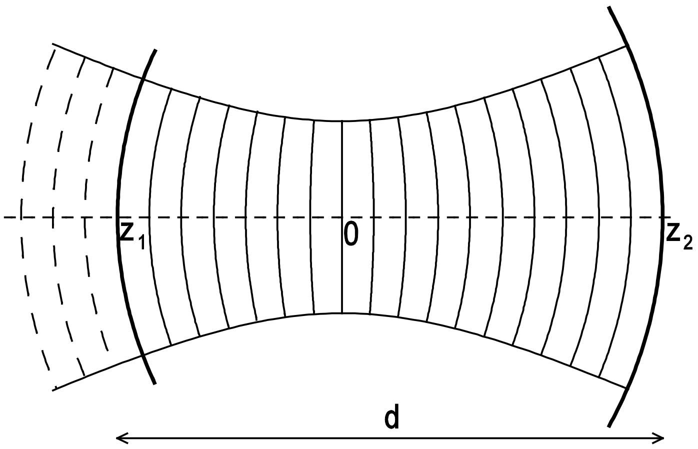
- $z_{1,2}$ is the position of mirror '1','2'; z=0 is the position of the focus!
- d is the distance between the two mirrors → $z_2 - z_1 = d$
- because $R(z) = z + \frac{z_0^2}{z}$ → radius of wave front <0 for z <0
- from Chapter 1: beam hits concave mirror → radius $R_i(i=1,2) < 0.$
- beam hits convex mirror → radius $R_i(i=1,2) > 0.$

Examples:

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



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



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

$$\boxed{R_1 = R(z_1), \qquad R_2 = -R(z_2)}$$

$$\curvearrowright \quad R_1 = z_1 + \frac{z_0^2}{z_1}, \qquad -R_2 = z_2 + \frac{z_0^2}{z_2}.$$

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

$$z_1(R_1 - z_1) = -z_2(R_2 + z_2)$$

with $z_2 = z_1 + d$ 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!

$$\rightarrow \; z_0^2 = R_1 z_1 - z_1^2 = -\frac{d(R_1 + d)(R_2 + d)(R_1 + R_2 + d)}{(R_1 + R_2 + 2d)^2} > 0$$

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

$$\rightarrow \; -d(R_1 + d)(R_2 + d)(R_1 + R_2 + d) > 0$$

If we introduce the so-called resonator parameters

$$\boxed{g_1 = \left(1 + \frac{d}{R_1}\right), \qquad g_2 = \left(1 + \frac{d}{R_2}\right)}$$

We can re-express the stability condition as

$$-d(R_1 + d)(R_2 + d)(R_1 + R_2 + d) = d g_1 g_2 R_1 R_2 \frac{(1 - g_1 g_2) R_1 R_2}{d}$$

$$= g_1 g_2 (1 - g_1 g_2)(R_1 R_2)^2 > 0.$$

This inequality is fulfilled for

(1) $g_1 g_2 > 0$ and $1 - g_1 g_2 > 0$

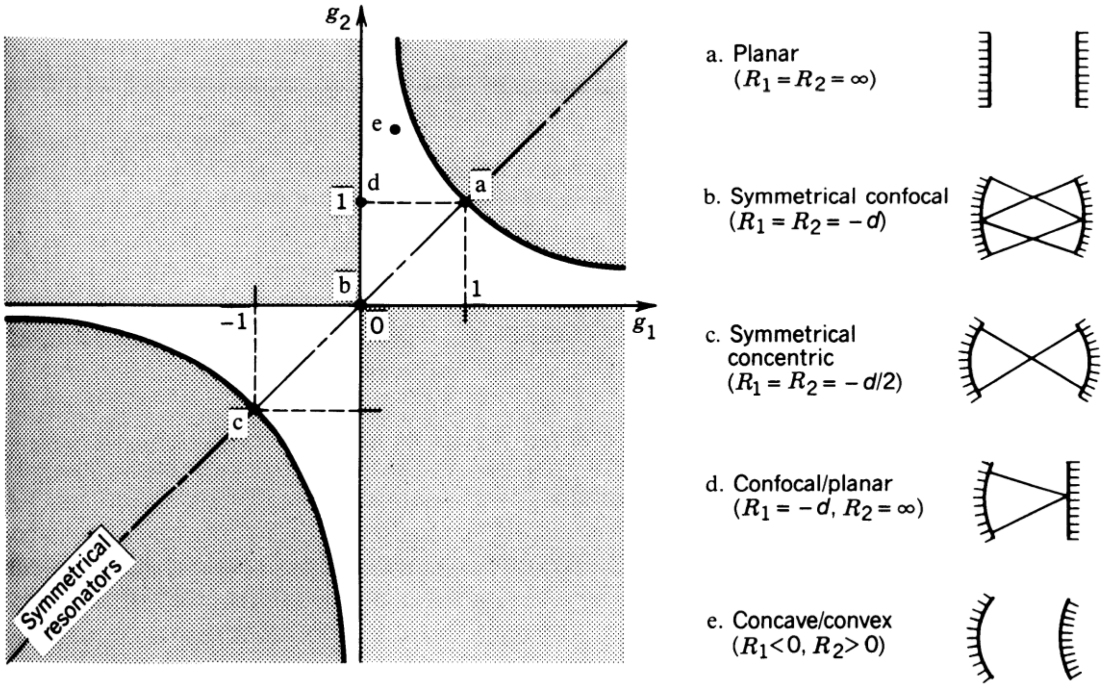
or for

(2) $g_1 g_2 < 0$ and $1 - g_1 g_2 < 0$

which results in the following condition for the placement and radii of the mirrors for a stable cavity

$$\boxed{0 < g_1 g_2 < 1 \text{ which is equivalent to } 0 < \left(1+\frac{d}{R_1}\right)\left(1+\frac{d}{R_2}\right) < 1}$$

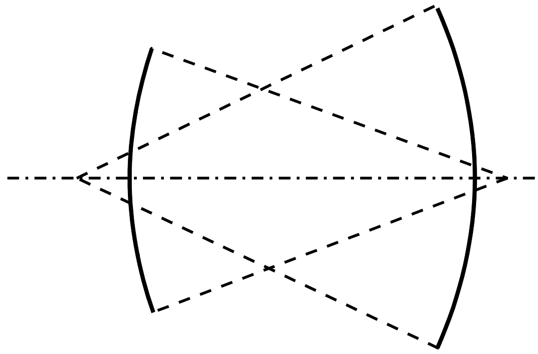
This final form of the stability condition can be visualized: The range of stability of a resonator lies between coordinate axes and hyperbolas:



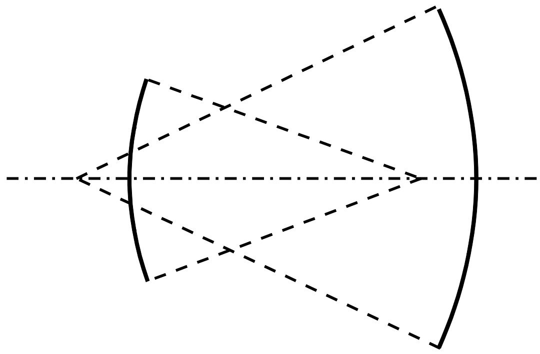
Resonator stability diagram. A spherical-mirror resonator is stable if the parameters $g_1 = 1 + d/R_1$ and $g_2 = 1 + d/R_2$ lie in the unshaded regions bounded by the lines $g_1 = 0$ and $g_2 = 0$, and the hyperbola $g_2 = 1/g_1$. R is negative for a concave mirror and positive for a convex mirror. 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) $R_1, R_2 < 0;\ |R_1| > d,\ |R_2| > d;\ \curvearrowright 0 \le g_1 \le 1,\ 0 \le g_2 \le 1; \curvearrowright 0 \le g_1 g_2 \le 1 \curvearrowright$ stable



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



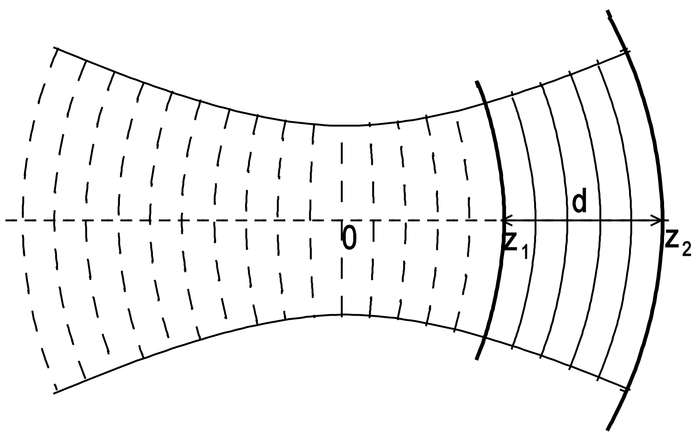
Remark connection to He-Ne-Labwork script (and Wikipedia):

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

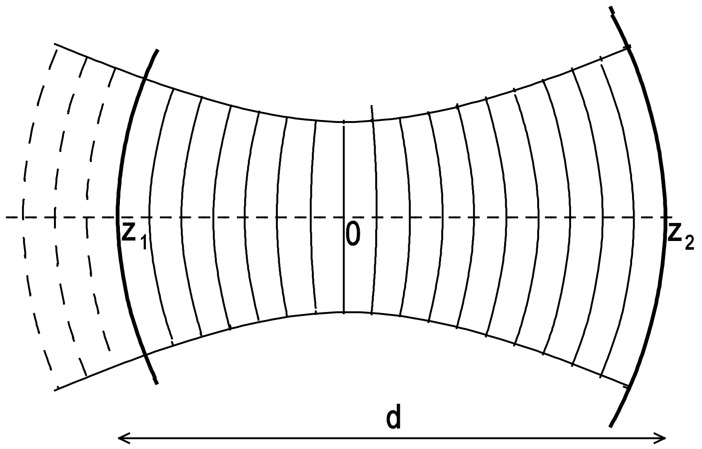
- Direction of z-axis reversed for the two mirrors
- beam hits concave mirror → radius $R_i(i=1,2) > 0$.
- beam hits convex mirror → radius $R_i(i=1,2) < 0$.
- $z_{1,2}$ is the distance of mirror '1','2' to the focus!
- d is the distance between the two mirrors → $z_2 + z_1 = d$

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:

$$\boxed{R_1 = R(z_1), \qquad R_2 = R(z_2)}$$

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

$$\boxed{g_1 = \left(1 - \frac{d}{R_1}\right), \qquad g_2 = \left(1 - \frac{d}{R_2}\right)}$$

the same stability condition

$$g_1 g_2 \left(1 - g_1 g_2\right)\left(R_1 R_2\right)^2 > 0, \qquad \boxed{0 < g_1 g_2 < 1.}$$

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:

$$v_{\mathrm{G}}(x,y,z) = A \frac{w_0}{w(z)} \exp\left[-\frac{x^2+y^2}{w^2(z)}\right] \exp\left[\mathbf{i}\frac{k}{2}\frac{x^2+y^2}{R(z)}\right] \exp\left[\mathbf{i}\varphi(z)\right].$$

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

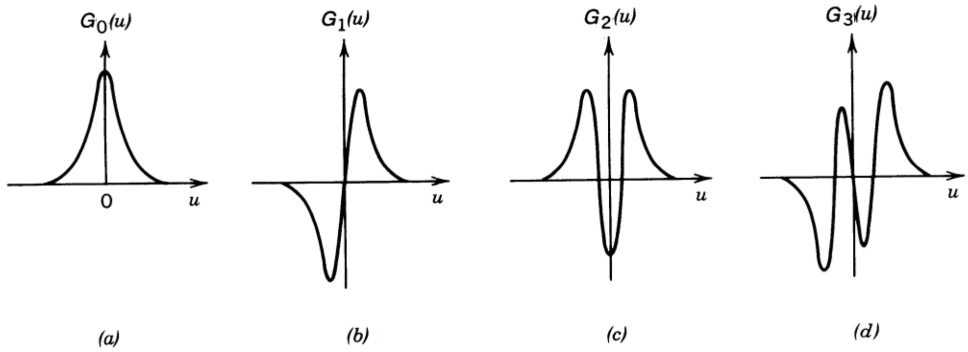
$$u_{l,m}(x,y,z) = A_{l,m} \frac{w_0}{w(z)} G_l\left[\sqrt{2}\frac{x}{w(z)}\right] G_m\left[\sqrt{2}\frac{y}{w(z)}\right] \times$$

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

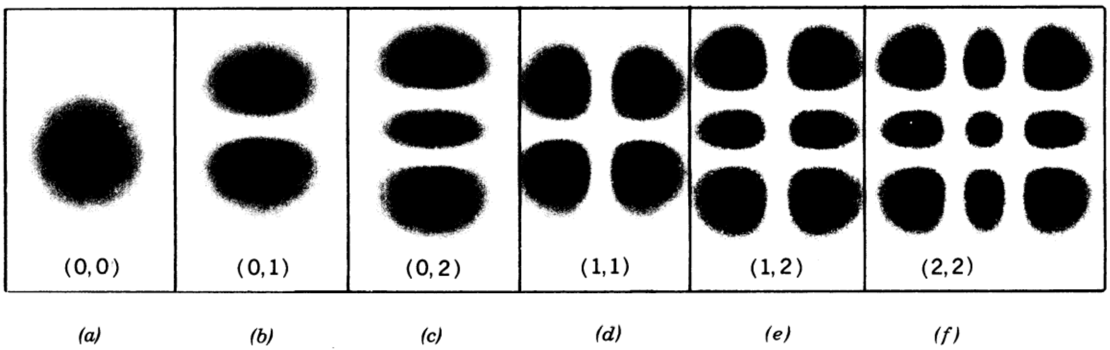
$$G_l(\xi) = H_l(\xi) \exp\left(-\frac{\xi^2}{2}\right)$$

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

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



Several low-order Hermite-Gaussian fuctions: (a) $G_0(u)$; (b) $G_1(u)$; (c) $G_2(u)$; and (d) $G_3(u)$.



Intensity distributions of several low-order Hermite-Gaussian beams in the transverse plane. The order (l,m) is indicated in each case.

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 the propagation of monochromatic beams, where the frequency ω 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 a typical envelope length T_0 of

$$10^{-13}\mathrm{s}(100\mathrm{fs}) \le T_0 \le 10^{-10}\mathrm{s}(100\mathrm{ps}).$$

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

$$f(t) = \exp(-\mathrm{i}\omega_0 t)\exp\left(-\frac{t^2}{T_0^2}\right)$$

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

→ spectral width: $\quad 4\cdot 10^{10} s^{-1} \le \omega_s \le 4\cdot 10^{13} s^{-1}$

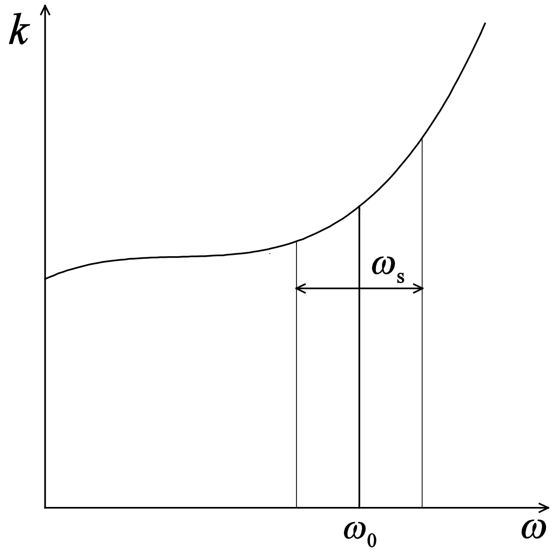
– center frequency of visible light: $\omega_0 = 2\pi\nu \sim 4\cdot 10^{15}$ s^{-1}

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

Hence, we have the following order of magnitudes:

$$\boxed{\omega_s << \omega_0 \rightarrow \omega - \omega_0 = \bar{\omega} << \omega_0}$$

In this situation it can be helpful to replace the complicated frequency dependence (dispersion relation) of the wave vector $\mathbf{k}(\omega)$ 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:

$$\boxed{k(\omega) \approx k(\omega_0) + \left.\frac{\partial k}{\partial \omega}\right|_{\omega_0} (\omega - \omega_0) + \frac{1}{2}\left.\frac{\partial^2 k}{\partial \omega^2}\right|_{\omega_0} (\omega - \omega_0)^2 + \ldots}$$

The three expansion coefficients and their physical significance

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.

A) Phase velocity v_{Ph}

$$k(\omega_0) = k_0, \rightarrow \frac{1}{v_{\mathbf{Ph}}} = \frac{k_0}{\omega_0} = \frac{n(\omega_0)}{c}$$

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

B) Group velocity v_g

$$\left.\frac{\partial k}{\partial\omega}\right|_{\omega_0} = \frac{1}{v_g}$$

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

$$\boxed{k(\omega) = \frac{\omega}{c}n(\omega) \rightarrow \frac{1}{v_g} = \left.\frac{\partial k}{\partial\omega}\right|_{\omega_0} = \frac{1}{c}\left[n(\omega_0) + \omega_0\left.\frac{\partial n}{\partial\omega}\right|_{\omega_0}\right]}$$

$$v_g = \frac{c}{\left[n(\omega_0) + \omega_0\left.\frac{\partial n}{\partial\omega}\right|_{\omega_0}\right]} = \frac{c}{n_g(\omega_0)} = v_{PH}\frac{n(\omega_0)}{n_g(\omega_0)}$$

$$\boxed{n_g(\omega_0) = n(\omega_0) + \omega_0\left.\frac{\partial n}{\partial\omega}\right|_{\omega_0}}$$ → group index

normal dispersion: $\partial n / \partial\omega > 0 \curvearrowright n_g > n, v_g < v_{PH}$

anomalous dispersion: $\partial n / \partial\omega < 0 \curvearrowright n_g < n, v_g > v_{PH}$

C) Group velocity dispersion (GVD) or simply dispersion D_ω

$$\left.\frac{\partial^2 k}{\partial\omega^2}\right|_{\omega_0} = D_\omega$$

→ GVD changes pulse shape upon propagation (see below)

$$\boxed{D = D_\omega = \left.\frac{\partial^2 k}{\partial\omega^2}\right|_{\omega_0} = \frac{\partial}{\partial\omega}\left(\frac{1}{v_g}\right)}$$

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

$$\rightarrow D > 0 \curvearrowright \frac{\partial v_g}{\partial\omega} < 0$$

$$\rightarrow D < 0 \curvearrowright \frac{\partial v_g}{\partial\omega} > 0$$

Alternatively in telecommunication one often uses

$$\boxed{D_\lambda = \frac{\partial}{\partial\lambda}\left(\frac{1}{v_g}\right) = -\frac{2\pi}{\lambda^2}cD_\omega}$$

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

$$\Delta\ \bar{u}(\mathbf{r},\omega)+\frac{\omega^2}{c^2}\varepsilon(\omega)\ \bar{u}(\mathbf{r},\omega)=0$$

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

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

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

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

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

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

$$U_0(\alpha,\beta,\omega)=\frac{1}{(2\pi)^3}\iiint_{-\infty}^{\infty}u_0(x,y,t)\exp\left[-\mathbf{i}\left(\alpha x+\beta y-\omega t\right)\right]dxdydt$$

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

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

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

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

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

$$k(\omega)\approx k\left(\omega_0\right)+\left.\frac{\partial k}{\partial\omega}\right|_{\omega_0}\left(\omega-\omega_0\right)+\left.\frac{1}{2}\frac{\partial^2 k}{\partial\omega^2}\right|_{\omega_0}\left(\omega-\omega_0\right)^2+..$$

Moreover, in the second term $\exp[-\mathbf{i}(\alpha^2+\beta^2)z/\{2k(\omega)\}]$ 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 to 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:

$$\boxed{\begin{aligned} H_{\mathbf{FP}}(\alpha,\beta,\omega;z) &\approx \exp\left[\mathbf{i}k_0 z\right]\exp\left[\mathbf{i}\frac{1}{v_{\mathbf{g}}}(\omega-\omega_0)\;z\right]\times \\ &\times\exp\left[\mathbf{i}\frac{D}{2}(\omega-\omega_0)^2\;z\right]\exp\left[-\mathbf{i}\frac{\alpha^2+\beta^2}{2k_0}z\right] \\ &= \exp\left[\mathbf{i}k_0 z\right]\exp\left[\mathbf{i}\;z\left(\frac{\bar{\omega}}{v_{\mathbf{g}}}+\frac{1}{2}D\bar{\omega}^2-\frac{1}{2}\frac{1}{k_0}\left(\alpha^2+\beta^2\right)\right)\right]\end{aligned}}$$

with $\bar{\omega}=\omega-\omega_0$

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 $\bar{\omega}$ from the center frequency ω_0

$$\boxed{H_{\mathbf{FP}}(\alpha,\beta,\bar{\omega};z)\approx\exp\left[\mathbf{i}k_0 z\right]\bar{H}_{\mathbf{FP}}(\alpha,\beta,\bar{\omega};z)}$$

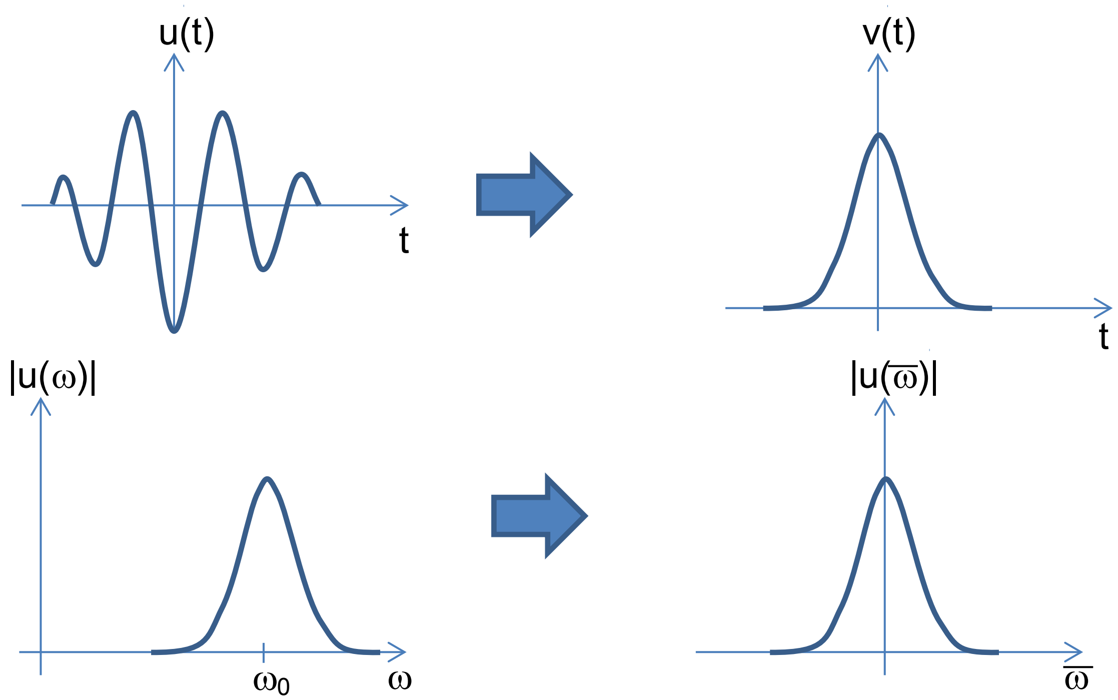
The new transfer function $\bar{H}_{\mathbf{FP}}(\alpha,\beta,\bar{\omega};z)$ is the propagation function for the slowly varying envelope $v(x,y,z,t)$:

$$\begin{aligned} u(x,y,z,t) &= \exp\left[\mathbf{i}k_0 z\right]\iiint_{-\infty}^{\infty}U_0(\alpha,\beta,\omega)\bar{H}_{\mathbf{FP}}(\alpha,\beta,\bar{\omega};z)\times \\ &\times\exp\left[\mathbf{i}\left(\alpha x+\beta y-\omega\,t\right)\right]d\alpha d\beta d\omega\end{aligned}$$

$$\boxed{\begin{aligned} u(x,y,z,t) &= \exp\left[\mathbf{i}\left(k_0 z-\omega_0 t\right)\right]\iiint_{-\infty}^{\infty}U_0(\alpha,\beta,\omega)\bar{H}_{\mathbf{FP}}(\alpha,\beta,\bar{\omega};z)\times \\ &\times\exp\left[\mathbf{i}\left(\alpha x+\beta y-\bar{\omega}\,t\right)\right]d\alpha d\beta d\bar{\omega}\end{aligned}}$$

Illustration of the slowly varying envelope in the spectral domain

$$u(x,y,z,t)=v(x,y,z,t)\exp\left[\mathbf{i}\left(k_0 z-\omega_0 t\right)\right]$$



$$v(x,y,z,t)=\iiint_{-\infty}^{\infty}U_0(\alpha,\beta,\omega)\bar{H}_{\mathrm{FP}}(\alpha,\beta,\bar{\omega};z)\exp\left[\mathbf{i}\left(\alpha x+\beta y-\bar{\omega}\,t\right)\right]d\alpha d\beta d\bar{\omega}$$

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

$$u_0(x,y,t)=v_0(x,y,t)\exp\left(-\mathbf{i}\omega_0 t\right)$$

$$\rightarrow\ V_0(\alpha,\beta,\bar{\omega})=\frac{1}{\left(2\pi\right)^3}\iiint_{-\infty}^{\infty}v_0(x,y,t)\exp\left[-\mathbf{i}\left(\alpha x+\beta y-\bar{\omega}t\right)\right]dxdydt$$

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

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

Co-moving reference frame

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

$$\bar{H}_{\mathrm{FP}}(\alpha,\beta,\bar{\omega};z)=\exp\left[\mathbf{i}\frac{\bar{\omega}}{v_g}z\right]\tilde{H}_{\mathrm{FP}}(\alpha,\beta,\bar{\omega};z)$$

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

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

$$\tau = t - \frac{z}{v_{\mathbf{g}}}$$

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

$$\boxed{\begin{aligned}\tilde{v}(x,y,z,\tau) = \iiint_{-\infty}^{\infty} V_0(\alpha,\beta,\overline{\omega})\exp\left[\mathbf{i}\,\frac{z}{2}\left(D\overline{\omega}^2 - \frac{\alpha^2+\beta^2}{k_0}\right)\right]\times \\ \times\exp\left\{\mathbf{i}\left[\alpha x + \beta y - \overline{\omega}\tau\right]\right\}d\alpha d\beta d\overline{\omega}\end{aligned}}$$

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

$$\boxed{\tilde{u}(x,y,z,\tau) = \tilde{v}(x,y,z,\tau)\exp\left[\mathbf{i}\left(k_0 z - \omega_0 t\right)\right] = \tilde{v}(x,y,z,\tau)\exp\left[\mathbf{i}\left(k_0 z - \frac{\omega_0}{v_{\mathbf{g}}}z - \omega_0\tau\right)\right]}$$

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

$$\tilde{V}(\alpha,\beta,\overline{\omega};z) = V_0(\alpha,\beta,\overline{\omega})\exp\left[\mathbf{i}\,\frac{z}{2}\left(D\overline{\omega}^2 - \frac{\alpha^2+\beta^2}{k_0}\right)\right]$$

Then we take the spatial derivative of the transfer function along the propagation direction z

$$\mathbf{i}\frac{\partial\tilde{V}(\alpha,\beta,\overline{\omega};z)}{\partial z} = -\frac{1}{2}\left(D\overline{\omega}^2 - \frac{\alpha^2+\beta^2}{k_0}\right)\tilde{V}(\alpha,\beta,\overline{\omega};z)$$

As before in the case of monochromatic beams, we use Fourier back-transformation to get the differential equation in the time-position domain

$$\boxed{\mathbf{i}\frac{\partial\tilde{v}(x,y,z,\tau)}{\partial z} - \frac{D}{2}\frac{\partial^2}{\partial\tau^2}\tilde{v}(x,y,z,\tau) + \frac{1}{2k_0}\Delta^{(2)}\tilde{v}(x,y,z,\tau) = 0}$$

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.

$$\boxed{\mathbf{i}\frac{\partial}{\partial z}\tilde{v}(x,y,z,\tau)-\frac{D}{2}\frac{\partial^2}{\partial\tau^2}\tilde{v}(x,y,z,\tau)+\frac{1}{2\bar{k}_0}\Delta^{(2)}\tilde{v}(x,y,z,\tau)+\left[\frac{k_0^2(\mathbf{r})-\bar{k}_0^2}{2\bar{k}_0}\right]\tilde{v}(x,y,z,\tau)=0}$$

with $\bar{k}_0\approx\langle k_0(\mathbf{r})\rangle$

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 $z\ll L_B$. For broad beams, the diffraction length L_B 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. Later in the lecture series we will see that this case is also valid for mode propagation in waveguides, as e.g. optical telecommunication fibers.

Description in frequency domain

1) initial condition: $u_0(t)=v_0(t)\exp(-\mathbf{i}\omega_0 t)$
2) initial spectrum: $V_0(\bar{\omega})=U_0(\omega)$
3) propagation of the spectrum: $V(\bar{\omega};z)=V_0(\bar{\omega})\exp\left[\mathbf{i}\ z\frac{D}{2}\bar{\omega}^2\right]$
4) back-transformation to τ leads to the following evolution of the slowly varying envelope in the co-moving frame:

$$\boxed{\tilde{v}(z,\tau)=\int_{-\infty}^{\infty}V_0(\bar{\omega})\exp\left[\mathbf{i}z\frac{D}{2}\bar{\omega}^2\right]\exp\left[-\mathbf{i}\bar{\omega}\tau\right]d\bar{\omega}}$$

Description in time domain

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

FT^{-1} of $\tilde{H}_P(\bar{\omega};z)=\exp\left[\mathbf{i}z\frac{D}{2}\bar{\omega}^2\right]\rightarrow\tilde{h}_P(\tau;z)=\sqrt{\frac{2}{-\mathbf{i}\pi Dz}}\exp\left[-\mathbf{i}\frac{\tau^2}{2Dz}\right]$

and the evolution is described by the convolution integral

$$\boxed{\tilde{v}(z,\tau)=\int_{-\infty}^{\infty}\tilde{h}_P(\tau-\tau';z)\tilde{v}_0(\tau')d\tau'}$$

B) The evolution equation for slowly varying envelope in the co-moving frame reads

$$\boxed{\mathbf{i}\frac{\partial\tilde{v}(z,\tau)}{\partial z}-\frac{D}{2}\frac{\partial^2}{\partial\tau^2}\tilde{v}(z,\tau)=0}$$

Analogy of diffraction and dispersion

DIFFRACTION DISPERSION

$$\mathbf{i}\frac{\partial}{\partial z}v(x,y,z)+\frac{1}{2k}\Delta^{(2)}v(x,y,z)=0 \quad \leftrightarrow \quad \mathbf{i}\frac{\partial}{\partial z}\tilde{v}(z,\tau)-\frac{D}{2}\frac{\partial^2}{\partial\tau^2}\tilde{v}(z,\tau)=0$$

$$(\alpha,\beta) \quad \leftrightarrow \quad \bar{\omega}$$

$$(x,y) \quad \leftrightarrow \quad \tau$$

$$\nabla \quad \leftrightarrow \quad \frac{\partial}{\partial\tau}$$

$$\frac{1}{k_0} \quad \leftrightarrow \quad -D \text{ but } D \lessgtr 0 \text{ can vary}$$

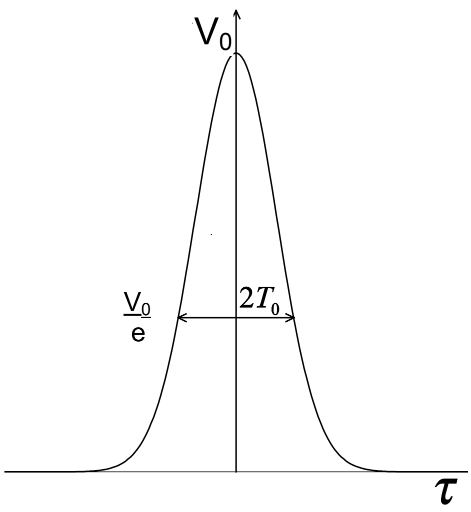
In the following we will study two typical examples of pulse propagation.

2.9.3 Example 1: Gaussian pulse without chirp

use analogies to spatial diffraction

1. Initial pulse shape

pulse without chirp → corresponds to Gaussian pulse in the waist (focus) with flat phase



$$u_0(t)=A_0\exp\left(-\frac{t^2}{T_0^2}\right)\exp\left(-\mathbf{i}\omega_0 t\right) \rightarrow v_0(\tau)=A_0\exp\left(-\frac{\tau^2}{T_0^2}\right)$$

2. Initial pulse spectrum

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

→ spectral width: $\omega_s^2=4/T_0^2$

Use results from propagation of Gaussian beams:

z_0 describes Gaussian pulse $z_0 = \left(\frac{k}{2} w_0^2\right)$ → $\boxed{z_0 = -\frac{1}{2}\frac{T_0^2}{D} \gtrless 0}$

Hence anomalious GVD is equivalent to 'normal' diffraction.

Dispersion length: $L_D = 2|z_0|$

3. Evolution of the amplitude

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

with

$$A(z) = A_0 \sqrt[4]{\frac{1}{1+\left(\frac{z}{z_0}\right)^2}}, \qquad T(z) = T_0 \sqrt{1+\left(\frac{z}{z_0}\right)^2}$$

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

'Phase curvature' is not fitting to the description of pulses → introduction of new parameter Chirp

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

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\Phi(x,y,z) = \frac{k}{R(z)}$$

For monochromatic fields the temporal dynamics of the phase is:

$$\Phi(\tau) = -\omega\tau \qquad \rightarrow \qquad -\frac{\partial\Phi(\tau)}{\partial\tau} = \omega$$

→ arbitrary time dependence of phase

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

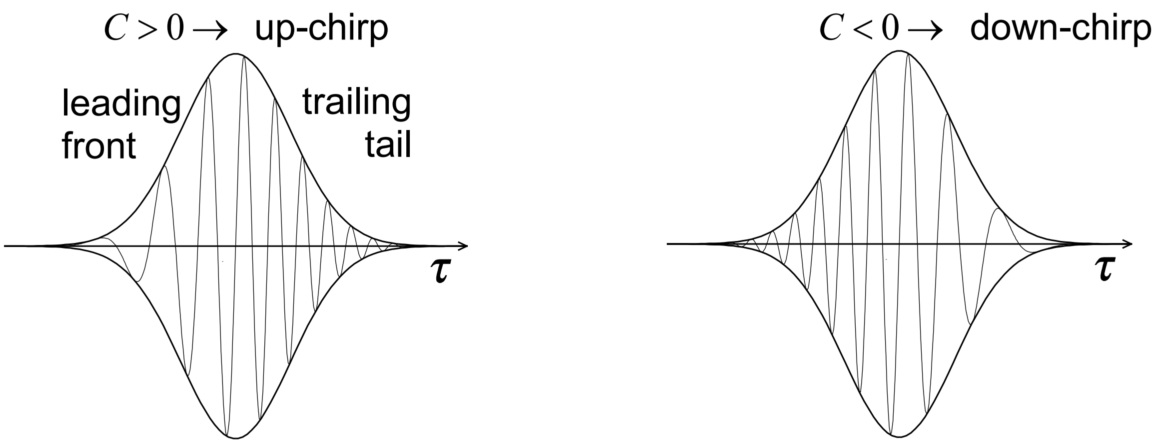
The chirp of a pulse describes the variation of the temporal frequency of the electric field in the pulse.

