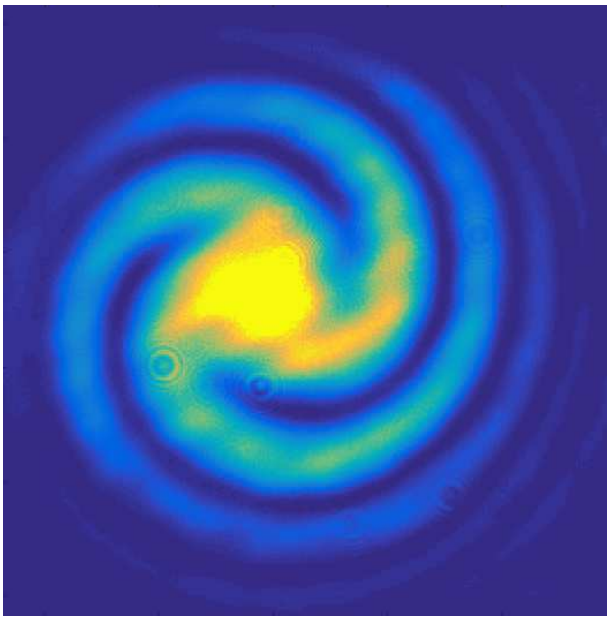
Photonic Devices

(SS 2020 - physics640)

Stefan Linden
Physikalisches Institut
Universität Bonn



Contents

1 Elements of electrodynamics and solid state optics

1.1 Macroscopic Maxwell equations

Classical electromagnetic phenomena in matter can be described with the macroscopic Maxwell's equations :

$$\nabla \cdot \mathbf{D}(\mathbf{r},t) = \varrho(\mathbf{r},t), \tag{1.1.1}$$

$$\nabla \times \mathbf{E}(\mathbf{r},t) = -\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}, \tag{1.1.2}$$

$$\nabla \cdot \mathbf{B}(\mathbf{r},t) = 0, \tag{1.1.3}$$

$$\nabla \times \mathbf{H}(\mathbf{r},t) = \mathbf{j}(\mathbf{r},t) + \frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t}. \tag{1.1.4}$$

Here, $\mathbf{E}$ denotes the electric field strength, $\mathbf{D}$ the electric displacement field, $\mathbf{H}$ the magnetic field strength, $\mathbf{B}$ the magnetic flux density, ϱ the free charge density, and $\mathbf{j}$ the free electric current density. In this description, each field is the spatial and temporal average of the corresponding microscopic field.

The electromagnetic properties of a material can be characterized by the polarization density $\mathbf{P}(\mathbf{r},t)$ and the magnetization $\mathbf{M}(\mathbf{r},t)$. These quantities represent the macroscopic density of electric and magnetic dipole moments, respectively. By definition, $\mathbf{D}$ and $\mathbf{H}$ can be written as

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t), \tag{1.1.5}$$

$$\mathbf{H}(\mathbf{r},t) = \frac{\mathbf{B}(\mathbf{r},t)}{\mu_0} - \mathbf{M}(\mathbf{r},t), \tag{1.1.6}$$

where ϵ_0 is the vacuum permittivity and μ_0 is the vacuum permeability . In general, $\mathbf{P}(\mathbf{r},t)$ and $\mathbf{M}(\mathbf{r},t)$ depend in a non-trivial way on the applied fields. For small field strengths, the response of a medium to an electric field $\mathbf{E}$ can often be approximated by:

$$\mathbf{P}(\mathbf{r},t) = \epsilon_0 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\boldsymbol{\chi}}_e(\mathbf{r},\mathbf{r}',t,t')\,\mathbf{E}(\mathbf{r}',t')\,dt'd\mathbf{r}', \tag{1.1.7}$$

where the second-order tensor $\hat{\boldsymbol{\chi}}_e(\mathbf{r}, \mathbf{r}', t, t')$ denotes the linear electric susceptibility of the medium. In an isotropic, homogeneous medium with local response, equation (1.1.7) becomes

$$\mathbf{P}(\mathbf{r}, t) = \epsilon_0 \int_{-\infty}^{t} \chi_e(t - t')\, \mathbf{E}(\mathbf{r}, t')\, dt'. \tag{1.1.8}$$

For the further analysis, it is often convenient to switch from the time domain to the frequency domain. For this purpose, we introduce the (inverse) Fourier transform:

$$\tilde{f}(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t) e^{\imath\omega t} dt, \tag{1.1.9}$$

$$f(t) = \mathcal{F}^{-1}\left\{\tilde{f}(\omega)\right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{-\imath\omega t} d\omega. \tag{1.1.10}$$

Applying the Fourier transform to $\mathbf{P}(\mathbf{r}, t)$ and using the convolution theorem yields:

$$\mathbf{P}(\mathbf{r}, \omega) = \mathcal{F}\{\mathbf{P}(\mathbf{r}, t)\} = \epsilon_0 \chi_e(\omega)\, \mathbf{E}(\mathbf{r}, \omega), \tag{1.1.11}$$

where $\mathbf{E}(\mathbf{r}, \omega)$ and $\chi_e(\omega)$ are the Fourier transforms of $\mathbf{E}(\mathbf{r}, t)$ and $\chi_e(t)$, respectively. Following the same arguments, we find for the magnetization $\mathbf{M}(\mathbf{r}, \omega)$:

$$\mathbf{M}(\mathbf{r}, \omega) = \mu_0 \chi_m(\omega)\, \mathbf{H}(\mathbf{r}, \omega), \tag{1.1.12}$$

where $\chi_m(\omega)$ is the magnetic susceptibility. The material equations have the following form in the frequency domain:

$$\mathbf{D}(\mathbf{r}, \omega) = \epsilon_0 \mathbf{E}(\mathbf{r}, \omega) + \mathbf{P}(\mathbf{r}, \omega) = \epsilon_0 \epsilon(\omega)\, \mathbf{E}(\mathbf{r}, \omega), \tag{1.1.13}$$

$$\mathbf{B}(\mathbf{r}, \omega) = \mu_0 \mathbf{H}(\mathbf{r}, \omega) + \mathbf{M}(\mathbf{r}, \omega) = \mu_0 \mu(\omega)\, \mathbf{H}(\mathbf{r}, \omega). \tag{1.1.14}$$

Here, we have introduced the electric permittivity

$$\epsilon(\omega) = 1 + \chi_e(\omega). \tag{1.1.15}$$

and the magnetic permeability

$$\mu(\omega) = 1 + \chi_m(\omega). \tag{1.1.16}$$

Mathematical interlude: The Fourier transform

The Fourier transform $\tilde{f}(\omega) = \mathcal{F}\{f(t)\}$ of a complex valued function $f(t)$ is defined as

$$\tilde{f}(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t) e^{\imath\omega t} dt. \tag{1.1.17}$$

Conversely, $f(t)$ can be calculated via the inverse Fourier transform

$$f(t) = \mathcal{F}^{-1}\left\{\tilde{f}(\omega)\right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{-\imath\omega t} d\omega. \tag{1.1.18}$$

Time translation

The Fourier transform of $f(t - t_0)$ (time translation) is given by

$$\mathcal{F}\{f(t-t_0)\} = \int_{-\infty}^{\infty} f(t-t_0)\, e^{\imath\omega t} dt \overset{t=t'+t_0}{=} \int_{-\infty}^{\infty} f(t')\, e^{\imath\omega(t'+t_0)} dt' = e^{\imath\omega t_0}\, \tilde{f}(\omega). \tag{1.1.19}$$

Frequency modulation

The Fourier transform of $e^{-\imath\omega_0 t} f(t)$ (frequency modulation) is given by

$$\mathcal{F}\left\{e^{-\imath\omega_0 t}\, f(t)\right\} = \int_{-\infty}^{\infty} e^{-\imath\omega_0 t}\, f(t)\, e^{\imath\omega t} dt = \int_{-\infty}^{\infty} f(t)\, e^{\imath(\omega-\omega_0)t} dt = \tilde{f}(\omega - \omega_0). \tag{1.1.20}$$

Time scaling and time reversal

The Fourier transform of $f(at)$, where a is a real constant (time scaling), can be calculated as

$$\mathcal{F}\{f(at)\} = \int_{-\infty}^{\infty} f(at)\, e^{\imath\omega t} dt \overset{t'=at}{=} \frac{1}{|a|} \int_{-\infty}^{\infty} f(t')\, e^{\imath(\omega/a)t'} dt' = \frac{1}{|a|} \tilde{f}(\omega/a). \tag{1.1.21}$$

Hence, with $a = -1$ (time reversal), we obtain

$$\mathcal{F}\{f(-t)\} = \tilde{f}(-\omega). \tag{1.1.22}$$

Fourier transform of real functions

Let $f(t)$ be a real function with $f(t) = f^*(t)$. The Fourier transform of $f^*(t)$ can be calculated as

$$\begin{aligned} f(t)^* &= \left[\mathcal{F}^{-1}\left\{\tilde{f}(\omega)\right\}\right]^* = \left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}(\omega)e^{-\imath\omega t}d\omega\right]^* = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}^*(\omega)e^{\imath\omega t}d\omega \\ &\overset{-\omega=\omega'}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}^*(-\omega')e^{-\imath\omega' t}d\omega' \overset{!}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}(\omega')e^{-\imath\omega' t}d\omega' = f(t). \end{aligned} \tag{1.1.23}$$

The comparison shows that

$$\tilde{f}^*(-\omega) = \tilde{f}(\omega). \tag{1.1.24}$$

δ-function and the Fourier transform

The δ-function is defined through

$$f(t_0) = \int_{-\infty}^{\infty}\delta(t-t_0)\,f(t)dt. \tag{1.1.25}$$

Using the inverse Fourier transform, $f(t_0)$ can be written as

$$\begin{aligned} f(t_0) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}(\omega)\,e^{-\imath\omega t_0}d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}f(t)e^{\imath\omega t}dt\right]e^{-\imath\omega t_0}d\omega \\ &= \int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{\imath\omega(t-t_0)}d\omega\right]f(t)dt. \end{aligned} \tag{1.1.26}$$

By comparing equation (1.1.25) with equation (1.1.26), we find that

$$\delta(t-t_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{\imath\omega(t-t_0)}d\omega. \tag{1.1.27}$$

Convolution theorem

Let $h(t)$ be the convolution of $f(t)$ and $g(t)$:

$$h(t) = \int_{-\infty}^{\infty} f(t-t')g(t')dt'. \tag{1.1.28}$$

The Fourier transform of $h(t)$ is then given by:

$$\begin{aligned}
\tilde{h}(\omega) &= \int_{-\infty}^{\infty} h(t)e^{\imath\omega t}dt = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(t-t')g(t')dt'\right]e^{\imath\omega t}dt \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(t-t')e^{\imath\omega t}dt\right]g(t')dt' = \int_{-\infty}^{\infty} e^{\imath\omega t'}\,\tilde{f}(\omega)g(t')dt' \\
&= \tilde{f}(\omega)\tilde{g}(\omega).
\end{aligned} \tag{1.1.29}$$

Some important Fourier transforms

- δ-impulse

$$f(t) = \delta(t) \Rightarrow \tilde{f}(\omega) = \int_{-\infty}^{\infty} \delta(t)e^{\imath\omega t}dt = 1. \tag{1.1.30}$$

- Rectangular pulse

$$f(t) = rect(at) \Rightarrow \tilde{f}(\omega) = \frac{1}{|a|}\,sinc\left(\frac{\omega}{2\pi a}\right). \tag{1.1.31}$$

- Gaussian function

$$f(t) = e^{-\frac{at^2}{2}} \Rightarrow \tilde{f}(\omega) = \sqrt{\frac{2\pi}{a}}e^{-\frac{\omega^2}{2a}}. \tag{1.1.32}$$

- Complex exponential function

$$f(t) = e^{-\imath\omega_0 t} \Rightarrow \tilde{f}(\omega) = \int_{-\infty}^{\infty} e^{\imath(\omega-\omega_0)t}dt = 2\pi\,\delta(\omega-\omega_0). \tag{1.1.33}$$

1.2 Electromagnetic waves

1.2.1 Wave equation

An important consequence of Maxwell equations is the existence of electromagnetic waves. To derive the wave equation for the electric field, we first take the curl of equation (1.1.2):

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r},t) = -\nabla \times \frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}. \tag{1.2.1}$$

Subsequently, we switch the order of $\partial/\partial t$ and $\nabla\times$:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r},t) = -\partial/\partial t\left(\nabla \times \mathbf{B}(\mathbf{r},t)\right). \tag{1.2.2}$$

Upon substitution of equations (1.1.6) and (1.1.4) in (1.2.2), we obtain:

$$\nabla\times\nabla\times\mathbf{E}(\mathbf{r},t)+\mu_0\epsilon_0\frac{\partial^2\mathbf{E}(\mathbf{r},t)}{\partial t^2} = -\mu_0\frac{\partial^2\mathbf{P}(\mathbf{r},t)}{\partial t^2}-\mu_0\frac{\partial}{\partial t}\mathbf{j}(\mathbf{r},t)-\nabla\times\frac{\partial\mathbf{M}(\mathbf{r},t)}{\partial t}. \tag{1.2.3}$$

An analogous derivation yields the wave equation for the magnetic field:

$$\nabla\times\nabla\times\mathbf{H}(\mathbf{r},t)+\mu_0\epsilon_0\frac{\partial^2\mathbf{H}(\mathbf{r},t)}{\partial t^2} = \nabla\times\mathbf{j}(\mathbf{r},t)+\nabla\times\frac{\partial\mathbf{P}(\mathbf{r},t)}{\partial t}-\epsilon_0\frac{\partial^2\mathbf{M}(\mathbf{r},t)}{\partial t^2}. \tag{1.2.4}$$

The wave equations (1.2.3) and (1.2.4) are valid for all materials. The terms on the right hand sides of the wave equations describe the sources of electromagnetic radiation.

1.2.2 Plane waves in vacuum

As a first example, we consider electromagnetic waves in vacuum. In this case, the material properties are given by $\mathbf{P}=0, \mathbf{M}=0, \mathbf{j}=0$, and $\varrho=0$. By using the vector identity $\nabla\times\nabla\times = -\nabla^2+\nabla\nabla\cdot$, the wave equations (1.2.3) and (1.2.4) can be written as:

$$\nabla^2\mathbf{E}(\mathbf{r},t)-\mu_0\epsilon_0\frac{\partial^2\mathbf{E}(\mathbf{r},t)}{\partial t^2} = 0, \tag{1.2.5}$$

$$\nabla^2\mathbf{B}(\mathbf{r},t)-\mu_0\epsilon_0\frac{\partial^2\mathbf{B}(\mathbf{r},t)}{\partial t^2} = 0. \tag{1.2.6}$$

In the following, we concentrate on electromagnetic waves that have the form of a monochromatic plane wave[1]:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0\, e^{\imath(\mathbf{k}_0\cdot\mathbf{r}-\omega_0 t)}, \tag{1.2.7}$$
$$\mathbf{B}(\mathbf{r},t) = \mathbf{B}_0\, e^{\imath(\mathbf{k}_0\cdot\mathbf{r}-\omega_0 t)}, \tag{1.2.8}$$

where, $\mathbf{E}_0$ and $\mathbf{B}_0$ are constant (complex valued) amplitude vectors, $\mathbf{k}_0$ is the vacuum wave vector, and ω is the angular frequency. Upon substitution of $\mathbf{E}(\mathbf{r},t)$ in the wave equation (1.2.5), we obtain the vacuum dispersion relation

$$\mathbf{k}_0\cdot\mathbf{k}_0 = \frac{\omega_0^2}{c_0^2} \tag{1.2.9}$$

with

$$c_0^2 = \frac{1}{\mu_0\epsilon_0}. \tag{1.2.10}$$

By definition, the vacuum speed of light is $c_0 = 299792458\,\mathrm{m}\,s^{-1}$.

The electric field strength $\mathbf{E}(\mathbf{r},t)$ is constant on the planes defined by $\mathbf{k}_0\cdot\mathbf{r}-\omega_0 t = const.$ These planes are oriented perpendicular to $\mathbf{k}_0$ and propagate with the phase velocity

$$\mathbf{v}_{\mathrm{P}hase} = \frac{\omega_0}{|\mathbf{k}_0|}\hat{\mathbf{e}}_k = c_0\,\hat{\mathbf{e}}_k, \tag{1.2.11}$$

where $\hat{\mathbf{e}}_k$ is the unit vector parallel to $\mathbf{k}_0$. The distance between two consecutive planes with the same phase defines the vacuum wavelength

$$\lambda_0 = \frac{2\pi}{|\mathbf{k}_0|}. \tag{1.2.12}$$

In addition to satisfying the wave equation, $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$ must also be valid solutions of Maxwell's equations. It follows from $\nabla\cdot\mathbf{E}(\mathbf{r},t) = 0$ that the amplitude vector $\mathbf{E}_0$ is perpendicular to the wavevector $\mathbf{k}_0$, i.e.,

$$\mathbf{k}_0\cdot\mathbf{E}_0 = 0. \tag{1.2.13}$$

Accordingly, $\nabla\cdot\mathbf{B}(\mathbf{r},t) = 0$ implies that also the amplitude vector $\mathbf{B}_0$ is perpendicular to the wavevector $\mathbf{k}_0$, i.e.,

$$\mathbf{k}_0\cdot\mathbf{B}_0 = 0. \tag{1.2.14}$$

[1] Here, we define the plane waves as complex vector fields in order to simplify the mathematical treatment. However, only the real parts of the complex fields are relevant from a physical point of view.

Furthermore, we find from $\nabla \times \mathbf{E}(\mathbf{r},t) = -\dot{\mathbf{B}}(\mathbf{r},t)$ that

$$\mathbf{k}_0 \times \mathbf{E}_0 = \omega_0 \mathbf{B}_0. \tag{1.2.15}$$

Hence, $\mathbf{k}_0$, $\mathbf{E}_0$ and $\mathbf{B}_0$ are mutually orthogonal to each other and define a right-handed set of vectors. Equation (1.2.15) allows to express $\mathbf{H}(\mathbf{r},t)$ as

$$\mathbf{H}(\mathbf{r},t) = \frac{1}{Z_0}\hat{\mathbf{e}}_k \times \mathbf{E}(\mathbf{r},t), \tag{1.2.16}$$

with the vacuum impedance

$$Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} \approx 377\Omega. \tag{1.2.17}$$

According to Poynting's theorem, the instantaneous energy flux density associated with an electromagnetic wave is given by the Poynting vector [2].

$$\mathbf{S}(\mathbf{r},t) = \mathbf{E}(\mathbf{r},t) \times \mathbf{H}(\mathbf{r},t). \tag{1.2.18}$$

For optical fields, the Poynting vector is a rapidly varying function of time. In most applications, we are, however, not interested in the fast fluctuations of $\mathbf{S}(\mathbf{r},t)$ but rather in its time-averaged value:

$$\langle \mathbf{S}(\mathbf{r}) \rangle = \frac{1}{T}\int_0^T \mathbf{S}(\mathbf{r},t)\, dt, \tag{1.2.19}$$

where T is the integration time. The time average of the product of two time-harmonic functions $A(t) = \Re[A_0 e^{-\imath\omega t}]$ and $B(t) = \Re[B_0 e^{-\imath\omega t}]$ is given by (Proof: Exercise):

$$\langle A(t)B(t) \rangle = \frac{1}{2}\Re[A_0 B_0^*]. \tag{1.2.20}$$

With this, we can express the time-averaged Poynting vector of a plane wave as

$$\langle \mathbf{S} \rangle = \frac{1}{2}\sqrt{\frac{\epsilon_0}{\mu_0}}|\mathbf{E}_0|^2 \mathbf{e}_k = \frac{1}{2Z_0}|\mathbf{E}_0|^2 \mathbf{e}_k. \tag{1.2.21}$$

The intensity of a plane wave, i.e, its irradiance, is given by the magnitude of thetime-averaged Poynting vector:

$$I = |\langle \mathbf{S} \rangle|. \tag{1.2.22}$$

[2] Important: One must not use the complex valued fields in this equation but only the corresponding real parts!

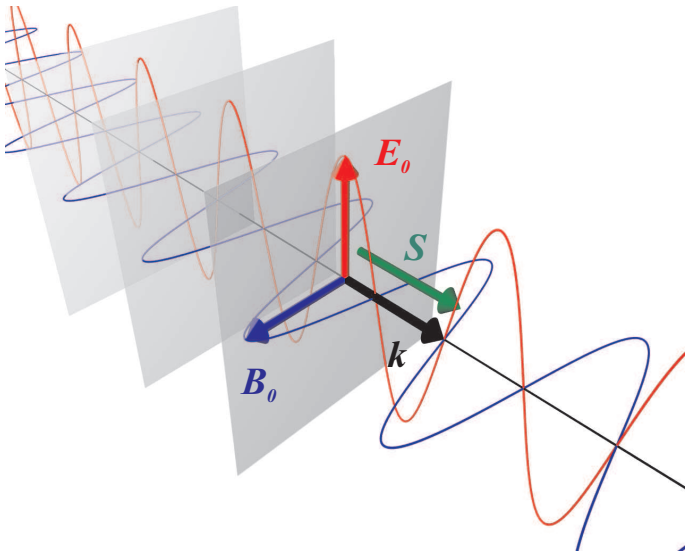


Figure 1.1: Electromagnetic plane wave in vacuum.

1.2.3 Monochromatic waves in magnetodielectric media

Light in matter is not a pure electromagnetic wave but rather a mixture of an electromagnetic wave and a material excitation characterized by $\mathbf{P}(\mathbf{r},t)$ and $\mathbf{M}(\mathbf{r},t)$. In the following, we consider a magnetodielectric medium with $\mathbf{D}(\mathbf{r},\omega)=\epsilon_0\epsilon(\omega)\mathbf{E}(\mathbf{r},\omega)$, $\mathbf{B}(\mathbf{r},\omega)=\mu_0\mu(\omega)\mathbf{H}(\mathbf{r},\omega)$, $\mathbf{j}=0$, $\varrho=0$.