→ parabolic approximation → 'chirp' constant → dimensionless chirp parameter (often just chirp)

$$\boxed{C = -\frac{T_0^2}{2}\frac{\partial^2\Phi(\tau)}{\partial\tau^2}}$$

integration leads to:

$$-\frac{\partial\Phi(\tau)}{\partial\tau} = \omega(\tau) = \omega_0 + 2C\frac{\tau}{T_0^2}, \qquad -\Phi(\tau) = \omega_0\tau + C\frac{\tau^2}{T_0^2}$$



phase curvature $R(z)$ → Chirp $C(z)$

Complete phase:

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

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

with

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

→ $C(0)=0,\ C(|z_0|)=-\frac{1}{2}\mathrm{sgn}(z_0),\ C(z\rightarrow\infty)=-\frac{z_0}{z}$ with $z_0=-\frac{T_0^2}{2D}$

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

Complete field:

$$\boxed{u(z,\tau)=A_0\sqrt{\frac{T_0}{T(z)}}\exp\left[-\frac{\tau^2}{T(z)^2}\right]\exp\left[-\mathbf{i}C(z)\ \frac{\tau^2}{T_0^2}\right]\exp\left[\mathbf{i}\varphi(z)\right]\exp\left[\mathbf{i}\left(k_0 z-\omega_0 t\right)\right]}$$

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

important parameter → dispersion parameter $z_0=-\frac{T_0^2}{2D}$

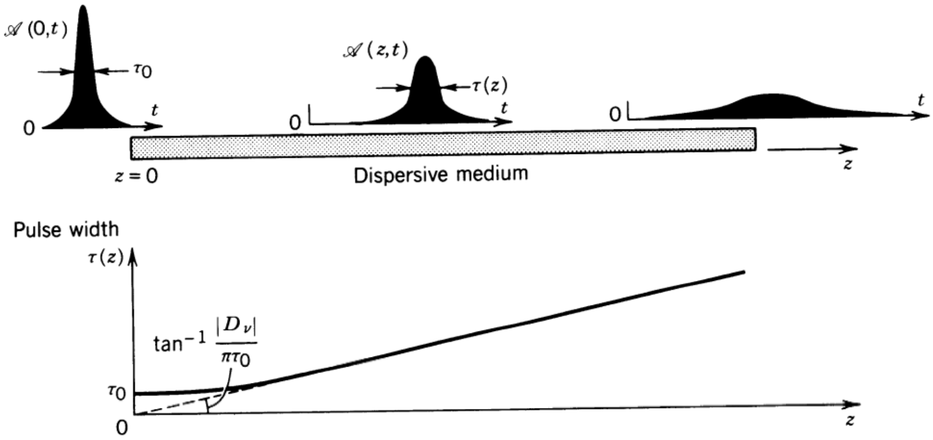
1) $z<<|z_0|$: no effect

2) $z\simeq|z_0|$: similar to beam diffraction

3) $z\gg|z_0|$: asymptotic dependence

$$T(z)\approx T_0\frac{z}{|z_0|}\rightarrow T(z)/z=T_0/|z_0|=\frac{2|D|}{T_0}=const.$$

$$T(z) \approx \frac{2|D|}{T_0} z$$



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.

$D \lessgtr 0$ 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 → chirped pulses
- Chirp is introduced on purpose, for subsequent pulse compression
- analogy to curved phase → focusing
- chirped pulse amplification (CPA) → Petawatt lasers

1. Input pulse shape

$$v_0(\tau) = A_0 \exp\left[-\frac{\tau^2(1+\mathrm{i}C_0)}{T_0^2}\right] \qquad C_0 \text{ – initial chirp}$$

2. Input pulse spectrum

$$V_0(\bar{\omega}) = A_0 \exp\left[-\frac{\bar{\omega}^2 T_0^2(1-\mathrm{i}C_0)}{4(1+C_0^2)}\right]$$

→ spectral width: $\bar{\omega}_s^2 = \dfrac{4(1+C_0^2)}{T_0^2}$

➔ spectral width of chirped pulse is larger than that of unchirped pulse

$\left(\bar{\omega}_s^2 = 4/T_0^2\right)$ only for transform limited pulses

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

Gaussian beam → q -parameter → similar to Gaussian pulse

Use analogy: however it is limited to homogeneous space

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

Remember beams:

$$\frac{1}{q(z)}=\frac{1}{R(z)}+\mathrm{i}\frac{2}{kw^2(z)}.$$

$$k\rightarrow-\frac{1}{D},\ w^2(z)\rightarrow T^2(z),\ \frac{1}{R(z)}\rightarrow\frac{2DC(z)}{T_0^2}$$

$$\rightarrow\ \frac{1}{q(z)}=\frac{2DC(z)}{T_0^2}-\mathrm{i}\frac{2D}{T^2(z)}$$

$$\boxed{\frac{1}{q(z)}=\frac{2D}{T_0^2}\left[C(z)-\mathrm{i}\frac{T_0^2}{T^2(z)}\right]}\ (*)$$

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

hence at $z=0$:

$$\rightarrow\ \frac{1}{q(0)}=\frac{2D}{T_0^2}\left[C_0-\mathrm{i}\right]\ \text{with}\ C_0=C(0)$$

Idea:

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

b) $\frac{1}{q(z)}=\frac{2D}{T_0^2}\left[C(z)-\mathrm{i}\frac{T_0^2}{T^2(z)}\right].$

set a) and b) equal → $T(z)$, $C(z)$

generally: 2 equations for $C_0, T_0, z, C(z), T(z)$ → 3 values predetermined

here: $z=d$

1) Determination of q parameter at input

$$q(0)=\frac{T_0^2}{2D}\frac{\left(C_0+\mathrm{i}\right)}{\left(1+C_0^2\right)}$$

2) Evolution of q parameter

$$q(d)=q(0)+d=\frac{T_0^2}{2D}\frac{\left(C_0+\mathrm{i}\right)}{\left(1+C_0^2\right)}+d=\frac{1}{2D\left(1+C_0^2\right)}\left[2Dd\left(1+C_0^2\right)+C_0T_0^2+\mathrm{i}T_0^2\right]$$

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

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

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

4) Set two equations equal

$$\frac{\left[2Dd\left(1+C_0^2\right)+C_0T_0^2+\mathbf{i}T_0^2\right]}{2D\left(1+C_0^2\right)}=\frac{T_0^2T^2(d)\left[C(d)T^2(d)+\mathbf{i}T_0^2\right]}{2D\left[C^2(d)T^4(d)+T_0^4\right]}$$

a) real part $\boxed{\frac{\left[2Dd\left(1+C_0^2\right)+C_0T_0^2\right]}{\left(1+C_0^2\right)}=\frac{C(d)T_0^2T^4(d)}{\left[C^2(d)T^4(d)+T_0^4\right]}}$

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

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 $\left[2Dd\left(1+C_0^2\right)+C_0T_0^2\right]=0$

$\rightarrow$ $\boxed{d=-\frac{1}{2}\frac{T_0^2C_0}{D\left(1+C_0^2\right)}=-\frac{1}{2}\operatorname{sgn}(D)\frac{C_0}{\left(1+C_0^2\right)}L_D}$

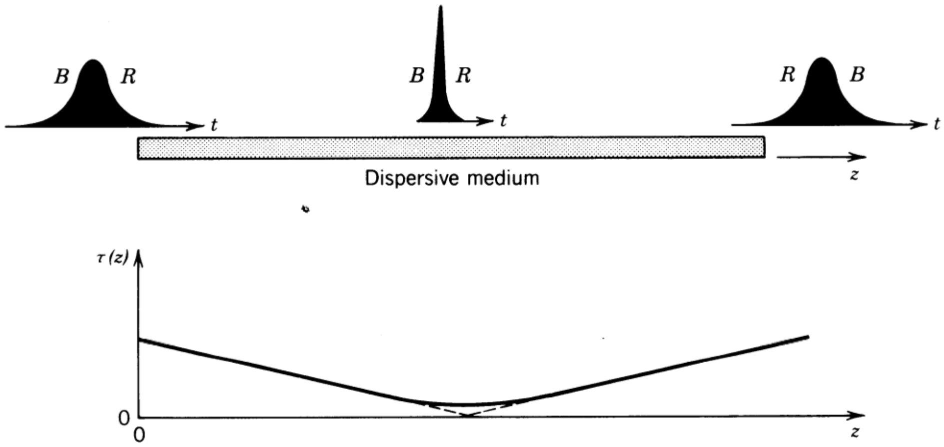
b) $\boxed{T^2(d)=\frac{T_0^2}{\left(1+C_0^2\right)}}$

Resulting properties

1) A pulse can be compressed when the product of initial chirp and dispersion is negative $\rightarrow C_0D<0$.
2) The possible compression increases with initial chirp.

Physical interpretation

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



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

1) First the 'red tail' of the pulse catches up with the 'blue front' until $C(z)=0$ (waist), i.e. the pulse is compressed. 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.

$$C(z)=-\frac{z}{z_0\left(1+\frac{z^2}{z_0^2}\right)} \qquad z_0=-\frac{T_0^2}{2D}$$

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

$$t(x,y) \text{ with } t(x,y)=0 \text{ for } |x|,|y|>a \text{ (aperture)}$$

Here we consider the description based on scalar diffraction theory. Then we can split our diffraction problem into three processes:

i) propagation from light source to aperture
$\rightarrow$ not important, generally plane wave (no diffraction)

ii) multiply field distribution of illuminating wave by transfer function

$$u_+(x,y,z_{\mathrm{A}})=t(x,y)u_-(x,y,z_{\mathrm{A}})$$

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

$$u(x,y,z)=\iint_{-\infty}^{\infty} H\left(\alpha,\beta;z-z_{\mathrm{A}}\right)U_+(\alpha,\beta;z_{\mathrm{A}})\exp\left[\mathbf{i}\left(\alpha x+\beta y\right)\right]d\alpha d\beta$$

or

$$u(x,y,z)=\iint_{-\infty}^{\infty} h\left(x-x',y-y',z-z_{\mathrm{A}}\right)u_+(x',y',z_{\mathrm{A}})dx'dy'$$

$$\text{with } h=\frac{1}{(2\pi)^2}\mathrm{FT}^{-1}\left[H\right]$$

In the following we will use the notation $z_{\mathrm{B}}=z-z_{\mathrm{A}}$. 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_{B} 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

$$H(\alpha,\beta;z_B) = \exp\left(\mathrm{i}\gamma(\alpha,\beta)z_B\right) \text{ where } \gamma^2 = k^2(\omega) - \alpha^2 - \beta^2.$$

Then we have no constraints with respect to spatial frequencies α,β. We get homogeneous and evanescent waves and can treat arbitrary small structures in the aperture by:

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

$$\text{where } U_+(\alpha,\beta) = \mathrm{FT}\left[u_+(x,y)\right]$$

Derivation of the response function

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

$$\boxed{\frac{1}{r}\exp(\mathrm{i}kr) = \frac{\mathrm{i}}{2\pi}\iint_{-\infty}^{\infty}\frac{1}{\gamma}\exp\left[\mathrm{i}(\alpha x + \beta y + \gamma z)\right]d\alpha d\beta}$$

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

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

and therefore

$$\boxed{h(x,y,z) = -\frac{1}{2\pi}\frac{\partial}{\partial z}\left[\frac{1}{r}\exp(\mathrm{i}kr)\right]} \quad \text{with } r = \sqrt{x^2+y^2+z^2}.$$

The resulting expression in position space for the propagation of mono-chromatic beams is also called 'Rayleigh-formula':

$$\boxed{u(x,y,z) = \iint_{-\infty}^{\infty} h(x-x',y-y',z_B)u_+(x',y',z_A)dx'dy'}$$

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

$$H_F(\alpha,\beta;z_B) = \exp(\mathrm{i}kz_B)\exp\left[-\mathrm{i}\frac{z_B}{2k}\left(\alpha^2+\beta^2\right)\right]$$

$$h_F(x,y,z_B) = -\frac{\mathrm{i}}{\lambda z_B}\exp(\mathrm{i}kz_B)\exp\left[\mathrm{i}\frac{k}{2z_B}(x^2+y^2)\right]$$

3.2.3 Paraxial Fraunhofer approximation (far field approximation)

A further simplification of the beam propagation is possible for many diffraction problems. Let us assume a narrow angular spectrum

$$\alpha^2+\beta^2 << k^2$$

and the additional condition for the so-called Fresnel number N_F

$$N_F \lesssim 0.1 \quad \text{with} \quad N_F = \frac{a}{\lambda}\frac{a}{z_B}$$

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 $N_F \lesssim 0.1$.
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:

$$u_F(x,y,z_B) = \iint_{-\infty}^{\infty} h_F(x-x',y-y';z_B)u_+(x',y')dx'dy'$$
$$= -\frac{\mathrm{i}}{\lambda z_B}\exp(\mathrm{i}kz_B)\iint_{-\infty}^{\infty} u_+(x',y')\exp\left\{\mathrm{i}\frac{k}{2z_B}\left[(x-x')^2+(y-y')^2\right]\right\}dx'dy'$$

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

$u_+(x,y)=t(x,y)u_-(x,y)$, and $t(x,y)=0$ for $|x|,|y|>a$ (aperture)

$\rightarrow u_+(x,y)=0$ for $|x|,|y|>a$

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

$$u_F(x,y,z_B) = -\frac{\mathrm{i}}{\lambda z_B}\exp(\mathrm{i}kz_B)\iint_{-a}^{a} u_+(x',y')\exp\left\{\mathrm{i}\frac{k}{2z_B}\left[(x-x')^2+(y-y')^2\right]\right\}dx'dy'$$

Now, let us have a closer look at the integral:

$$\iint_{-a}^{a} u_+(x',y')\exp\left\{\mathbf{i}\frac{k}{2z_B}\left[(x-x')^2+(y-y')^2\right]\right\}dx'dy'$$
$$=\iint_{-a}^{a} u_+(x',y')\exp\left\{\mathbf{i}\frac{k}{2z_B}\left[x^2-2xx'+x'^2+y^2-2yy'+y'^2\right]\right\}dx'dy'$$
$$=\iint_{-a}^{a} u_+(x',y')\exp\left\{\mathbf{i}\frac{k}{2z_B}\left[x^2+y^2\right]\right\}\exp\left\{-\mathbf{i}\frac{kx}{z_B}\left[x'+y'\right]\right\}\exp\left\{\mathbf{i}\frac{k}{2z_B}\left[x'^2+y'^2\right]\right\}dx'dy'$$

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

$$\rightarrow \frac{k}{2z_B}\left[x'^2+y'^2\right]<\frac{ka^2}{z_B}=2\pi N_F$$

$$\rightarrow \text{for } N_F \lesssim 0.1 \rightarrow \exp\left\{\mathbf{i}\frac{k}{2z_B}\left[x'^2+y'^2\right]\right\}\approx 1$$

This means that we can neglect the quadratic phase term in x',y' and we get for the field far from the object, the so-called far field:

$$\boxed{\begin{aligned} u_{FR}(x,y,z_B) &= -\frac{\mathbf{i}}{\lambda z_B}\exp(\mathbf{i}kz_B)\exp\left[\mathbf{i}\frac{k}{2z_B}(x^2+y^2)\right] \\ &\quad\times\iint_{-\infty}^{\infty} u_+(x',y')\exp\left\{-\mathbf{i}\left(\frac{kx}{z_B}x'+\frac{ky}{z_B}y'\right)\right\}dx'dy' \\ &= -\mathbf{i}\frac{(2\pi)^2}{\lambda z_B}\exp(\mathbf{i}kz_B)U_+(k\frac{x}{z_B},k\frac{y}{z_B})\exp\left[\mathbf{i}\frac{k}{2z_B}(x^2+y^2)\right]\end{aligned}}$$

This is the far-field in paraxial Fraunhofer approximation. Surprisingly, the intensity distribution of the far field in position space is just given by the Fourier transform of the field distribution at the aperture

$$I_{FR}(x,y,z_B)\sim\frac{1}{(\lambda z_B)^2}\left|U_+(k\frac{x}{z_B},k\frac{y}{z_B};z_A)\right|^2$$

Interpretation

For a plane in the far field at $z=z_B$ in each point x,y only one angular frequency $(\alpha=kx/z_B;\beta=ky/z_B)$ with spectral amplitude $U_+(kx/z_B,ky/z_B)$ contributes to the field distribution. This is in contrast to the previously considered cases, where all angular frequencies contributed to the response in a single position point.

In summary, we have shown that in (paraxial) Fraunhofer approximation the propagated field, or diffraction pattern, is very simple to calculate. We just need to Fourier transform the field at the aperture. In order to apply this approximation we have to check that:

A) $\alpha^2+\beta^2 << k^2$ ➔ smallest features $\Delta x, \Delta y >> \lambda$ ➔ narrow angular spectrum (paraxiality)

B) $N_F = \frac{a^2}{\lambda z_B} << 1$ ➔ largest feature a determines $z_B >> \frac{a^2}{\lambda}$ ➔ far field

Example: $\Delta x, \Delta y = 10\lambda, \quad a = 100\lambda, \ \lambda = 1\mu$m $\quad \rightarrow \quad z_B >> 10^4\lambda \approx 1$ 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 $\alpha^2+\beta^2 \le k^2$ contribute to the far field distribution.

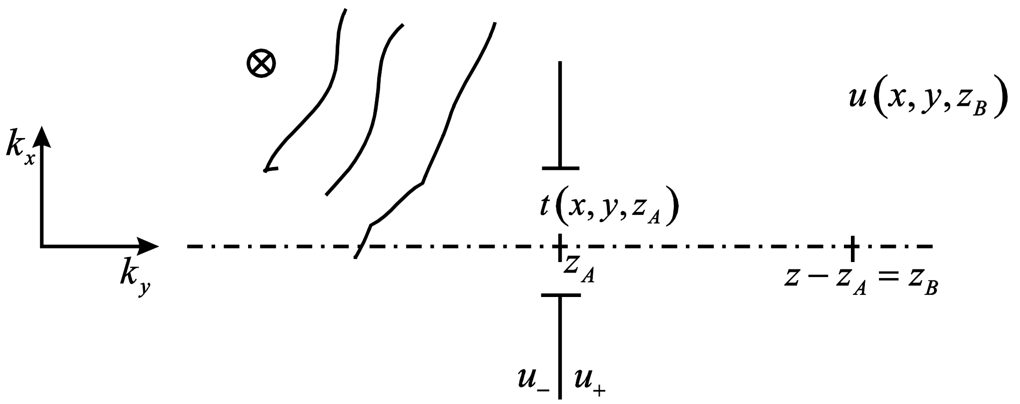
Evanescent waves decay for $\quad kz_B >> 1 \rightarrow z_B >> \lambda$.

$$N_F \lesssim 0.1 \qquad \text{with } N_F = \frac{a}{\lambda}\frac{a}{z_B} = \frac{1}{\pi}\frac{z_0}{z_B}$$

$$u_{\text{FR non-paraxial}}(x,y,z_B) = -\mathbf{i}\frac{(2\pi)^2}{\lambda\sqrt{x^2+y^2+z_B{}^2}}\frac{z_B}{\sqrt{x^2+y^2+z_B{}^2}}$$
$$\times U_+\left(\frac{kx}{\sqrt{x^2+y^2+z_B{}^2}},\frac{ky}{\sqrt{x^2+y^2+z_B{}^2}};z_A\right)\exp\left(\mathbf{i}k\sqrt{x^2+y^2+z_B{}^2}\right)$$

3.3 *Fraunhofer diffraction at plane masks (paraxial)*

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



Let us consider an incident plane wave, with a wave vector perpendicular $(k_x = k_y = 0)$ to the mask t

$$u_-(x,y,z_A) = A\exp\left(\mathbf{i}k_z z_A\right)$$

or, more general, inclined with a certain angle

$$u_-(x,y,z_A) = A\exp\left[\mathbf{i}\left(k_x x + k_y y + k_z z_A\right)\right]$$

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

$$u_+(x,y,z_A)=u_-(x,y,z_A)t(x,y)=A\exp\left[\mathbf{i}\left(k_x x+k_y y+k_z z_A\right)\right]t(x,y)$$

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

$$\boxed{I(x,y,z_\mathrm{B})\sim\left|u(x,y,z_\mathrm{B})\right|^2\sim\frac{1}{\left(\lambda z_\mathrm{B}\right)^2}\left|U_+(k\frac{x}{z_\mathrm{B}},k\frac{y}{z_\mathrm{B}})\right|^2}$$

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

$$\alpha=k\frac{x}{z_\mathrm{B}},\beta=k\frac{y}{z_\mathrm{B}}.$$

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

$$\begin{aligned}
&U_+(k\frac{x}{z_\mathrm{B}},k\frac{y}{z_\mathrm{B}})\\
&=\frac{A}{\left(2\pi\right)^2}\exp\left(\mathbf{i}k_z z_A\right)\iint_{-\infty}^{\infty}t(x',y')\exp\left[-\mathbf{i}\left(k\frac{x}{z_\mathrm{B}}-k_x\right)x'-\mathbf{i}\left(k\frac{y}{z_\mathrm{B}}-k_y\right)y'\right]dx'dy'\\
&=A\exp\left(\mathbf{i}k_z z_A\right)T\left(k\frac{x}{z_\mathrm{B}}-k_x,k\frac{y}{z_\mathrm{B}}-k_y\right)
\end{aligned}$$

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

$$\boxed{I(x,y,z_\mathrm{B})\sim\frac{1}{\left(\lambda z_\mathrm{B}\right)^2}\left|T\left(k\frac{x}{z_\mathrm{B}}-k_x,k\frac{y}{z_\mathrm{B}}-k_y\right)\right|^2}$$

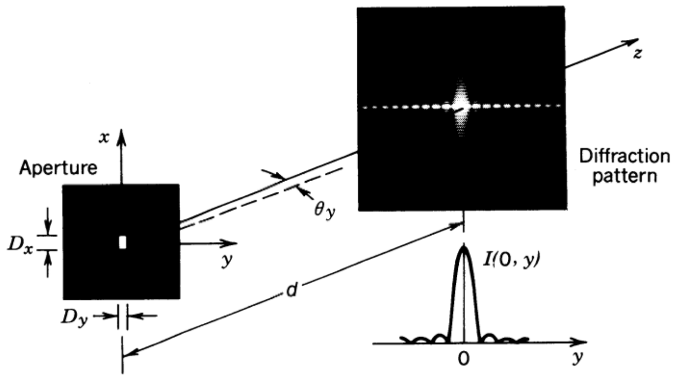
This is the absolute square of the Fourier transform of the aperture function. In paraxial approximation an inclination of the illuminating plane wave just shifts the pattern transversely.

Examples

A) Rectangular aperture illuminated by normal plane wave

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

$$I(x,y,z_\mathrm{B})\sim\mathrm{sinc}^2\left(ka\frac{x}{z_\mathrm{B}}\right)\mathrm{sinc}^2\left(kb\frac{y}{z_\mathrm{B}}\right)$$

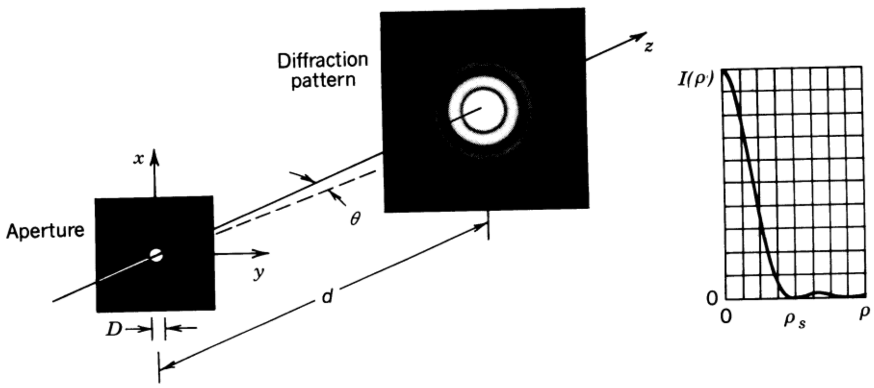


Fraunhofer diffraction pattern from a rectangular aperture. The central lobe of the pattern has half-angular widths $\theta_x = \lambda / D_x$ and $\theta_y = \lambda / D_y$.

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

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

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



The Fraunhofer diffraction pattern from a circular aperture produces the Airy pattern with the radius of the central disk subtending an angle $\theta \approx 1.22\lambda / D$.

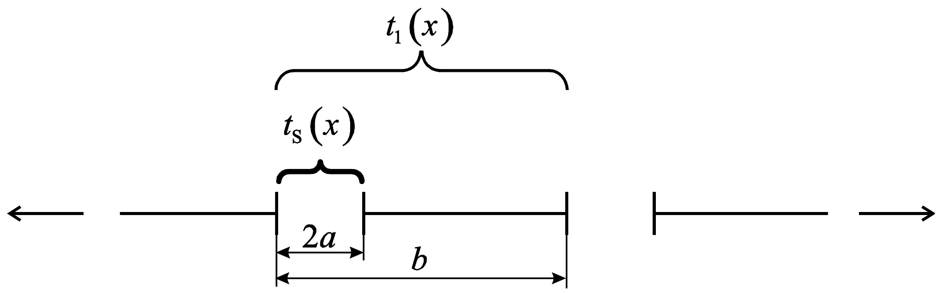
The first zero of the Bessel function (size of Airy disk):

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

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

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 $2a$:



Then, we can express the mask function t as:

$$t(x)=\sum_{n=0}^{N-1} t_1(x-nb) \text{ with } t_1(x)=\begin{cases} t_s(x) & \text{for } |x|\leqslant a \\ 0 & \text{elsewhere}\end{cases}$$

The Fourier transform of the mask is given as

$$T\left(k\frac{x}{z_{\mathrm{B}}}\right)\sim\sum_{n=0}^{N-1}\int_{-\infty}^{\infty} t_1(x'-nb)\exp\left(-\mathrm{i}k\frac{x}{z_{\mathrm{B}}}x'\right)dx'$$

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

$$T\left(k\frac{x}{z_{\mathrm{B}}}\right)\sim\sum_{n=0}^{N-1}\int_{-a}^{a} t_{\mathrm{S}}(X')\exp\left(-\mathrm{i}k\frac{x}{z_{\mathrm{B}}}X'\right)\exp\left(-\mathrm{i}k\frac{x}{z_{\mathrm{B}}}nb\right)dX'$$

$$T\left(k\frac{x}{z_{\mathrm{B}}}\right)\sim T_{\mathrm{S}}\left(k\frac{x}{z_{\mathrm{B}}}\right)\sum_{n=0}^{N-1}\exp\left(-\mathrm{i}k\frac{x}{z_{\mathrm{B}}}nb\right)$$

We see that the Fourier transform T_{S} of the elementary slit t_{S} appears. The second factor has its origin in the periodic arrangement. With some math we can identify this second expression as a geometrical series and perform the summation by using the following formula for the definition of the sinc-funtion:

$$\left|\sum_{n=0}^{N-1}\exp(-\mathrm{i}\delta n)\right|=\left|\frac{\sin\left(N\frac{\delta}{2}\right)}{\sin\left(\frac{\delta}{2}\right)}\right|$$

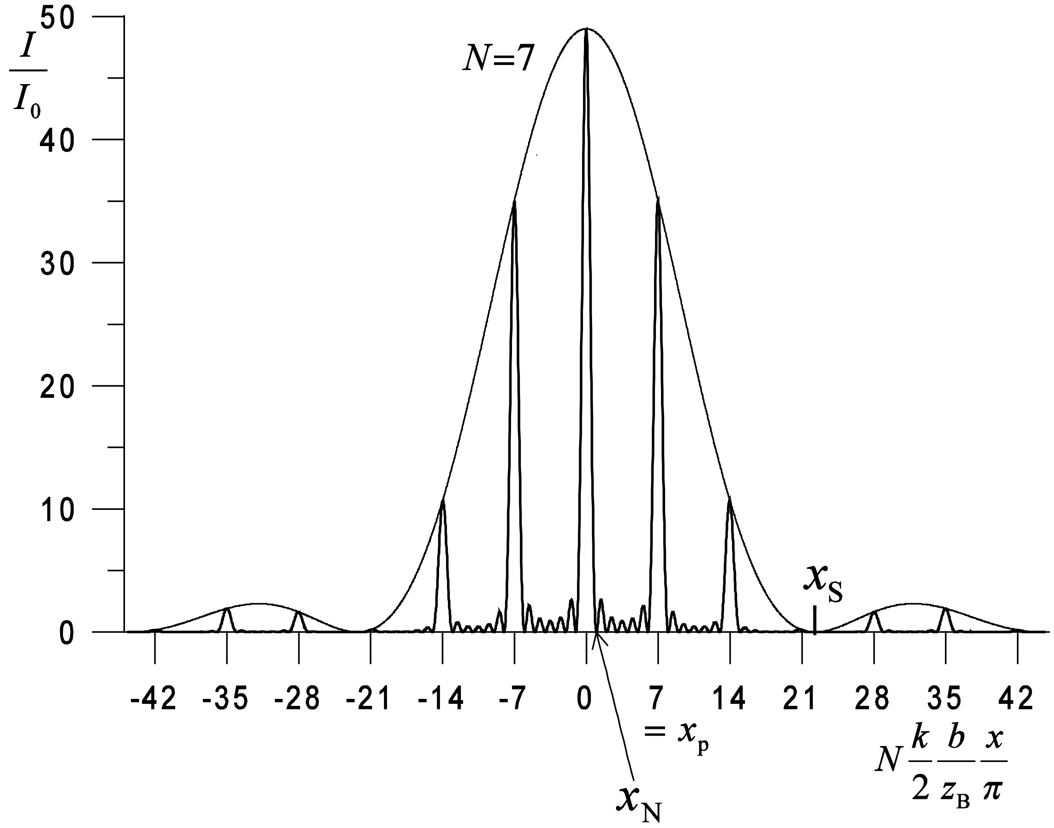
Thus we finally write:

$$\boxed{T\left(k\frac{x}{z_{\mathrm{B}}}\right)\sim T_{\mathrm{S}}\left(k\frac{x}{z_{\mathrm{B}}}\right)\frac{\sin\left(N\frac{k}{2}\frac{x}{z_{\mathrm{B}}}b\right)}{\sin\left(\frac{k}{2}\frac{x}{z_{\mathrm{B}}}b\right)}}$$

For the particular case of a simple grating of slit apertures with $t_s(x)=1$ we have

$$T_1\left(k\frac{x}{z_B}\right)=\mathrm{sinc}\left(k\frac{x}{z_B}a\right)$$

and therefore $$\boxed{I\sim\mathrm{sinc}^2\left(k\frac{x}{z_B}a\right)\frac{\sin^2\left(N\frac{k}{2}\frac{x}{z_B}b\right)}{\sin^2\left(\frac{k}{2}\frac{x}{z_B}b\right)}}$$



We find three important parameters for the diffraction pattern of a grating:

– **Global width of diffraction pattern** → first zero of slit function T_s

$$k\frac{x_s}{z_B}a=\pi \;\Rightarrow\; \boxed{x_s=\frac{\lambda z_B}{2a}}$$

The width of the total far-field diffraction pattern x_S (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

$$\max\left(\frac{\sin^2\left(N\frac{k}{2}\frac{x}{z_B}b\right)}{\sin^2\left(\frac{k}{2}\frac{x}{z_B}b\right)}\right) \rightarrow \frac{k}{2}\frac{x_P}{z_B}b=n\pi \rightarrow \boxed{x_P=n\frac{\lambda z_B}{b}}$$

These are the so-called diffraction orders, which are determined exclusively by the grating period.

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

$$N\frac{k}{2}\frac{x_{\mathbf{N}}}{z_{\mathbf{B}}}b=\pi \rightarrow \boxed{x_{\mathbf{N}}=\frac{\lambda z_{\mathbf{B}}}{Nb}}$$

The width of a maximum in the far-field diffraction patter 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).

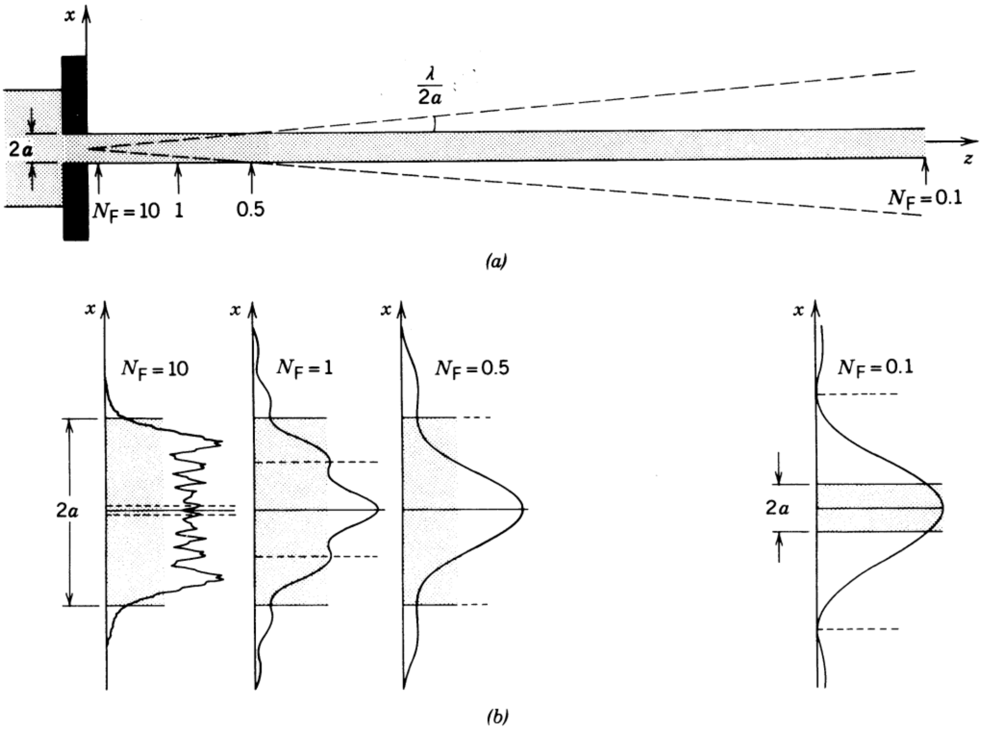
These observations are consistent with the general property of the Fourier-transform: small scales in position space give rise to a broad angular spectrum and vice versa.

3.4 Remarks on Fresnel diffraction

Fresnel number $N_{\mathbf{F}}=\frac{a}{\lambda}\frac{a}{z_{\mathbf{B}}}$

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

- $N_{\mathrm{F}}\gtrsim 10$ (a large, $\lambda z_{\mathbf{B}}$ small, $z_{\mathbf{B}}<1/30\ z_0$) → shadow
- $N_{\mathrm{F}}\lesssim 0.1$ ($z_{\mathrm{B}}>3z_0$) → Fraunhofer → FT of aperture
- $10\gtrsim N_{\mathbf{F}}\gtrsim 0.1$ ($1/30\ z_0<z_{\mathbf{B}}<3z_0$) → 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 $N_{\mathbf{F}}=10,1,0.5$ and 0.1. The dashed area represents the geometrical shadow of the slit. The dashed lines at $|x|=(\lambda/D)d$ 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* $N_{\mathrm{F}} = a^2 / \lambda d = 0.5$.

4. Fourier optics - optical filtering

From previous chapters we know how to propagate the optical field through homogeneous space, and we also know the transfer function of a thin lens. Thus, we have all tools at hand to describe optical imaging. 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 → calculated in Fourier domain
- interaction with lens or filter → 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

$$\underbrace{-\frac{\mathrm{i}}{\lambda f}\exp\left(\mathrm{i}kf\right)\exp\left[\mathrm{i}\frac{k}{2f}\left(x^2+y^2\right)\right]}_{\text{spherical wave}}t_{\mathrm{L}}\left(x,y\right)=\underbrace{-\frac{\mathrm{i}}{\lambda f}\exp\left(\mathrm{i}kf\right)}_{\text{plane wave}}$$

And therefore the response function for a thin lens is given as (see chapter 2.8.3):

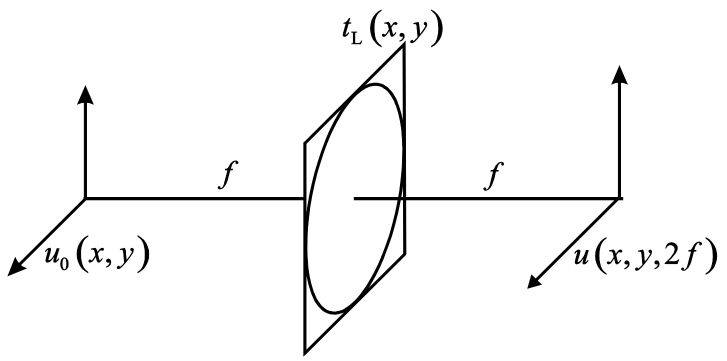
$$\boxed{t_{\mathbf{L}}(x,y)=\exp\left[-\mathbf{i}\frac{k}{2f}\left(x^2+y^2\right)\right]}$$

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

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

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

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

B) Propagation from object to lens
(lens positioned at distance f)

$$U_-(\alpha,\beta;f)=H_{\mathbf{F}}(\alpha,\beta;f)U_0(\alpha,\beta)$$

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

C) Interaction with lens
(multiplication in position space or convolution in Fourier domain)

$$u_+(x,y,f)=t_{\mathbf{L}}(x,y)u_-(x,y,f)$$

$$U_+(\alpha,\beta;f)=T_L(\alpha,\beta)*U_-(\alpha,\beta;f)$$
$$=-\mathbf{i}\frac{\lambda f}{(2\pi)^2}\exp(\mathbf{i}kf)\iint_{-\infty}^{\infty}\exp\left\{\mathbf{i}\frac{f}{2k}\left[(\alpha-\alpha')^2+(\beta-\beta')^2\right]\right\}\cdot$$
$$\cdot\exp\left[-\frac{\mathbf{i}}{2k}\left(\alpha'^2+\beta'^2\right)f\right]U_0(\alpha',\beta')d\alpha' d\beta'$$

D) Propagation from lens to image plane

$$U(\alpha,\beta;2f)=H_F(\alpha,\beta;f)U_+(\alpha,\beta;f)$$

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

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

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

$$=-\mathbf{i}\frac{\lambda f}{(2\pi)^2}\exp(2\mathbf{i}kf)u_0\left(-\frac{f}{k}\alpha,-\frac{f}{k}\beta\right)$$

We see that the spectrum in the image plane is given by the optical field in the object plane.

E) Fourier back transform in image plane

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

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

With the coordinate transformation

$$x' = -\frac{f}{k}\alpha, \qquad y' = -\frac{f}{k}\beta \qquad \rightarrow \qquad d\alpha = -\frac{2\pi}{\lambda f}dx', \qquad d\beta = -\frac{2\pi}{\lambda f}dy'$$

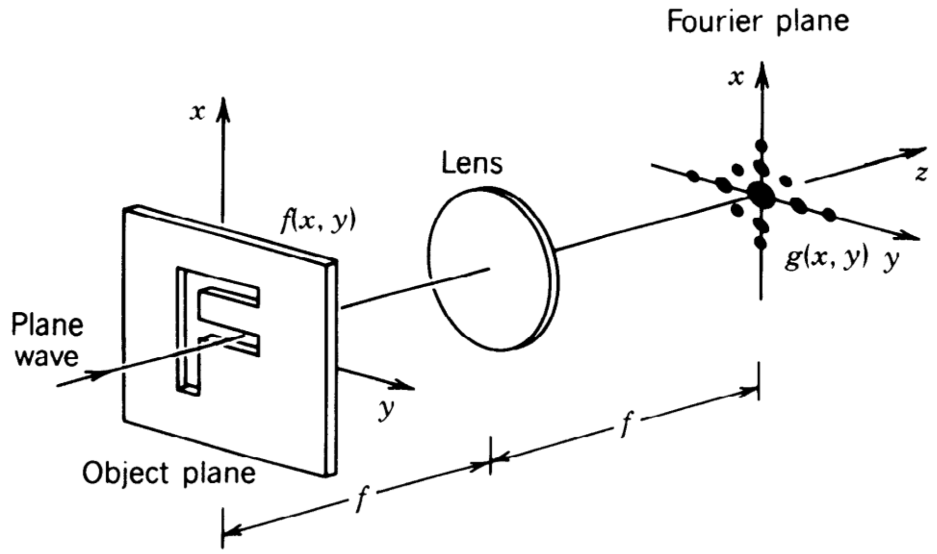
we get:

$$\rightarrow u(x,y,2f) = -\mathrm{i}\frac{1}{\lambda f}\exp(2\mathrm{i}kf)\int\int_{-\infty}^{\infty} u_0(x',y')\exp\left[-\mathrm{i}\frac{k}{f}(xx'+yy')\right]dx'dy'$$

$$\boxed{u(x,y,2f) = -\mathrm{i}\frac{(2\pi)^2}{\lambda f}\exp(2\mathrm{i}kf)U_0\left(\frac{k}{f}x,\frac{k}{f}y\right)}$$

The image in the second focal plane is the Fourier transform of the optical field in the object plane. This is simmilar to the far field in Fraunhofer approximation, but for $z_B \leftrightarrow f$. This finding allows us to perform an optical Fourier transform over shorter distances.

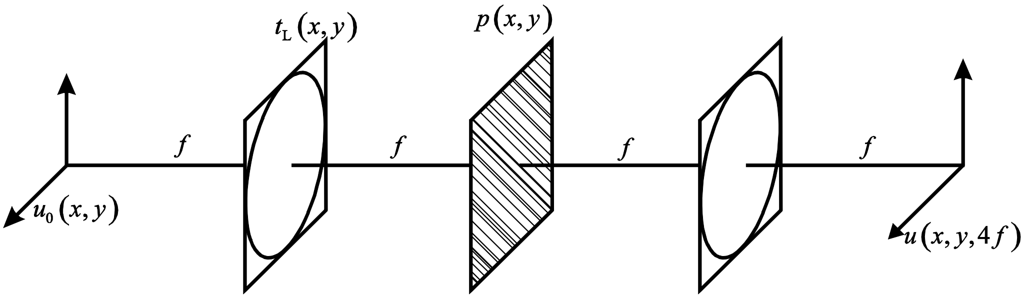
And in the Fourier plane it is possible to manipulate the spectrum.



4.2 Optical filtering and image processing

4.2.1 The 4f-setup

For image manipulation (filtering) it would be advantageous if we could perform a Fourier back-transform by means of an optical imaging setup as well. It turns out that this leads to the so-called 4f-setup:



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 the Fourier plane is the FT of the optical field in the object plane.

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

We have to compute the imaging with the second lens after the manipulation of the spectrum in the Fourier plane.

Our final goal is the transmission function $H_{\mathbf{A}}(\alpha,\beta;4f)$ of the complete imaging system:

$$\boxed{u(-x,-y,4f) = \int\int_{-\infty}^{\infty} H_{\mathbf{A}}(\alpha,\beta;4f)U_0(\alpha,\beta)\exp\left[\mathbf{i}(\alpha x+\beta y)\right]d\alpha d\beta}$$

Note: We will see in the following calculation that the second lens does a Fourier transform $\sim \exp\left[-\mathbf{i}(\alpha x+\beta y)\right]$. In order to obtain a proper back transform we have to pass to mirrored coordinates $x \to -x,\ \ y \to -y$.

The transmission mask $p(x,y)$ contains all constrains of the system (e.g. a limited aperture) and optical filtering (which we can design).

A) Field behind transmission mask

$$u_+(x,y,2f) = u(x,y,2f)p(x,y) \sim AU_0\left(\frac{k}{f}x, \frac{k}{f}y\right)p(x,y)$$

B) Second lens → Fourier back-transform of field distribution

$$u(x,y,4f) = -\mathbf{i}\frac{(2\pi)^2}{\lambda f}\exp(2\mathbf{i}kf)U_+\left(\frac{k}{f}x, \frac{k}{f}y; 2f\right)$$

Now we can make the link to the initial spectrum in the object plane U_0:

$$u(x,y,4f)\sim\iint_{-\infty}^{\infty}u_{+}(x',y',2f)\exp\left[-\mathbf{i}\frac{k}{f}(xx'+yy')\right]dx'dy'$$
$$\sim\iint_{-\infty}^{\infty}U_0\left(\frac{k}{f}x',\frac{k}{f}y'\right)p(x',y')\exp\left[-\mathbf{i}\frac{k}{f}(xx'+yy')\right]dx'dy'$$

Here we do not care about the amplitudes and just write ~:
To get the anticipated form we need to perform a coordinate transformation:

$$\alpha=\frac{k}{f}x',\beta=\frac{k}{f}y'$$

Then we can write:

$$u(x,y,4f)\sim\iint_{-\infty}^{\infty}U_0(\alpha,\beta)\,p(\frac{f}{k}\alpha,\frac{f}{k}\beta)\exp\left[-\mathbf{i}(\alpha x+\beta y)\right]d\alpha d\beta$$

By passing to mirrored coordinates $x\to -x,\;\; y\to -y$ we get

$$u(-x,-y,4f)\sim\iint_{-\infty}^{\infty}p(\frac{f}{k}\alpha,\frac{f}{k}\beta)U_0(\alpha,\beta)\exp\left[\mathbf{i}(\alpha x+\beta y)\right]d\alpha d\beta$$

Hence we can identify the transmission function of the system

$$\boxed{H_{\mathbf{A}}(\alpha,\beta;4f)\sim p(\frac{f}{k}\alpha,\frac{f}{k}\beta)}$$

Summary

- Fourier amplitudes get multiplied by transmission mask
- transmission mask → transfer function
- coordinates of image → mirrored coordinates of object

In position space we can formulate propagation through a 4f-system by using the response function $h_{\mathbf{A}}(x,y)$

$$\boxed{u(-x,-y,4f)=\iint_{-\infty}^{\infty}h_{\mathbf{A}}(x-x',y-y')u_0(x',y')dx'dy'}$$

As usual, the response function is given as the Fourier transform of the transfer function:

$$h_{\mathbf{A}}(x,y)=\frac{1}{(2\pi)^2}\iint_{-\infty}^{\infty}H_{\mathbf{A}}(\alpha,\beta;4f)\exp\left\{\mathbf{i}\left[\alpha x+\beta y\right]\right\}d\alpha d\beta$$

From above we have $H_{\mathbf{A}}(\alpha,\beta;4f)\sim p(\frac{f}{k}\alpha,\frac{f}{k}\beta)$

$$h_{\mathbf{A}}(x,y)\sim\iint_{-\infty}^{\infty}p(\frac{f}{k}\alpha,\frac{f}{k}\beta)\exp\left\{\mathbf{i}\left[\alpha x+\beta y\right]\right\}d\alpha d\beta$$

We introduce the coordinate transform $\left(\bar{x}=\frac{f}{k}\alpha,\quad \bar{y}=\frac{f}{k}\beta\right)$

$$h_A(x,y)\sim\int\int_{-\infty}^{\infty}p(\bar{x},\bar{y})\exp\left\{\mathrm{i}\frac{k}{f}\left[\bar{x}x+\bar{y}y\right]\right\}d\bar{x}d\bar{y}\sim P\left[-\frac{k}{f}x,-\frac{k}{f}y\right]$$

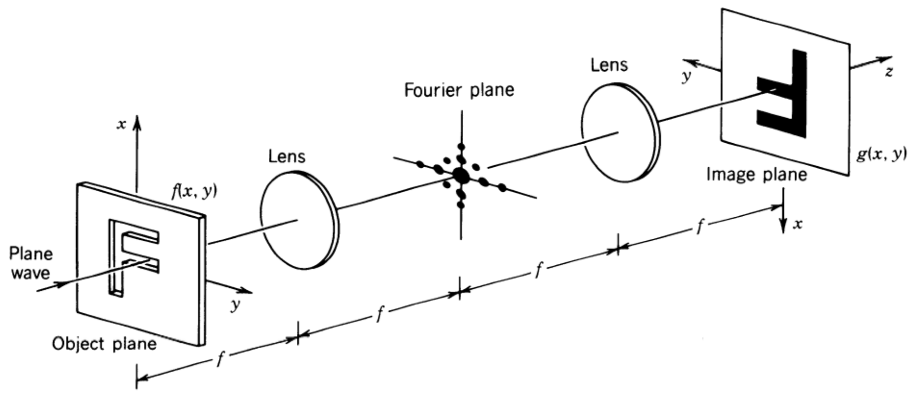
$$\boxed{u(-x,-y,4f)\sim\int\int_{-\infty}^{\infty}P\left[\frac{k}{f}(x'-x),\frac{k}{f}(y'-y)\right]u_0(x',y')dx'dy'}$$

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

4.2.2 Examples of aperture functions

Example 1: The ideal image (infinite aperture)

Be careful, we use paraxial approximation → limited angular spectrum

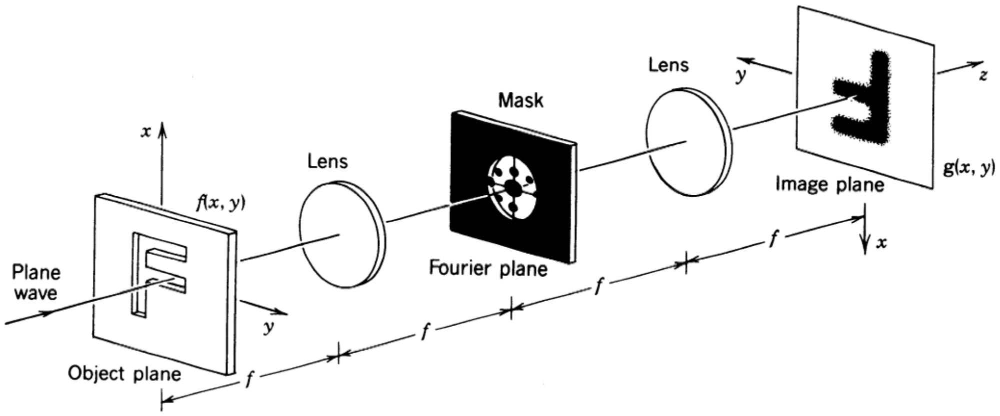


The 4-f system performs a Fourier transform followed by an inverse Fourier transform, so that the image is a perfect replica of the object (perfect only within the Fresnel approximation).

$p=1\rightarrow P\sim\delta(x-x')\delta(y-y')\rightarrow\boxed{u(-x,-y,4f)\sim u_0(x,y)}$ → mirrored original

Example 2: Finite aperture

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



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 $H(\nu_x,\nu_y)=p(\lambda f\nu_x,\lambda f\nu_y)$.

Transmission function:

$$\boxed{H_A(\alpha,\beta;4f)\sim\begin{cases}1 & \text{for } \left(\frac{f}{k}\alpha\right)^2+\left(\frac{f}{k}\beta\right)^2\le(D/2)^2\\ 0 & \text{elsewhere}\end{cases}}$$

– 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.

$$\left(\frac{f}{k}\alpha\right)^2+\left(\frac{f}{k}\beta\right)^2\le(D/2)^2\rightarrow(\alpha)^2+(\beta)^2\le\left(\frac{k}{f}D/2\right)^2$$

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

$$\rho_{\text{max}}^2=\frac{k^2}{f^2}\left(\frac{D}{2}\right)^2 \rightarrow \boxed{\rho_{\text{max}}=\frac{2\pi n}{\lambda f}\frac{D}{2}}$$

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

$$\Delta r_{\min}\approx\frac{2\pi}{\rho_{\text{max}}}=\frac{2\lambda f}{nD}$$

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

$$u(-x,-y,4f) \sim \iint_{-\infty}^{\infty} P\left[\frac{k}{f}(x'-x),\frac{k}{f}(y'-y)\right]u_0(x',y')dx'dy'$$

$$\sim \iint_{-\infty}^{\infty} \frac{\mathrm{J}_1\left[\frac{kD}{2f}\sqrt{(x'-x)^2+(y'-y)^2}\right]}{\frac{kD}{2f}\sqrt{(x'-x)^2+(y'-y)^2}}u_0(x',y')dx'dy,'$$

One point of the object x_0, y_0 gives an Airy disk (pixel) in the image:

$$\sim \left[\frac{\mathrm{J}_1\left(\frac{kD}{2f}\sqrt{(x_0-x)^2+(y_0-y)^2}\right)}{\frac{kD}{2f}\sqrt{(x_0-x)^2+(y_0-y)^2}}\right]^2$$

We can define the limit of optical resolution:

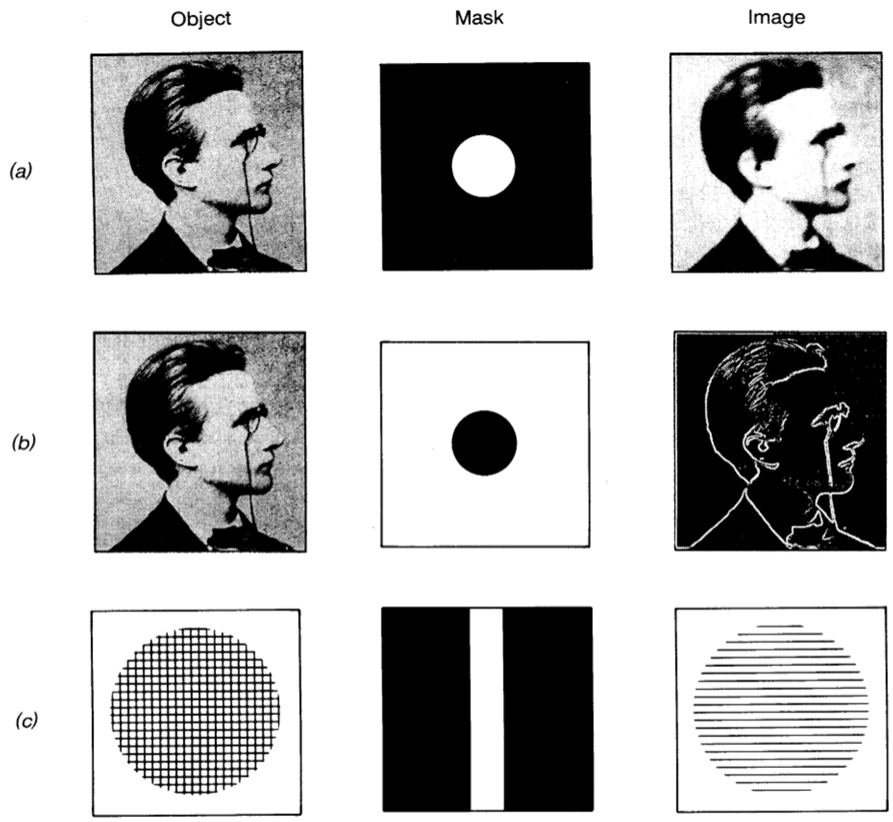
Two objects in the object plane can be independently resolved in the image plane as long as the intensity maximum of one of the objects is not closer to the other object than its first intensity minimum:

$$\frac{kD}{2f}\Delta r_{\min} = 1.22\pi$$

Hence we find:

$$\boxed{\Delta r_{\min} = \frac{1.22\lambda f}{nD}}$$

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

If we assume propagation in z direction (k-vector points in z-direction), $\mathbf{div}\mathbf{E}(\mathbf{r},t)=0$ implies that we can have two nonzero transversal field components → $x-$ and $y-$components $E_\mathbf{x},E_\mathbf{y}$

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

$$\mathbf{k}=\begin{pmatrix}0\\0\\k\end{pmatrix}\qquad \rightarrow \text{ propagation in } z \text{ direction}$$

The evolution of the real electric field vector is given as

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

Because the field is transversal we have two free complex field components

$$\mathbf{E}=\begin{pmatrix}E_\mathbf{x}\exp\left(\mathbf{i}\varphi_\mathbf{x}\right)\\E_\mathbf{y}\exp\left(\mathbf{i}\varphi_\mathbf{y}\right)\\0\end{pmatrix}\qquad \text{with } E_{\mathbf{x},\mathbf{y}} \text{ and } \varphi_{\mathbf{x},\mathbf{y}} \text{ being real}$$

Then the real electric field vector is given as

$$\mathbf{E}_\mathbf{r}(\mathbf{r},t)=\begin{pmatrix}E_\mathbf{x}\cos\left(\omega t-kz-\varphi_\mathbf{x}\right)\\E_\mathbf{y}\cos\left(\omega t-kz-\varphi_\mathbf{y}\right)\\0\end{pmatrix}$$

Here, only the relative phase is interesting → $\delta=\varphi_\mathbf{y}-\varphi_\mathbf{x}$

Conclusion:

Normal modes in isotropic, dispersive media are in general elliptically polarized. The field amplitudes $E_{\mathbf{x}}, E_{\mathbf{y}}$ and the phase difference $\delta = \left(\varphi_{\mathbf{y}} - \varphi_{\mathbf{x}}\right)$ are free parameters

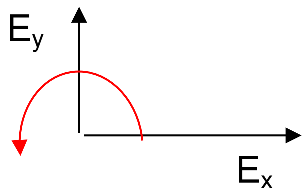
5.3 Polarization states

Let us have a look at different possible parameter settings:

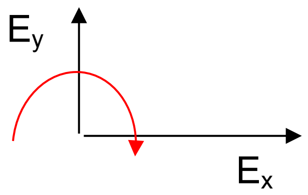
A) linear polarization → $\delta = n\pi$ (or $E_{\mathbf{x}} = 0$ or $E_y = 0$)

B) circular polarization → $E_{\mathbf{x}} = E_{\mathbf{y}} = E, \delta = \pm\pi/2$

$\delta = +\pi/2$ → counterclockwise rotation



$\delta = -\pi/2$ → clockwise rotation



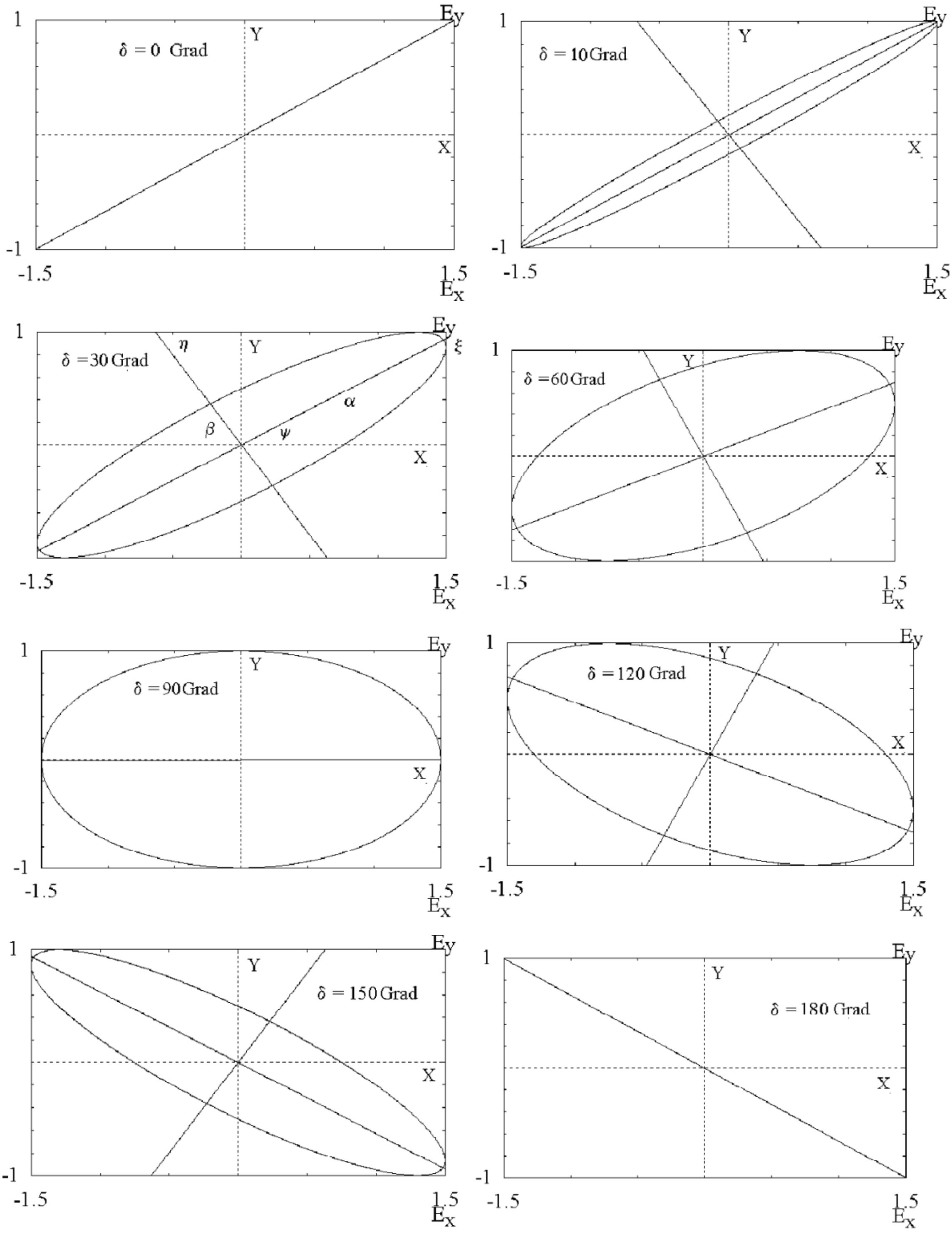
These pictures are for an observer looking contrary to the propagation direction.

C) elliptic polarization → $E_{\mathbf{x}} \neq E_{\mathbf{y}} \neq 0, \delta \neq n\pi$

$0 < \delta < \pi$ → counterclockwise

$\pi < \delta < 2\pi$ → clockwise

Examples



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 $\delta=\pm\pi/2$.

$$E\begin{pmatrix}\cos(\omega t)\\ \sin(\omega t)\\ 0\end{pmatrix}+E\begin{pmatrix}\cos(\omega t)\\ -\sin(\omega t)\\ 0\end{pmatrix}=2E\begin{pmatrix}\cos(\omega t)\\ 0\\ 0\end{pmatrix}$$

6. Principles of optics in crystals

In this chapter we will treat light propagation in anisotropic media (the worst case). Like in the isotropic case before we will seek for the normal modes, and in order to keep things simple we assume homogeneous media.

6.1 Susceptibility and dielectric tensor

before: isotropy (optical properties independent of direction)

now: anisotropy (optical properties depend on direction)

The common reason for anisotropy in many optical media (in particular crystals) is that the polarization $\mathbf{P}$ depends on the direction of the 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 ➔ electro-optical material
- Quartz ➔ polarizer
- liquid crystals ➔ 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 → ???

Before (isotropic)

$$\bar{\mathbf{P}}(\mathbf{r},\omega) = \varepsilon_0 \chi(\omega)\ \bar{\mathbf{E}}(\mathbf{r},\omega)$$

$$\bar{\mathbf{D}}(\mathbf{r},\omega) = \varepsilon_0 \varepsilon(\omega)\ \bar{\mathbf{E}}(\mathbf{r},\omega)$$

In the following we will write $\bar{\mathbf{E}} \rightarrow \mathbf{E}$, because we assume monochromatic light and the frequency ω is just a parameter.

Now (anisotropic)

$$P_i(\mathbf{r},\omega) = \varepsilon_0 \sum_{j=1}^{3} \underbrace{\chi_{ij}(\omega)}_{\text{tensor components}} E_j(\mathbf{r},\omega)$$

The linear susceptibility tensor has 3x3=9 tensor components. Direct consequences of this relation between polarization P and electric field E are:

- $\mathbf{P} \nparallel \mathbf{E}$: the polarization is not necessarily parallel to the electric field
- The tensor elements χ_{ij} depend on the structure of crystal. However, we do not need to know the microscopic structure because of the different length scales involved (optics: $5\cdot10^{-7}$ m; crystal: $5\cdot10^{-10}$ m),

but the field is influenced by the symmetries of the crystal (see next section).

- In complete analogy we find for the D field:

$$\boxed{\begin{aligned} &D_i(\mathbf{r},\omega) = \varepsilon_0 \sum_{j=1}^{3} \varepsilon_{ij}(\omega) E_j(\mathbf{r},\omega) \\ &\mathbf{D}(\mathbf{r},\omega) = \varepsilon_0 \hat{\varepsilon}(\boldsymbol{\omega}) \mathbf{E}(\mathbf{r},\omega) \end{aligned}}$$

As for the polarization we find:

- $\mathbf{D} \nparallel \mathbf{E}$

We introduce the following notation:

- $\hat{\boldsymbol{\chi}} = (\chi_{ij})$ → susceptibility tensor
- $\hat{\boldsymbol{\varepsilon}} = (\varepsilon_{ij})$ → dielectric tensor
- $\hat{\boldsymbol{\sigma}} = (\hat{\boldsymbol{\varepsilon}})^{-1} = (\sigma_{ij})$ → inverse dielectric tensor $\sum_{j=1}^{3} \sigma_{ij}(\omega) D_j(\mathbf{r},\omega) = \varepsilon_0 E_i(\mathbf{r},\omega)$

The following properties of the dielectric and inverse dielectric tensor are important:

- $\sigma_{ij}, \varepsilon_{ij}$ are real in the transparent region (omit ω), we have no losses (see our assumptions above)
- The tensors are symmetric (hermitian), only 6 components are independent $\varepsilon_{ij} = \varepsilon_{ji}, \sigma_{ij} = \sigma_{ji}$.
- 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 σ_{ij}, it means that we are looking for directions where $\mathbf{D} \parallel \mathbf{E}$, i.e., our principal axes:

$$\boxed{\varepsilon_0 E_i = \sum_{j=1}^{3} \sigma_{ij} D_j \doteq \lambda D_i}$$

This is a so-called eigenvalue problem, with eigenvalues λ. If we want to solve for the eigenvalues we get

$$\det\left[\sigma_{ij} - \lambda I_{ij}\right] = 0, \quad \text{with } I_{ij} = \delta_{ij}$$

This leads to a third order equation in λ, hence we expect three solutions (roots) $\lambda^{(\alpha)}$. The corresponding eigenvectors can be computed from

$$\sum_{j=1}^{3} \sigma_{ij} D_j^{(\alpha)} = \lambda^{(\alpha)} D_i^{(\alpha)}.$$

The eigenvectors are orthogonal:

$$D_i^{(\beta)} D_i^{(\alpha)} = 0 \quad \text{for } \lambda^{(\alpha)} \neq \lambda^{(\beta)}$$

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:

$$\varepsilon_{ij} = \varepsilon_i \delta_{ij}, \qquad \sigma_{ij} = \sigma_i \delta_{ij} = \frac{1}{\varepsilon_i}\delta_{ij}$$

$$\boxed{\left(\varepsilon_{ij}\right) = \begin{bmatrix} \varepsilon_1(\omega) & 0 & 0 \\ 0 & \varepsilon_2(\omega) & 0 \\ 0 & 0 & \varepsilon_3(\omega) \end{bmatrix}}$$

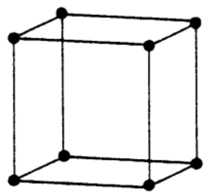
The above reasoning shows that anisotropic media are characterized in general by three independent dielectric functions (in the principal coordinate system).
It is easier to do all calculations in the principal coordinate system (coordinate system of the crystal) and back-transform the final results 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....)

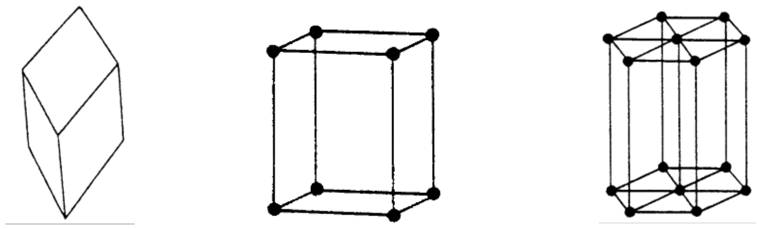


$$\boxed{\varepsilon_1(\omega) = \varepsilon_2(\omega) = \varepsilon_3(\omega)} \Rightarrow D_i = \varepsilon_0 \varepsilon(\omega) E_i$$

Cubic crystals behave like gas, amorphous solids, liquids, and have no anisotropy.