In dispersive media with $\epsilon=\epsilon(\omega)$ and $\mu=\mu(\omega)$, wave propagation is most conveniently discussed in terms of monochromatic fields:

$$\mathbf{E}(\mathbf{r},t)=\mathbf{E}(\mathbf{r})e^{-\imath\omega t}, \tag{1.2.23}$$

$$\mathbf{H}(\mathbf{r},t)=\mathbf{H}(\mathbf{r})e^{-\imath\omega t}. \tag{1.2.24}$$

For such monochromatic fields, the wave equation reduces to the Helmholtz equation ($\partial/\partial t\rightarrow-\imath\omega$):

$$\left[\nabla^2+\mu_0\epsilon_0\epsilon(\omega)\mu(\omega)\omega^2\right]\mathbf{E}(\mathbf{r})=0. \tag{1.2.25}$$

In homogeneous and isotropic media, plane electromagnetic waves are solutions of the Helmholtz equation:

$$\mathbf{E}(\mathbf{r})=\mathbf{E}_0e^{\imath\mathbf{k}\cdot\mathbf{r}}, \tag{1.2.26}$$

$$\mathbf{H}(\mathbf{r})=\mathbf{H}_0e^{\imath\mathbf{k}\cdot\mathbf{r}}. \tag{1.2.27}$$

Upon substituting $\mathbf{E}(\mathbf{r})$ into the Helmholtz equation, we obtain the dispersion relation of the magnetodielectric medium:

$$\mathbf{k}\cdot\mathbf{k}=\frac{\omega^2}{c_0^2}\epsilon(\omega)\mu(\omega). \tag{1.2.28}$$

The wave vector in the medium is related to the vacuum wave vector by

$$\mathbf{k} = n(\omega)\,\mathbf{k}_0, \tag{1.2.29}$$

with the refractive index[3]

$$n(\omega) = \sqrt{\epsilon(\omega)\,\mu(\omega)}. \tag{1.2.30}$$

In general, the refractive index $n(\omega) = n'(\omega) + \imath\, n''(\omega)$ and the wave vector $\mathbf{k} = \mathbf{k}' + \imath\,\mathbf{k}''$ are complex quantities. In absorbing media, the amplitude of the wave falls off exponentially in the propagation direction:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0 e^{\imath(\mathbf{k}'\cdot\mathbf{r}-\omega t)}\, e^{-\mathbf{k}''\cdot\mathbf{r}}, \tag{1.2.31}$$

$$\mathbf{B}(\mathbf{r},t) = \mathbf{B}_0 e^{\imath(\mathbf{k}'\cdot\mathbf{r}-\omega t)}\, e^{-\mathbf{k}''\cdot\mathbf{r}}. \tag{1.2.32}$$

As in vacuum, the wavelength in the medium is defined by the separation of two consecutive planes with the same phase:

$$\lambda = \frac{2\pi}{|\mathbf{k}'|} = \frac{\lambda_0}{|n'(\omega)|}. \tag{1.2.33}$$

The phase velocity in the medium is related to the vacuum speed of light by:

$$|\mathbf{v}_{\mathrm{P}hase}| = \frac{\omega}{|\mathbf{k}'|} = \frac{c_0}{|n'(\omega)|}. \tag{1.2.34}$$

The impedance of the medium can be calculated as:

$$Z(\omega) = \sqrt{\frac{\mu_0\mu(\omega)}{\epsilon_0\epsilon(\omega)}} \tag{1.2.35}$$

The time average of the Poynting vector in the medium can be written as:

$$\begin{aligned}
\langle \mathbf{S}(\mathbf{r})\rangle &= \frac{1}{2}\Re\left[\mathbf{E}_0\, e^{\imath\mathbf{k}'\cdot\mathbf{r}}\, e^{-\mathbf{k}''\cdot\mathbf{r}} \times \mathbf{H}_0^*\, e^{-\imath\mathbf{k}'\cdot\mathbf{r}} e^{-\mathbf{k}''\cdot\mathbf{r}}\right] \\
&= \frac{1}{2}\Re\left[\mathbf{E}_0 \times \frac{1}{Z^*}\left(\hat{\mathbf{e}}_{\mathbf{k}} \times \mathbf{E}_0^*\right)\right] e^{-2\mathbf{k}''\cdot\mathbf{r}} \\
&= \frac{1}{2}\,\frac{1}{\Re(Z)}\,|\mathbf{E}_0|^2\,\hat{\mathbf{e}}_{\mathbf{k}}\, e^{-2\mathbf{k}''\cdot\mathbf{r}}.
\end{aligned} \tag{1.2.36}$$

For passive media, $\langle \mathbf{S}(\mathbf{r})\rangle$ falls off in the propagation direction. By comparison with Beer's law

$$I(z) = I(0)e^{-\alpha z}, \tag{1.2.37}$$

we can connect the absorption coefficient with the imaginary part of the refractive index:

$$\alpha = \frac{4\pi n''(\omega)}{\lambda_0}. \tag{1.2.38}$$

[3] Here, we have tacitly assumed that $\Re(\epsilon) > 0, \Re(\mu) > 0$.

1.3 Electromagnetic fields at interfaces

1.3.1 Boundary conditions

In the following, we consider an interface between two isotropic media. We assume that there are no free charges or free currents present at the boundary ($\rho = 0$, $\mathbf{j} = 0$).

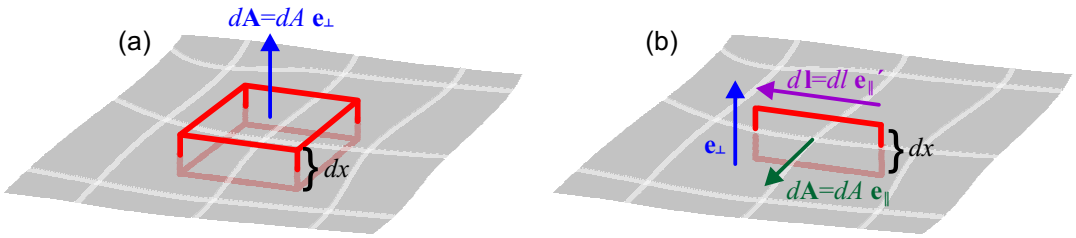


Figure 1.2: Schemes to derive the boundary conditions for (a) the normal components of $\mathbf{D}$ and $\mathbf{B}$ and (b) the tangential components of $\mathbf{E}$ and $\mathbf{H}$, respectively.

To derive the boundary condition for the normal component of $\mathbf{D}$, we consider a small box that is intersected by the interface. The top and bottom surface of the box are parallel to the interface (see Fig. 1.2 (a)). Applying the divergence theorem to $\nabla \cdot \mathbf{D}(\mathbf{r}, t) = 0$ yields:

$$\int_{dV} \nabla \cdot \mathbf{D}(\mathbf{r}, t)\, d^3r = \int_{\partial(dV)} \mathbf{D}(\mathbf{r}, t) \cdot d\mathbf{A} \stackrel{!}{=} 0. \tag{1.3.1}$$

Here dV is the volume of the box and $\partial(dV)$ its surface. In the limit $dx \to 0$, we obtain:

$$\int_{\partial(dV)} \mathbf{D}(\mathbf{r}, t) \cdot d\mathbf{A} \stackrel{dx\to 0}{\longrightarrow} dA\, \mathbf{e}_\perp \cdot (\mathbf{D}_1(\mathbf{r}_0, t) - \mathbf{D}_2(\mathbf{r}_0, t)) \stackrel{!}{=} 0, \tag{1.3.2}$$

where $\mathbf{D}_1(\mathbf{r}_0, t)$ and $\mathbf{D}_2(\mathbf{r}_0, t)$ are the electric displacement fields in the two media in the close vicinity of the interface. Equation (1.3.2) shows, that the normal component $\mathbf{D}_\perp(\mathbf{r}_0, t) = \mathbf{e}_\perp \cdot \mathbf{D}(\mathbf{r}_0, t)$ is continuous across the interface. An analogous consideration shows that the normal component $\mathbf{B}_\perp(\mathbf{r}_0, t) = \mathbf{e}_\perp \cdot \mathbf{B}(\mathbf{r}_0, t)$ is also continuous across the interface.

Next, we consider a small rectangle that intersects the interface at right angles. The top and bottom side of the rectangle are assumed to be parallel to the interface (see Fig. 1.2 (b)). Applying Stoke's theorem to $\nabla \times \mathbf{E}(\mathbf{r}, t) = -\dot{\mathbf{B}}(\mathbf{r}, t)$ yields:

$$\int_{dA} \nabla \times \mathbf{E}(\mathbf{r}, t) \cdot \mathbf{e}_\parallel dA = \int_{\partial dA} \mathbf{E}(\mathbf{r}, t) \cdot d\mathbf{r} = -\int_{dA} \dot{\mathbf{B}}(\mathbf{r}, t) \cdot \mathbf{e}_\parallel dA, \tag{1.3.3}$$

where dA is the area of the rectangle and ∂dA is its circumference. For $dx \to 0$, we obtain:

$$\int_{\partial dA} \mathbf{E}(\mathbf{r},t) \cdot d\mathbf{r} \overset{dx\to 0}{\longrightarrow} dl \underbrace{\left(\mathbf{e}_{\parallel} \times \mathbf{e}_{\perp}\right)}_{\mathbf{e}_{\parallel'}} \cdot \left(\mathbf{E}_1(\mathbf{r}_0,t) - \mathbf{E}_2(\mathbf{r}_0,t)\right) = 0. \tag{1.3.4}$$

Here, we have used that the surface integral of $\dot{\mathbf{B}}(\mathbf{r},t) \cdot \mathbf{e}_{\parallel}$ vanishes for $dx \to 0$. Equation (1.3.4) shows that the tangential component of the electric field strength $\mathbf{E}_{\parallel}(\mathbf{r}_0,t) = \mathbf{E}(\mathbf{r}_0,t) \cdot \mathbf{e}_{\parallel'}$ is continuous across the interface. An analogous consideration shows that the tangential component of the magnetic field strength $\mathbf{H}_{\parallel}(\mathbf{r}_0,t) = \mathbf{H}(\mathbf{r}_0,t) \cdot \mathbf{e}_{\parallel'}$ is also continuous across the interface.

The results of this subsection are summarized in the following table:

Tangential components	Normal components
$\mathbf{E}_{1,\parallel}(\mathbf{r}_0,t) = \mathbf{E}_{2,\parallel}(\mathbf{r}_0,t)$	$\mathbf{E}_{1,\perp}(\mathbf{r}_0,t) = \frac{\epsilon_2}{\epsilon_1}\mathbf{E}_{2,\perp}(\mathbf{r}_0,t)$
$\mathbf{D}_{1,\parallel}(\mathbf{r}_0,t) = \frac{\epsilon_1}{\epsilon_2}\mathbf{D}_{2,\parallel}(\mathbf{r}_0,t)$	$\mathbf{D}_{1,\perp}(\mathbf{r}_0,t) = \mathbf{D}_{2,\perp}(\mathbf{r}_0,t)$
$\mathbf{H}_{1,\parallel}(\mathbf{r}_0,t) = \mathbf{H}_{2,\parallel}(\mathbf{r}_0,t)$	$\mathbf{H}_{1,\perp}(\mathbf{r}_0,) = \frac{\mu_2}{\mu_1}\mathbf{H}_{2,\perp}(\mathbf{r}_0,t)$
$\mathbf{B}_{1,\parallel}(\mathbf{r}_0,t) = \frac{\mu_1}{\mu_2}\mathbf{B}_{2,\parallel}(\mathbf{r}_0,t)$	$\mathbf{B}_{1,\perp}(\mathbf{r}_0,t) = \mathbf{B}_{2,\perp}(\mathbf{r}_0,t)$