B) uniaxial

- two crystallographic equivalent directions
- trigonal (quartz, lithium niobate), tetragonal, hexagonal



$$\boxed{\varepsilon_1(\omega) = \varepsilon_2(\omega) \neq \varepsilon_3(\omega)}$$

C) biaxial

- no crystallographic equivalent directions
- orthorhombic, monoclinic, triclinic

$$\boxed{\varepsilon_1(\omega) \neq \varepsilon_2(\omega) \neq \varepsilon_3(\omega)}$$

6.3 The index ellipsoid

The index ellipsoid offers a simple geometrical interpretation of the inverse dielectric tensor $\hat{\boldsymbol{\sigma}} = [\hat{\boldsymbol{\varepsilon}}]^{-1}$. The defining equation for the index ellipsoid is

$$\sum_{i,j=1}^{3} \sigma_{ij} x_i x_j = 1$$

which describes a surface in three dimensional space.

Remark on the physics of the index ellipsoid:
The index ellipsoid defines a surface of constant electric energy density:

$$\sum_{i,j=1}^{3} \sigma_{ij} D_i D_j = \varepsilon_0 \sum_{i=1}^{3} E_i D_i = 2w_{\mathrm{el}}$$

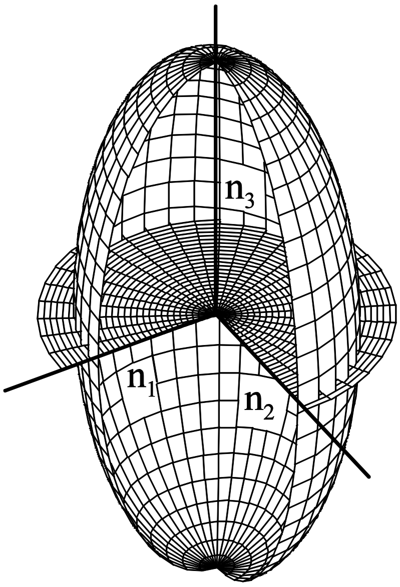
In the principal coordinate system the defining equation of the index ellipsoid reads:

$$\boxed{\sigma_1 x_1^2 + \sigma_2 x_2^2 + \sigma_3 x_3^2 = \frac{x_1^2}{\varepsilon_1} + \frac{x_2^2}{\varepsilon_2} + \frac{x_3^2}{\varepsilon_3} = 1}$$

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 $\sqrt{\varepsilon}$ 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.



Graphical representation of the epsilon tensor of an anisotropic crystal by the so-called index ellipsoid.

Summary:

- anisotropic media → tensor instead of scalar
  → in principal system: $n_i = \sqrt{\varepsilon_i}$
- The index ellipsoid is degenerate for special cases:
  - isotropic crystal: sphere
  - uniaxial crystal: rotational symmetric with respect to z-axis and $n_1 = n_2$

6.4 Normal modes in anisotropic media

Let us now look for the normal modes in crystals. A normal mode is:

- a solution to the wave equation, which shows only a phase dynamics during propagation while amplitude and polarization remain constant
  → most simple solution $\sim \exp\left\{\mathrm{i}\left[\mathbf{k}(\omega)\mathbf{r}-\omega t\right]\right\}$
- a solution where the spatial and temporal evolution of the phase are connected by a dispersion relation $\omega = \omega(\mathbf{k})$ or $\mathbf{k} = \mathbf{k}(\omega)$

Before – isotropic media

In isotropic media the normal modes are monochromatic plane waves

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

with the dispersion relation

$$\boxed{\mathbf{k}^2(\omega) = k^2(\omega) = \frac{\omega^2}{c^2}\varepsilon(\omega)}$$

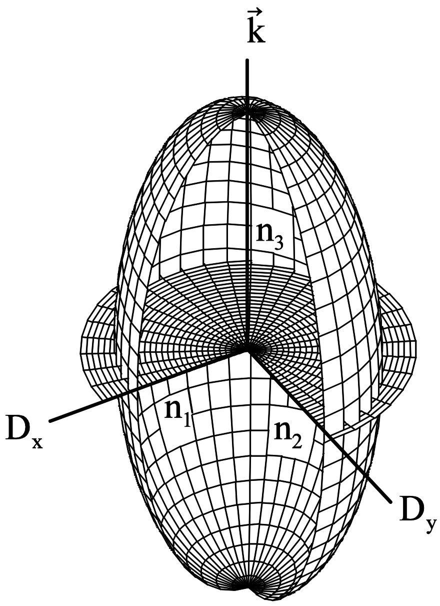
with $\varepsilon(\omega) > 0$ and real as well as $\mathbf{k}\cdot\mathbf{E} = \mathbf{k}\cdot\mathbf{D} = 0$. The normal modes are elliptically polarized, and the polarization is conserved during propagation.

Now – anisotropic media

What are the normal modes in anisotropic media?

6.4.1 Normal modes propagating in principal directions

Let us first calculate the normal modes for propagation in the direction of the principal axes of the index ellipsoid, which is the simple case.



We assume without loss of generality that the principal axes are in x, y, z direction and the light propagates in z direction $(\mathbf{k} \to k_\mathbf{z})$. Then, the fields are arbitrary in the x,y-plane

$$D_x, D_y \neq 0$$

and

$$D_i = \varepsilon_0 \varepsilon_\mathbf{i} E_i$$

In general we have $\mathbf{E} \nparallel \mathbf{D}$, but here $\mathbf{k} \cdot \mathbf{D} = 0 \Rightarrow \mathbf{k} \cdot \mathbf{E} = 0$, and the two polarization directions $(x = 1, y = 2)$ are decoupled:

$$D_1, \varepsilon_1 \curvearrowright D_1 \exp\left[\mathbf{i}\left(k_1 z - \omega t\right)\right] = D_1 \exp\left[\mathbf{i}\varphi_1(z)\right] \exp\left(-\mathbf{i}\omega t\right) \text{ with } k_1^2 = \frac{\omega^2}{c^2}\varepsilon_1(\omega)$$

$$D_2, \varepsilon_2 \curvearrowright D_2 \exp\left[\mathbf{i}\left(k_2 z - \omega t\right)\right] = D_2 \exp\left[\mathbf{i}\varphi_2(z)\right] \exp\left(-\mathbf{i}\omega t\right) \text{ with } k_2^2 = \frac{\omega^2}{c^2}\varepsilon_2(\omega)$$

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:

$$\boxed{\begin{aligned}&\mathbf{D}^{(\mathtt{a})} = \left\{ D_1 \exp\left[\mathtt{i}\left(\mathbf{k}_{\mathtt{a}}\mathbf{r} - \omega t \right) \right] \right\} \mathbf{e}_1 \quad \rightarrow \mathbf{k}_{\mathtt{a}}^2 = \frac{\omega^2}{c^2} n_{\mathtt{a}}^2 = \mathbf{k}_1^2 \rightarrow \quad \text{normal mode a}\\ &\mathbf{D}^{(\mathtt{b})} = \left\{ D_2 \exp\left[\mathtt{i}\left(\mathbf{k}_{\mathtt{b}}\mathbf{r} - \omega t \right) \right] \right\} \mathbf{e}_2 \quad \rightarrow \mathbf{k}_{\mathtt{b}}^2 = \frac{\omega^2}{c^2} n_{\mathtt{b}}^2 = \mathbf{k}_2^2 \rightarrow \text{normal mode b}\end{aligned}}$$

→ For light propagation in principle direction we find two perpendicular linearly polarized normal modes with $\mathbf{E} \parallel \mathbf{D}$.

Remark on the indices in the index ellipsoid

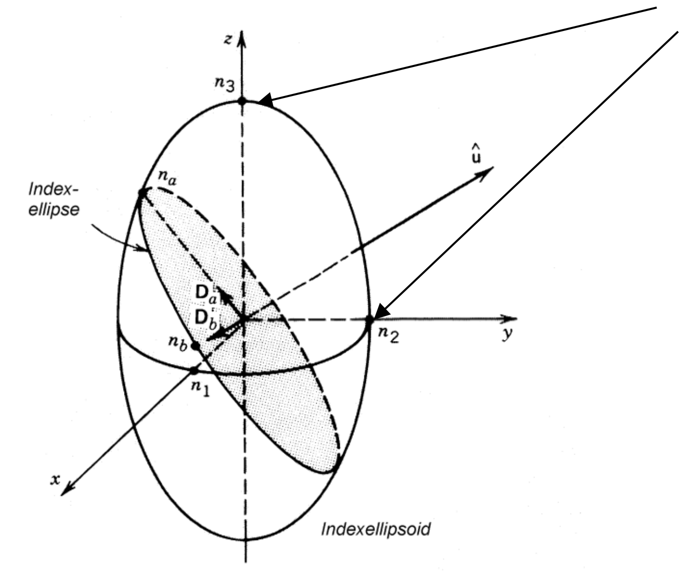
The indices $n_i = \sqrt{\varepsilon_i}$ in the index ellipsoid are connected to the indices n_a and n_b of the normal modes propagating along the principal axis. However please be careful about the direction correspondence. For example, the two normal modes propagating along the z direction have phase velocities determined by the indices $n_1 = n_x$ and $n_2 = n_y$ 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 ω we take the ε_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 → $\mathbf{k}/k = \mathbf{u}$.
- We draw a plane through the origin of index ellipsoid which is perpendicular to $\mathbf{k}$.



- The resulting intersection is an ellipse, the so-called index ellipse.
- The half-lengths of the principle axes of this ellipse equal the refractive indices $n_\mathbf{a}, n_\mathbf{b}$ of the normal modes for the propagation direction $\mathbf{u} = \mathbf{k}/k$

$$\rightarrow k_a = \frac{\omega}{c} n_\mathbf{a} \text{ and } k_\mathbf{b} = \frac{\omega}{c} n_\mathbf{b}$$

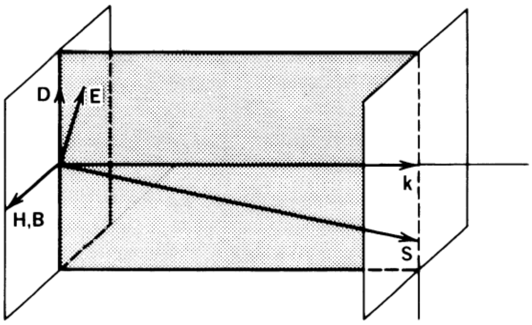
- The directions of the principal axes of the index ellipse are the polarization direction of the normal modes $\mathbf{D}^{(\mathbf{a})}$ and $\mathbf{D}^{(\mathbf{b})}$.
- The electric field vectors of the normal modes $\mathbf{E}^{(\mathbf{a})}$ and $\mathbf{E}^{(\mathbf{b})}$ follow from

$$E_i^{(\mathbf{a})} = \frac{D_i^{(\mathbf{a})}}{\varepsilon_0 \varepsilon_i}, \qquad E_i^{(\mathbf{b})} = \frac{D_i^{(\mathbf{b})}}{\varepsilon_0 \varepsilon_i}$$

- Thus, $\mathbf{D}^{(\mathbf{a,b})} \nparallel \mathbf{E}^{(\mathbf{a,b})}$, and $\mathbf{E}^{(\mathbf{a,b})}$ are not perpendicular to $\mathbf{k}$.
- This has a direct consequence on the pointing vector:

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

hence $\mathbf{k}$ is not parallel to $\langle \mathbf{S} \rangle$ because $\langle \mathbf{S} \rangle \perp \mathbf{E}$



- 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

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

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

In the isotropic case we found the dispersion relation

$$\boxed{\mathbf{k}^2(\omega)=k^2(\omega)=\frac{\omega^2}{c^2}\varepsilon(\omega)}$$

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 $\sim\exp\left\{\mathrm{i}\left[\mathbf{k}(\omega)\mathbf{r}-\omega t\right]\right\}$, but the wavenumber depends on the direction $\mathbf{u}$ of propagation, where $\mathbf{u}=\mathbf{k}/k$. Hence

$$k=k(\omega,\mathbf{u})$$

and the polarization of the normal modes is not elliptic.

In the following, we start again from Maxwell's equations and plug in the plane wave ansatz. We will use the following notation for the directional dependence of $\mathbf{k}$:

$$\mathbf{k}=\begin{pmatrix}k_1\\k_2\\k_3\end{pmatrix}=k\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}\quad\text{with}\quad u_1^2+u_2^2+u_3^2=1$$

Our aim is to derive $\omega=\omega(k_1,k_2,k_3)$ or $\omega=\omega(k,u_1,u_2,u_3)$ or $k=k(\omega,u_1,u_2,u_3)$.

We start from Maxwell's equations for the plane wave Ansatz:

$$\mathbf{k}\bullet\mathbf{D}=0\qquad\qquad \mathbf{k}\times\mathbf{E}=\omega\mu_0\mathbf{H}$$

$$\mathbf{k}\bullet\mathbf{H}=0\qquad\qquad \mathbf{k}\times\mathbf{H}=-\omega\mathbf{D}$$

Now we follow the usual derivation of the wave equation:

$$-\left[\mathbf{k}\times\left(\mathbf{k}\times\mathbf{E}\right)\right]=\frac{\omega^2}{c^2}\frac{1}{\varepsilon_0}\mathbf{D}\;\rightarrow\;-\mathbf{k}\left(\mathbf{k}\cdot\mathbf{E}\right)+\mathbf{k}^2\mathbf{E}=\frac{\omega^2}{c^2}\frac{1}{\varepsilon_0}\mathbf{D}$$

- Here $\mathbf{k}\cdot\mathbf{E}$ does not vanish as it would have in the isotropic case!
- In the principal coordinate system and with $D_i=\varepsilon_0\varepsilon_i E_i$ we find

$$-k_i\sum_j k_jE_j\;+\;k^2E_i=\frac{\omega^2}{c^2}\varepsilon_iE_i$$

$$\boxed{\left(\frac{\omega^2}{c^2}\varepsilon_i-k^2\right)E_i=-k_i\sum_j k_jE_j}$$

Remark: for isotropic media the r.h.s. of this equation would vanish ($\mathbf{k}\cdot\mathbf{E}=\mathbf{0}$). Now, we have the following problem to solve:

$$\begin{bmatrix} \frac{\omega^2}{c^2}\varepsilon_1 - k_2^2 - k_3^2 & k_1k_2 & k_1k_3 \\ k_2k_1 & \frac{\omega^2}{c^2}\varepsilon_2 - k_1^2 - k_3^2 & k_2k_3 \\ k_3k_1 & k_3k_2 & \frac{\omega^2}{c^2}\varepsilon_3 - k_1^2 - k_2^2 \end{bmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The general way to solve this problem is using $\det[..]=0$, 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:

$$\left(\frac{\omega^2}{c^2}\varepsilon_i - k^2\right)E_i = -k_i\sum_j k_jE_j \;\rightarrow\; E_i = -\frac{k_i}{\left(\frac{\omega^2}{c^2}\varepsilon_i - k^2\right)}\sum_j k_jE_j$$

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:

$$\sum_j k_jE_j = -\sum_i \frac{k_i^2}{\left(\frac{\omega^2}{c^2}\varepsilon_i - k^2\right)}\sum_j k_jE_j.$$

Because $\operatorname{div}\mathbf{E} = \sum_j k_jE_j \neq 0$ we can divide and get the (implicit) dispersion relation:

$$\boxed{\sum_i \frac{k_i^2}{\left(k^2 - \frac{\omega^2}{c^2}\varepsilon_i\right)} = 1}$$

With $\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = k(\omega)\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \frac{\omega}{c}n(\omega)\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ we can write

$$\sum_i \frac{k^2u_i^2}{\left(k^2 - \frac{\omega^2}{c^2}\varepsilon_i\right)} = 1 \rightarrow \sum_i \frac{u_i^2}{\left(1 - \frac{\varepsilon_i}{n^2}\right)} = 1$$

$$\rightarrow \boxed{\sum_i \frac{u_i^2}{\left[n^2 - \varepsilon_i\right]} = \frac{1}{n^2}} \text{ final form of DR}$$

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 $\mathbf{u}$ to determine the direction.

A more explicit form of the dispersion relation can be obtained by multiplying with denominators:

$$u_1^2\left(n^2-\varepsilon_2\right)\left(n^2-\varepsilon_3\right)n^2+u_2^2\left(n^2-\varepsilon_1\right)\left(n^2-\varepsilon_3\right)n^2+u_3^2\left(n^2-\varepsilon_1\right)\left(n^2-\varepsilon_2\right)n^2=\left(n^2-\varepsilon_1\right)\left(n^2-\varepsilon_2\right)\left(n^2-\varepsilon_3\right)$$

The resulting equation is quadratic in $(n^2)^2$ since the n^6-terms cancel. Hence, we get two (positive) solutions $n_\mathtt{a}, n_\mathtt{b}$ and therefore $k_\mathtt{a}=n_\mathtt{a}(\omega/c)$ and $k_\mathtt{b}=n_\mathtt{b}(\omega/c)$ for the two orthogonally polarized normal modes $\mathbf{D}^{(\mathtt{a})}$ and $\mathbf{D}^{(\mathtt{b})}$.
In particular, for the propagation in direction of the principal axis ($u_3=1$ and $u_1=u_2=0$, see 6.4.1) we find:

$$\rightarrow \left(n^2-\varepsilon_1\right)\left(n^2-\varepsilon_2\right)n^2=\left(n^2-\varepsilon_1\right)\left(n^2-\varepsilon_2\right)\left(n^2-\varepsilon_3\right)$$

$$\curvearrowright \left(n^2-\varepsilon_1\right)\left(n^2-\varepsilon_2\right)\varepsilon_3=0$$

$$\curvearrowright \boxed{n_\mathtt{a}^2=\varepsilon_1 \quad , \quad n_\mathtt{b}^2=\varepsilon_2}$$

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

$$\left(\frac{\omega^2}{c^2}\varepsilon_i-k^2\right)E_i=-k_i\sum_j k_jE_j.$$

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

$$E_i=-\frac{k_i}{\left(\frac{\omega^2}{c^2}\varepsilon_i-k^2\right)}\sum_j k_jE_j.$$

The sum does not depend on the index i. Hence the last term of the equation must be constant

$$\rightarrow \sum_j k_jE_j=\text{const.}$$

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

$$\curvearrowright E_1:E_2:E_3=\frac{k_1}{\frac{\omega^2}{c^2}\varepsilon_1-k^2}:\frac{k_2}{\frac{\omega^2}{c^2}\varepsilon_2-k^2}:\frac{k_3}{\frac{\omega^2}{c^2}\varepsilon_3-k^2}$$

and with $D_i=\varepsilon_0\varepsilon_iE_i$

$$\curvearrowright D_1:D_2:D_3=\frac{\varepsilon_1k_1}{\frac{\omega^2}{c^2}\varepsilon_1-k^2}:\frac{\varepsilon_2k_2}{\frac{\omega^2}{c^2}\varepsilon_2-k^2}:\frac{\varepsilon_3k_3}{\frac{\omega^2}{c^2}\varepsilon_3-k^2}$$

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
  → phase difference 0 $\rightarrow$ linear polarization

How do we see the orthogonality $\mathbf{D}^{(a)}\bullet\mathbf{D}^{(b)}=0$? (be careful: $\mathbf{E}^{(a)}\bullet\mathbf{E}^{(b)}\neq 0$)

$$\mathbf{D}^{(a)}\bullet\mathbf{D}^{(b)}\sim\sum_i\frac{k_a k_b \varepsilon_i^2 u_i^2}{\left(k_a^2-\frac{\omega^2}{c^2}\varepsilon_i\right)\left(k_b^2-\frac{\omega^2}{c^2}\varepsilon_i\right)}$$

$$=\frac{c^2}{\omega^2}\frac{k_a k_b}{\left(k_b^2-k_a^2\right)}\left[k_a^2\sum_i\frac{\varepsilon_i u_i^2}{\left(k_a^2-\frac{\omega^2}{c^2}\varepsilon_i\right)}-k_b^2\sum_i\frac{\varepsilon_i u_i^2}{\left(k_b^2-\frac{\omega^2}{c^2}\varepsilon_i\right)}\right]$$

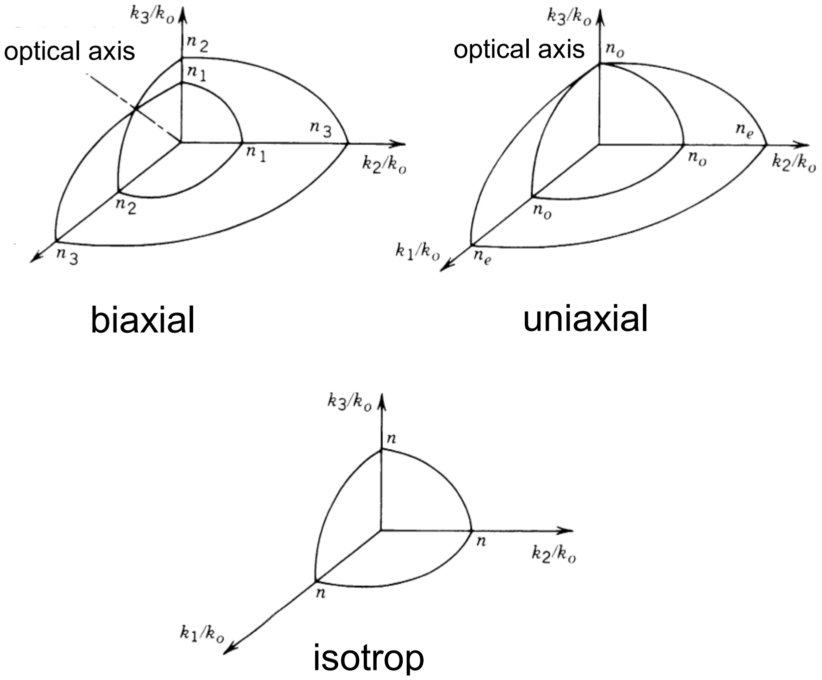
Since the two red terms vanish due to the dispersion relation, it follows that $\mathbf{D}^{(a)}\bullet\mathbf{D}^{(b)}=0$. The vanishing of the red terms can be seen when rewriting the dispersion relation:

$$1=\sum_i\frac{k_{a,b}^2 u_i^2}{\left(k_{a,b}^2-\frac{\omega^2}{c^2}\varepsilon_i\right)}=\sum_i\frac{\left(k_{a,b}^2-\frac{\omega^2}{c^2}\varepsilon_i+\frac{\omega^2}{c^2}\varepsilon_i\right)u_i^2}{\left(k_{a,b}^2-\frac{\omega^2}{c^2}\varepsilon_i\right)}=1+\frac{\omega^2}{c^2}\sum_i\frac{\varepsilon_i u_i^2}{\left(k_{a,b}^2-\frac{\omega^2}{c^2}\varepsilon_i\right)}$$

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 $n_{\mathbf{a}} = n_{\mathbf{b}}$ in the poles → connecting line defines the optical axis (for $\varepsilon_1 = \varepsilon_2 = \varepsilon_{\mathbf{or}}$, $\varepsilon_3 = \varepsilon_{\mathbf{e}}$ the z-axis is the optical axis)

biaxial: 4 points with $n_{\mathbf{a}} = n_{\mathbf{b}}$ → connecting lines define two optical axes

How to read the figure:

- fix propagation direction (u_1, u_2) → intersection with surfaces
- distances from origin to intersections with surfaces correspond to refractive indices of normal modes
- definition of optical axis → $n_{\mathbf{a}} = n_{\mathbf{b}}$

Summary: there are two geometrical constructions:

A) index ellipsoid (visualization of dielectric tensor)

- fix propagation direction $\curvearrowright$ index ellipse $\curvearrowright$ half lengths of principal axes give $n_{\mathbf{a}}, n_{\mathbf{b}}$ (refractive indices of the normal modes)
- optical axis → 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 $\curvearrowright$ intersection with surfaces $\curvearrowright$ distances from origin give $n_{\mathbf{a}}, n_{\mathbf{b}}$
- optical axis connects points with $n_{\mathbf{a}} = n_{\mathbf{b}}$

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 $\mathbf{D}^{(\mathrm{a})}$, $\mathbf{D}^{(\mathrm{b})}$.

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 one 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

$$\varepsilon_1 = \varepsilon_2 = \varepsilon_{\mathrm{or}}, \quad \varepsilon_3 = \varepsilon_{\mathrm{e}},$$

which are the ordinary and extraordinary dielectric constants.

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 $D^{(\mathrm{or})}$ is polarized perpendicular to the z-axis and the k-vector and it does not interact with ε_e.

→ The extraordinary wave $D^{(\mathrm{e})}$ is polarized perpendicular to the k-vector and $D^{(\mathrm{or})}$.

Let us now derive the dispersion relation: From above we know the implicit form

$$\rightarrow \boxed{\sum_i \frac{u_i^2}{\left[n^2 - \varepsilon_i\right]} = \frac{1}{n^2}}$$

For uniaxial crystals this leads to

$$\frac{u_1^2}{\left[n^2 - \varepsilon_{\mathrm{or}}\right]} + \frac{u_2^2}{\left[n^2 - \varepsilon_{\mathrm{or}}\right]} + \frac{u_3^2}{\left[n^2 - \varepsilon_{\mathrm{e}}\right]} = \frac{1}{n^2}$$

$$n^2\left[n^2 - \varepsilon_{\mathrm{e}}\right]\left[n^2 - \varepsilon_{\mathrm{or}}\right]\left(u_1^2 + u_2^2\right) + n^2\left[n^2 - \varepsilon_{\mathrm{or}}\right]^2 u_3^2 = \left[n^2 - \varepsilon_{\mathrm{e}}\right]\left[n^2 - \varepsilon_{\mathrm{or}}\right]^2$$

A) ordinary wave: independent of direction

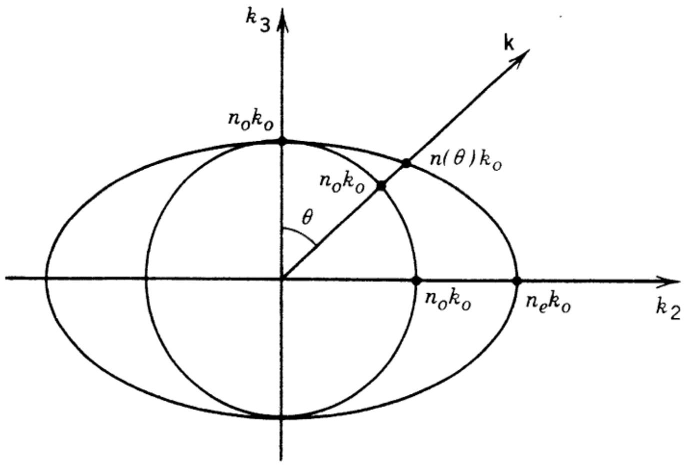
$$n_{\mathbf{a}}^2 = \varepsilon_{\mathbf{or}} \Rightarrow \boxed{k_{\mathbf{a}}^2 = \frac{\omega^2}{c^2} n_{\mathbf{a}}^2 = k_0^2 \varepsilon_{\mathbf{or}}}$$

B) extraordinary wave (derivation is your exercise): dependent on direction

$$\boxed{\frac{\left(u_1^2 + u_2^2\right)}{\varepsilon_{\mathbf{e}}} + \frac{u_3^2}{\varepsilon_{\mathbf{or}}} = \frac{1}{n_b^2}, \qquad k_b^2 = \frac{\omega^2}{c^2} n_b^2\left(u_1, u_2, u_3\right)}$$

Hence for a given direction u_i one gets the two refractive indexes $n_{\mathbf{a}}, n_{\mathbf{b}}$.

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 ($u_1 = 0$).



Normal surfaces for a uniaxial crystal.

The shape of the normal surfaces can be derived in the following way.

We have with

$$\boxed{k_i^2 = k_0^2 n^2 u_{\mathbf{i}}^2}$$

A) ordinary wave

$$k_{\mathbf{a}}^2 = k_1^2 + k_2^2 + k_3^2 = k_0^2 \varepsilon_{\mathbf{or}}$$

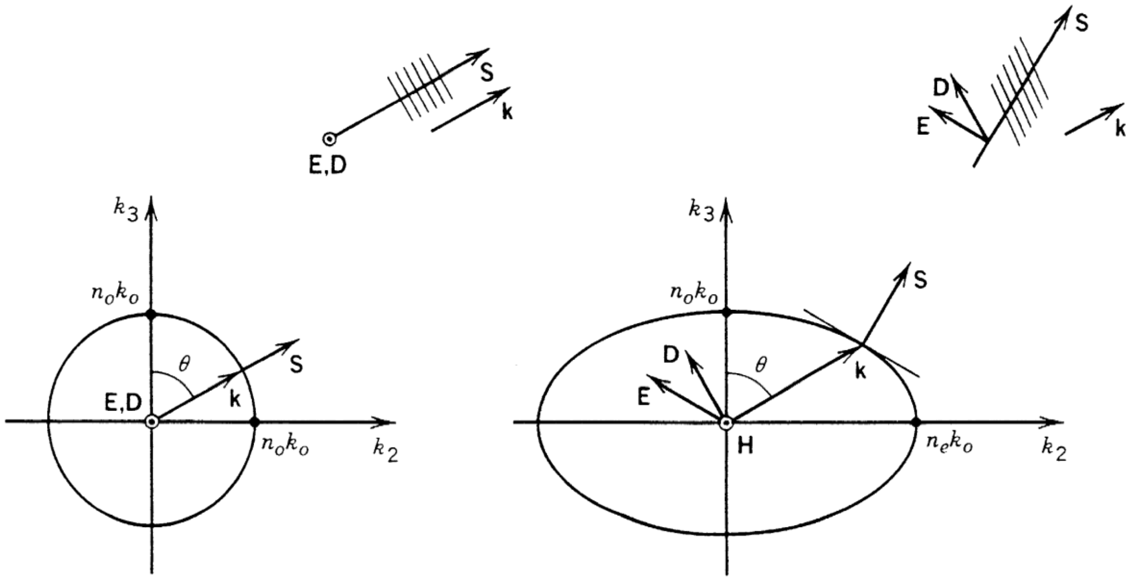
B) extraordinary wave

$$\boxed{\frac{1}{\varepsilon_{\mathbf{e}}} \frac{\left(k_1^2 + k_2^2\right)}{k_0^2} + \frac{1}{\varepsilon_{\mathbf{or}}} \frac{k_3^2}{k_0^2} = 1}$$

What about the fields? We know from before that

$$D_1 : D_2 : D_3 = \frac{\varepsilon_{\mathrm{or}} k_1}{\frac{\omega^2}{c^2}\varepsilon_{\mathrm{or}} - k^2} : \frac{\varepsilon_{\mathrm{or}} k_2}{\frac{\omega^2}{c^2}\varepsilon_{\mathrm{or}} - k^2} : \frac{\varepsilon_{\mathrm{e}} k_3}{\frac{\omega^2}{c^2}\varepsilon_{\mathrm{e}} - k^2}$$

For the extraordinary wave all denominators are finite, and in particular $k_1 = 0$ implies $D_1^{(\mathrm{e})} = 0$, hence $\mathbf{D}^{(\mathrm{e})}$ is polarized in the y-z plane. Then, $\mathbf{D}^{(\mathrm{or})} \perp \mathbf{D}^{(\mathrm{e})}$ implies that $\mathbf{D}^{(\mathrm{or})}$ is polarized in x-direction.



(a) Ordinary (b) Extraordinary

In summary, we find for the polarizations of the fields:

A) ordinary: $\mathbf{D}$ perpendicular to optical axis and $\mathbf{k}$,

$$\mathbf{D} \perp \mathbf{k}, \mathbf{D} \,||\, \mathbf{E}$$

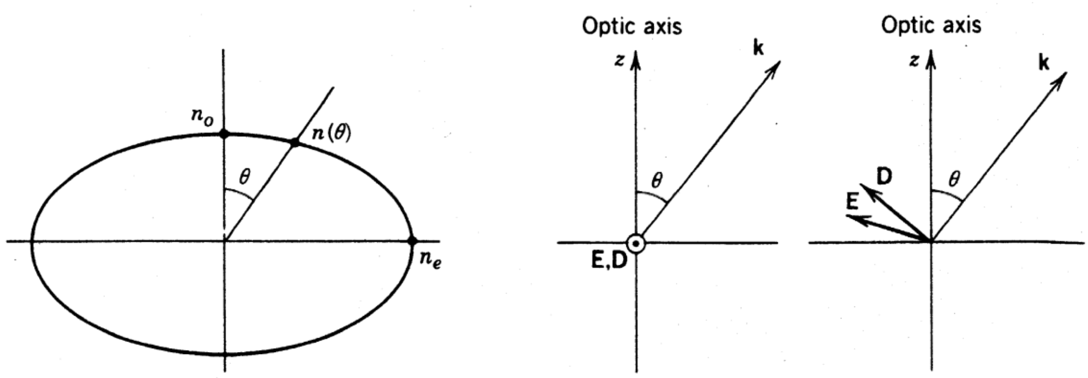
B) extraordinary: $\mathbf{D}$ perpendicular to $\mathbf{k}$ and in the plane $\mathbf{k}$-optical axis

$$\mathbf{D} \perp \mathbf{k}, \mathbf{D} \nparallel \mathbf{E}, \text{ because } D_1 = \varepsilon_0 \varepsilon_{\mathrm{or}} E_1,\ D_3 = \varepsilon_0 \varepsilon_{\mathrm{e}} E_3$$

If we introduce an angle Θ, as in the figures below, to describe the propagation direction, a simple computation of $n_b^2(\Theta)$ for the extraordinary wave is possible (exercise):

$$u_2 = \sin\Theta, \qquad u_3 = \cos\Theta$$

$$\boxed{n_{\mathrm{b}}^{2}(\Theta) = \frac{\varepsilon_{\mathrm{e}} \varepsilon_{\mathrm{or}}}{\varepsilon_{\mathrm{or}} \sin^2\Theta + \varepsilon_{\mathrm{e}} \cos^2\Theta}}$$



o wave e wave

The following classification for uniaxial crystals is commonly used

$\varepsilon_{\mathrm{or}} > \varepsilon_{\mathrm{e}}$ → negative uniaxial

$\varepsilon_{\mathrm{or}} < \varepsilon_{\mathrm{e}}$ → 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 → 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 → new types of normal modes in inhomogeneous space → “guided” waves (propagation of confined light beams without diffraction)

Approach

- take Maxwell's transition condition for interfaces
- calculate field in inhomogeneous media → 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
  → no interaction of normal modes in homogeneous space
- inhomogeneity breaks this orthogonality
  → 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

$$\mathbf{rotrot}\ \overline{\mathbf{E}}(\mathbf{r},\omega)-\frac{\omega^2}{c^2}\overline{\mathbf{E}}(\mathbf{r},\omega)=\mathbf{i}\omega\mu_0\overline{\mathbf{j}}(\mathbf{r},\omega)+\mu_0\omega^2\overline{\mathbf{P}}(\mathbf{r},\omega)$$

In general, for isotropic media all field components are coupled due to the $\mathrm{rot}\,\mathrm{rot}$ operator. However, for problems with translational invariance in at least one direction (homogeneous infinite media, layers or interfaces) a simplification is possible. Let us assume, e.g. translational invariance of the system in y direction and propagation in the x-z-plane → $\partial/\partial y=0$

$$\mathbf{rot}\ \mathbf{rot}\ \overline{\mathbf{E}}=\mathbf{grad}\ \mathbf{div}\ \overline{\mathbf{E}}-\Delta\overline{\mathbf{E}}=\begin{bmatrix}\frac{\partial}{\partial x}\left(\frac{\partial \overline{E}_x}{\partial x}+\frac{\partial \overline{E}_z}{\partial z}\right)\\ 0\\ \frac{\partial}{\partial z}\left(\frac{\partial \overline{E}_x}{\partial x}+\frac{\partial \overline{E}_z}{\partial z}\right)\end{bmatrix}-\begin{bmatrix}\Delta^{(2)}\overline{E}_x\\ \Delta^{(2)}\overline{E}_y\\ \Delta^{(2)}\overline{E}_z\end{bmatrix}$$

Then, we can split the electric field as $\overline{\mathbf{E}}=\overline{\mathbf{E}}_\perp+\overline{\mathbf{E}}_\parallel$ with

$$\overline{\mathbf{E}}_\perp=\begin{pmatrix}0\\ \overline{E}_y\\ 0\end{pmatrix},\qquad \overline{\mathbf{E}}_\parallel=\begin{pmatrix}\overline{E}_x\\ 0\\ \overline{E}_z\end{pmatrix},\nabla^{(2)}=\begin{pmatrix}\partial/\partial x\\ 0\\ \partial/\partial z\end{pmatrix},\Delta^{(2)}=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}$$

$\overline{\mathbf{E}}_\perp$ is polarized perpendicular to the plane of propagation, $\overline{\mathbf{E}}_\parallel$ is polarized parallel to this plane.

Common notations are:

perpendicular: $\perp$ → s → TE (transversal electric)

parallel: $\parallel$ → p → TM (transversal magnetic)

Both components are decoupled and can be treated independently:

$$\boxed{\begin{aligned}\Delta^{(2)}\overline{\mathbf{E}}_\perp(\mathbf{r},\omega)+\frac{\omega^2}{c^2}\overline{\mathbf{E}}_\perp(\mathbf{r},\omega)&=-\mathbf{i}\omega\mu_0\overline{\mathbf{j}}_\perp(\mathbf{r},\omega)-\mu_0\omega^2\overline{\mathbf{P}}_\perp(\mathbf{r},\omega)\\ \Delta^{(2)}\overline{\mathbf{E}}_\parallel(\mathbf{r},\omega)+\frac{\omega^2}{c^2}\overline{\mathbf{E}}_\parallel(\mathbf{r},\omega)-\mathbf{grad}^{(2)}\ \mathbf{div}^{(2)}\ \overline{\mathbf{E}}_\parallel&=-\mathbf{i}\omega\mu_0\overline{\mathbf{j}}_\parallel(\mathbf{r},\omega)-\mu_0\omega^2\overline{\mathbf{P}}_\parallel(\mathbf{r},\omega)\end{aligned}}$$

From

$$\overline{\mathbf{H}}(\mathbf{r},\omega)=-\frac{\mathbf{i}}{\omega\mu_0}\mathbf{rot}\ \overline{\mathbf{E}}(\mathbf{r},\omega)$$

we can conclude that the corresponding magnetic fields are

$$\mathbf{E}_{\mathrm{TE}}=\begin{pmatrix}0\\E_{\mathrm{y}}\\0\end{pmatrix}=\begin{pmatrix}0\\E\\0\end{pmatrix},\qquad \mathbf{H}_{\mathrm{TE}}=\begin{pmatrix}H_{\mathrm{x}}\\0\\H_{\mathrm{z}}\end{pmatrix}$$

$$\mathbf{E}_{\mathrm{TM}}=\begin{pmatrix}E_{\mathrm{x}}\\0\\E_{\mathrm{z}}\end{pmatrix},\qquad \mathbf{H}_{\mathrm{TM}}=\begin{pmatrix}0\\H_{\mathrm{y}}\\0\end{pmatrix}=\begin{pmatrix}0\\H\\0\end{pmatrix}$$

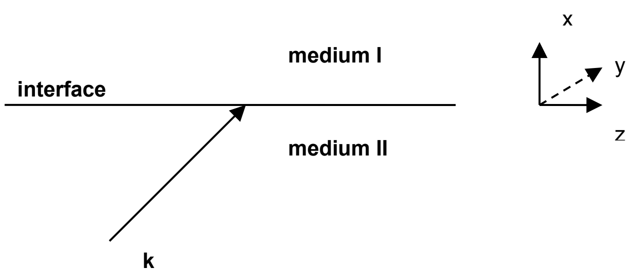
7.1.3 Interfaces and symmetries

Up to now we treated plane waves of the form

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

- homogeneous space implies: $\exp\left(\mathrm{i}\mathbf{kr}\right)$
- monochromaticity leads to: $\exp\left(-\mathrm{i}\omega t\right)$

Now, we will break the homogenity in x-direction by considering an interface in *y-z* – plane which is infinite in y and z



W.l.o.g. we can assume

$$\mathbf{k}=\left(k_{\mathrm{x}},\quad 0,\quad k_{\mathrm{z}}\right)$$

by choosing an appropriate coordinate system (plane of incidence is the x,z-plane), and then the problem does not depend on the y-coordinate.

As pointed out before, we can split the fields in TE and TM polarization $\mathbf{E}=\mathbf{E}_{\mathrm{TE}}+\mathbf{E}_{\mathrm{TM}}$ and treat them separately.

We still have homogeneity in z-direction, and therefore we expect solutions ~ $\exp\left(\mathrm{i}k_{\mathrm{z}}z\right)$. The wave vector component k_{z} has to be continuous at the interface (follows strictly from continuity of transverse field components, see 7.1.4.). Therefore, we can write for the electric field:

$$\mathbf{E}(x,z,t)=\mathbf{E}_{TE}\left(x\right)\exp\left[\mathrm{i}\left(k_{\mathrm{z}}z-\omega t\right)\right]+\mathbf{E}_{TM}\left(x\right)\exp\left[\mathrm{i}\left(k_{\mathrm{z}}z-\omega t\right)\right]$$

7.1.4 Transition conditions

From Maxwell's equations follow transition conditions for the field components. Here we will use that E_t, H_t (transverse components) are continuous at an interface between two media. This implies for the:

A) Continuity of fields

TE: $E = E_y$ and H_z continuous

TM: E_z and $H = H_y$ continuous

B) Continuity of wave vectors

homogeneous in *z*-direction → phase $\mathrm{e}^{ik_z z}$ → k_z continuous

7.2 Fields in a layer system → 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 $\varepsilon_\mathrm{f}(\omega)$

- aim: for given fields at $x = 0$ → 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 ω in the notation.

TE-polarization

We have to solve the wave equation (no y-dependence because of translational invariance):

$$\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}+\frac{\omega^2}{c^2}\varepsilon_\mathrm{f}\right]\mathbf{E}_{TE}(x,z)=0$$

We use the ansatz from above:

$$\mathbf{E}_{TE}(x,z)=\mathbf{E}_{TE}(x)\exp\left(ik_\mathrm{z}z\right) \text{ and } \mathbf{H}_{TE}(x,z)=\mathbf{H}_{TE}(x)\exp\left(ik_\mathrm{z}z\right)$$

$$\left[\frac{\partial^2}{\partial x^2}+\frac{\omega^2}{c^2}\varepsilon_\mathrm{f}-k_\mathrm{z}^2\right]\mathbf{E}_{TE}(x)=0$$

with: $$\mathbf{H}_\mathrm{TE}(x,z)=-\frac{i}{\omega\mu_0}\mathbf{rot}\mathbf{E}_\mathrm{TE}(x,z)$$

Now let us extract the equations for transversal fields $E_{\mathrm{y}} = E,\ H_{\mathrm{z}}$:

$$\left[\frac{\partial^2}{\partial x^2} + k_{\mathrm{fx}}^2\left(k_{\mathrm{z}},\omega\right)\right]E(x) = 0 \text{ with } k_{\mathrm{fx}}^2\left(k_{\mathrm{z}},\omega\right) = \frac{\omega^2}{c^2}\varepsilon_{\mathrm{f}}\left(\omega\right) - k_{\mathrm{z}}^2$$

$$H_{\mathrm{z}}(x) = -\frac{i}{\omega\mu_0}\frac{\partial}{\partial x}E(x)$$

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 $H_{\mathrm{y}} = H,\ E_{\mathrm{z}}$:

$$\left[\frac{\partial^2}{\partial x^2} + k_{\mathrm{fx}}^2\left(k_{\mathrm{z}},\omega\right)\right]H(x) = 0$$

$$E_{\mathrm{z}}(x) = \frac{i}{\omega\varepsilon_0\varepsilon_{\mathrm{f}}}\frac{\partial}{\partial x}H(x)$$

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: $H_{\mathrm{TM}} \rightarrow E_{\mathrm{TM}} \rightarrow E = E_{\mathrm{TM}} + E_{\mathrm{TE}}$

Because the equations for TE and TM have identical structure, we can treat them simultaneously. We rename

$E, H \rightarrow F$ generalized field 1

$i\omega\mu_0 H_{\mathrm{z}},\ -i\omega\varepsilon_0 E_{\mathrm{z}} \rightarrow G$ generalized field 2

and write down the problem to solve:

$$\boxed{\left[\frac{\partial^2}{\partial x^2} + k_{\mathrm{fx}}^2\left(k_{\mathrm{z}},\omega\right)\right]F(x) = 0}$$

$$\boxed{G(x) = \alpha_{\mathrm{f}}\frac{\partial}{\partial x}F(x)} \qquad \text{with } \alpha_{\mathrm{fTE}} = 1,\ \alpha_{\mathrm{fTM}} = \frac{1}{\varepsilon_{\mathrm{f}}}$$

We know the general solution of this system (harmonic oscillator equation):

$$F(x) = C_1 \exp\left(ik_{\mathrm{fx}} x\right) + C_2 \exp\left(-ik_{\mathrm{fx}} x\right)$$

$$G(x) = \alpha_{\mathrm{f}} \frac{\partial}{\partial x} F(x) = i\alpha_{\mathrm{f}} k_{\mathrm{fx}} \left[C_1 \exp\left(ik_{\mathrm{fx}} x\right) - C_2 \exp\left(-ik_{\mathrm{fx}} x\right) \right]$$

We have as initial conditions $F(0), G(0)$ given:

$$F(0) = C_1 + C_2$$

$$G(0) = i\alpha_{\mathrm{f}} k_{\mathrm{fx}} \left[C_1 - C_2 \right]$$

from which we can compute the constants C_1, C_2:

$$C_1 = \frac{1}{2}\left[F(0) - \frac{i}{\alpha_{\mathrm{f}} k_{\mathrm{fx}}} G(0) \right]$$

$$C_2 = \frac{1}{2}\left[F(0) + \frac{i}{\alpha_{\mathrm{f}} k_{\mathrm{fx}}} G(0) \right]$$

The final solution of the initial value problem is therefore:

$$\boxed{\begin{aligned} F(x) &= \cos\left(k_{\mathrm{fx}} x\right) F(0) + \frac{1}{\alpha_{\mathrm{f}} k_{\mathrm{fx}}} \sin\left(k_{\mathrm{fx}} x\right) G(0) \\ G(x) &= -\alpha_{\mathrm{f}} k_{\mathrm{fx}} \sin\left(k_{\mathrm{fx}} x\right) F(0) + \cos\left(k_{\mathrm{fx}} x\right) G(0) \end{aligned}}$$

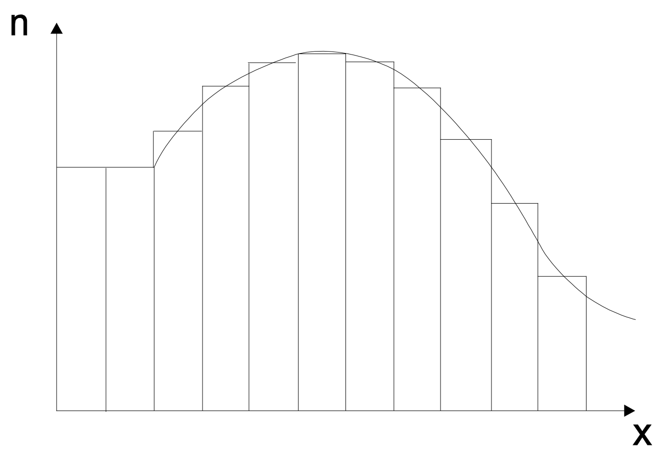
By resubstituting we have the electromagnetic field in the layer $0 \le x \le d$.

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:

$$F(x)=\cos\left(k_{\mathrm{fx}}x\right)F(0)+\frac{1}{\alpha_{\mathrm{f}}k_{\mathrm{fx}}}\sin\left(k_{\mathrm{fx}}x\right)G(0)$$

$$G(x)=-\alpha_{\mathrm{f}}k_{\mathrm{fx}}\sin\left(k_{\mathrm{fx}}x\right)F(0)+\cos\left(k_{\mathrm{fx}}x\right)G(0)$$

We can write this formally in matrix notation as

$$\boxed{\begin{Bmatrix}F(x)\\G(x)\end{Bmatrix}=\hat{\mathbf{m}}(x)\begin{Bmatrix}F(0)\\G(0)\end{Bmatrix},}$$

where the 2x2-matrix $\hat{\mathbf{m}}$ describes propagation of the fields:

$$\boxed{\hat{\mathbf{m}}(x)=\begin{Bmatrix}\cos\left(k_{\mathrm{fx}}x\right) & \frac{1}{k_{\mathbf{fx}}\alpha_{\mathbf{f}}}\sin\left(k_{\mathbf{fx}}x\right)\\ -k_{\mathbf{fx}}\alpha_{\mathbf{f}}\sin\left(k_{\mathbf{fx}}x\right) & \cos\left(k_{\mathbf{fx}}x\right)\end{Bmatrix}}$$

- To compute the field at the end of the layer we set $x=d$.
- We assume no absorption in the layer → $\left\|\hat{\mathbf{m}}(x)\right\|=1$.
- A system of layers is characterized by ε_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

$$\begin{pmatrix}F\\G\end{pmatrix}_{d_1+d_2}=\hat{\mathbf{m}}_2(d_2)\begin{pmatrix}F\\G\end{pmatrix}_{d_1}=\hat{\mathbf{m}}_2(d_2)\hat{\mathbf{m}}_1(d_1)\begin{pmatrix}F\\G\end{pmatrix}_0$$

B) N layers

$$\begin{pmatrix}F\\G\end{pmatrix}_{d_1+d_2+..+d_{\mathbf{N}}=D}=\prod_{i=1}^{N}\hat{\mathbf{m}}_i(d_i)\begin{pmatrix}F\\G\end{pmatrix}_0=\hat{\mathbf{M}}\begin{pmatrix}F\\G\end{pmatrix}_0 \quad \text{with } \hat{\mathbf{M}}=\prod_{i=1}^{N}\hat{\mathbf{m}}_i(d_i)$$

All matrices $\hat{\mathbf{m}}_i$ have the same form, but different $\alpha_f^i, d_i, k_{\mathbf{fx}}^i = \sqrt{\frac{\omega^2}{c^2}\varepsilon_i(\omega) - k_{\mathrm{z}}^2}$.

Summary of matrix method

- $F(0)$ and $G(0)$ given (E, H_z for TE, E_z, H for TM)
- $k_{\mathbf{z}}, \alpha_f^i, \varepsilon_i, d_i$ given → matrix elements
- multiplication of matrices (in the right order) → total matrix
- fields $F(D)$ and $G(D)$

7.3 Reflection – transmission problem for layer systems

7.3.1 General layer systems

7.3.1.1 Reflection- and transmission coefficients → generalized Fresnel formulas

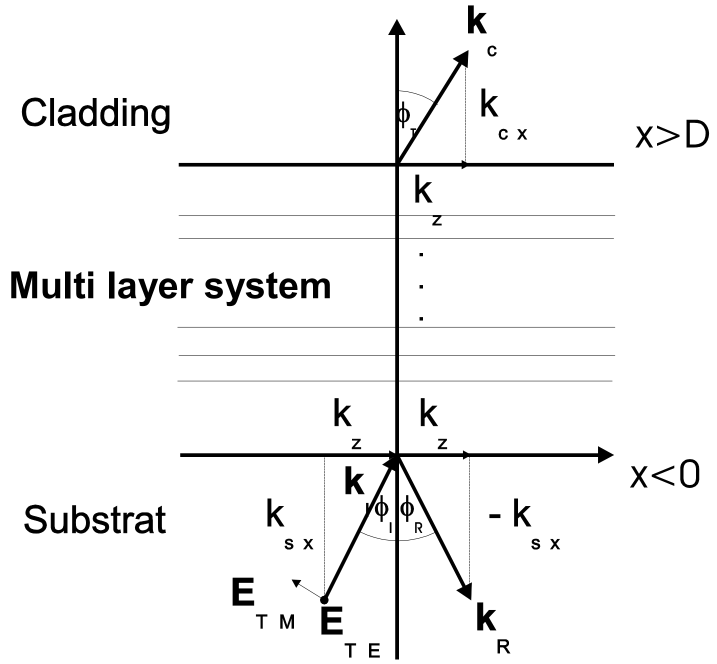
In the previous chapter, we have learned how to link the electromagnetic field on one side of an arbitrary multilayer system with the field on the other side. We have seen that after splitting in the TE/TM-polarizations, continuous (transversal) field components are sufficient to describe the whole field. 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 ($\mathbf{k}_{\mathbf{I}}$), reflected ($\mathbf{k}_{\mathbf{R}}$) and transmitted ($\mathbf{k}_{\mathbf{T}}$) fields:

$$\mathbf{k}_{\mathbf{I}} = \begin{pmatrix} k_{\mathbf{sx}} \\ 0 \\ k_{\mathbf{z}} \end{pmatrix}, \qquad \mathbf{k}_{\mathbf{R}} = \begin{pmatrix} -k_{\mathbf{sx}} \\ 0 \\ k_{\mathbf{z}} \end{pmatrix}, \qquad \mathbf{k}_{\mathbf{T}} = \begin{pmatrix} k_{\mathbf{cx}} \\ 0 \\ k_{\mathbf{z}} \end{pmatrix}$$

with $k_{\mathbf{sx}} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_{\mathrm{s}} - k_{\mathbf{z}}^2} = \sqrt{k_{\mathbf{s}}^2(\omega) - k_{\mathbf{z}}^2}$, $\qquad k_{\mathbf{cx}} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_{\mathbf{c}} - k_{\mathbf{z}}^2} = \sqrt{k_{\mathbf{c}}^2(\omega) - k_{\mathbf{z}}^2}$,

where $\varepsilon_{\mathbf{s}}(\omega)$ and $\varepsilon_{\mathbf{c}}(\omega)$ 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 $\pm k_{\mathbf{x}}$ 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 $k_{\mathbf{z}}$ is a conserved quantity:

1. $k_{\mathbf{s}} \sin \varphi_{\mathbf{I}} = k_{\mathbf{s}} \sin \varphi_{\mathbf{R}} \curvearrowright \varphi_{\mathbf{I}} = \varphi_{\mathbf{R}}$ (reflection)
2. $k_{\mathbf{s}} \sin \varphi_{\mathbf{I}} = k_{\mathbf{c}} \sin \varphi_{\mathbf{T}} \curvearrowright n_{\mathbf{s}} \sin \varphi_{\mathbf{I}} = n_{\mathbf{c}} \sin \varphi_{\mathbf{T}}$ (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 $F_{\mathbf{I}}$ and reflected field $F_{\mathbf{R}}$ as:

$$F_{\mathbf{s}}(x,z) = \exp\left(\mathbf{i}k_{\mathbf{z}}z\right)\left[F_{\mathbf{I}} \exp\left(\mathbf{i}k_{\mathbf{sx}}x\right) + F_{\mathbf{R}} \exp\left(-\mathbf{i}k_{\mathbf{sx}}x\right)\right]$$
$$G_{\mathbf{s}}(x,z) = \mathbf{i}\alpha_{\mathbf{s}}k_{\mathbf{sx}} \exp\left(\mathbf{i}k_{\mathbf{z}}z\right)\left[F_{\mathbf{I}} \exp\left(\mathbf{i}k_{\mathbf{sx}}x\right) - F_{\mathbf{R}} \exp\left(-\mathbf{i}k_{\mathbf{sx}}x\right)\right]$$

B) Field in layer system

The fields inside the layer system can be expressed as

$$F_{\mathbf{f}}(x,z) = \exp\left(\mathbf{i}k_{\mathbf{z}}z\right)F(x)$$
$$G_{\mathbf{f}}(x,z) = \exp\left(\mathbf{i}k_{\mathbf{z}}z\right)G(x)$$

where the amplitudes $F(x)$ and $G(x)$ are given by matrix method as

$$\begin{pmatrix} F \\ G \end{pmatrix}_x = \hat{\mathbf{M}}(x) \begin{pmatrix} F \\ G \end{pmatrix}_0$$

C) field in cladding

The fields in the cladding can be expressed based on the complex amplitude of the transmitted field F_T as:

$$F_c(x,z) = \exp\left(\mathrm{i}k_z z\right) F_T \exp\left[\mathrm{i}k_{cx}\left(x - D\right)\right]$$
$$G_c(x,z) = \mathrm{i}\alpha_c k_{cx} \exp\left(\mathrm{i}k_z z\right) F_T \exp\left[\mathrm{i}k_{cx}\left(x - D\right)\right]$$

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 $(\sim \sin\varphi_I), \varepsilon_i, d_i$. We know that F and G are continuous at the interfaces, in particular at $x=0$ and $x=D$. We have:

$$\begin{pmatrix} F \\ G \end{pmatrix}_D = \hat{\mathbf{M}}(D) \begin{pmatrix} F \\ G \end{pmatrix}_0 .$$

Field in cladding at $x = D$ *field in substrate at $x = 0$*

On the other hand, we have expressions for the fields at $x=0$ and $x=D$ from our decomposition in incident, reflected and transmitted field from above. Hence, we can write:

$$\begin{pmatrix} F_T \\ \mathrm{i}\alpha_c k_{cx} F_T \end{pmatrix} = \begin{Bmatrix} M_{11}(D) & M_{12}(D) \\ M_{21}(D) & M_{22}(D) \end{Bmatrix} \begin{pmatrix} F_I + F_R \\ \mathrm{i}\alpha_s k_{sx}\left(F_I - F_R\right) \end{pmatrix}.$$

We consider F_I as known, and F_R and F_T as unknown and get:

$$\boxed{F_R = \frac{\left(\alpha_s k_{sx} M_{22} - \alpha_c k_{cx} M_{11}\right) - \mathrm{i}\left(M_{21} + \alpha_s k_{sx}\alpha_c k_{cx} M_{12}\right)}{\underbrace{\left(\alpha_s k_{sx} M_{22} + \alpha_c k_{cx} M_{11}\right) + \mathrm{i}\left(M_{21} - \alpha_s k_{sx}\alpha_c k_{cx} M_{12}\right)}_{N}} F_I}$$

$$F_T = \frac{2\alpha_s k_{sx}(M_{11}M_{22} - M_{12}M_{21})}{(\alpha_s k_{sx} M_{22} + \alpha_c k_{cx} M_{11}) + i(M_{21} - \alpha_s k_{sx}\alpha_c k_{cx} M_{12})} F_I$$

$$\boxed{F_T = \frac{2\alpha_s k_{sx}}{N} F_I}$$

These are the general formulas for reflected and transmitted amplitudes. Please remember that the matrix elements depend on the polarization direction → $M_{ij}^{TE} \neq M_{ij}^{TM}$.

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

$$F = E = E_{\mathtt{y}}, \qquad \alpha_{\mathtt{TE}} = 1$$

i) reflected field

$$E_{\mathtt{R}}^{\mathtt{TE}} = R_{\mathtt{TE}} E_{\mathtt{I}}^{\mathtt{TE}}$$

with the reflection coefficient

$$\boxed{R_{\mathtt{TE}} = \frac{\left(k_{\mathtt{sx}} M_{22}^{\mathtt{TE}} - k_{\mathtt{cx}} M_{11}^{\mathtt{TE}}\right) - \mathtt{i}\left(M_{21}^{\mathtt{TE}} + k_{\mathtt{sx}} k_{\mathtt{cx}} M_{12}^{\mathtt{TE}}\right)}{\underbrace{\left(k_{\mathtt{sx}} M_{22}^{\mathtt{TE}} + k_{\mathtt{cx}} M_{11}^{\mathtt{TE}}\right) + \mathtt{i}\left(M_{21}^{\mathtt{TE}} - k_{\mathtt{sx}} k_{\mathtt{cx}} M_{12}^{\mathtt{TE}}\right)}_{N_{\mathrm{TE}}}}}$$

ii) transmitted field

$$E_{\mathtt{T}}^{\mathtt{TE}} = T_{\mathtt{TE}} E_{\mathtt{I}}^{\mathtt{TE}}$$

with the transmission coefficient

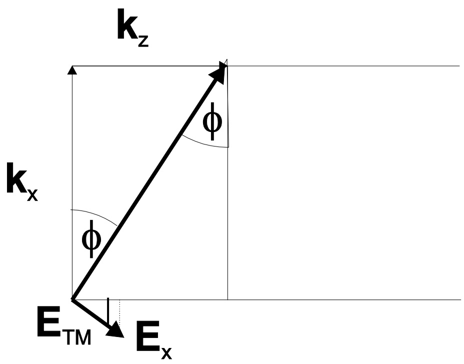
$$\boxed{T_{\mathtt{TE}} = \frac{2k_{\mathtt{sx}}}{\left(k_{\mathtt{sx}} M_{22}^{\mathtt{TE}} + k_{\mathtt{cx}} M_{11}^{\mathtt{TE}}\right) + \mathtt{i}\left(M_{21}^{\mathtt{TE}} - k_{\mathtt{sx}} k_{\mathtt{cx}} M_{12}^{\mathtt{TE}}\right)} = \frac{2k_{\mathtt{sx}}}{N_{\mathtt{TE}}},}$$

→We get complex coefficients for reflection and transmission, which determine the amplitude and phase of the reflected and transmitted light.

B) TM-polarization

$$F = H = H_{\mathtt{y}}, \qquad \alpha_{\mathtt{TM}} = \frac{1}{\varepsilon}.$$

In the case of TM polarization we have the problem that an analog calculation to TE would lead to $H_{\mathtt{R,T}}/H_{\mathtt{I}}$, i.e., relations between the magnetic field. However, we want $E_{\mathtt{R,T}}^{\mathtt{TM}}/E_{\mathtt{I}}^{\mathtt{TM}}$. Therefore, we have to convert the H-field to the $E^{\mathtt{TM}}$-field:



As can be seen in the figure, we can express the amplitude of the E_{TM} field in terms of the E_x component:

$$\frac{E_{\mathrm{x}}}{E^{\mathrm{TM}}} = -\sin\varphi = -\frac{k_{\mathrm{z}}}{k}$$

$$\curvearrowright E^{\mathrm{TM}} = -\frac{k}{k_{\mathrm{z}}} E_{\mathrm{x}},$$

With Maxwell we can link E_x to H_y:

$$\mathbf{E} = -\frac{1}{\omega\varepsilon_0\varepsilon}(\mathbf{k}\times\mathbf{H}) \rightarrow E_{\mathrm{x}} = \frac{1}{\omega\varepsilon_0\varepsilon} k_{\mathrm{z}} H_{\mathrm{y}} \rightarrow E^{\mathrm{TM}} = -\frac{k}{\omega\varepsilon_0\varepsilon} H_{\mathrm{y}} = -\frac{1}{c\varepsilon_0\sqrt{\varepsilon}} H_{\mathrm{y}}$$

result: $$\frac{E^{\mathrm{TM}}_{\mathrm{R,T}}}{E^{\mathrm{TM}}_{\mathrm{I}}} = \sqrt{\frac{\varepsilon_{\mathrm{s}}}{\varepsilon_{\mathrm{s,c}}}} \frac{H_{\mathrm{R,T}}}{H_{\mathrm{I}}}, \quad \rightarrow \sqrt{\varepsilon_{\mathrm{s}}/\varepsilon_{\mathrm{c}}}$$ relevant for transmission only

Hence we find the following for TM polarization:

$$E^{\mathrm{TM}}_{\mathrm{R}} = R_{\mathrm{TM}} E^{\mathrm{TM}}_{\mathrm{I}}$$

with the reflection coefficient

$$\boxed{R_{\mathrm{TM}} = \frac{\left(\varepsilon_{\mathrm{c}} k_{\mathrm{sx}} M^{\mathrm{TM}}_{22} - \varepsilon_{\mathrm{s}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{11}\right) - \mathrm{i}\left(\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}} M^{\mathrm{TM}}_{21} + k_{\mathrm{sx}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{12}\right)}{\underbrace{\left(\varepsilon_{\mathrm{c}} k_{\mathrm{sx}} M^{\mathrm{TM}}_{22} + \varepsilon_{\mathrm{s}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{11}\right) + \mathrm{i}\left(\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}} M^{\mathrm{TM}}_{21} - k_{\mathrm{sx}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{12}\right)}_{N_{\mathrm{TM}}}}}$$

$$E^{\mathrm{TM}}_{\mathrm{T}} = T_{\mathrm{TM}} E^{\mathrm{TM}}_{\mathrm{I}}$$

with the transmission coefficient

$$\boxed{T_{\mathrm{TM}} = \frac{2\sqrt{\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}}}\, k_{\mathrm{sx}}}{\left(\varepsilon_{\mathrm{c}} k_{\mathrm{sx}} M^{\mathrm{TM}}_{22} + \varepsilon_{\mathrm{s}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{11}\right) + \mathrm{i}\left(\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}} M^{\mathrm{TM}}_{21} - k_{\mathrm{sx}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{12}\right)} = \frac{2\sqrt{\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}}}\, k_{\mathrm{sx}}}{N_{\mathrm{TM}}}}$$

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

$$\boxed{R_{\mathrm{TE}} = \frac{\left(k_{\mathrm{sx}} M^{\mathrm{TE}}_{22} - k_{\mathrm{cx}} M^{\mathrm{TE}}_{11}\right) - \mathrm{i}\left(M^{\mathrm{TE}}_{21} + k_{\mathrm{sx}} k_{\mathrm{cx}} M^{\mathrm{TE}}_{12}\right)}{\left(k_{\mathrm{sx}} M^{\mathrm{TE}}_{22} + k_{\mathrm{cx}} M^{\mathrm{TE}}_{11}\right) + \mathrm{i}\left(M^{\mathrm{TE}}_{21} - k_{\mathrm{sx}} k_{\mathrm{cx}} M^{\mathrm{TE}}_{12}\right)}}$$

$$\boxed{T_{\mathrm{TE}} = \frac{2k_{\mathrm{sx}}}{N_{\mathrm{TE}}}}$$

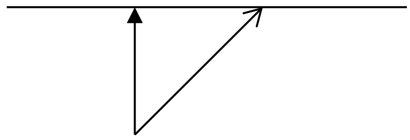
$$\boxed{R_{\mathrm{TM}} = \frac{\left(\varepsilon_{\mathrm{c}} k_{\mathrm{sx}} M^{\mathrm{TM}}_{22} - \varepsilon_{\mathrm{s}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{11}\right) - \mathrm{i}\left(\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}} M^{\mathrm{TM}}_{21} + k_{\mathrm{sx}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{12}\right)}{\left(\varepsilon_{\mathrm{c}} k_{\mathrm{sx}} M^{\mathrm{TM}}_{22} + \varepsilon_{\mathrm{s}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{11}\right) + \mathrm{i}\left(\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}} M^{\mathrm{TM}}_{21} - k_{\mathrm{sx}} k_{\mathrm{cx}} M^{\mathrm{TM}}_{12}\right)}}$$

$$\boxed{T_{\mathrm{TM}} = \frac{2\sqrt{\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}}}\, k_{\mathrm{sx}}}{N_{\mathrm{TM}}}.}$$

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:

- flux through a surface with $x = \text{const}$



$$\langle \mathbf{S} \rangle \mathbf{e}_{\mathbf{x}} = \frac{1}{2} \Re\left(\mathbf{E} \times \mathbf{H}^{*}\right) \mathbf{e}_{\mathbf{x}}$$

With
$$\mathbf{H}^{*} = \frac{1}{\omega\mu_0}\left(\mathbf{k}^{*} \times \mathbf{E}^{*}\right)$$

we find

$$\boxed{\langle \mathbf{S} \rangle \mathbf{e}_{\mathbf{x}} = \frac{1}{2\omega\mu_0} \Re\left(\mathbf{k}^{*}\mathbf{e}_{\mathbf{x}}\right) |\mathbf{E}|^2 = \frac{1}{2\omega\mu_0} \Re\left(k_{\mathbf{x}}\right) |\mathbf{E}|^2 .}$$

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

$$k_{\mathbf{sx}} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_{\mathrm{s}} - k_{\mathbf{z}}^2} = \sqrt{k_{\mathbf{s}}^2(\omega) - k_{\mathbf{z}}^2}$$

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

$$\langle \mathbf{S} \rangle_{\mathbf{s}} \mathbf{e}_{\mathbf{x}} = \frac{1}{2\omega\mu_0}\left[k_{\mathbf{sx}} |\mathbf{E}_{\mathbf{I}}|^2 - k_{\mathbf{sx}} |\mathbf{E}_{\mathbf{R}}|^2 \right]$$

In contrast, in the cladding

$$k_{\mathbf{cx}} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_{\mathrm{c}} - k_{\mathbf{z}}^2} = \sqrt{k_{\mathbf{c}}^2(\omega) - k_{\mathbf{z}}^2}$$

may be complex-valued. The energy flux from the layer system into the cladding is

$$\langle \mathbf{S} \rangle_{\mathbf{c}} \mathbf{e}_{\mathbf{x}} = \frac{1}{2\omega\mu_0} \Re\left(k_{\mathbf{cx}}\right) |\mathbf{E}_{\mathbf{T}}|^2 .$$

Because we have energy conservation: $\langle \mathbf{S} \rangle_{\mathbf{s}} \mathbf{e}_{\mathbf{x}} = \langle \mathbf{S} \rangle_{\mathbf{c}} \mathbf{e}_{\mathbf{x}} \curvearrowright$

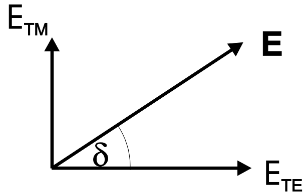
$$\boxed{|\mathbf{E}_{\mathbf{I}}|^2 = |\mathbf{E}_{\mathbf{R}}|^2 + \frac{\Re\left(k_{\mathbf{cx}}\right)}{k_{\mathbf{sx}}} |\mathbf{E}_{\mathbf{T}}|^2}$$

Now we will compute the global reflectivity ρ and transmissivity τ of a layer system. Of course, we will decompose into TE and TM polarizations and relate to the reflectivities $\rho_{\mathrm{TE,TM}}$ and transmissivities $\tau_{\mathrm{TE,TM}}$. We know:

$$\mathbf{E}_{\mathrm{R}} = \mathbf{E}_{\mathrm{R}}^{\mathrm{TE}} + \mathbf{E}_{\mathrm{R}}^{\mathrm{TM}}, \qquad \mathbf{E}_{\mathrm{T}} = \mathbf{E}_{\mathrm{T}}^{\mathrm{TE}} + \mathbf{E}_{\mathrm{T}}^{\mathrm{TM}}$$

$$\begin{aligned}\left|\mathbf{E}_{\mathrm{I}}\right|^2 &= \left|\mathbf{E}_{\mathrm{R}}^{\mathrm{TE}}\right|^2 + \left|\mathbf{E}_{\mathrm{R}}^{\mathrm{TM}}\right|^2 + \frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left(\left|\mathbf{E}_{\mathrm{T}}^{\mathrm{TE}}\right|^2 + \left|\mathbf{E}_{\mathrm{T}}^{\mathrm{TM}}\right|^2\right) \\ &= \left\{\left|R_{\mathrm{TE}}\right|^2 + \frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left|T_{\mathrm{TE}}\right|^2\right\}\left|\mathbf{E}_{\mathrm{I}}^{\mathrm{TE}}\right|^2 + \left\{\left|R_{\mathrm{TM}}\right|^2 + \frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left|T_{\mathrm{TM}}\right|^2\right\}\left|\mathbf{E}_{\mathrm{I}}^{\mathrm{TM}}\right|^2.\end{aligned}$$

Here, we just substituted the reflected and transmitted field amplitudes by incident amplitudes times Fresnel coefficients. Now, we decompose the incident field as follows:



$$E_{\mathrm{I}}^{\mathrm{TE}} = \left|\mathbf{E}_{\mathrm{I}}\right|\cos\delta, \qquad E_{\mathrm{I}}^{\mathrm{TM}} = \left|\mathbf{E}_{\mathrm{I}}\right|\sin\delta.$$

Then, we can divide by the (arbitrary) amplitude $\left|\mathbf{E}_{\mathrm{I}}\right|^2$ and write

$$1 = \left\{\left|R_{\mathrm{TE}}\right|^2 + \frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left|T_{\mathrm{TE}}\right|^2\right\}\cos^2\delta + \left\{\left|R_{\mathrm{TM}}\right|^2 + \frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left|T_{\mathrm{TM}}\right|^2\right\}\sin^2\delta$$

$$1 = \left(\left|R_{\mathrm{TE}}\right|^2\cos^2\delta + \left|R_{\mathrm{TM}}\right|^2\sin^2\delta\right) + \frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left(\left|T_{\mathrm{TE}}\right|^2\cos^2\delta + \left|T_{\mathrm{TM}}\right|^2\sin^2\delta\right)$$

The red and blue terms can be identified as $\boxed{1 = \rho + \tau}$

The global reflectivity and transmissivity are therefore given as

$$\boxed{\begin{aligned}\rho &= \rho_{\mathrm{TE}}\cos^2\delta + \rho_{\mathrm{TM}}\sin^2\delta \\ \tau &= \tau_{\mathrm{TE}}\cos^2\delta + \tau_{\mathrm{TM}}\sin^2\delta\end{aligned}}$$

with the reflectivities

$$\boxed{\rho_{\mathrm{TE,TM}} = \left|R_{\mathrm{TE,TM}}\right|^2, \qquad \tau_{\mathrm{TE,TM}} = \frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left|T_{\mathrm{TE,TM}}\right|^2.}$$

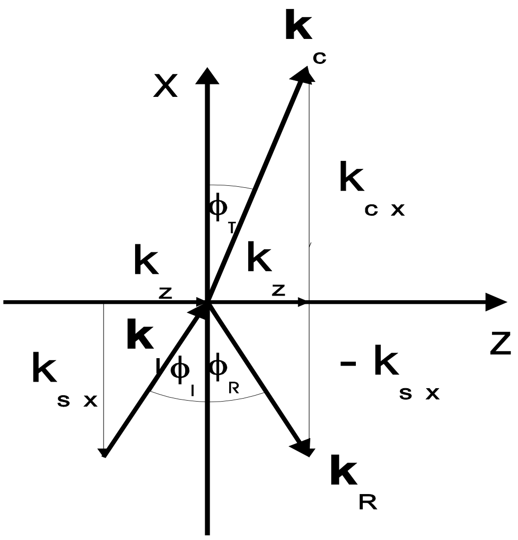
for the two polarization states TE and TM.