1.3.2 Fresnel equations

Next, we examine the reflection and refraction of a monochromatic plane wave that impinges on the planar interface between two media. The media are assumed to be isotropic with electric permittivities ϵ_1 and ϵ_2, and magnetic permeabilities μ_1 and μ_2, respectively. The incident wave (medium 1) is given by:

$$\mathbf{E}_i(\mathbf{r},t) = \mathbf{E}_i e^{\imath(\mathbf{k}_i\cdot\mathbf{r}-\omega_0 t)},\ \mathbf{B}_i(\mathbf{r},t) = \mathbf{B}_i e^{\imath(\mathbf{k}_i\cdot\mathbf{r}-\omega_0 t)}. \tag{1.3.5}$$

One part of the incident wave is reflected back at the interface into medium 1. The reflected wave can be expressed as:

$$\mathbf{E}_r(\mathbf{r},t) = \mathbf{E}_r e^{\imath(\mathbf{k}_r\cdot\mathbf{r}-\omega_0 t)},\ \mathbf{B}_r(\mathbf{r},t) = \mathbf{B}_r e^{\imath(\mathbf{k}_r\cdot\mathbf{r}-\omega_0 t)}. \tag{1.3.6}$$

The other part is transmitted into medium 2:

$$\mathbf{E}_t(\mathbf{r},t) = \mathbf{E}_t e^{\imath(\mathbf{k}_t\cdot\mathbf{r}-\omega_0 t)},\ \mathbf{B}_t(\mathbf{r},t) = \mathbf{B}_t e^{\imath(\mathbf{k}_t\cdot\mathbf{r}-\omega_0 t)}. \tag{1.3.7}$$

To satisfy the boundary conditions at all points of the interface, the tangential components of the three wave vectors must be identical:

$$|\mathbf{k}_i| \sin(\theta_i) = |\mathbf{k}_r| \sin(\theta_r) = |\mathbf{k}_t| \sin(\theta_t) . \tag{1.3.8}$$

Here, the angles θ_i, θ_r, and θ_t are defined in Fig. 1.3. Snell's law of refraction is a direct consequence of the conservation of the tangential component of the wave vector at the interface:

$$n_i(\omega_0) \sin(\theta_i) = n_t(\omega_0) \sin(\theta_t) . \tag{1.3.9}$$

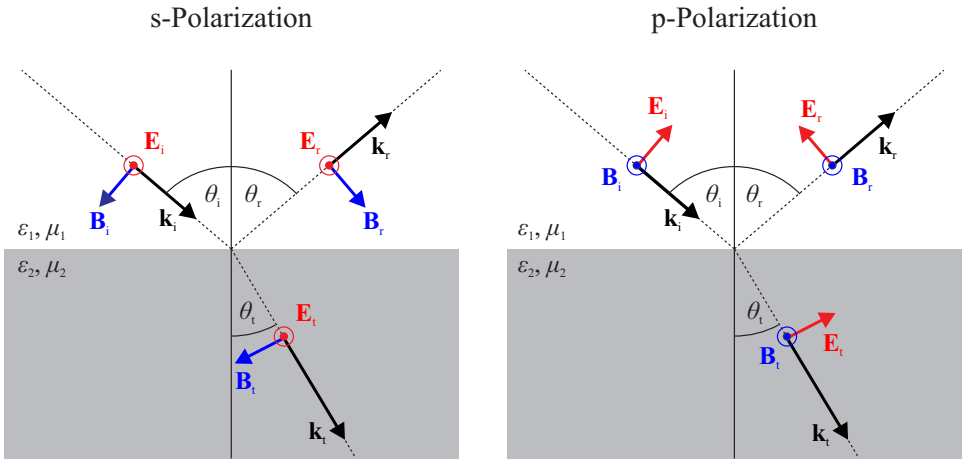


Figure 1.3: Reflexion and refraction of a plane wave at a planar interface. Left: s-polarization. Right: p-polarization.

The complex amplitudes of the transmitted and reflected wave can be calculated by applying the boundary conditions. To simplify the math, it is useful to consider waves which are either polarized perpendicular (s-polarized) or parallel (p-polarized) to the plane of incidence[4]. After some algebra, we obtain the so-called Fresnel equations for these two polarization states.

[4]The plane of incidence is spanned by the surface normal and the wave vector of the incident wave.

s-Polarization

The amplitude transmission coefficient t_s and the amplitude reflection coefficient r_s for s-polarized light are given by:

$$t_s = \left(\frac{\mathbf{E}_t}{\mathbf{E}_i}\right)_s = \frac{2\,(n_i/\mu_i)\cos(\theta_i)}{(n_i/\mu_i)\cos(\theta_i) + (n_t/\mu_t)\cos(\theta_t)} \tag{1.3.10}$$

$$= \frac{2\,(1/\eta_i)\cos(\theta_i)}{(1/\eta_i)\cos(\theta_i) + (1/\eta_t)\cos(\theta_t)}. \tag{1.3.11}$$

$$r_s = \left(\frac{\mathbf{E}_r}{\mathbf{E}_i}\right)_s = \frac{(n_i/\mu_i)\cos(\theta_i) - (n_t/\mu_t)\cos(\theta_t)}{(n_i/\mu_i)\cos(\theta_i) + (n_t/\mu_t)\cos(\theta_t)} \tag{1.3.12}$$

$$= \frac{(1/\eta_i)\cos(\theta_i) - (1/\eta_t)\cos(\theta_t)}{(1/\eta_i)\cos(\theta_i) + (1/\eta_t)\cos(\theta_t)}. \tag{1.3.13}$$

p-Polarization

The amplitude transmission coefficient t_p and the amplitude reflection coefficient r_p for p-polarized light can be written as:

$$t_p = \left(\frac{\mathbf{E}_t}{\mathbf{E}_i}\right)_p = \frac{2\,(n_i/\mu_i)\cos(\theta_i)}{(n_i/\mu_i)\cos(\theta_t) + (n_t/\mu_t)\cos(\theta_i)}, \tag{1.3.14}$$

$$= \frac{2\,(1/\eta_i)\cos(\theta_i)}{(1/\eta_i)\cos(\theta_t) + (1/\eta_t)\cos(\theta_i)}, \tag{1.3.15}$$

$$\tag{1.3.16}$$

$$r_p = \left(\frac{\mathbf{E}_r}{\mathbf{E}_i}\right)_p = \frac{(n_t/\mu_t)\cos(\theta_i) - (n_i/\mu_i)\cos(\theta_t)}{(n_i/\mu_i)\cos(\theta_t) + (n_t/\mu_t)\cos(\theta_i)} \tag{1.3.17}$$

$$= \frac{(1/\eta_t)\cos(\theta_i) - (1/\eta_i)\cos(\theta_t)}{(1/\eta_i)\cos(\theta_t) + (1/\eta_t)\cos(\theta_i)}. \tag{1.3.18}$$

Example: Fresnel equations for normal incidence

For normal incidence ($\theta_i = \theta_r = \theta_t = 0$), the reflection coefficient becomes:

$$r(\theta = 0^\circ) = \frac{(1/\eta_t) - (1/\eta_i)}{(1/\eta_t) + (1/\eta_i)} \tag{1.3.19}$$

$$= \frac{\eta_i - \eta_t}{\eta_i + \eta_t}. \tag{1.3.20}$$

Inspection of this equation shows that the amplitude reflection coefficient for normal incidence is determined by the impedances of the two media!

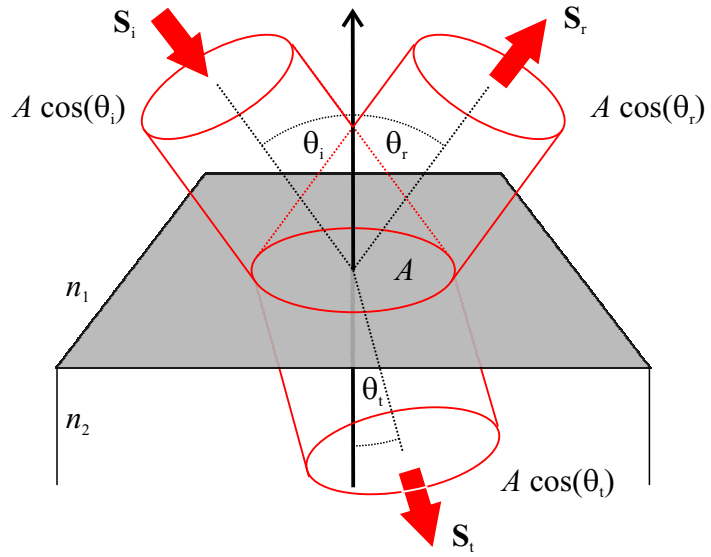


Figure 1.4: Energy transport through an interface.

In the following, we consider the energy transport associated with the incident, the reflected, and the transmitted wave. The corresponding time averaged Poynting vectors are given by:

$$\bar{\mathbf{S}}_i = \frac{1}{2}\frac{1}{Z_1}|\mathbf{E}_0|^2 \frac{\mathbf{k}_i}{|\mathbf{k}_i|}, \tag{1.3.21}$$

$$\bar{\mathbf{S}}_r = \frac{1}{2}\frac{1}{Z_1}|r\mathbf{E}_0|^2 \frac{\mathbf{k}_r}{|\mathbf{k}_r|}, \tag{1.3.22}$$

and

$$\bar{\mathbf{S}}_t = \frac{1}{2}\frac{1}{Z_2}|t\mathbf{E}_0|^2 \frac{\mathbf{k}_t}{|\mathbf{k}_t|}. \tag{1.3.23}$$

respectively. The reflectance R is defined as the fraction of the incident power that is reflected at the interface:

$$R = \frac{|\bar{\mathbf{S}}_r \cdot \mathbf{e}_\perp|}{|\bar{\mathbf{S}}_i \cdot \mathbf{e}_\perp|} = |r|^2. \tag{1.3.24}$$

Accordingly, the transmittance T is the fraction of the incident power which is transmitted through the interface:

$$T = \frac{|\bar{\mathbf{S}}_t \cdot \mathbf{e}_\perp|}{|\bar{\mathbf{S}}_i \cdot \mathbf{e}_\perp|} = \frac{Z_1 \cos(\theta_t)}{Z_2 \cos(\theta_i)} |t|^2. \tag{1.3.25}$$

For non-absorbing media, we obtain (energy conservation):

$$R + T = 1. \tag{1.3.26}$$

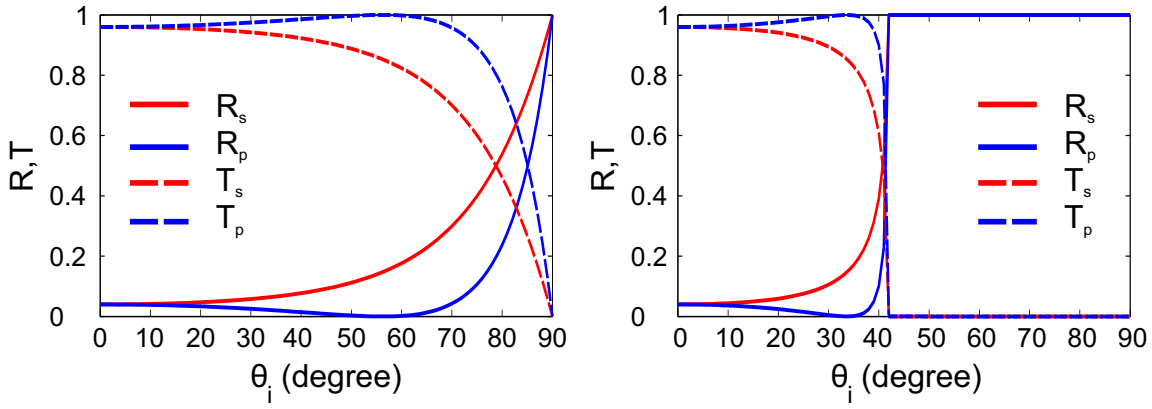


Figure 1.5: Reflectance and transmittance of a glass/air interface. Left: Incidence from air. Right: Incidence from glass.