7.3.2 Single interface

7.3.2.1 (classical) Fresnel formulas

Let us now consider the important example of the most simple layer system, namely the single interface. The relevant wave vectors are (as usual):

$$\mathbf{k}_{\mathrm{I}} = \begin{pmatrix} k_{\mathrm{sx}} \\ 0 \\ k_{\mathrm{z}} \end{pmatrix}, \qquad \mathbf{k}_{\mathrm{R}} = \begin{pmatrix} -k_{\mathrm{sx}} \\ 0 \\ k_{\mathrm{z}} \end{pmatrix}, \qquad \mathbf{k}_{\mathrm{T}} = \begin{pmatrix} k_{\mathrm{cx}} \\ 0 \\ k_{\mathrm{z}} \end{pmatrix}.$$



The continuous component of the wave vector, expressed in terms of the angel of incidence, is

$$k_{\mathrm{z}} = \frac{\omega}{c}\sqrt{\varepsilon_{\mathrm{s}}}\sin\varphi_{\mathrm{I}} = \frac{\omega}{c} n_{\mathrm{s}} \sin\varphi_{\mathrm{I}}.$$

Then, the discontinuous component is given as

$$\rightarrow k_{i\mathrm{x}} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_i - k_{\mathrm{z}}^2} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_i - \frac{\omega^2}{c^2}\varepsilon_{\mathrm{s}}\sin^2\varphi_{\mathrm{I}}} = \frac{\omega}{c}\sqrt{n_i^2 - n_{\mathrm{s}}^2\sin^2\varphi_{\mathrm{I}}}$$

$$\boxed{\curvearrowright k_{\mathrm{sx}} = \frac{\omega}{c} n_{\mathrm{s}} \cos\varphi_{\mathrm{I}}, \qquad k_{\mathrm{cx}} = \frac{\omega}{c}\sqrt{n_{\mathrm{c}}^2 - n_{\mathrm{s}}^2\sin^2\varphi_{\mathrm{I}}} = \frac{\omega}{c} n_c \cos\varphi_{\mathrm{T}},}$$

As above, we can assume that k_{sx} is always real, because otherwise we have no incident wave. k_{cx} is real for $n_{\mathrm{c}} > n_{\mathrm{s}}\sin\varphi_{\mathrm{I}}$, but imaginary for $n_{\mathrm{c}} < n_{\mathrm{s}}\sin\varphi_{\mathrm{I}}$ (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

$$R_{\mathrm{TE}}=\frac{\left(k_{\mathrm{sx}}M_{22}-k_{\mathrm{cx}}M_{11}\right)-\mathrm{i}\left(M_{21}+k_{\mathrm{sx}}k_{\mathrm{cx}}M_{12}\right)}{\underbrace{\left(k_{\mathrm{sx}}M_{22}+k_{\mathrm{cx}}M_{11}\right)+\mathrm{i}\left(M_{21}-k_{\mathrm{sx}}k_{\mathrm{cx}}M_{12}\right)}},\qquad T_{\mathrm{TE}}=\frac{2k_{\mathrm{sx}}}{N_{\mathrm{TE}}}$$

$$\boxed{\begin{aligned}R_{\mathrm{TE}}&=\frac{\left(k_{\mathrm{sx}}-k_{\mathrm{cx}}\right)}{\left(k_{\mathrm{sx}}+k_{\mathrm{cx}}\right)}=\frac{n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}-\sqrt{n_{\mathrm{c}}^2-n_{\mathrm{s}}^2\sin^2\varphi_{\mathrm{I}}}}{n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}+\sqrt{n_{\mathrm{c}}^2-n_{\mathrm{s}}^2\sin^2\varphi_{\mathrm{I}}}}=\frac{n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}-n_{\mathrm{c}}\cos\varphi_{\mathrm{T}}}{n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}+n_{\mathrm{c}}\cos\varphi_{\mathrm{T}}}\\ T_{\mathrm{TE}}&=\frac{2k_{\mathrm{sx}}}{\left(k_{\mathrm{sx}}+k_{\mathrm{cx}}\right)}=\frac{2n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}}{n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}+\sqrt{n_{\mathrm{c}}^2-n_{\mathrm{s}}^2\sin^2\varphi_{\mathrm{I}}}}=\frac{2n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}}{n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}+n_{\mathrm{c}}\cos\varphi_{\mathrm{T}}}\end{aligned}}$$

$$\rho_{\mathrm{TE}}=\left|R_{\mathrm{TE}}\right|^2=\frac{\left|k_{\mathrm{sx}}-k_{\mathrm{cx}}\right|^2}{\left|k_{\mathrm{sx}}+k_{\mathrm{cx}}\right|^2}$$

$$\tau_{\mathrm{TE}}=\frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left|T_{\mathrm{TE}}\right|^2=\frac{4k_{\mathrm{sx}}\Re\left(k_{\mathrm{cx}}\right)}{\left|k_{\mathrm{sx}}+k_{\mathrm{cx}}\right|^2}.$$

$$\curvearrowright\ \rho_{\mathrm{TE}}+\tau_{\mathrm{TE}}=1$$

B) TM-Polarisation

$$R_{\mathrm{TM}}=\frac{\left(\varepsilon_{\mathrm{c}}k_{\mathrm{sx}}M_{22}-\varepsilon_{\mathrm{s}}k_{\mathrm{cx}}M_{11}\right)-\mathrm{i}\left(\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}}M_{21}+k_{\mathrm{sx}}k_{\mathrm{cx}}M_{12}\right)}{\underbrace{\left(\varepsilon_{\mathrm{c}}k_{\mathrm{sx}}M_{22}+\varepsilon_{\mathrm{s}}k_{\mathrm{cx}}M_{11}\right)+\mathrm{i}\left(\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}}M_{21}-k_{\mathrm{sx}}k_{\mathrm{cx}}M_{12}\right)}}\quad T_{\mathrm{TM}}=\frac{2\sqrt{\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}}}k_{\mathrm{sx}}}{N_{\mathrm{TM}}}.$$

with: $\hat{\mathbf{M}}=\hat{\mathbf{m}}(d=0)=\begin{pmatrix}1&0\\0&1\end{pmatrix}$

$$\boxed{\begin{aligned}R_{\mathrm{TM}}&=\frac{\left(k_{\mathrm{sx}}\varepsilon_{\mathrm{c}}-k_{\mathrm{cx}}\varepsilon_{\mathrm{s}}\right)}{\left(k_{\mathrm{sx}}\varepsilon_{\mathrm{c}}+k_{\mathrm{cx}}\varepsilon_{\mathrm{s}}\right)}=\frac{n_{\mathrm{s}}n_{\mathrm{c}}^2\cos\varphi_{\mathrm{I}}-n_{\mathrm{s}}^2\sqrt{n_{\mathrm{c}}^2-n_{\mathrm{s}}^2\sin^2\varphi_{\mathrm{I}}}}{n_{\mathrm{s}}n_{\mathrm{c}}^2\cos\varphi_{\mathrm{I}}+n_{\mathrm{s}}^2\sqrt{n_{\mathrm{c}}^2-n_{\mathrm{s}}^2\sin^2\varphi_{\mathrm{I}}}}=\frac{n_{\mathrm{c}}\cos\varphi_{\mathrm{I}}-n_{\mathrm{s}}\cos\varphi_{\mathrm{T}}}{n_{\mathrm{c}}\cos\varphi_{\mathrm{I}}+n_{\mathrm{s}}\cos\varphi_{\mathrm{T}}}\\ T_{\mathrm{TM}}&=\frac{2k_{\mathrm{sx}}\sqrt{\varepsilon_{\mathrm{c}}\varepsilon_{\mathrm{s}}}}{\left(k_{\mathrm{sx}}\varepsilon_{\mathrm{c}}+k_{\mathrm{cx}}\varepsilon_{\mathrm{s}}\right)}=\frac{2n_{\mathrm{s}}^2n_{\mathrm{c}}\cos\varphi_{\mathrm{I}}}{n_{\mathrm{s}}n_{\mathrm{c}}^2\cos\varphi_{\mathrm{I}}+n_{\mathrm{s}}^2\sqrt{n_{\mathrm{c}}^2-n_{\mathrm{s}}^2\sin^2\varphi_{\mathrm{I}}}}=\frac{2n_{\mathrm{s}}\cos\varphi_{\mathrm{I}}}{n_{\mathrm{c}}\cos\varphi_{\mathrm{I}}+n_{\mathrm{s}}\cos\varphi_{\mathrm{T}}},\end{aligned}}$$

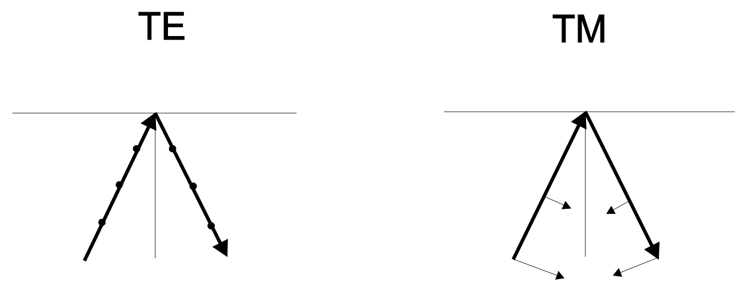
$$\rho_{\mathrm{TM}}=\left|R_{\mathrm{TM}}\right|^2=\frac{\left|k_{\mathrm{sx}}\varepsilon_{\mathrm{c}}-k_{\mathrm{cx}}\varepsilon_{\mathrm{s}}\right|^2}{\left|k_{\mathrm{sx}}\varepsilon_{\mathrm{c}}+k_{\mathrm{cx}}\varepsilon_{\mathrm{s}}\right|^2},$$

$$\tau_{\mathrm{TM}}=\frac{\Re\left(k_{\mathrm{cx}}\right)}{k_{\mathrm{sx}}}\left|T_{\mathrm{TM}}\right|^2=\frac{4k_{\mathrm{sx}}\Re\left(k_{\mathrm{cx}}\right)\varepsilon_{\mathrm{s}}\varepsilon_{\mathrm{c}}}{\left|k_{\mathrm{sx}}\varepsilon_{\mathrm{c}}+k_{\mathrm{cx}}\varepsilon_{\mathrm{s}}\right|^2}$$

$$\curvearrowright\ \rho_{\mathrm{TM}}+\tau_{\mathrm{TM}}=1$$

Remark

It may seem that we have a problem for $\varphi_{\mathrm{I}} = 0$. For $\varphi_{\mathrm{I}} = 0$, TE and TM polarization should be equivalent, because the fields are always polarized parallel to the interface. However, formally we have $R_{\mathrm{TE}} = -R_{\mathrm{TM}}$, $T_{\mathrm{TE}} = T_{\mathrm{TM}}$. 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.

$$\rho_{\mathrm{TE}} = \frac{|k_{\mathrm{sx}} - k_{\mathrm{cx}}|^2}{|k_{\mathrm{sx}} + k_{\mathrm{cx}}|^2} \qquad \rho_{\mathrm{TM}} = \frac{|k_{\mathrm{sx}}\varepsilon_{\mathrm{c}} - k_{\mathrm{cx}}\varepsilon_{\mathrm{s}}|^2}{|k_{\mathrm{sx}}\varepsilon_{\mathrm{c}} + k_{\mathrm{cx}}\varepsilon_{\mathrm{s}}|^2}$$

With $k_{\mathrm{cx}} = \frac{\omega}{c}\sqrt{n_{\mathrm{c}}^2 - n_{\mathrm{s}}^2 \sin^2 \varphi_{\mathrm{I}}}$ we can compute the smallest angle of incidence with $\rho_{\mathrm{TE,TM}} = 1$:

$$k_{\mathrm{cx}} = 0 \curvearrowright n_{\mathrm{c}} = n_{\mathrm{s}} \sin\varphi_{\mathrm{Itot}}$$

$$\boxed{\sin\varphi_{\mathrm{Itot}} = \frac{n_{\mathrm{c}}}{n_{\mathrm{s}}}.}$$

For angles of incidence larger than this limit angle, $\varphi_{\mathrm{I}} > \varphi_{\mathrm{Itot}}$ we have

$$k_{\mathrm{cx}} = \mathbf{i}\frac{\omega}{c}\sqrt{n_{\mathrm{s}}^2 \sin^2\varphi_{\mathrm{I}} - n_{\mathrm{c}}^2} = \mathbf{i}\mu_{\mathrm{c}} = \mathbf{i}\sqrt{k_z^2 - \frac{\omega^2}{c^2}\varepsilon_{\mathrm{c}}} \curvearrowright \text{imaginary}$$

$$\rightarrow \Re(k_{\mathrm{cx}}) = 0 \rightarrow \text{TIR}$$

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):

$$\rho_{\mathrm{TE}} = \frac{|k_{\mathrm{sx}} - \mathbf{i}\mu_{\mathrm{c}}|^2}{|k_{\mathrm{sx}} + \mathbf{i}\mu_{\mathrm{c}}|^2} = 1 \qquad \tau_{\mathrm{TE}} = \frac{4k_{\mathrm{sx}}\Re(k_{\mathrm{cx}})}{|k_{\mathrm{sx}} + k_{\mathrm{cx}}|^2} = 0.$$

Remark

For metals in visible range (below the plasma frequency) we have always TIR, because: $\Re(\varepsilon_c)<0 \rightarrow k_{cx}=\frac{\omega}{c}\sqrt{\varepsilon_c-n_s^2\sin^2\varphi_I}$ always imaginary.

In the case of TIR the modulus of the coefficient of reflection is one, but the coefficient itself is complex → nontrivial phase shift for reflected light:

A) TE-polarization

$$R_{TE}=1\cdot\exp(\mathbf{i}\Theta_{TE})=\frac{k_{sx}-\mathbf{i}\mu_c}{k_{sx}+\mathbf{i}\mu_c}=\frac{Z}{Z^*}=\frac{\exp(\mathbf{i}\alpha)}{\exp(-\mathbf{i}\alpha)}=\exp(2\mathbf{i}\alpha)$$

$$\curvearrowright \tan\alpha=\tan\frac{\Theta_{TE}}{2}=-\frac{\mu_c}{k_{sx}}=-\frac{\sqrt{n_s^2\sin^2\varphi_I-n_c^2}}{n_s\cos\varphi_I}=-\frac{\sqrt{\sin^2\varphi_I-\sin^2\varphi_{Itot}}}{\cos\varphi_I}.$$

B) TM-polarization

$$R_{TM}=1\cdot\exp(\mathbf{i}\Theta_{TM})=\frac{k_{sx}\varepsilon_c-\mathbf{i}\mu_c\varepsilon_s}{k_{sx}\varepsilon_c+\mathbf{i}\mu_c\varepsilon_s}=\frac{Z}{Z^*}=\frac{\exp(\mathbf{i}\alpha)}{\exp(-\mathbf{i}\alpha)}=\exp(2\mathbf{i}\alpha)$$

$$\curvearrowright \tan\alpha=\tan\frac{\Theta_{TM}}{2}=-\frac{\mu_c\varepsilon_s}{k_{sx}\varepsilon_c}=\frac{\varepsilon_s}{\varepsilon_c}\tan\frac{\Theta_{TE}}{2},$$

In conclusion, we have seen that the phase shifts of the reflected light at TIR is different for TE and TM polarization, and because $\varepsilon_s>\varepsilon_c$

$$\left|\Theta_{TM}\right|>\left|\Theta_{TE}\right|,$$

As a consequence, incident linearly polarized light gets generally elliptically polarized after TIR $\curvearrowright$ Fresnel prism

Remark

The field in the cladding is evanescent $\sim\exp(\mathbf{i}k_{xc}x)=\exp(-\mu_c x)$.

$\curvearrowright$ The averaged energy flux in the cladding normal to the interface vanishes.

$$\langle\mathbf{S}\rangle_x=\frac{1}{2\omega\mu_0}\Re\left(\mathbf{k}|\mathbf{E}|^2\right)_x=\frac{1}{2\omega\mu_0}\underbrace{\Re(k_x)}_{=0}|\mathbf{E}|^2=0.$$

7.3.2.3 The Brewster angle

There exists another special angle with particular reflection properties. For TM-polarization, for incident light at the Brewster angle φ_B we find $R_{TM}=0$:

$$\rho_{TM}=\frac{\left|k_{sx}\varepsilon_c-k_{cx}\varepsilon_s\right|^2}{\left|k_{sx}\varepsilon_c+k_{cx}\varepsilon_s\right|^2}=0,$$

$$\rightarrow \; k_{\mathbf{sx}}\varepsilon_{\mathbf{c}} = k_{\mathbf{cx}}\varepsilon_{\mathbf{s}}$$

$$\varepsilon_{\mathbf{c}}^2\left(\varepsilon_{\mathbf{s}} - \sin^2\varphi_{\mathbf{B}}\varepsilon_{\mathbf{s}}\right) = \varepsilon_{\mathbf{s}}^2\left(\varepsilon_{\mathbf{c}} - \sin^2\varphi_{\mathbf{B}}\varepsilon_{\mathbf{s}}\right)$$

$$\sin^2\varphi_{\mathbf{B}} = \frac{\varepsilon_{\mathbf{s}}\varepsilon_{\mathbf{c}}\left(\varepsilon_{\mathbf{s}} - \varepsilon_{\mathbf{c}}\right)}{\varepsilon_{\mathbf{s}}\left(\varepsilon_{\mathbf{s}}^2 - \varepsilon_{\mathbf{c}}^2\right)} = \frac{\varepsilon_{\mathbf{c}}}{\left(\varepsilon_{\mathbf{s}} + \varepsilon_{\mathbf{c}}\right)}$$

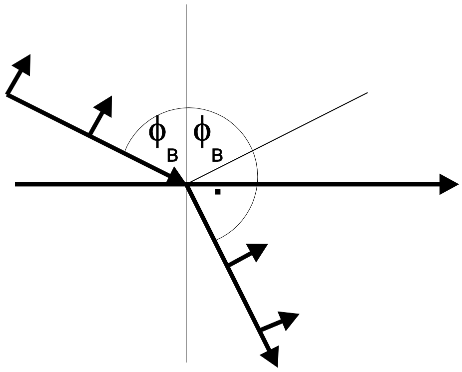
$$\cos^2\varphi_{\mathbf{B}} = 1 - \sin^2\varphi_{\mathbf{B}} = 1 - \frac{\varepsilon_{\mathbf{c}}}{\left(\varepsilon_{\mathbf{s}} + \varepsilon_{\mathbf{c}}\right)} = \frac{\varepsilon_{\mathbf{s}}}{\left(\varepsilon_{\mathbf{s}} + \varepsilon_{\mathbf{c}}\right)}$$

With the last two lines we can write the final result for the Brewster angle:

$$\boxed{\tan\varphi_{\mathbf{B}} = \sqrt{\frac{\varepsilon_{\mathbf{c}}}{\varepsilon_{\mathbf{s}}}}.}$$

The Brewster angle exists only for TM polarization, but for any $n_{\mathbf{s}} \lessgtr n_{\mathbf{c}}$.

There is a simple physical interpretation, why there is no reflection at the interface for the Brewster angle.



$$\tan\varphi_{\mathbf{B}} = \frac{\sin\varphi_{\mathbf{B}}}{\cos\varphi_{\mathbf{B}}} = \frac{n_{\mathbf{c}}}{n_{\mathbf{s}}}$$

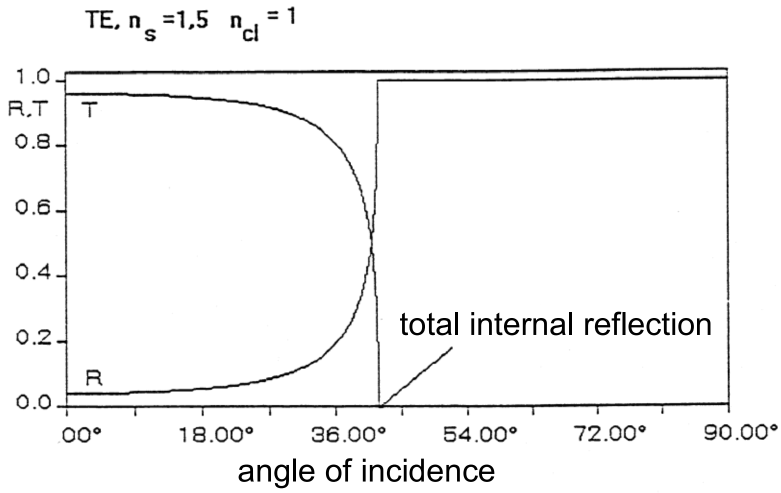
$$\curvearrowright n_{\mathbf{s}}\sin\varphi_{\mathbf{B}} = n_{\mathbf{c}}\cos\varphi_{\mathbf{B}} = n_{\mathbf{c}}\sin\left(\frac{\pi}{2} - \varphi_{\mathbf{B}}\right)$$

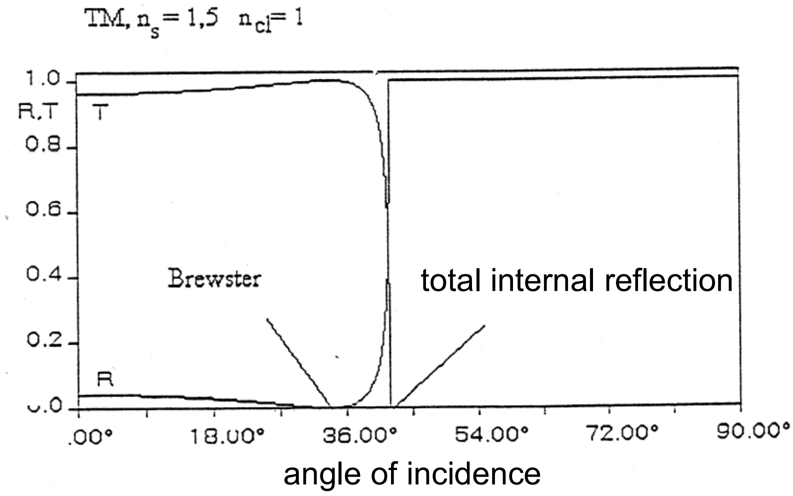
At the same time the angle of the transmitted light is always

$$n_{\mathbf{s}}\sin\varphi_{\mathbf{B}} = n_{\mathbf{c}}\sin\varphi_{\mathbf{T}} \curvearrowright \varphi_{\mathbf{T}} = \frac{\pi}{2} - \varphi_{\mathbf{B}},$$

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).

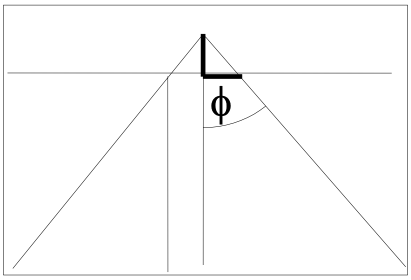
In summary, we have the following results for reflectivity and transmittivity at a single interface with $\varepsilon_s > \varepsilon_c$.





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:

$$E_{\mathrm{I}}(x,z)=E_{\mathrm{I}}\exp\left[\mathbf{i}\left(k_{\mathbf{sx}}x+k_{\mathbf{z}}z\right)\right]\rightarrow E_{\mathrm{I}}(x,z)=E_{\mathrm{I}}\exp\left[\mathbf{i}\left(\alpha z+\gamma_{\mathbf{s}}x\right)\right]$$

with

$$\alpha = \frac{\omega}{c} n_{\mathbf{s}} \sin\varphi_{\mathbf{I}}, \qquad \gamma_{\mathbf{s}} = \sqrt{\frac{\omega^2}{c^2} n_{\mathbf{s}}^2 - \alpha^2}.$$

This gives rise to a reflected plane wave as

$$E_{\mathbf{R}}(x,z) = E_{\mathbf{I}} \exp\left[\mathbf{i}\left(\alpha z - \gamma_{\mathbf{s}} x\right)\right] \exp\left[\mathbf{i}\Theta\left(\alpha\right)\right].$$

The reflected plane wave gets a phase shift $\Theta(\alpha)$, which depends on the angle of incidence (here characterized by the transverse wave number α).
Now we want to treat beams, which we can write as a superposition of plane waves (Fourier amplitude $e_{\mathbf{I}}(\alpha)$):

$$E_{\mathbf{I}}(x,z) = \int d\alpha e_{\mathbf{I}}(\alpha) \exp\left[\mathbf{i}\left(\alpha z + \gamma_{\mathbf{s}}(\alpha) x\right)\right]$$

We assume a mean angle of incidence $\varphi_{\mathbf{I}0}$:

$\alpha_0 = \frac{\omega}{c} n_{\mathbf{s}} \sin\varphi_{\mathbf{I}0}$ mean angular frequency

$\alpha = \alpha_0 + \varepsilon$

In the Fourier integral, we have to integrate over angular frequencies with non-zero amplitudes $e_{\mathbf{I}}(\alpha)$ only ($e_{\mathbf{I}}(\alpha_0 + \varepsilon) \neq 0$ for $-\Delta \leq \varepsilon \leq \Delta$ only)

$$E_{\mathbf{I}}(x,z) = \exp\left(\mathbf{i}\alpha_0 z\right) \int_{-\Delta}^{\Delta} d\varepsilon e_{\mathbf{I}}\left(\alpha_0 + \varepsilon\right) \exp\left[\mathbf{i}\left(\varepsilon z + \gamma_{\mathbf{s}} x\right)\right]$$

$$E_{\mathbf{R}}(x,z) = \exp\left(\mathbf{i}\alpha_0 z\right) \int_{-\Delta}^{\Delta} d\varepsilon e_{\mathbf{I}}\left(\alpha_0 + \varepsilon\right) \exp\left[\mathbf{i}\left(\varepsilon z - \gamma_{\mathbf{s}} x\right)\right] \exp\left[\mathbf{i}\Theta\left(\alpha_0 + \varepsilon\right)\right].$$

Let us make the following assumptions:

– small divergence of beam (narrow spectrum, $\Delta \ll \frac{\omega}{c} n_{\mathbf{s}}$)

– all Fourier component undergo TIR ($\Theta(\alpha) > \Theta_{Itot}$)

Then, it is justified to expand the phase shift $\Theta(\alpha)$ of the reflected wave into a Taylor series up to first order:

$$\Theta\left(\alpha_0 + \varepsilon\right) \approx \Theta\left(\alpha_0\right) + \left.\frac{\partial\Theta}{\partial\alpha}\right|_{\alpha_0} \varepsilon = \Theta\left(\alpha_0\right) + \Theta'\varepsilon$$

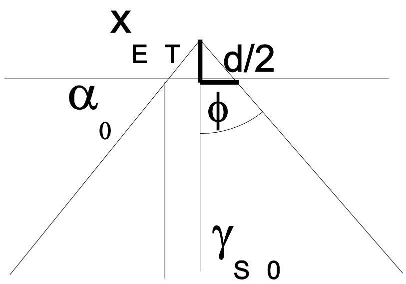
Then, the reflected beam at the interface at $x = 0$ is given as:

$$E_{\mathbf{R}}(0,z) = \exp\left\{\mathbf{i}\left[\alpha_0 z + \Theta\left(\alpha_0\right)\right]\right\} \int_{-\Delta}^{\Delta} d\varepsilon e_{\mathbf{I}}\left(\alpha_0 + \varepsilon\right) \exp\left[\mathbf{i}\left(z + \Theta'\right)\varepsilon\right]$$

We can identify the remaining integral as a shifted version of the incident beam profile at $x = 0$

$$\boxed{E_{\mathbf{R}}(0,z) = \exp\left\{\mathbf{i}\left[\Theta\left(\alpha_0\right) - \alpha_0\Theta'\right]\right\} E_{\mathbf{I}}\left(z + \Theta'\right).}$$

Thus, the reflected beam appears shifted by $d = -\Theta'$ (Goos-Hänchen Shift).



Let us finally compute the shift $d = -\Theta'$. We know from before that the phase shift for TIR is given as:

$$\tan\frac{\Theta_{\mathrm{TE}}}{2} = -\frac{\mu_{\mathrm{c}}}{k_{\mathrm{sx}}} = -\frac{\mu_{\mathrm{c}}}{\gamma_{\mathrm{s}}}$$

$$\Theta = -2\arctan\frac{\mu_c}{k_{sx}} = -2\arctan\frac{\sqrt{\alpha^2 - \frac{\omega^2}{c^2}n_{\mathrm{c}}^2}}{\sqrt{\frac{\omega^2}{c^2}n_{\mathrm{s}}^2 - \alpha^2}}$$

$$\Theta' = \frac{\partial\Theta}{\partial\alpha} = -2\times\frac{1}{1+\frac{\mu_c^2}{k_{sx}^2}}\times\frac{\frac{2\alpha}{2\mu_c}k_{sx} - \frac{-2\alpha}{2k_{sx}}\mu_c}{k_{sx}^2} = -2\frac{\frac{\alpha\left(k_{sx}^2+\mu_c^2\right)}{\mu_c k_{sx}}}{k_{sx}^2+\mu_c^2} = -2\frac{1}{\mu_c}\frac{\alpha}{k_{sx}}$$

$\boxed{\Theta'\big|_{\alpha_0} = -d = -2x_{\mathrm{ET}}\tan\varphi_{\mathrm{IO}}}$ with $x_{\mathrm{ET}} = \frac{1}{\mu_{\mathrm{c}}} = \frac{1}{\sqrt{\alpha_0^2 - \frac{\omega^2}{c^2}n_{\mathrm{c}}^2}}$ and $\tan\varphi_{\mathrm{IO}} = \frac{\alpha}{k_{sx}} = \frac{\alpha}{\gamma_s}$

→ x_{ET} 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 these multi-layer systems become periodic, 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). Furthermore periodic structures are of general importance in physics (lattices, crystals, atomic chains …). We can learn many things about the general features of such periodic systems by looking at the optical properties of periodic (dielectric) multi-layer systems

In our theoretical approach, we will assume these layer systems as infinite, i.e. consisting of an infinite number of layers, and we treat them as so-called one-dimensional photonic crystals. We will discuss effects like band gaps, dispersion and diffraction in such periodic media, and gain understanding of the basics of Bragg reflection and the physics of photonic crystals.