1.3.3 Dielectric mirrors

The majority of lasers utilizes dielectric mirrors (Bragg mirrors), which can have a reflectance of better than $R = 0.99$ for the corresponding design wavelength. A dielectric mirror is composed of a sequence of layers of (at least) two different dielectric materials. Its working principle is shown in Fig. 1.6 (a). The light wave impinges on the sequence of layers. At each interface, a (small) fraction of the light wave is reflected. The total reflected wave is given by the superposition of all reflected waves. A strong reflected wave develops if all the partial waves interfere constructively, i.e., if the total phase difference between two successive partial wave is a multiple of 2π. Here, one has to take different optical path-lengths and phase changes due to reflections into account.

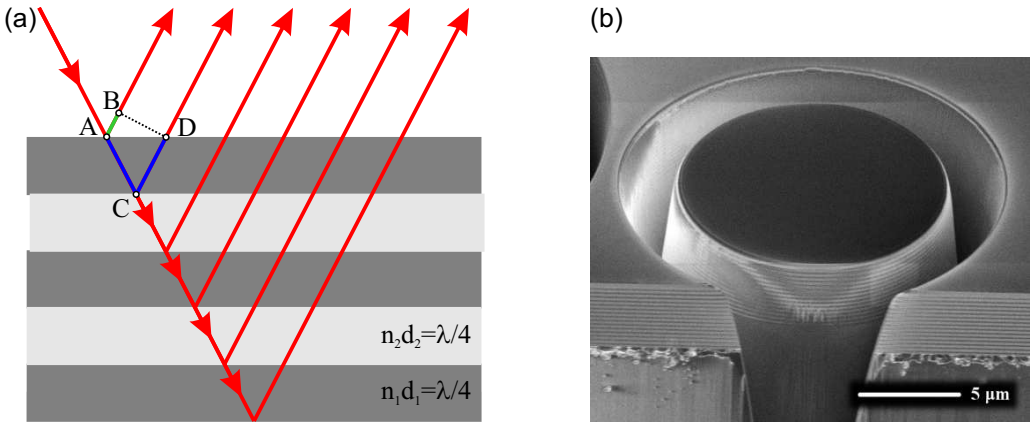


Figure 1.6: (a) Working principle of a dielectric mirror. (b) Electron micrograph of a dielectric mirror made from alternating layers of SiO_2 and Ta_2O_5. Image taken from Wikipedia.

For normal incidence, constructive interference of all partial waves is achieved for a so-called quarter wave stack. Here, the layers have an optical path length

$$S = n_i d_i = \lambda_0/4, \tag{1.3.27}$$

where, i is the index of the layer, n_i is its refractive index, d_i is its geometrical thickness, and λ_0 is the vacuum wavelength. The difference in the optical path length of two successive partial wave is $\Delta S = 2S = \lambda/2$. This corresponds to a difference in phase due to wave propagation of $2k_0 n_i d_i = \pi$. Additionally, there is another π difference in phase due to reflection from the different boundaries (low-to-high index boundary vs. high-to-low index boundary).

1.3.4 Matrix theory of stratified media

In this section, we discuss an elegant matrix approach for the calculation of the optical properties of stratified media that consist of a sequence of dielectric layers. Multiple reflections of a wave at the different interfaces make a "summation by hand" of all the partial waves quite cumbersome. To overcome this problem, we lump all the forward and backward propagating partial waves at a given plane m into the complex amplitudes $U_m^{(+)}$ and $U_m^{(-)}$, respectively:

$$U_m^{(+)} = \sum \text{all forward propgating partial waves at plane } m, \tag{1.3.28}$$

$$U_m^{(-)} = \sum \text{all backward propgating partial waves at plane } m, \tag{1.3.29}$$

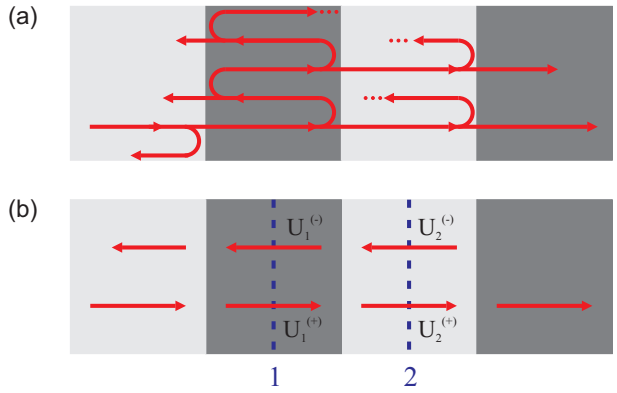


Figure 1.7: (a) Multiple reflections in a stratified medium are often to cumbersome for "summation by hand". (b) All forward and backward propagating partial waves at a given plane can be aggregated into collected waves with amplitudes $U_m^{(+)}$ and $U_m^{(-)}$, respectively.

Next, we consider two different planes of the structure (see Fig. 1.7 (b)) . The amplitudes of the collected waves in these two planes are related by a transfer matrix (T-matrix) $\mathbf{M}$:

$$\begin{pmatrix} U_2^{(+)} \\ U_2^{(-)} \end{pmatrix} = \underbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}_{\mathbf{M}} \begin{pmatrix} U_1^{(+)} \\ U_1^{(-)} \end{pmatrix}. \tag{1.3.30}$$

The four coefficients A, B, C, and D encode the properties of the structure between the two planes.

A big advantage of the matrix approach is that the T-matrix of a sequence of planes can be easily calculated by matrix multiplication of the individual T-matrices :

$$\mathbf{M} = \mathbf{M}_N \cdots \mathbf{M}_2 \mathbf{M}_1. \tag{1.3.31}$$

Note that matrix multiplication is in general not commutative. Thus, it is important to adhere to the proper order of the matrizes (see Fig. 1.8).

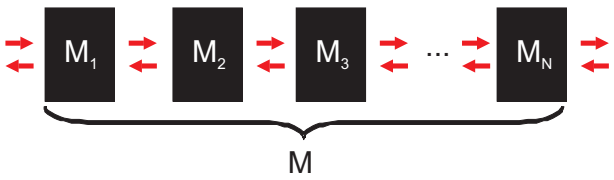


Figure 1.8: The T-matrix of a sequence ofplanes can be calculated by matrix multiplication of the individual T-matrices .

So far, we have used the T-matrix to connect the forward and backward propagating waves on the left side of the structure with the forward and backward propagating waves

on the right side of the structure (see Fig. 1.9 (a)). In order to determine the coefficients of the T-matrix, it is advantageous to look at the problem from a different point of view. For this purpose, we introduce the so-called scattering matrix **S** that relates the outgoing waves to the ingoing waves (see Fig. 1.9 (b)):

$$\begin{pmatrix} U_2^{(+)} \\ U_1^{(-)} \end{pmatrix} = \underbrace{\begin{bmatrix} t_{12} & r_{21} \\ r_{12} & t_{21} \end{bmatrix}}_{\mathbf{S}} \begin{pmatrix} U_1^{(+)} \\ U_2^{(-)} \end{pmatrix} \tag{1.3.32}$$

The coefficients of the scattering matrix **S** have the following physical meaning:

- t_{12}: Amplitude transmission coefficient for wave propagation in forward direction.
- t_{21}: Amplitude transmission coefficient for wave propagation in backward direction.
- r_{12}: Amplitude reflection coefficient for light incidence from left (forward direction).
- r_{21}: Amplitude reflection coefficient for light incidence from right (backward direction).

Note that the S-matrix of a composite structure is **not** the product of the individual S-matrices.

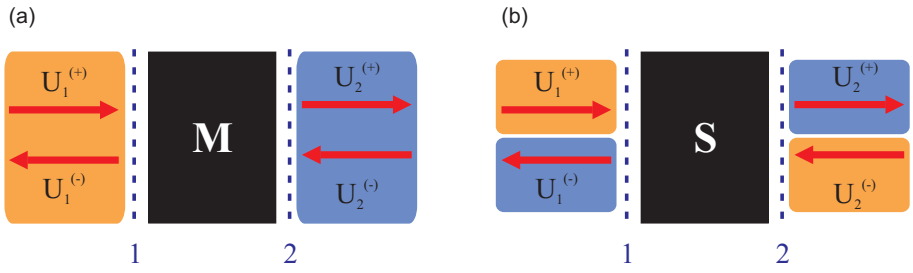


Figure 1.9: Comparisson between T-matrix (a) and S-matrix (b).

It can be easily shown (proof: exercise) that the coeffiecients of a T-matrix **M** and the corresponding S-matrix **S** are related:

$$\mathbf{S} = \begin{bmatrix} t_{12} & r_{21} \\ r_{12} & t_{21} \end{bmatrix} = \frac{1}{D} \begin{bmatrix} AD - BC & B \\ -C & 1 \end{bmatrix} \tag{1.3.33}$$

$$\mathbf{M} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \frac{1}{t_{21}} \begin{bmatrix} t_{12}t_{21} - r_{12}r_{21} & r_{21} \\ -r_{12} & 1 \end{bmatrix}. \tag{1.3.34}$$

As a first example, we consider two planes in the same medium with refractive index n separated by a distance d (see Fig. 1.10 (a)). As a result of the wave propagation, the

waves pick up an additional phase:

$$U_2^{(+)} = e^{\imath\varphi} U_1^{(+)} \tag{1.3.35}$$
$$U_1^{(-)} = e^{\imath\varphi} U_2^{(-)} \tag{1.3.36}$$

with

$$\varphi = n k_0 d. \tag{1.3.37}$$

From this, we immediately find

$$\mathbf{S} = \begin{bmatrix} e^{\imath\varphi} & 0 \\ 0 & e^{\imath\varphi} \end{bmatrix}, \tag{1.3.38}$$
$$\mathbf{M} = \begin{bmatrix} e^{\imath\varphi} & 0 \\ 0 & e^{-\imath\varphi} \end{bmatrix}. \tag{1.3.39}$$

As a second example, we consider the interface between two media (see Fig. 1.10 (b)). In this case, the coefficients of the S-matrix can be easily calculated with the Fresnel-equations for normal incidence:

$$\mathbf{S} = \begin{bmatrix} t_{12} & r_{21} \\ r_{12} & t_{21} \end{bmatrix} = \frac{1}{n_1 + n_2} \begin{bmatrix} 2n_1 & n_2 - n_1 \\ n_1 - n_2 & 2n_2 \end{bmatrix}. \tag{1.3.40}$$

Using equation (1.3.34), we obtain the corresponding T-matrix:

$$\mathbf{M} = \frac{1}{2n_2} \begin{bmatrix} n_2 + n_1 & n_2 - n_1 \\ n_2 - n_1 & n_2 + n_1 \end{bmatrix}. \tag{1.3.41}$$

These two simple T-matrices in combination with equation (1.3.31) are sufficient to calculate the T-matrix of any multilayer-structure. For instance, the T-matrix of a dielectric layer with refractive index n_2 and thickness d sandwiched between two media with refractive n_1 (see Fig. 1.10 (c)) reads:

$$\begin{aligned} \mathbf{M} &= \frac{1}{4n_1n_2} \begin{bmatrix} n_1 + n_2 & n_1 - n_2 \\ n_1 - n_2 & n_1 + n_2 \end{bmatrix} \begin{bmatrix} e^{i\varphi_2} & 0 \\ 0 & e^{-i\varphi_2} \end{bmatrix} \begin{bmatrix} n_2 + n_1 & n_2 - n_1 \\ n_2 - n_1 & n_2 + n_1 \end{bmatrix} \\ &= \frac{1}{4n_1n_2} \begin{bmatrix} 4\cos(\varphi_2)n_1n_2 + 2\imath sin(\varphi_2)\,(n_1^2 + n_2^2) & 2\imath sin(\varphi_2)\,(n_1^2 - n_2^2) \\ 2\imath sin(\varphi_2)\,(n_1^2 - n_2^2) & 4\cos(\varphi_2)n_1n_2 - 2\imath sin(\varphi_2)\,(n_1^2 + n_2^2) \end{bmatrix} \end{aligned} \tag{1.3.42}$$

with

$$\varphi_2 = n_2 k_0 d. \tag{1.3.43}$$

After comparison with equation (1.3.33) we find that the amplitude reflection coefficient r_{12} of the layer is given by

$$r_{12} = -\frac{2\imath sin(\varphi_2)\,(n_1^2 - n_2^2)}{4\cos(\varphi_2)n_1n_2 - 2\imath sin(\varphi_2)\,(n_1^2 + n_2^2)}. \tag{1.3.44}$$

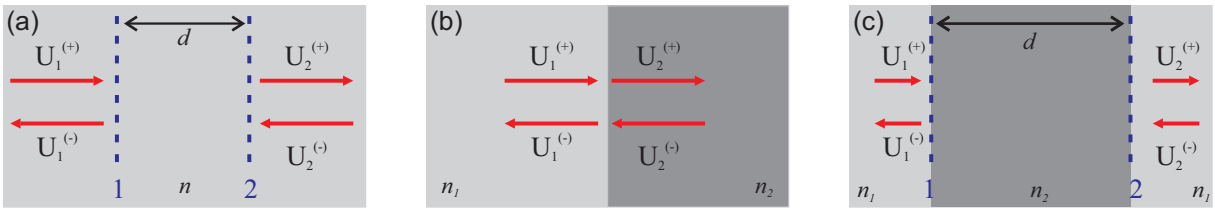


Figure 1.10: (a) Propagation in a homogeneous medium. (b) Interface between two dielectric materials. (c) Dielectric layer of thickness d.

1.4 Classical theory of material dispersion

In the final section of this chapter, we briefly discuss some generic optical properties of dielectrics and metals. In doing so, we restrict ourselves to a classical description of the materials. Moreover, we concentrate on the electric permittivity $\epsilon(\omega)$ since all known natural materials exhibit only a very weak magnetic response at optical frequencies ($\omega > 10\,\mathrm{T}Hz$) such that $\mu(\omega) \simeq 1$. The deviations of $\mu(\omega)$ from unity are typically in the order of 10^{-4}.

1.4.1 Lorentz oscillator model

The optical properties of dielectric media can be described with the Lorentz oscillator model. Here, we assume that the electrons are bound to the nucleus by an harmonic potential. An external electric field can deflect the electron cloud from its equilibrium position. The resulting restoring force is proportional to the displacement $\mathbf{x}$ of the electron cloud from the equilibrium position.

The dynamics of the system can be described with the classical equation of motion of a driven harmonic oscillator:

$$m\ddot{\mathbf{x}} + \gamma m\dot{\mathbf{x}} + m\omega_e^2\mathbf{x} = q\mathbf{E}e^{-\imath\omega t}. \tag{1.4.1}$$

Here, m is the mass of the electron cloud and q is its charge, γ is a phenomenological damping constant, ω_e is the eigenfrequency of the oscillator. The external electric field

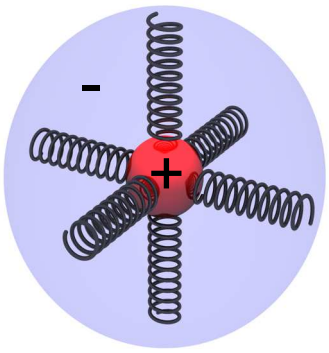


Figure 1.11: Classical atom in which the electron cloud is harmonically bound to the nucleus.

with amplitude $\mathbf{E}$ is oscillating with the frequency ω. In steady state, one can solve the equation of motion with the ansatz:

$$\mathbf{x}(t) = \mathbf{x_0} e^{-\imath\omega t}. \tag{1.4.2}$$

Inserting $\mathbf{x}(t)$ into equation (1.4.1) yields:

$$\mathbf{x_0} = \frac{q\mathbf{E}}{m}\left(\frac{1}{\omega_e^2 - \omega^2 - \imath\gamma\omega}\right). \tag{1.4.3}$$

The forced oscillations of the electron cloud relative to the nucleus results in an electric dipole moment:

$$\mathbf{p} = q\mathbf{x_0}. \tag{1.4.4}$$

The corresponding polarization density is given by:

$$\mathbf{P} = N\mathbf{p} = \frac{Nq^2\mathbf{E}}{m}\left(\frac{1}{\omega_e^2 - \omega^2 - \imath\gamma\omega}\right), \tag{1.4.5}$$

where N is the number of atoms per unit volume. By comparing the last equation with the material equation (1.1.11), we obtain the electric permittivity of the Lorentz oscillator model:

$$\epsilon_{LO}(\omega) = 1 + \frac{f}{\omega_e^2 - \omega^2 - \imath\gamma\omega} \tag{1.4.6}$$

with

$$f = \frac{Nq^2}{m\epsilon_0}. \tag{1.4.7}$$

Solving for the real and the imaginary part of $\epsilon_{LO}(\omega)$, we find:

$$\epsilon_{LO}(\omega) = 1 + \frac{f(\omega_e^2 - \omega^2)}{(\omega_e^2 - \omega^2)^2 + \omega^2\gamma^2} + \imath\frac{f\omega\gamma}{(\omega_e^2 - \omega^2)^2 + \omega^2\gamma^2}. \tag{1.4.8}$$

The corresponding refractive index $n(\omega)$ is given by

$$n(\omega) = \sqrt{\epsilon(\omega)}. \tag{1.4.9}$$

The Lorentz oscillator model can be easily extended to account for multiple resonances by summing the contributions of the individual resonances:

$$\epsilon_{LO}(\omega) = 1 + \sum_j \frac{f_j}{\omega_{e,j}^2 - \omega^2 - \imath\gamma_j\omega}. \tag{1.4.10}$$

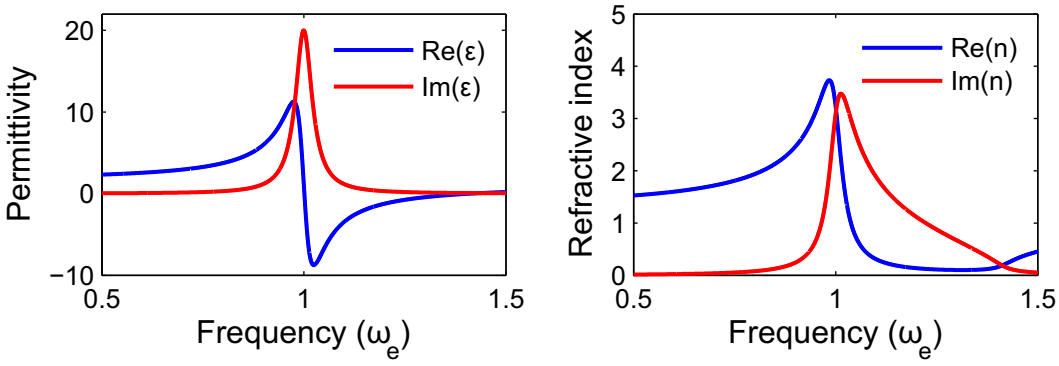


Figure 1.12: Electric permittivity and refractive index of the Lorentz oscillator model.

1.4.2 Sellmeier equation

In technical optics, one often employs the empirical Sellmeier equation instead of the Lorentz oszillator model to calculate the refractive index n of a transparent medium as a function of the vacuum wavelength λ:

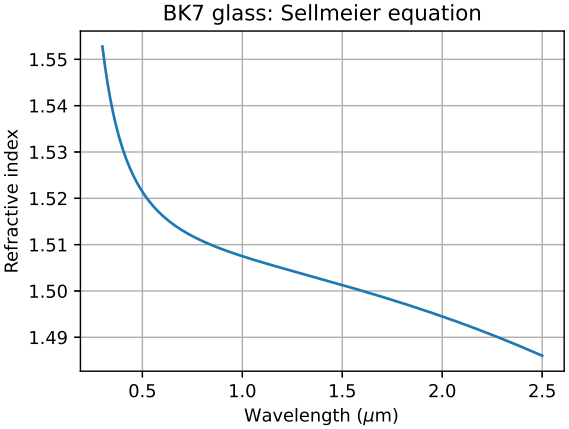
$$n^2(\lambda) = 1 + \sum_j \frac{B_j\lambda^2}{\lambda^2 - C_J}. \tag{1.4.11}$$

The coefficients B_j and C_j are the so-called Sellmeier coefficients that are used to fit the meassured refractive index data. Sellmeier coefficients for many common optical glasses and materials can be found in the online database *https://refractiveindex.info/* .

Sellmeier equation for BK7 glass

The refractive index of BK7 glass can be calculated in the wavelength range from 0.3 µm to 2.5 µm using the Sellmeier equation (source: Refractiveindex.info)

$$n^2 - 1 = \frac{1.03961212\lambda^2}{\lambda^2 - 0.00600069867} + \frac{0.231792344\lambda^2}{\lambda^2 - 0.0200179144} + \frac{1.01046945\lambda^2}{\lambda^2 - 103.560653}.$$



1.4.3 Drude model

Next, we consider a classical model to characterize the optical properties of metals. This so-called Drude model is based on the assumption that the conduction electrons of the metal form a free electron gas. The heavy atomic nuclei are immobile and give rise to a positively charged background. Interactions among the electrons and between the electrons and the positive background are only taken implicitly into account via the effective mass m.

The classical equation of motion of a conduction electron reads:

$$m\ddot{\mathbf{x}} + \gamma m\dot{\mathbf{x}} = q\mathbf{E}e^{-\imath\omega t}. \tag{1.4.12}$$

Formally, this is the equation of motion of the Lorentz oscillator model with $\omega_e = 0$. Thus, we obtain:

$$\epsilon_{\mathrm{D}}(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + \imath\gamma\omega} \tag{1.4.13}$$

with the plasma frequency

$$\omega_p = \sqrt{\frac{Ne^2}{m\epsilon_0}}. \tag{1.4.14}$$

Solving for the real and the imaginary part of $\epsilon_{\mathrm{D}}(\omega)$, we find:

$$\epsilon_{\mathrm{D}}(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + \gamma^2} + \frac{\imath\omega_p^2\gamma}{\omega^3 + \gamma^2\omega}. \tag{1.4.15}$$

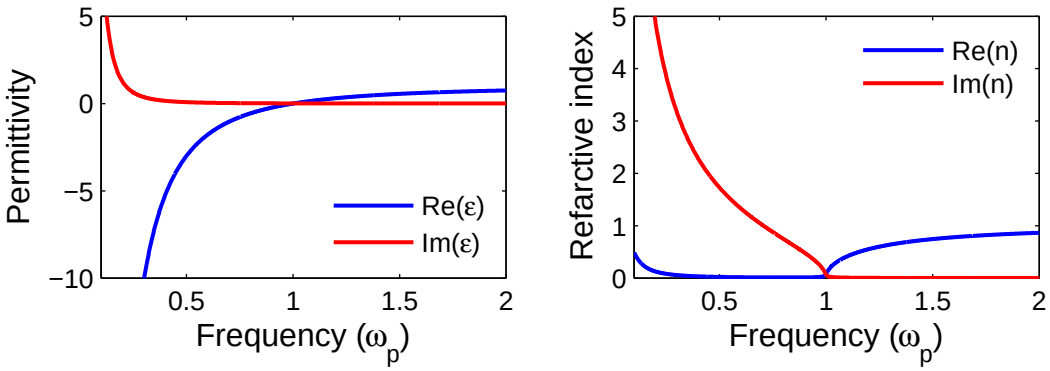


Figure 1.13: Electric permittivity and refractive index of the Drude model.