In order to keep things simple we will treat:

- semi-infinite periodic multi-layer systems $\left[x>0,\left(\varepsilon_1,d_1\right),\left(\varepsilon_2,d_2\right)\right]$
- TE-polarization only
- monochromatic light

At the interface between substrate and Bragg-mirror $\left(x=0\right)$ we have incident and reflected electric field:

$$E_0=E_\mathrm{R}+E_\mathrm{I} \quad \text{and} \quad \left.\frac{\partial E}{\partial x}\right|_0=E_0'=ik_\mathrm{sx}\left(E_I-E_R\right)$$

Alternatively the incident and reflected field can be expressed by the field at the interface and its derivative as

$$E_\mathrm{I}=\frac{E_0}{2}-\frac{iE_0'}{2k_\mathrm{sx}} \quad \text{and} \quad E_\mathrm{R}=\frac{E_0}{2}+\frac{iE_0'}{2k_\mathrm{sx}}$$

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

$$i\omega\mu_0H_\mathrm{z}=G=\frac{\partial E}{\partial x}=E'.$$

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 ε_1 and ε_2 having a thickness of d_1 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:

$$E(x,z;\omega)=\exp\left\{\mathbf{i}\left[k_\mathbf{x}\left(k_\mathbf{z},\omega\right)x+k_\mathbf{z}z\right]\right\}E_{k_\mathbf{x}}(x)$$

with

$$E_{k_\mathbf{x}}(x+\Lambda)=E_{k_\mathbf{x}}(x)$$

In other words, $E_{k_\mathbf{x}}(x)$ is a periodic function and we are looking for solutions $E(x,z;\omega)$ which have the same amplitude after one period of the medium, but we allow for a different phase $\sim\exp\left\{\mathbf{i}\left[k_\mathbf{x}\left(k_\mathbf{z},\omega\right)x\right]\right\}$. Here $k_\mathbf{x}$ 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 $k_{\mathbf{x}}(k_{\mathbf{z}},\omega)$, in complete analogy to the DR for plane waves $k_{\mathbf{x}}^2 = \frac{\omega^2}{c^2}\varepsilon(\omega) - k_{\mathbf{z}}^2$ in homogeneous media. In order to make the difference to the homogeneous case more obvious, we change the notation for the Bloch vector: $k_x \rightarrow K$.

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 $N+1$):

$$\begin{pmatrix} E \\ E' \end{pmatrix}_{(\mathbf{N+1})\Lambda} = \exp(\mathbf{i}K\Lambda)\begin{pmatrix} E \\ E' \end{pmatrix}_{\mathbf{N}\Lambda}.$$

On the other hand, we know from our matrix method:

$$\begin{pmatrix} E \\ E' \end{pmatrix}_{(\mathbf{N+1})\Lambda} = \hat{\mathbf{M}}\begin{pmatrix} E \\ E' \end{pmatrix}_{\mathbf{N}\Lambda}$$

with $\hat{\mathbf{M}} = \hat{\mathbf{m}}(d_2)\hat{\mathbf{m}}(d_1) \rightarrow M_{ij} = \sum_k m_{ik}^{(2)} m_{kj}^{(1)}$

If the Bloch wave is a solution to our problem, we can set the two expressions equal:

$$\left\{\hat{\mathbf{M}} - \exp(\mathbf{i}K\Lambda)\hat{\mathbf{I}}\right\}\begin{pmatrix} E \\ E' \end{pmatrix}_{\mathbf{N}\Lambda} = 0$$

and with $\mu = \exp(\mathbf{i}K\Lambda)$ we have to solve the following eigenvalue problem:

$$\left\{\hat{\mathbf{M}} - \mu(K)\hat{\mathbf{I}}\right\}\begin{pmatrix} E \\ E' \end{pmatrix}_{\mathbf{N}\Lambda} = 0$$

This eigenvalue problem determines the Bloch vector K and will finally give our dispersion relation.

As usual, we use the solvability condition $\det\{\hat{\mathbf{M}} - \mu\hat{\mathbf{I}}\} = 0$ to compute the dispersion relation expressed in $\mu = \exp(\mathbf{i}K\Lambda)$. Hence we still need to compute K afterwards.

$$\boxed{\mu_{\pm} = \exp(\mathbf{i}K_{\pm}\Lambda) = \frac{(M_{11}+M_{22})}{2} \pm \sqrt{\left\{\frac{(M_{11}+M_{22})}{2}\right\}^2 - 1}.}$$

Note that we used $\det\left\{\hat{\mathbf{M}}\right\} = 1$, which explains why the off-diagonal elements of the matrix do not appear in the formula. Moreover, because of $\det\left\{\hat{\mathbf{M}}\right\} = 1$ we have $\mu_+\mu_- = 1$.

The corresponding eigenvectors (field and its derivative at $x = N\Lambda$) can be computed from

$$\left\{\hat{\mathbf{M}} - \exp\left(\mathrm{i}K\Lambda\right)\hat{\mathbf{I}}\right\}\begin{pmatrix} E \\ E' \end{pmatrix}_{\mathrm{N}\Lambda} = 0$$

$$\rightarrow \begin{pmatrix} M_{11} - \mu & M_{12} \\ M_{21} & M_{22} - \mu \end{pmatrix}\begin{pmatrix} E \\ E' \end{pmatrix}_{N\Lambda} = 0$$

From the first row the following condition can be derived:

$$\left(M_{11} - \mu\right)E + M_{12}E' = 0$$

Since the investigated system is linear and invariant for the phase, the absolute amplitude and phased of the E-component of the eigenvector can be chosen arbitrarily. Here we take $E = 1$ and get for the full eigenvector

$$\begin{pmatrix} E \\ E' \end{pmatrix}_{\mathrm{N}\Lambda} = \begin{Bmatrix} 1 \\ \left(\mu - M_{11}\right)/M_{12} \end{Bmatrix} E_{\mathrm{N}\Lambda}.$$

If field values of the Bloch mode, i.e. the function $E_{k_x}(x+\Lambda) = E_{k_x}(x)$, inside the layers are desired, they can be computed by using the matrix formalism and the above eigenvector $(E, E')_{N\Lambda}$.

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

$$\boxed{\rho = \left|\frac{E_{\mathrm{R}}}{E_{\mathrm{I}}}\right|^2} \quad \text{with } E_{\mathrm{R}} = \frac{E_0}{2} + \frac{iE_0'}{2k_{\mathrm{sx}}} \quad \text{and } E_{\mathrm{I}} = \frac{E_0}{2} - \frac{iE_0'}{2k_{\mathrm{sx}}} \quad \text{from before}$$

$$\rightarrow \boxed{\rho = \left|\frac{k_{\mathrm{sx}}E_0 + \mathrm{i}E_0'}{k_{\mathrm{sx}}E_0 - \mathrm{i}E_0'}\right|^2}$$

With our knowledge of the eigenvector from above we can compute the reflectivity:

$$E'_0 = \frac{\mu - M_{11}}{M_{12}}E_0 \rightarrow \boxed{\rho = \left|\frac{k_{\mathrm{sx}}E_0 + \mathrm{i}E_0'}{k_{\mathrm{sx}}E_0 - \mathrm{i}E_0'}\right|^2 = \left|\frac{k_{\mathrm{sx}} + \mathrm{i}\dfrac{\mu - M_{11}}{M_{12}}}{k_{\mathrm{sx}} - \mathrm{i}\dfrac{\mu - M_{11}}{M_{12}}}\right|^2}$$

According to this formula two scenarios are possible:

A) total internal reflection ρ = 1

Hence μ has to be real which results in the condition

$$\rightarrow \boxed{\left|\frac{\left(M_{11}+M_{22}\right)}{2}\right| \geq 1}$$

with [for our example $(\varepsilon_1,\varepsilon_2,d_1,d_2)$]

$$M_{11} = \cos\left(k_{1\mathbf{x}}d_1\right)\cos\left(k_{2\mathbf{x}}d_2\right) - \frac{k_{2\mathbf{x}}}{k_{1\mathbf{x}}}\sin\left(k_{1\mathbf{x}}d_1\right)\sin\left(k_{2\mathbf{x}}d_2\right)$$

$$M_{22} = \cos\left(k_{1\mathbf{x}}d_1\right)\cos\left(k_{2\mathbf{x}}d_2\right) - \frac{k_{1\mathbf{x}}}{k_{2\mathbf{x}}}\sin\left(k_{1\mathbf{x}}d_1\right)\sin\left(k_{2\mathbf{x}}d_2\right).$$

This defines the so-called band gap, i.e. frequencies of excitation for which no propagating solutions exist.

B) propagating normal modes

Hence μ must be complex which results in the condition

$$\rightarrow \boxed{\left|\frac{\left(M_{11}+M_{22}\right)}{2}\right| < 1}$$

We can compute the explicit dispersion relation:

$$\mu = \exp\left(\mathbf{i}K\Lambda\right) = \frac{\left(M_{11}+M_{22}\right)}{2} \pm \sqrt{\left\{\frac{\left(M_{11}+M_{22}\right)}{2}\right\}^2 - 1}.$$

$$\mu = \exp\left(\mathbf{i}K\Lambda\right) = \cos\left(K\Lambda\right) + \mathbf{i}\sin\left(K\Lambda\right) = \cos\left(K\Lambda\right) \pm \sqrt{\left\{\cos\left(K\Lambda\right)\right\}^2 - 1}$$

$$\rightarrow \boxed{\cos\left\{K\left(k_{\mathbf{z}},\omega\right)\Lambda\right\} = \frac{\left(M_{11}+M_{22}\right)}{2}}$$

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 $k_{\mathbf{x}}^2 = \frac{\omega^2}{c^2}\varepsilon(\omega) - k_{\mathbf{z}}^2$.

Interpretation

- For the case of total internal reflection (μ real, $\left|\frac{\left(M_{11}+M_{22}\right)}{2}\right| \geq 1$) the Bloch vector K is complex, $\mu = \exp\left(\mathbf{i}K\Lambda\right) = \exp\left(\mathbf{i}\Re\left(K\right)\Lambda\right)\exp\left(-\Im\left(K\right)\Lambda\right)$.
  Hence $\Re\left\{K\left(k_{\mathbf{z}},\omega\right)\Lambda\right\} = n\pi$ and

$$\Im\left\{K\left(k_{\mathbf{z}},\omega\right)\Lambda\right\} = -\ln\left(\left(-1\right)^n\left\{\frac{\left(M_{11}+M_{22}\right)}{2} \pm \sqrt{\left\{\frac{\left(M_{11}+M_{22}\right)}{2}\right\}^2 - 1}\right\}\right).$$

The ± 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 $n=1...\infty$. 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

$$\Re\left\{K\left(k_z,\omega\right)\Lambda\right\}=n\pi \text{ and } \Im\left\{K\left(k_z,\omega\right)\Lambda\right\}=0$$

$$\curvearrowright K\left(k_z,\omega\right)=n\pi/\Lambda$$

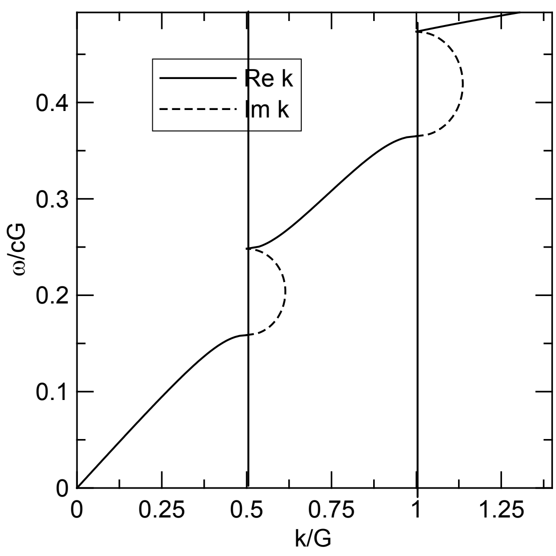
- Outside the band gaps, i.e., inside the bands, we find propagating solutions which have different properties than the normal modes in homogeneous media (different dispersion relation).
  → We can exploit the strong curvature, i.e. frequency dependence, of DR for, e.g., dispersion compensation or diffraction free propagation

Special case: normal incidence

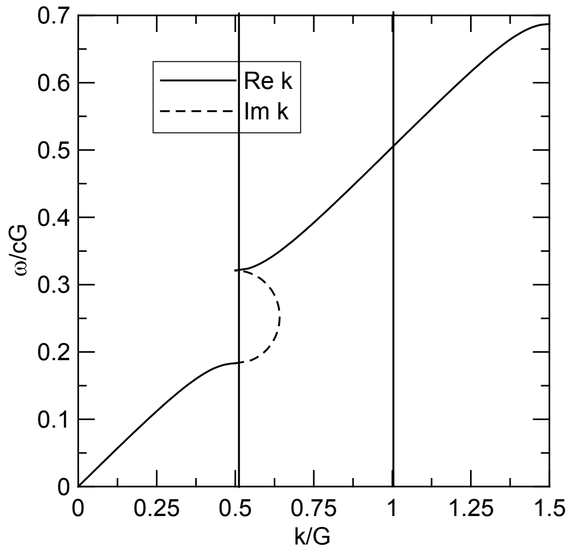
In general there is a complex interplay between the angle of incidence and frequency of light determining the reflection properties of multilayer systems. Therefore let us have a look at the simpler case of normal incidence $\left(k_z=0\right)$.
In a graphical representation of the dispersion relation for $k_z=0$ it is common to use the following dimensionless quantities

$$\frac{\omega}{cG} \text{ and } \frac{K}{G} \text{ with the scaling constant } G=\frac{2\pi}{\Lambda}$$

Examples for normal incidence

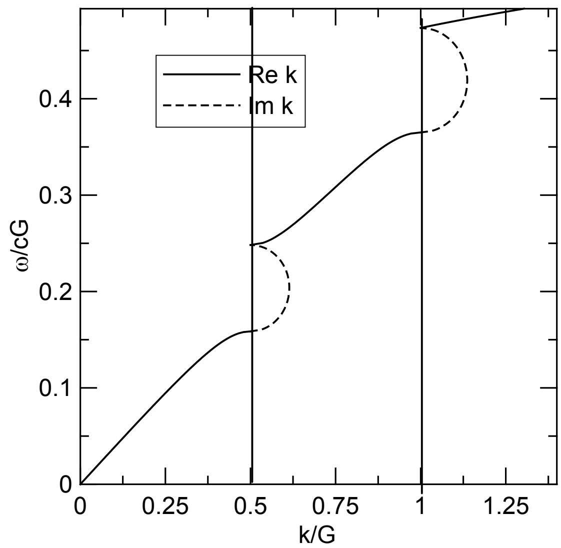


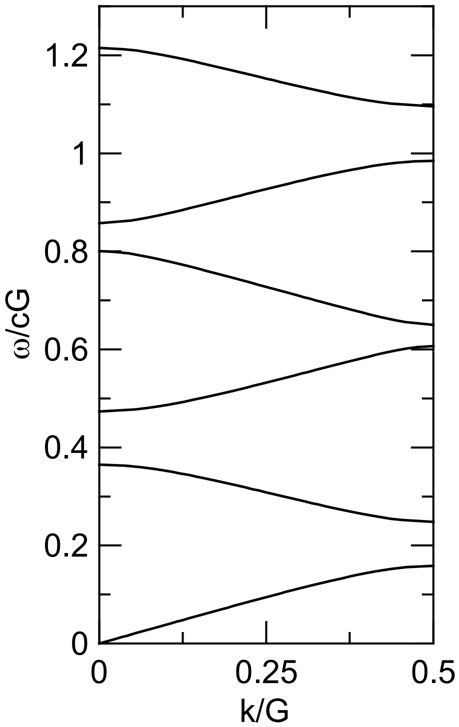
n_1=1.4, d_1=0.5Λ
n_2=3.4, d_2=0.5Λ



n_1=1.4, d_1= 17/24 Λ
n_2=3.4 , d_2 =7/24 Λ

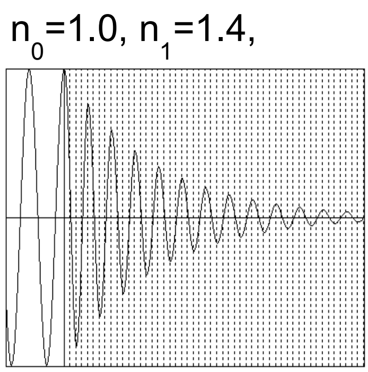
It is common to use the reduced band structure - Brillouin zone, where the information for all possible Bloch vectors is mapped onto the Bloch vectors in the following interval $-0.5\leq(k/G)\leq 0.5$.



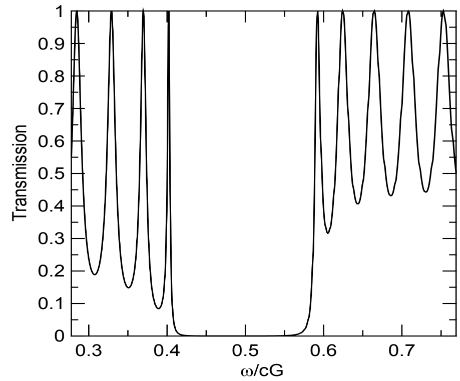


n_1=1.4, d_1=0.5Λ
n_2=3.4, d_2=0.5Λ

Because of $e^{iK\Lambda}$ we need only $-\pi \le K\Lambda \le \pi$ → $\left|\frac{K}{G}\right| \le 0.5$ to describe the dispersion relation.
Inside the band gap, we find damped solutions:



field in a Bragg-mirror



Transmission

Let us quantify the damping. In our example (n_1,n_2,d_1,d_2) we have

$$\frac{(M_{11}+M_{22})}{2} = \cos\left(\frac{\omega}{c}n_1 d_1\right)\cos\left(\frac{\omega}{c}n_2 d_2\right) - \frac{1}{2}\left(\frac{n_2}{n_1}+\frac{n_1}{n_2}\right)\sin\left(\frac{\omega}{c}n_1 d_1\right)\sin\left(\frac{\omega}{c}n_2 d_2\right)$$

In the middle of the first band gap (optimum configuration for high reflection) we have $\frac{\omega_B}{c}n_1 d_1 = \frac{\omega_B}{c}n_2 d_2 = \frac{\pi}{2}$, with ω_B being the Bragg frequency, and therefore $\frac{(M_{11}+M_{22})}{2} = -\frac{1}{2}\left(\frac{n_2}{n_1}+\frac{n_1}{n_2}\right) < -1$. If we plug this (for $n=1$) in our expression for $\Im(K)$ and assume a small index contrast $|n_2 - n_1| << (n_2 + n_1)$ we find

$$\boxed{\Lambda\Im(K)_{\max} \approx 2\frac{n_2 - n_1}{n_2 + n_1}}$$ (do derivation as an exercise)

$\curvearrowright$ Damping is proportional to index contrast of the subsequent layers $|n_2 - n_1|$

The spectral width of the gap $\left(\left|\frac{(M_{11} + M_{22})}{2}\right| \geq 1\right)$ is then

$$\boxed{\Delta\omega_{\mathrm{gap}} \approx \frac{2\omega_{\mathrm{B}}}{\pi}\Lambda\Im(K)_{\max}}$$ (do derivation as an exercise)

$\curvearrowright$ Spectral width is proportional to index contrast as well.

7.3.4 Fabry-Perot-resonators

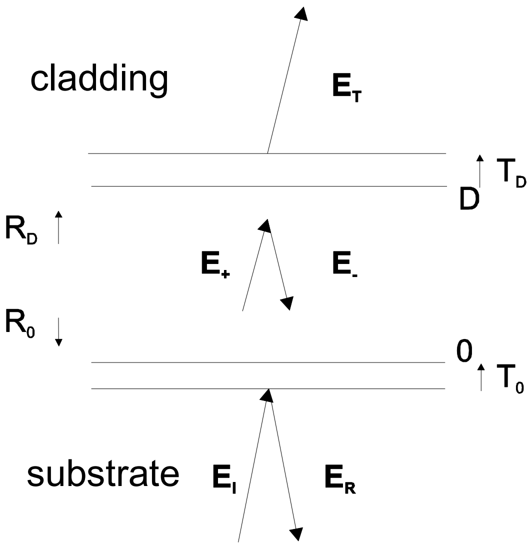
In this chapter we will treat a special multi-layer system, the so-called Fabry-Perot-resonator. To construct a Fabry-Perot-resonator, one can start from highly reflecting periodic multi-layer system (Bragg reflector). If one changes just a single layer somewhere in the middle of the otherwise periodic 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 → Fabry-Perot-Resonator with active medium inside the cavity
- nonlinear optics → high intensities inside the cavity → 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.



- $E_{\mathtt{I}}, E_{\mathtt{R}}$ and $E_{\mathtt{T}}$, → amplitudes of incident, reflected and transmitted external fields in substrate and cladding
- E_+, E_- → amplitudes of internal fields running forward and backward inside the cavity

Using the known coefficients of reflection and transmission of the two mirrors, we can eliminate E_+ and E_- by connecting the field amplitudes:

A) At the lower mirror inside the cavity:

$$T_0 E_{\mathtt{I}} + R_0 E_-(0) = E_+(0)$$

B) At the upper mirror outside the cavity:

$$E_{\mathtt{T}} = T_{\mathtt{D}} E_+(D)$$

And with $E_+(D) = E_+(0)\exp\left(\mathrm{i}k_{\mathtt{fx}}D\right)$ $\boxed{E_+(0) = \frac{E_{\mathtt{T}}}{T_{\mathtt{D}}}\exp\left(-\mathrm{i}k_{\mathtt{fx}}D\right)}$

C) At the upper mirror inside the cavity:

$$E_-(D) = R_{\mathtt{D}} E_+(D) = \frac{R_{\mathtt{D}}}{T_{\mathtt{D}}} E_{\mathtt{T}}$$

and with $E_-(0) = E_-(D)\exp\left(\mathrm{i}k_{\mathtt{fx}}D\right)$

$$\boxed{E_-(0) = \frac{R_{\mathtt{D}}}{T_{\mathtt{D}}} E_{\mathtt{T}} \exp\left(\mathrm{i}k_{\mathtt{fx}}D\right)}$$

D) we substitute $E_+(0)$ and $E_-(0)$ in A)

$$T_0 E_{\mathtt{I}} + R_0 E_-(0) = E_+(0)$$

$$T_0 E_{\mathtt{I}} + R_0 \frac{R_{\mathtt{D}}}{T_{\mathtt{D}}} E_{\mathtt{T}} \exp\left(\mathrm{i}k_{\mathtt{fx}}D\right) = \frac{E_{\mathtt{T}}}{T_{\mathtt{D}}}\exp\left(-\mathrm{i}k_{\mathtt{fx}}D\right)$$

$$\rightarrow\ E_{\mathtt{I}} = \frac{1}{T_0 T_{\mathtt{D}}}\left\{\exp\left(-\mathrm{i}k_{\mathtt{fx}}D\right) - R_0 R_{\mathtt{D}} \exp\left(\mathrm{i}k_{\mathtt{fx}}D\right)\right\} E_{\mathtt{T}}.$$

Thus, the coefficient of transmission for the whole FP-resonator expressed by the coefficients of the mirrors (R_0, T_0, R_D, and T_D), the cavity properties and the angle of incidence (D, $k_{\mathrm{fx}}=\sqrt{\frac{\omega^2}{c^2}\varepsilon_f-k_z^2}$) reads:

$$\boxed{T_{\mathrm{TE}}=\frac{E_{\mathrm{T}}}{E_{\mathrm{I}}}=\frac{T_0T_{\mathrm{D}}\exp\left(\mathrm{i}k_{\mathrm{fx}}D\right)}{1-R_0R_{\mathrm{D}}\exp\left(2\mathrm{i}k_{\mathrm{fx}}D\right)}.}$$

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 $(R,T)_{0,\mathrm{D}}=(|R|,|T|)_{0,\mathrm{D}}\exp\left(\mathrm{i}\varphi_{0,\mathrm{D}}^{\mathrm{R,T}}\right)$. Obviously, only the phase shifts induced by the coefficients of reflection R_0,R_{D} are important for the transmissivity of the FP resonator $\tau_{\mathrm{TE}}\sim|T_{\mathrm{TE}}|^2$.

For given $|R_0|,|R_{\mathrm{D}}|$, $|T_0|,|T_{\mathrm{D}}|$ and $\varphi_0^{\mathrm{R}},\varphi_D^{\mathrm{R}}$, the general transmissivity of a lossless Fabry-Perot resonator reads:

$$\boxed{\begin{aligned}|T_{\mathrm{TE}}|^2=|T|^2&=\frac{|T_0|^2|T_{\mathrm{D}}|^2}{1+|R_0|^2|R_{\mathrm{D}}|^2-2|R_0||R_{\mathrm{D}}|\cos(\underbrace{2k_{fx}D+\varphi_0^{\mathrm{R}}+\varphi_D^{\mathrm{R}}}_{\delta})}\\ \tau&=\frac{k_{\mathrm{cx}}}{k_{\mathrm{sx}}}|T|^2.\end{aligned}}$$

Here we introduced the phase-shift δ, 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

$$\tau=\frac{k_{\mathrm{cx}}}{k_{\mathrm{sx}}}\frac{|T_0|^2|T_{\mathrm{D}}|^2}{1+|R_0|^2|R_{\mathrm{D}}|^2-2|R_0||R_{\mathrm{D}}|\cos\delta}$$

Because we assume no losses we can use energy conservation at each mirror to eliminate $T_{0,\mathrm{D}}$:

$$\begin{aligned}|T_0|^2|T_{\mathrm{D}}|^2&=\frac{k_{\mathrm{sx}}}{k_{fx}}\left(1-|R_0|^2\right)\frac{k_{fx}}{k_{\mathrm{cx}}}\left(1-|R_{\mathrm{D}}|^2\right)\\&=\frac{k_{\mathrm{sx}}}{k_{\mathrm{cx}}}\left(1-|R_0|^2\right)\left(1-|R_{\mathrm{D}}|^2\right)\end{aligned}$$

Note: τ and ρ for a lossless mirror are the same for both sides of the mirror. For lossy mirrors only τ is the same, ρ is then side-dependent.

For discussing the effect of the phase shift δ 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

$$\tau = \frac{\left(1-|R_0|^2\right)\left(1-|R_D|^2\right)}{1+|R_0|^2|R_D|^2 - 2|R_0||R_D|\left(1-2\sin^2\frac{\delta}{2}\right)}$$

$$= \left\{\frac{\left(1-|R_0||R_D|\right)^2}{\left(1-|R_0|^2\right)\left(1-|R_D|^2\right)} + \frac{4|R_0||R_D|}{\left(1-|R_0|^2\right)\left(1-|R_D|^2\right)}\sin^2\frac{\delta}{2}\right\}^{-1}.$$

and with

$$\rho_0 = |R_0|^2, \rho_D = |R_D|^2$$

$$\boxed{\tau = \left\{\frac{\left(1-\sqrt{\rho_0\rho_D}\right)^2}{(1-\rho_0)(1-\rho_D)} + \frac{4\sqrt{\rho_0\rho_D}}{(1-\rho_0)(1-\rho_D)}\sin^2\frac{\delta}{2}\right\}^{-1}.}$$

b) symmetric FP-resonator (Airy-formula for transmissivity)

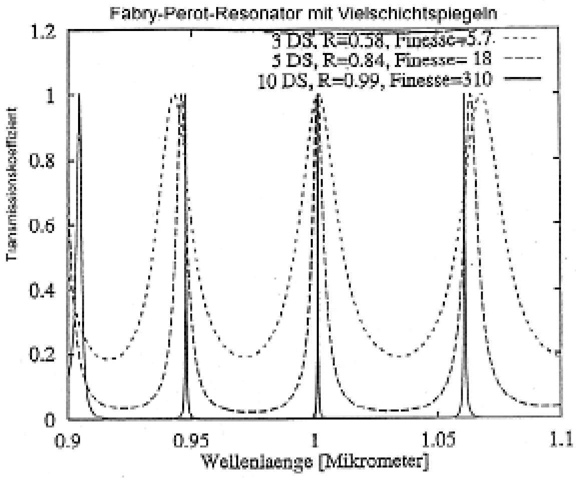
$$\rho_m = |R_0|^2 = |R_D|^2 = \rho_0 = \rho_D, \qquad \varphi^R = \varphi_0^R = \varphi_D^R$$

$$\tau = \frac{(1-\rho_m)^2}{(1-\rho_m)^2 + 4\rho_m\sin^2\frac{\delta}{2}} = \left\{\underbrace{\frac{(1-\rho_m)^2}{(1-\rho_m)^2}}_{=1} + \frac{4\rho}{(1-\rho_m)^2}\sin^2\frac{\delta}{2}\right\}^{-1}$$

$$\boxed{\tau = \left\{1 + F\sin^2\frac{\delta}{2}\right\}^{-1}}$$

with $F = \dfrac{4\rho_m}{(1-\rho_m)^2}$ and $\dfrac{\delta}{2} = k_{fx}D + \varphi$

The Airy-formula (see also Labworks script, where phase shifts φ 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^{TM} resp. ρ^{TM}
- Resonances of the cavity with $\tau_{\mathbf{MAX}} = 1$ occur for $\delta/2 = k_{fx} D_{\mathbf{MAX}} + \varphi = m\pi$ with

$$k_{\mathbf{fx}} = \frac{2\pi}{\lambda}\sqrt{n_{\mathbf{F}}^2 - n_{\mathbf{S}}^2 \sin^2 \varphi_{\mathbf{I}}}$$

$$\curvearrowright D_{\mathbf{MAX}} = \frac{m\pi - \varphi}{k_{\mathbf{fx}}} = \frac{\lambda}{2}\frac{\left(m - \frac{\varphi}{\pi}\right)}{\sqrt{n_{\mathbf{F}}^2 - n_{\mathbf{S}}^2 \sin^2 \varphi_{\mathbf{I}}}}$$

$$\curvearrowright \frac{\lambda}{2}\ \text{cavity}$$

where φ_{I} is the angle of incidence in the substrate

- transmission properties of a given resonator depend on $\varphi_{\mathbf{I}}$ and λ.
- minimum transmission is given as $\tau_{\mathbf{MIN}} = \dfrac{1}{1+F}$
- it is favorable to have large F, e.g.:

$$100 = F = \frac{4\rho_{\mathrm{m}}}{\left(1-\rho_{\mathrm{m}}\right)^2}$$

$$\rho_{\mathrm{m}} = 1 - \tau_{\mathrm{m}} \curvearrowright \frac{4\left(1-\tau_{\mathrm{m}}\right)}{\tau_{\mathrm{m}}^{2}} \approx \frac{4}{\tau_{\mathrm{m}}^{2}} \approx 100 \curvearrowright \tau_{\mathrm{m}} = 0.2 \curvearrowright \rho_{\mathrm{m}} = 0.8.$$

- pulses and beams can be treated efficiently in Fourier domain:
  → e.g. TE: $E_{\mathbf{T}}(\alpha,\beta,\omega) = T_{\mathbf{TE}}(\alpha,\beta,\omega) E_{\mathbf{I}}(\alpha,\beta,\omega)$
  Fourier back transformation: $E_{\mathbf{T}}(x,y,t) = \mathrm{FT}^{-1}\left[E_{\mathbf{T}}(\alpha,\beta,\omega)\right]$
- beams, because they always contain a certain range of incident angles $\varphi_{\mathbf{I}}$, 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 ε the FWHM and $\Delta = \pi$ the distance in rad between two resonances. To calculate the FWHM ε we can start from:

$$\left\{1 + F\sin^2\left(m\pi \pm \frac{\varepsilon}{2}\right)\right\}^{-1} \doteq \frac{1}{2}$$

For narrow resonances (small line width ε) we can write

$$\left\{1 + F\left(\frac{\varepsilon}{2}\right)^2\right\}^{-1} \approx \frac{1}{2} \rightarrow F\left(\frac{\varepsilon}{2}\right)^2 = 1 \curvearrowright \varepsilon = \frac{2}{\sqrt{F}}$$

$$\boxed{\Phi = \frac{\Delta}{\varepsilon} = \frac{\pi}{2}\sqrt{F} = \pi\frac{\sqrt{\rho_m}}{1-\rho_m}}$$

- The line width ε (FWHM) is inversely proportional to the finesse Φ.
- The Airy-formula can be expressed in terms of the finesse

$$\boxed{\tau = \left\{1 + \left(\frac{2\Phi}{\pi}\right)^2 \sin^2\frac{\delta}{2}\right\}^{-1}.}$$

- 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

$$\rightarrow kD + \varphi \pm \frac{\Delta k}{2}D = m\pi \pm \frac{\varepsilon}{2} = m\pi \pm \frac{\pi}{2\Phi}$$

$$\curvearrowright |\Delta k| = \frac{2\pi}{\lambda^2} n_f |\Delta\lambda| = \frac{\pi}{\Phi D}$$

$$\curvearrowright \left|\frac{\Delta\lambda}{\lambda}\right| = \frac{\lambda}{n_f \Phi D} \sim \frac{\lambda}{\Phi D} \sim \lambda\frac{\varepsilon}{D}$$

example: $\lambda = 5\cdot 10^{-7}$m, $\Phi = 30$, $n_f D = 4\cdot 10^{-3}$m $\curvearrowright$ $\Delta\lambda = 2\cdot 10^{-12}$m

- The field amplitudes (here forward field) inside the cavity are given as:
  $E_T = T_D E_+(D)$ Because $|T|^{-2} \sim 1/(1-\rho) \sim \Phi$ these intra-cavity fields can be very high $\curvearrowright$ 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:

$$\left|\Delta k\right| = \frac{1}{c} n_{\mathrm{F}} \Delta\omega = \frac{\pi}{\Phi D}$$

$$\Delta\omega = \frac{c}{n_{\mathrm{F}}} \frac{\pi}{\Phi D} \quad \curvearrowright \quad T_{\mathrm{c}} = \frac{1}{\Delta\omega} = \frac{n_{\mathrm{F}} \Phi D}{c\pi} \sim \frac{D}{\varepsilon}.$$

7.4 Guided waves in layer systems

Finally, we want to explore our layer systems as waveguides. For many applications it is interesting to have waves which propagate without diffraction. This is crucial for integrated optics, where we want to guide light in very small (micrometric or smaller) dielectric layers (film, fiber), or optical communication technology where some light encoded information is transported over long distances. Moreover, waveguides are important in nonlinear optics, due to confinement and long propagation distances nonlinear effects become important. Here, we will treat wave-guiding in one dimension only because we restrict ourselves to layer-systems, but general concepts developed can be transferred to other settings, e.g. where the waveguide is a fiber.