The Drude model fails to describe the optical properties of gold and copper in the visible part of the spectrum. The characteristic yellow and redish colour of these metals can be attributed to so-called interband transistions. In a classical model, they can be accounted for by additional Lorentz oscillator terms in the dielectric function:

$$\epsilon_{\mathrm{DL}}(\omega) = \epsilon_\infty - \frac{\omega_p^2}{\omega^2 + \imath\gamma_D\omega} + \sum_j \frac{f_j}{\omega_{e,j}^2 - \omega^2 - \imath\gamma_{L,j}\omega} \tag{1.4.16}$$

2 Polarization optics

Light is an electromagentic wave and as such has a vectorial character. The direction of the electric field vector $\mathbf{E}(\mathbf{r}, t)$ defines the polarization of the wave. While polarization effects are neglected in geometrical optics, it turns out that they play an important role in many light-matter interaction phenomena. For instance, the reflectance of a plane wave at an interface depends on the orientation of the electric field vector relative to the plane of incidence. In this chapter, we will discuss different polarization states of light and analyze devices that allow to prepare and maniplulate these states.

2.1 Polarization of light

Consider a monochromatic plane wave that is propagating along the z-direction in an isotropic and homogeneous dielectric medium. Its electric field can be written as:

$$\mathbf{E}(z,t) = \Re\left\{\mathbf{A}e^{\imath(kz-\omega t)}\right\}, \tag{2.1.1}$$

where

$$\mathbf{A} = A_x\hat{\mathbf{e}}_x + A_y\hat{\mathbf{e}}_y = a_x\hat{\mathbf{e}}_x + a_y e^{\imath\varphi}\hat{\mathbf{e}}_y \tag{2.1.2}$$

is a complex amplitude vector. If the amplitude vector is constant, the wave is said to have a well defined polarization state. The latter is characterized by the curve that is traced by the endpoint of the electric field vector at a fixed position $z = z_0$:

$$\mathbf{E}(z = z_0, t) = E_x\hat{\mathbf{e}}_x + E_y\hat{\mathbf{e}}_y. \tag{2.1.3}$$

with

$$E_x = a_x\cos(kz_0 - \omega t), \tag{2.1.4}$$
$$E_y = a_y\cos(kz_0 - \omega t + \varphi). \tag{2.1.5}$$

2.1.1 Linear polarization

The wave is said to be linearly polarized if the electric field vector $\mathbf{E}(z = z_0, t)$ oscillates back and forth along a line in the xy-plane. This corresponds to the case that either one

component of the amplitude vector vanishes (either $a_x = 0$ or $a_y = 0$) or that both electric field components oscillate exactly in phase ($\varphi = 0$) or out of phase ($\varphi = \pi$).

We can always find a suitable Cartesian coordinate system ($x'y'z$), such that the electric field of a linearly polarized wave has the form:

$$\mathbf{E}(z,t) = a_{x'}\hat{\mathbf{e}}_{x'} \cos(kz - \omega t)\,. \tag{2.1.6}$$

The vector $\hat{\mathbf{e}}_{x'}$ defines the polarization direction of the linearly polarized plane wave.

2.1.2 Circular polarization

Next, we consider the case that the two components of the amplitude vector have the same magnitude ($a_x = a_y = a$) and exhibit a phase shift of $\varphi = \pm\pi/2$:

$$\begin{aligned} \mathbf{E}(z = z_0, t) &= a\hat{\mathbf{e}}_x \cos(kz_0 - \omega t) + a\hat{\mathbf{e}}_y \cos\left(kz_0 - \omega t \pm \frac{\pi}{2}\right), & (2.1.7)\\ &= a\hat{\mathbf{e}}_x \cos(kz_0 - \omega t) \mp a\hat{\mathbf{e}}_y \sin(kz_0 - \omega t)\,. & (2.1.8) \end{aligned}$$

Right circular polarization

In the case $\varphi = -\pi/2$, the plane wave is said to be right circularly polarized. Its electric field is given by

$$\mathbf{E}_R(z = z_0, t) = a\hat{\mathbf{e}}_x \cos(kz_0 - \omega t) + a\hat{\mathbf{e}}_y \sin(kz_0 - \omega t)\,. \tag{2.1.9}$$

For a fixed position $z = z_0$, the tip of the electric field vector spins in a circle in the xy-plane with the angular frequency ω. While looking towards the source (hence opposite to the propagation direction), the electric field vector of a right circularly polarized wave rotates clock-wise. At a fixed point in time $t = t_0$ the tip of the electric field vector of a right circularly polarized wave describes a right-handed helix if viewed along the direction of propagation.

Leftt circular polarization

For $\varphi = \pi/2$, the plane wave is said to be left circularly polarized and its electric field can be written as

$$\mathbf{E}_L(z = z_0, t) = a\hat{\mathbf{e}}_x \cos(kz_0 - \omega t) - a\hat{\mathbf{e}}_y \sin(kz_0 - \omega t)\,. \tag{2.1.10}$$

While looking towards the source, the tip of the electric field vector of the left circularly polarized wave spins at a fixed position $z = z_0$ counter-clockwise in a circle in the xy-plane. A snapshot of the tip of the electric field vector of a left circularly polarized wave shows a left-handed helix if viewed along the direction of propagation.

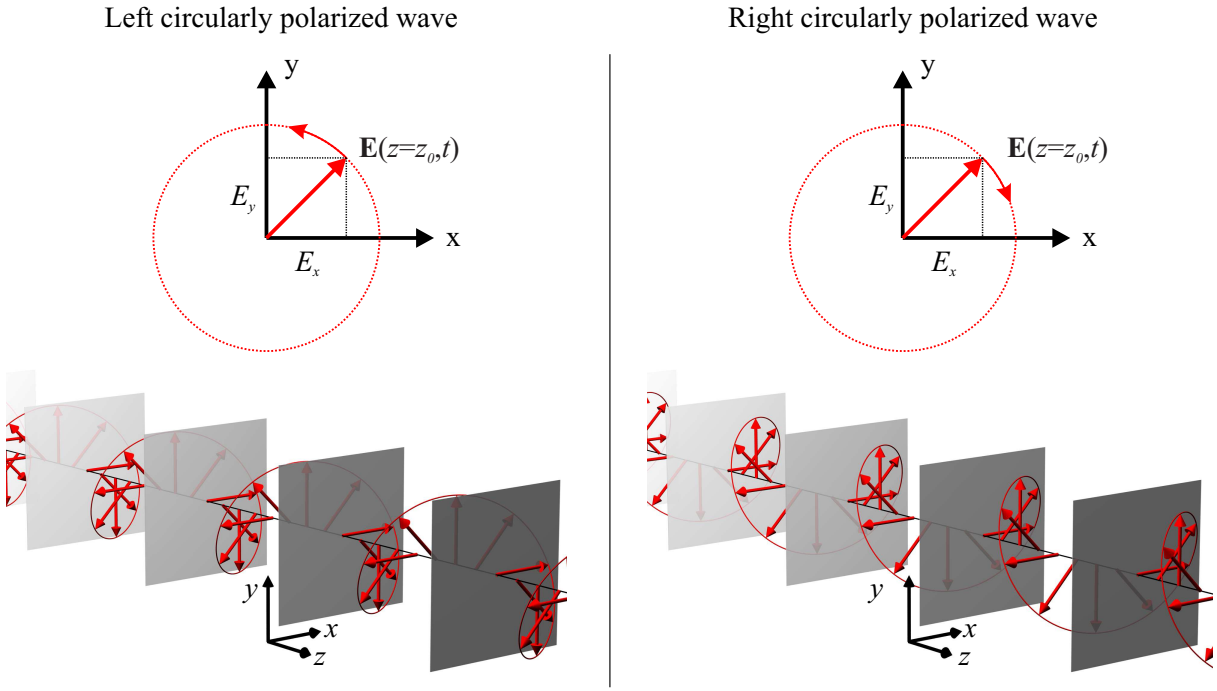


Figure 2.1: Electric field of a left circularly polarized wave and a right circularly polarized wave.

Superposition of circularly polarized waves

According to equations (2.1.9) and (2.1.10), a circularly polarized wave can be written as a superposition of two orthogonal linearly polarized waves that exhibit a relative 90° phase difference. Conversely, a superposition of a right circularly polarized wave and a left circularly polarized wave results in a linearly polarized wave, as the following example shows:

$$\begin{aligned} \mathbf{E}_R + \mathbf{E}_L &= a\,\hat{\mathbf{e}}_x \cos(kz - \omega t) + a\,\hat{\mathbf{e}}_y \sin(kz - \omega t) + \\ &\quad a\,\hat{\mathbf{e}}_x \cos(kz - \omega t) - a\,\hat{\mathbf{e}}_y \sin(kz - \omega t) \\ &= 2a\,\hat{\mathbf{e}}_x \cos(kz - \omega t)\,. \end{aligned} \tag{2.1.11}$$

2.1.3 Elliptical polarization

Except for the special cases discussed above, the tip of the electric field vector generally describes an ellipse. This can be seen by rewriting equations (2.1.4) and (2.1.5) for a fixed position $z = z_0$ as

$$\frac{E_x}{a_x} = \cos(kz_0 - \omega t) \tag{2.1.12}$$

$$\frac{E_y}{a_y} = \cos(kz_0 - \omega t + \varphi) = \cos(kz_0 - \omega t)\cos(\varphi) - \sin(kz_0 - \omega t)\sin(\varphi). \tag{2.1.13}$$

Next, we calculate the difference

$$\frac{E_y}{a_y} - \frac{E_x}{a_x}\cos(\varphi) = -\sin(kz_0 - \omega t)\sin(\varphi). \tag{2.1.14}$$

By taking the square of both sides of equation (2.1.14) and by using

$$\sin(kz_0 - \omega t) = \sqrt{1 - \left(\frac{E_x}{a_x}\right)^2}, \tag{2.1.15}$$

we obtain

$$\left(\frac{E_x}{a_x}\right)^2 + \left(\frac{E_y}{a_y}\right)^2 - 2\frac{E_x}{a_x}\frac{E_y}{a_y}\cos(\varphi) = \sin^2(\varphi). \tag{2.1.16}$$

Since this is the equation of an ellipse, one speaks in this case of elliptically polarized light. Figure 2.2 exemplifies the curves traced by the electric field vector in the xy-plane for $a_x = a_y$ and different values of φ.