7.4.1 Field structure of guided waves

Let us first do some general consideration about the field structure of guided waves. We want to find guided waves in a layer system. In such systems, till now, we have solved the reflection and transmission problem: For given $k_{\mathrm{z}}, \mathbf{E}_{\mathrm{I}}, d_{i,} \varepsilon_i$ we calculated $\mathbf{E}_{\mathrm{R},} \mathbf{E}_{\mathrm{T}}$.

Inside each layer we have plane waves

$$\sim \mathbf{E}_{\alpha}\left(k_{\mathrm{z}}, \omega\right) \exp\left[\mathrm{i}\left(k_{\alpha\mathrm{x}} x + k_{\mathrm{z}} z - \omega t\right)\right].$$

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:

$$\mathbf{E}_{\mathrm{T}}(x,z) = \mathbf{E}_{\mathrm{T}} \exp\left(\mathrm{i} k_{\mathrm{z}} z\right) \exp\left(-\mu_{\mathrm{c}} x\right)$$

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, i.e. in cladding and substrate? And what about a single interface?

We will concentrate our discussion first on a system of layers, which we will call the core of the waveguide, and which we place between semi-infinite substrate and cladding. The single interface, where the core is absent, will be considered at a later stage.

According to our discussion the guided waves have the following field structure:

- plane wave in propagation direction $\sim \exp(ik_z z)$
- evanescent waves in substrate and cladding

$$\sim \exp[-\mu_c(x-D)] \qquad \text{cladding}$$

$$\sim \exp(\mu_s x) \qquad \text{substrate}$$

$$\text{with } \mu_{s,c} = \sqrt{k_z^2 - \frac{\omega^2}{c^2}\varepsilon_{s,c}(\omega)} > 0$$

$\curvearrowright$ **1. condition:** $\boxed{k_z^2 > \frac{\omega^2}{c^2}\max\left\{\varepsilon_{s,c}(\omega)\right\}}$

- oscillating solution (standing wave) in the core (layers, fiber core, film)

$$\sim A\sin(k_{fx}x) + B\cos(k_{fx}x)$$

$$\text{with } k_{ix} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_i(\omega) - k_z^2} > 0$$

$\curvearrowright$ **2. condition:** $\boxed{k_z^2 < \frac{\omega^2}{c^2}\max_i\left\{\varepsilon_i(\omega)\right\}}$

Note that this 2. condition is not obvious at this stage. It appears due to transition conditions at the boundaries to substrate and cladding. The fact that the fields in the substrate and gladding are of evanescent nature requires that both fields are exponentially growing towards the interface to the core. The transition conditions at the interfaces impose continuity of the tangential components of the fields and their first derivative normal to the surface (This fact was shown during the derivation of the transmission reflection properties in a multilayer system). Since the field in the core must continuously connect the fields and derivatives of the fields in substrate and cladding, it must have some non-monotonic profile. This condition excludes evanescent field profiles in the core and leaves only oscillating radiation-type fields as a possible solution.

In summary, the z-component of the wave vector of guided waves has to fulfill:

$$\boxed{\max\left(\frac{\omega}{c}n_{s,c}\right) < k_z < \max_i\left(\frac{\omega}{c}n_i\right)}$$

The field structure in substrate and cladding is given as:

$$\mathbf{E}_{\mathbf{c}}(x,z)=\mathbf{E}_{\mathbf{T}}\exp\left(\mathbf{i}k_{\mathbf{z}}z\right)\exp\left[-\mu_{\mathbf{c}}\left(x-D\right)\right] \qquad x>D$$

$$\mathbf{E}_{\mathbf{s}}(x,z)=\mathbf{E}_{\mathbf{R}}\exp\left(\mathbf{i}k_{\mathbf{z}}z\right)\exp\left(\mu_{\mathbf{s}}x\right) \qquad x<0$$

7.4.2 Dispersion relation for guided waves

If we compare the principal shapes of the fields for the guide wave to our usual reflection transmission problem, we see that for guided waves the reflected and transmitted (evanescent) field exist for zero incident field $\mathbf{E}_{\mathbf{I}}\to 0$. In the following we will use this condition to derive the dispersion relation for guided waves.

Thus we have $E_{\mathbf{T}},E_{\mathbf{R}}\neq 0$ for $E_{\mathrm{I}}\to 0$. Consequently the coefficients for reflection and transmission are

$$T^{\text{TE,TM}}=\left(\frac{E_{\mathbf{T}}}{E_{\mathbf{I}}}\right)^{\text{TE,TM}},\ R^{\text{TE,TM}}=\left(\frac{E_{\mathbf{R}}}{E_{\mathbf{I}}}\right)^{\text{TE,TM}} \quad \text{with } E_{\mathrm{I}}\to 0$$

and we find $R,T\to\infty$ in the case of guided waves. In this sense, guided waves are resonances of the system. Let us compare to a driven harmonic oscillator

$$x=\frac{F}{\omega^2-\omega_0^2}\ ,\ x = \text{action},\ F = \text{cause}$$

In the case of resonance $(\omega=\omega_0)$ 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. Furthermore, a resonance of a system is equivalent to an eigenmode of the system, which is somehow localized.

We know the coefficients of reflection and transmission for a layer system from before, and they have the same denominator:

$$R=\frac{F_{\mathbf{R}}}{F_{\mathbf{I}}}=\frac{\left(\alpha_{\mathbf{s}}k_{\mathbf{sx}}M_{22}-\alpha_{\mathbf{c}}k_{\mathbf{cx}}M_{11}\right)-\mathbf{i}\left(M_{21}+\alpha_{\mathbf{s}}k_{\mathbf{sx}}\alpha_{\mathbf{c}}k_{\mathbf{cx}}M_{12}\right)}{\left(\alpha_{\mathbf{s}}k_{\mathbf{sx}}M_{22}+\alpha_{\mathbf{c}}k_{\mathbf{cx}}M_{11}\right)+\mathbf{i}\left(M_{21}-\alpha_{\mathbf{s}}k_{\mathbf{sx}}\alpha_{\mathbf{c}}k_{\mathbf{cx}}M_{12}\right)}$$

The pole is then given as:

$$\rightarrow\ \left(\alpha_{\mathbf{s}}k_{\mathbf{sx}}M_{22}+\alpha_{\mathbf{c}}k_{\mathbf{cx}}M_{11}\right)+\mathbf{i}\left(M_{21}-\alpha_{\mathbf{s}}k_{\mathbf{sx}}\alpha_{\mathbf{c}}k_{\mathbf{cx}}M_{12}\right)\doteq 0$$

Because we have evanescent waves in substrate and cladding the imaginary valued k_{sx} and k_{cx} can be expressed as real numbers with

$$k_{\mathbf{sx}}=i\mu_{\mathbf{s}}=i\sqrt{k_{\mathbf{z}}^2-\frac{\omega^2}{c^2}\varepsilon_{\mathbf{s}}(\omega)},\qquad k_{\mathbf{cx}}=i\mu_{\mathbf{c}}=i\sqrt{k_{\mathbf{z}}^2-\frac{\omega^2}{c^2}\varepsilon_{\mathbf{c}}(\omega)}$$

We can write the general dispersion relation of guided modes in an arbitrary layer system

$$\boxed{M_{11}^{\text{TE,TM}} + \alpha_s \mu_s M_{12}^{\text{TE,TM}} + \frac{1}{\alpha_c \mu_c} M_{21}^{\text{TE,TM}} + \frac{\alpha_s \mu_s}{\alpha_c \mu_c} M_{22}^{\text{TE,TM}} \doteq 0}$$

Here, as usual, we have $\alpha_{\text{TE}} = 1$, $\alpha_{\text{TM}} = 1/\varepsilon$ for the two independent polarizations.

The waveguide dispersion relation gives a discrete set of solutions, so-called waveguide modes. For given $\varepsilon_i, d_i, \omega$ we get $k_{z\nu}(\omega)$.

We can see that the dispersion relation of guided modes depends on the material's dispersion in the individual layers as well as in substrate and cladding according to the dispersion relation of homogeneous space as

$$k_i^2(\omega) = \frac{\omega^2}{c^2}\varepsilon_i(\omega) = k_{ix}^2 + k_z^2 .$$

In addition to the material's dispersion in the layers the dispersion relation of guided modes depends on the geometry of the waveguide's layer system

$$k_z(\omega, \mathbf{geometry})$$

In the case of guided waves it is easy to compute the field (mode profile) inside the layer system:

- take k_z from dispersion relation
- in the substrate we have:

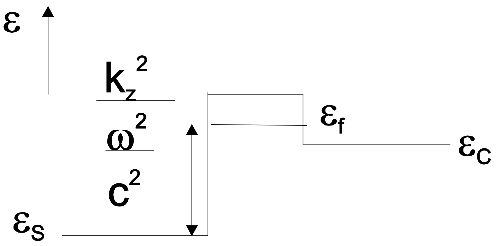
$$F(x) = F\exp(\mu_s x), \qquad G(x) = \alpha \frac{\partial}{\partial x}\left\{F\exp(\mu_s x)\right\}$$

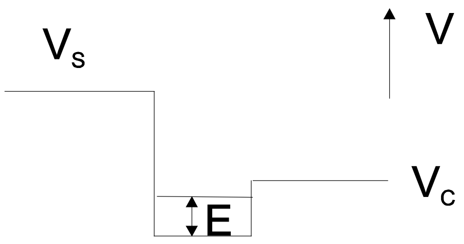
Hence, for given $F(0)$ we get $G(0) = \alpha\mu_s F(0)$ ($F(0) \to$ free parameter)

$$\boxed{\begin{pmatrix} F \\ G \end{pmatrix}_x = \hat{\mathbf{M}}(x)\begin{pmatrix} F \\ G \end{pmatrix}_0 = \hat{\mathbf{M}}(x)\begin{pmatrix} 1 \\ \alpha\mu_s \end{pmatrix} F(0).}$$

Analogy of optics to the stationary Schrödinger equation in QM

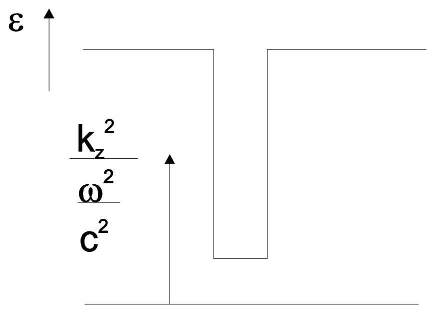
Optics (e.g. TE-polarization)		**Quantum mechanics**
$\left\{\frac{d^2}{dx^2} + \frac{\omega^2}{c^2}\varepsilon(x)\right\}E(x) = k_z^2 E(x)$	$\leftrightarrow$	$\left\{\frac{d^2}{dx^2} - \frac{2m}{\hbar^2}V(x)\right\}\psi(x) = -\frac{2m}{\hbar^2}E\psi(x)$
guided waves	$\leftrightarrow$	discrete energy eigenvalues
$k_z^2 > \frac{\omega^2}{c^2}\max\left\{\varepsilon_{s,c}\right\}$	$\leftrightarrow$	$E < V_{\text{ext}}$

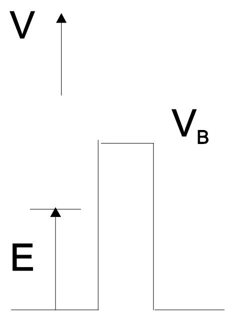




Tunnel effect

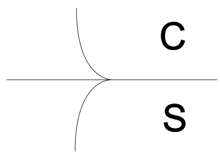
$$k_z^2 > \frac{\omega^2}{c^2}\varepsilon_{\text{Film}} \qquad \leftrightarrow \qquad E < V_{\text{Barriere}}$$





7.4.3 Guided waves at interface - surface polariton

Let us first have a look at the most simple case where the guiding layer structure is just an interface.



Our condition for guided waves is that on both sides of the interface we have evanescent waves: $k_z^2 > \frac{\omega^2}{c^2}\varepsilon_{c,s}$, because

$$\mu_{s,c} = \sqrt{k_z^2 - \frac{\omega^2}{c^2}\varepsilon_{c,s}} > 0$$

The general dispersion relation we derived before reads

$$M_{11}^{\text{TE,TM}} + \alpha_s\mu_s M_{12}^{\text{TE,TM}} + \frac{1}{\alpha_c\mu_c} M_{21}^{\text{TE,TM}} + \frac{\alpha_s\mu_s}{\alpha_c\mu_c} M_{22}^{\text{TE,TM}} \doteq 0$$

and with the matrix for a single interface: $\hat{\mathbf{M}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

we get the dispersion relation $\boxed{1 + \frac{\alpha_s\mu_s}{\alpha_c\mu_c} = 0}$

A) TE-polarization $(\alpha = 1)$

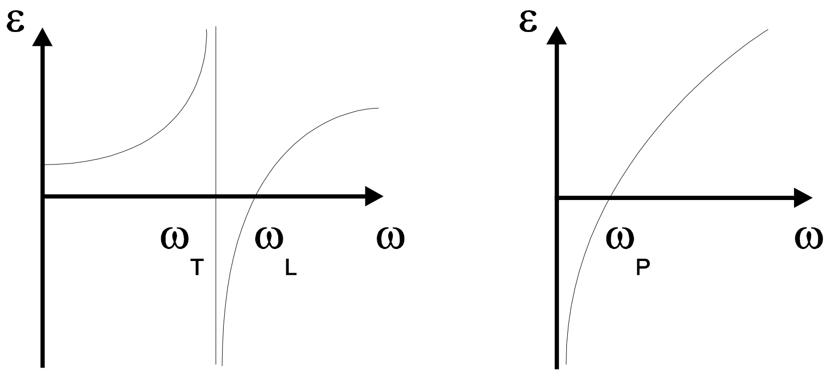
$\mu_s + \mu_c = 0, \rightarrow$ no solution because $\mu_c, \mu_s > 0$

B) TM-polarization $(\alpha = 1/\varepsilon)$

$$\boxed{\frac{\mu_c}{\varepsilon_c}+\frac{\mu_s}{\varepsilon_s}=0}$$

with $\mu_c,\mu_s>0 \curvearrowright \varepsilon_c\cdot\varepsilon_s<0$,

$\rightarrow$ on of the media has to have negative ε (dielectric near resonance or metal)



In dielectrics $\omega_{0(T)}<\omega<\omega_L$ we can find surface-phonon-polaritons.

In metals $\omega<\omega_p$ we can find surface-plasmon-polaritons.

Remark

Surface polaritons occur in TM polarization only, similar to the phenomenon of Brewster-angle (no reflection for $\frac{k_{cx}}{\varepsilon_c}-\frac{k_{sx}}{\varepsilon_s}=0,$)

Let us now compute the explicit dispersion relation for surface polaritons. W.o.l.g. we assume $\varepsilon_s(\omega)<0$, and because $\varepsilon_s(\omega)$ is near a resonance it will show a much stronger ω dependence than ε_c.

$$(\mu_c\varepsilon_s)^2=(\mu_s\varepsilon_c)^2$$

$$\varepsilon_s^2(\omega)\left\{k_z^2-\frac{\omega^2}{c^2}\varepsilon_c^2\right\}=\varepsilon_c^2\left\{k_z^2-\frac{\omega^2}{c^2}\varepsilon_s(\omega)\right\}$$

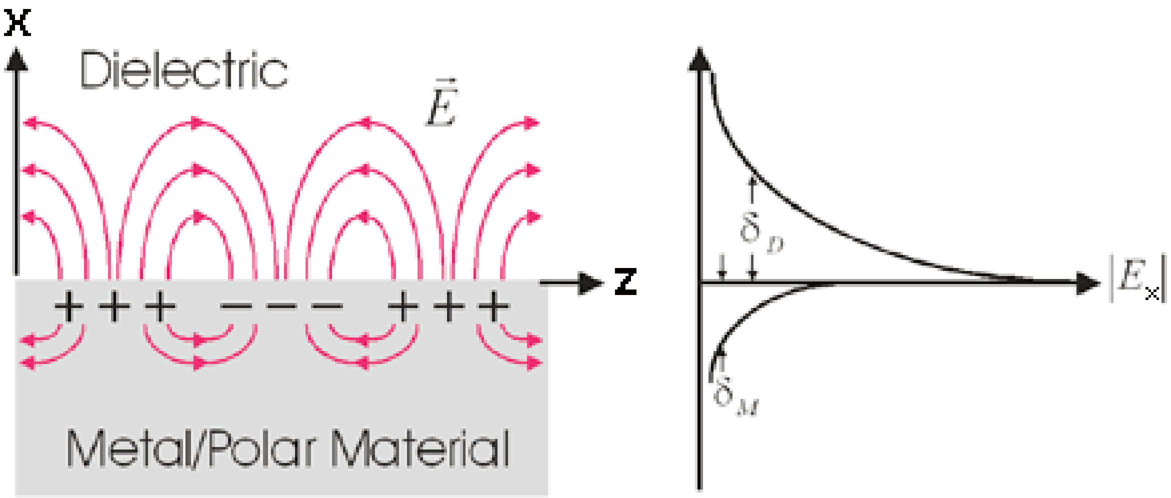
$$\boxed{k_z(\omega)=\frac{\omega}{c}\sqrt{\frac{\varepsilon_s(\omega)\varepsilon_c}{\varepsilon_c+\varepsilon_s(\omega)}}}$$

There is a second condition for existence of surface polaritons: $\varepsilon_c+\varepsilon_s(\omega)<0$

Conclusion:

$$k_z(\omega)=\frac{\omega}{c}\sqrt{\frac{\varepsilon_s(\omega)\varepsilon_c}{\varepsilon_s(\omega)+\varepsilon_c}}$$

TM polarization $\quad \varepsilon_s(\omega)<0 \quad \varepsilon_c>0 \quad |\varepsilon_s(\omega)|>\varepsilon_c$

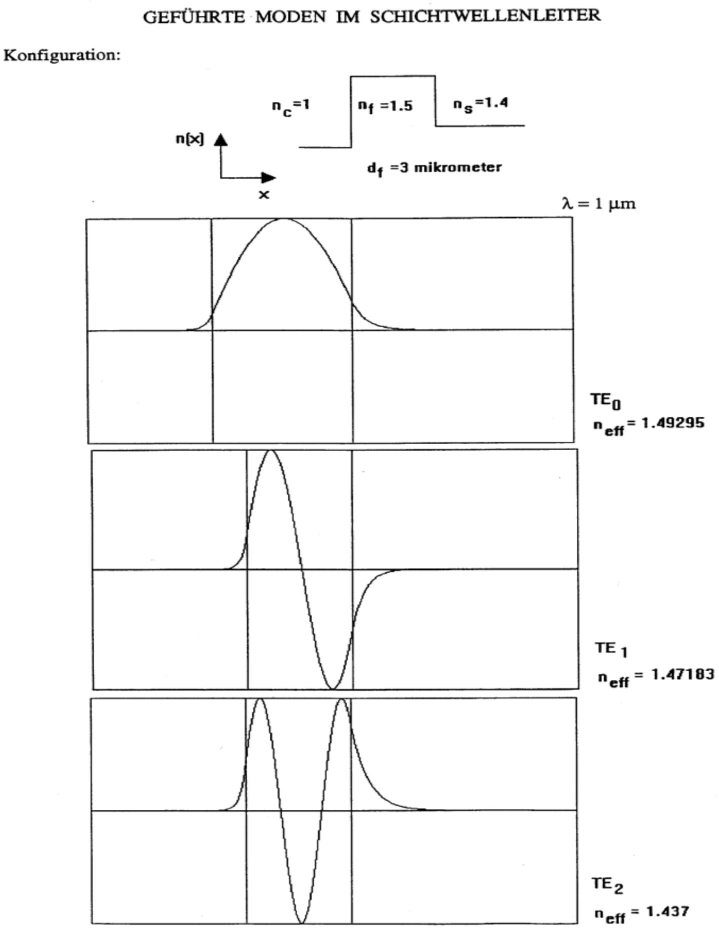


Surface polaritons may have very small effective wavelengths in z-direction:

$$\varepsilon_c = 1, |\varepsilon_s(\omega)| \approx 1 \to k_z(\omega) = \frac{2\pi}{\lambda}\sqrt{\frac{\varepsilon_s(\omega)\varepsilon_c}{\varepsilon_s(\omega)+\varepsilon_c}} = \frac{2\pi}{\lambda_{\text{eff}}} \to \lambda_{\text{eff}} << \lambda$$

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 $\frac{\omega^2}{c^2}\varepsilon_f(\omega) > k_z^2(\omega)$.



Such film waveguides are the basis of integrated optics. Typical parameters are:

$d \approx$ a few wavelengths

$\Delta\varepsilon = \varepsilon_f - \varepsilon_s \approx 10^{-3} - 10^{-1}$

Fabrication of film waveguides can be achieved by coating, diffusion or ion implantation.

The matrix of a single layer (film) is given as:

$$\hat{\mathbf{M}}^{\mathrm{TE,TM}} = \hat{\mathbf{m}}^{\mathrm{TE,TM}}(d) = \begin{pmatrix} \cos(k_{\mathrm{fx}}d) & \frac{1}{\alpha_{\mathrm{f}}k_{\mathrm{fx}}}\sin(k_{\mathrm{fx}}d) \\ -\alpha_{\mathrm{f}}k_{\mathrm{fx}}\sin(k_{\mathrm{fx}}d) & \cos(k_{\mathrm{fx}}d) \end{pmatrix}$$

From this matrix we can compute the dispersion relation for guided modes:

$$M_{11}+\alpha_s\mu_s M_{12}+\frac{1}{\alpha_c\mu_c}M_{21}+\frac{\alpha_s\mu_s}{\alpha_c\mu_c}M_{22}\doteq 0$$

$$\cos(k_{fx}d)+\frac{\alpha_s\mu_s}{\alpha_f k_{fx}}\sin(k_{fx}d)-\frac{\alpha_f k_{fx}}{\alpha_c\mu_c}\sin(k_{fx}d)+\frac{\alpha_s\mu_s}{\alpha_c\mu_c}\cos(k_{fx}d)=0$$

$$\frac{\sin(k_{fx}d)}{\cos(k_{fx}d)}=\tan(k_{fx}d)=\frac{1+\dfrac{\alpha_s\mu_s}{\alpha_c\mu_c}}{\dfrac{\alpha_f k_{fx}}{\alpha_c\mu_c}-\dfrac{\alpha_s\mu_s}{\alpha_f k_{fx}}}=\frac{\alpha_f k_{fx}(\alpha_s\mu_s+\alpha_c\mu_c)}{\alpha_f^2 k_{fx}^2-\alpha_c\alpha_s\mu_c\mu_s}$$

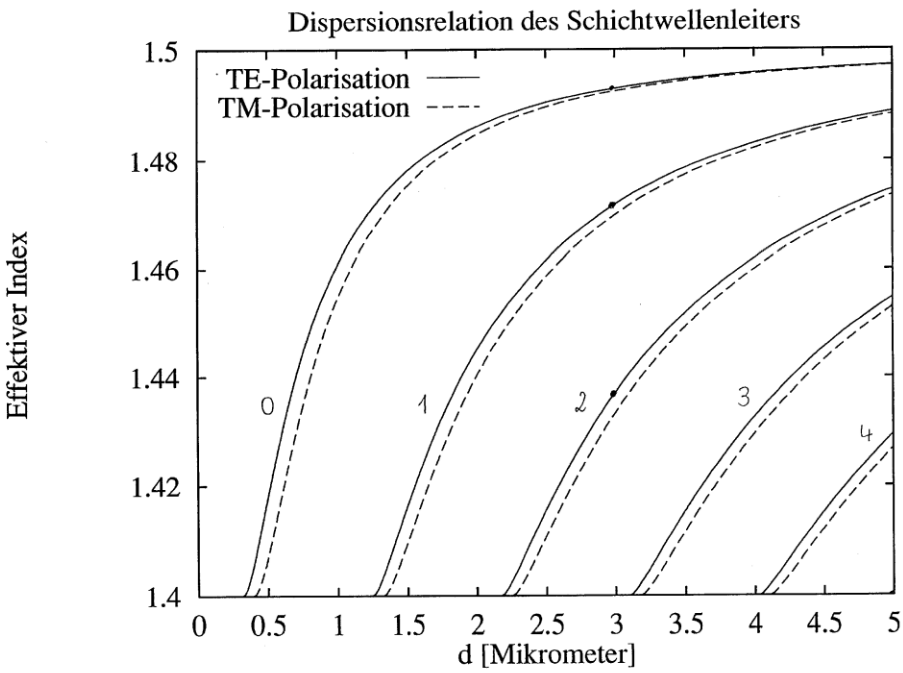
$$\boxed{\tan(k_{fx}d)=\frac{\dfrac{k_{fx}}{\alpha_f}\left(\dfrac{\mu_s}{\alpha_c}+\dfrac{\mu_c}{\alpha_s}\right)}{\dfrac{k_{fx}^2}{\alpha_c\alpha_s}-\dfrac{\mu_c\mu_s}{\alpha_f^2}}}$$

Here: TE-Polarisation $(\alpha=1)$

$$\boxed{\tan(k_{fx}d)=\frac{k_{fx}(\mu_s+\mu_c)}{k_{fx}^2-\mu_c\mu_s},}$$

This waveguide dispersion relation is an implicit equation for k_z. For given frequency ω and thickness d we get several solutions with index $k_{z\nu}$

Here is an example for fixed frequency ω, the effective index $n_{\mathrm{eff}}=k_z/\left(\dfrac{\omega}{c}\right)$ 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 → *cut-off* (here w.o.l.g. $\varepsilon_c < \varepsilon_s$)

The idea of the cut-off is that a mode is not guided anymore. Guiding means evanescent fields in the substrate and cladding, so cut-off means

$$\mu_s = \sqrt{k_z^2 - \frac{\omega^2}{c^2}\varepsilon_s} = 0 \rightarrow \text{no guiding} \curvearrowright \quad k_z^2 = \frac{\omega^2}{c^2}\varepsilon_s$$

We can plug this cut-off condition in the DR: $\tan(k_{fx}d) = \frac{k_{fx}(\mu_s + \mu_c)}{k_{fx}^2 - \mu_c\mu_s}$,

$$\tan\left(\frac{\omega}{c}\sqrt{\varepsilon_f - \varepsilon_s}d\right) = \frac{\sqrt{\varepsilon_f - \varepsilon_s}\sqrt{\varepsilon_s - \varepsilon_c}}{\varepsilon_f - \varepsilon_s} = \sqrt{\frac{\varepsilon_s - \varepsilon_c}{\varepsilon_f - \varepsilon_s}}$$

$$(\omega d)_{co}^{TE} = \frac{c}{\sqrt{\varepsilon_f - \varepsilon_s}}\left\{\arctan\sqrt{\frac{\varepsilon_s - \varepsilon_c}{\varepsilon_f - \varepsilon_s}} + \nu\pi\right\}$$

$$\boxed{\rightarrow (\omega d)_{co}^{TE} = \frac{c}{\sqrt{\varepsilon_f - \varepsilon_s}}\left\{\underbrace{\arctan a}_{\max \pi/2} + \nu\pi\right\}}$$

$$\varepsilon_s \approx \varepsilon_c \qquad a \rightarrow 0$$

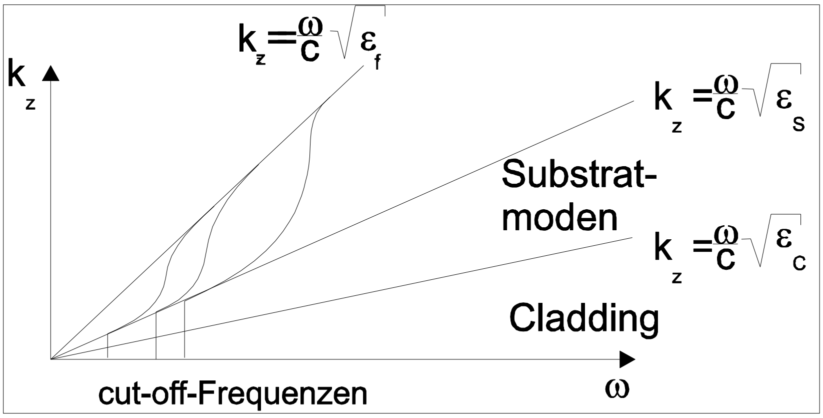
with parameter of asymmetry a: $\varepsilon_s \approx \varepsilon_f \qquad a \rightarrow \infty$

$$\varepsilon_s \approx \varepsilon_c$$

→ we can define a cut-off frequency for $k_z(\omega)$ when we keep d fix

→ we can define a cut-off thickness for $k_z(d)$ when we keep ω fix

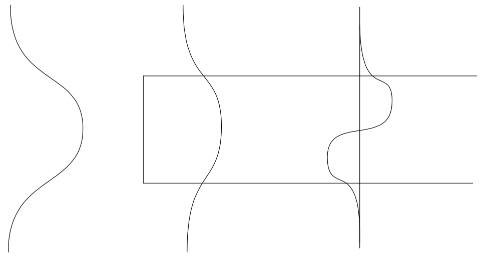
In a symmetric waveguide the fundamental more ($\nu = 0$) has cut-off = 0!
If we plot the dispersion curves for each mode we get a graphical representation of the dispersion relation:



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_z)

A) adaption of field → front face coupling



Then, inside the waveguide we have (without radiative modes):

$$E(x,z) = \sum_\nu a_\nu E_\nu(x)\exp\left(\mathbf{i}k_{z\nu}z\right)$$

$$\curvearrowright E(x,0) \approx \sum_\nu a_\nu E_\nu(x) \qquad \left|\int E_\mu(x)\right.$$

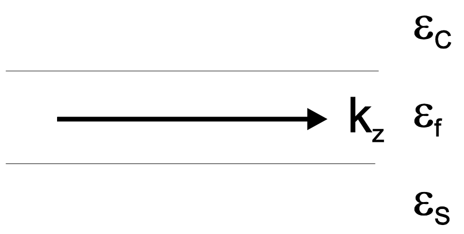
$$\texttt{with}:\ P_\nu = \frac{k_{z\nu}}{2\omega\mu_0}\int_{-\infty}^{\infty}\left|E_\nu(x)\right|^2 dx$$

$$a_\nu = \frac{k_{z\nu}}{2\omega\mu_0 P_\nu}\int_{-\infty}^{\infty} E_{\mathtt{in}}(x)E_\nu(x)dx,$$

→ mode ν couples to the incident field $E_{\mathtt{in}}(0)$ with amplitude a_ν .

→ Gauss-beam couples very good to the fundamental mode

B) adaption of wave vector → coupling through the interface



We know that k_z is continuous at interface. The condition for the existence of guided modes is

$$k_{\mathbf{z}} > \frac{\omega}{c}\sqrt{\varepsilon_{\mathbf{c,s}}}$$

but dispersion relation for waves in bulk media dictates

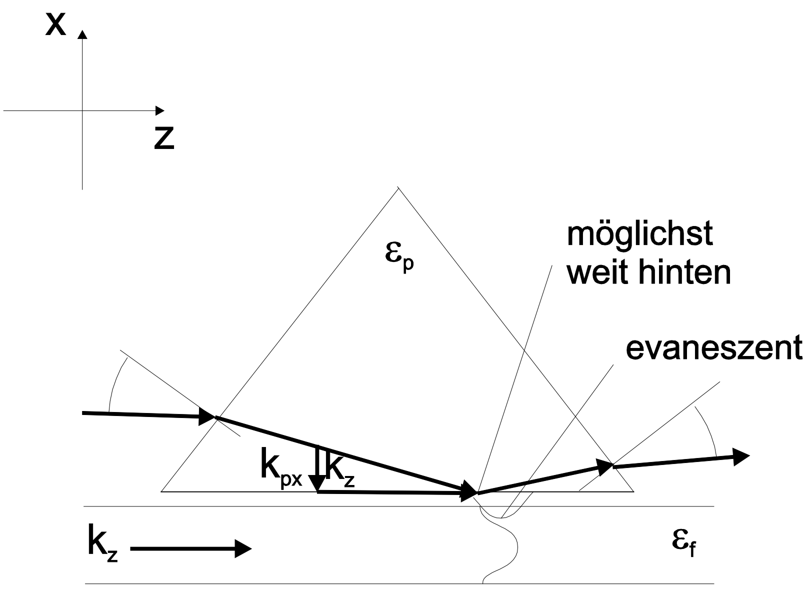
$$k_{\mathbf{z}} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_{\mathbf{c,s}} - k_{\mathbf{s,cx}}^2} < \frac{\omega}{c}\sqrt{\varepsilon_{\mathbf{c,s}}}$$

$\curvearrowright$ We got a problem!
There are two solutions:

i) coupling by prism

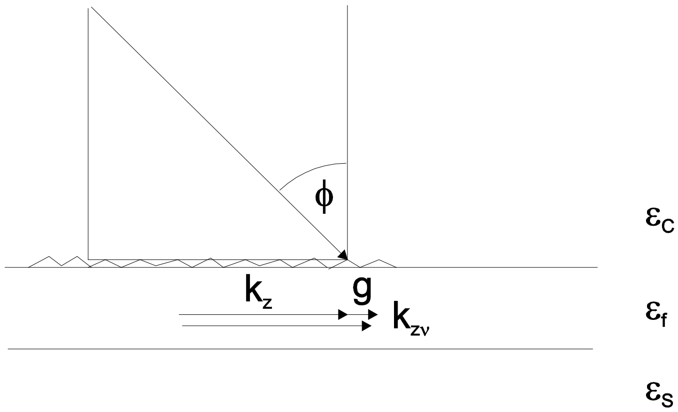
we bring a medium with $\varepsilon_{\mathbf{p}} > \varepsilon_{\mathbf{f}}$ (prism) near the waveguide

$$\curvearrowright k_z < \frac{\omega}{c}\sqrt{\varepsilon_{\mathrm{f}}} \curvearrowright k_z < \frac{\omega}{c}\sqrt{\varepsilon_{\mathrm{p}}} \curvearrowright k_{\mathrm{px}} = \sqrt{\frac{\omega^2}{c^2}\varepsilon_{\mathrm{p}} - k_z^2} > 0 \,.$$



→ 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):

$$d(z) = d + \varsigma(z)$$

$$\varsigma(z) = A\sin(gz) \quad mit \quad g = \frac{2\pi}{P} \ldots_{P-period}$$

coupling works for m'th diffraction order:

$$k_{z\nu} = k_z + mg$$

$$= \frac{\omega}{c} n_s \sin\phi + mg.$$