2.2 Waves in anisotropic media

So far we have restricted ourselves to isotropic media for which the condition $\mathbf{D} \| \mathbf{E}$ holds. In the following we consider wave propagation in media in which the incident electric field induces a polarization that is not oriented parallel to the driving field, i.e., $\mathbf{D} \nparallel \mathbf{E}$. These optically anisotropic materials are said to be birefringent. Their optical properties can be described with the Lorentz oscillator model, if we characterize the restoring force acting on the electron cloud in different directions by different springs.

The electric permittivity $\boldsymbol{\epsilon}$ of an anisotropic medium is a second-order tensor that connects $\mathbf{D}$ with $\mathbf{E}$:

$$\begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} = \epsilon_0 \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix}. \tag{2.2.1}$$

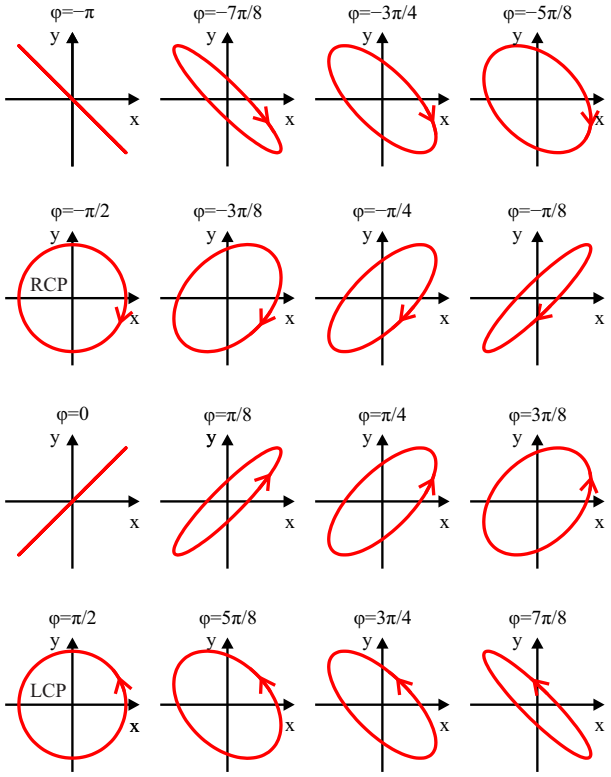


Figure 2.2: Trajectories of the tip of the electric field vector in the xy-plane of different elliptically polarized waves with $a_x = a_y$.

For any given anisotropic crystal, there is a unique coordinate system (x, y, z) in which $\boldsymbol{\epsilon}$ becomes diagonal:

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_{xx} & 0 & 0 \\ 0 & \epsilon_{yy} & 0 \\ 0 & 0 & \epsilon_{zz} \end{pmatrix} \tag{2.2.2}$$

The corresponding coordinate axes $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_y$, and $\hat{\mathbf{e}}_z$ are the so-called principal axes of the crystal. If the electric field $\mathbf{E}$ is oriented along one of the principle axes, the electric displacement field $\mathbf{D}$ points in the same direction. The principal indices of refraction are given by:

$$n_x = \sqrt{\epsilon_{xx}}, \quad n_y = \sqrt{\epsilon_{yy}}, \quad n_z = \sqrt{\epsilon_{zz}}. \tag{2.2.3}$$

In a uniaxial crystal, two of the principal refractive indices are equal ($n_x = n_y = n_o$) but different from the third one ($n_z = n_e$). The indices n_o and n_e are usually called the ordinary and the extraordinary refractive index, respectively. In the case $n_e > n_o$ the crystal is said to be positive uniaxial while crystals with $n_e < n_o$ are called negative uniaxial. The distinguished $\hat{\mathbf{e}}_z$-axis is denoted as the optical axis of the uniaxial crystal.

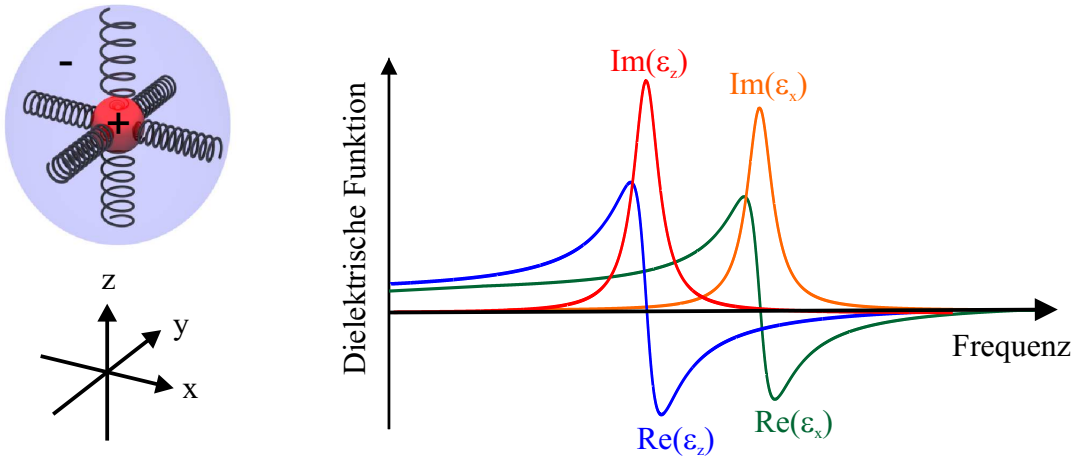


Figure 2.3: Anisotropic Lorentz oscillator model.

Ordinary and extraordinary refractive indices of selected uniaxial crystals ($\lambda = 589\,\mathrm{n}m$).

Crystal	n_o	n_e
Calcite	1.6584	1.4864
Quarz	1.5443	1.5534
Rutile (TiO_2)	2.616	2.903
Tourmaline	1.669	1.638

In the following, we consider a plane wave with wave vector $\mathbf{k}$ traveling in a uniaxial crystal without free charges and free currents ($\rho = 0, \mathbf{j} = 0$). Applying Maxwell's equations, we find:

$$\mathbf{k} \cdot \mathbf{D} = 0, \tag{2.2.4}$$

$$\mathbf{k} \cdot \mathbf{B} = 0, \tag{2.2.5}$$

$$\mathbf{k} \times \mathbf{E} = \mu_0 \omega \mathbf{H}, \tag{2.2.6}$$

$$\mathbf{k} \times \mathbf{H} = -\omega \mathbf{D}. \tag{2.2.7}$$

By definition, the Poynting vector has the following properties:

$$\mathbf{S} \cdot \mathbf{E} = 0, \tag{2.2.8}$$

$$\mathbf{S} \cdot \mathbf{H} = 0, \tag{2.2.9}$$

Next, we determine how the wave number $k = |\mathbf{k}|$ depends on the polarization of the wave and its propagation direction. For this purpose, it is convenient to express the electric field strength $\mathbf{E}$ as a function of the electric displacement field $\mathbf{D}$:

$$\mathbf{E} = \frac{\boldsymbol{\epsilon}^{-1}}{\epsilon_0}\mathbf{D} = \frac{\boldsymbol{\eta}}{\epsilon_0}\mathbf{D}. \tag{2.2.10}$$

Here, we have introduced in the second step the electric impermeability tensor $\boldsymbol{\eta}$. In the principal axis coordinate system, $\boldsymbol{\eta}$ is a diagonal tensor with elements:

$$\boldsymbol{\eta} = \begin{pmatrix} \frac{1}{n_o^2} & 0 & 0 \\ 0 & \frac{1}{n_o^2} & 0 \\ 0 & 0 & \frac{1}{n_e^2} \end{pmatrix}. \tag{2.2.11}$$

Without loss of generality, we assume that the wave is traveling in the yz-plane. We first consider a so-called ordinary wave with $\mathbf{D} = (D_1, 0, 0)$. It follows from equations (2.2.10) and (2.2.11) that in this case $\mathbf{E} \| \mathbf{D}$. As a result, we find that $\mathbf{S} \,\|\, \mathbf{k}$ and $|\mathbf{k}| = k_0 n_o$, where k_0 is the wave number in vacuum. Hence, the wave number of the ordinary wave is independent from the propagation direction.

Next, we consider a so-called extraordinary wave with $\mathbf{D} = (0, D_1, D_2)$. Inspection of equation (2.2.11) reveals that in this case generally $\mathbf{D} \nparallel \mathbf{E}$ and hence $\mathbf{k} \nparallel \mathbf{S}$. Using equations (2.2.6) and (2.2.7) we find that

$$\mathbf{k} \times (\mathbf{k} \times \mathbf{E}) = -\frac{\omega^2}{\epsilon_0 c_0^2}\mathbf{D}. \tag{2.2.12}$$

This can be written as

$$\mathbf{D} = \epsilon_0 n^2(\theta) \left(\mathbf{E} - \frac{\mathbf{k}\,(\mathbf{k} \cdot \mathbf{E})}{k^2} \right). \tag{2.2.13}$$

Here, θ is the angle enclosed by $\mathbf{k}$ and the optical axis. Next, we decompose the vectors in each case in a component perpendicular to the optical axis and a component parallel to the optical axis:

$$\mathbf{k} = k_\perp \hat{\mathbf{e}}_\perp + k_\parallel \hat{\mathbf{e}}_\parallel, \tag{2.2.14}$$

$$\mathbf{D} = D_\perp \hat{\mathbf{e}}_\perp + D_\parallel \hat{\mathbf{e}}_\parallel, \tag{2.2.15}$$

$$\mathbf{E} = E_\perp \hat{\mathbf{e}}_\perp + E_\parallel \hat{\mathbf{e}}_\parallel. \tag{2.2.16}$$

With $D_\perp = \epsilon_0 n_o^2 E_\perp$ and $D_\parallel = \epsilon_0 n_e^2 E_\parallel$, we find

$$k_\perp E_\perp = \frac{n^2(\theta) k_\perp^2\,(\mathbf{k} \cdot \mathbf{E})}{(n^2(\theta) - n_o^2)\,k^2}, \tag{2.2.17}$$

$$k_\parallel E_\parallel = \frac{n^2(\theta) k_\parallel^2\,(\mathbf{k} \cdot \mathbf{E})}{(n^2(\theta) - n_e^2)\,k^2}. \tag{2.2.18}$$

By adding the two contributions we obtain

$$k_\perp E_\perp + k_\parallel E_\parallel = \mathbf{k}\cdot\mathbf{E} = \left(\frac{n^2(\theta)k_\perp^2}{(n^2(\theta)-n_o^2)\,k^2} + \frac{n^2(\theta)k_\parallel^2}{(n^2(\theta)-n_e^2)\,k^2} \right) (\mathbf{k}\cdot\mathbf{E})\,. \tag{2.2.19}$$

This equation can be rewritten as

$$\frac{1}{n^2(\theta)} = \frac{k_\perp^2}{(n^2(\theta)-n_o^2)\,k^2} + \frac{k_\parallel^2}{(n^2(\theta)-n_e^2)\,k^2}. \tag{2.2.20}$$

With $k_\perp/k = \sin(\theta)$ and $k_\parallel/k = \cos(\theta)$, we obtain after a short calculation that

$$\frac{1}{n^2(\theta)} = \frac{\cos^2(\theta)}{n_o^2} + \frac{\sin^2(\theta)}{n_e^2}. \tag{2.2.21}$$

It can be shown that the Poynting vector $\mathbf{S}$ is oriented normal to the curve $\mathbf{k}(\theta)$.

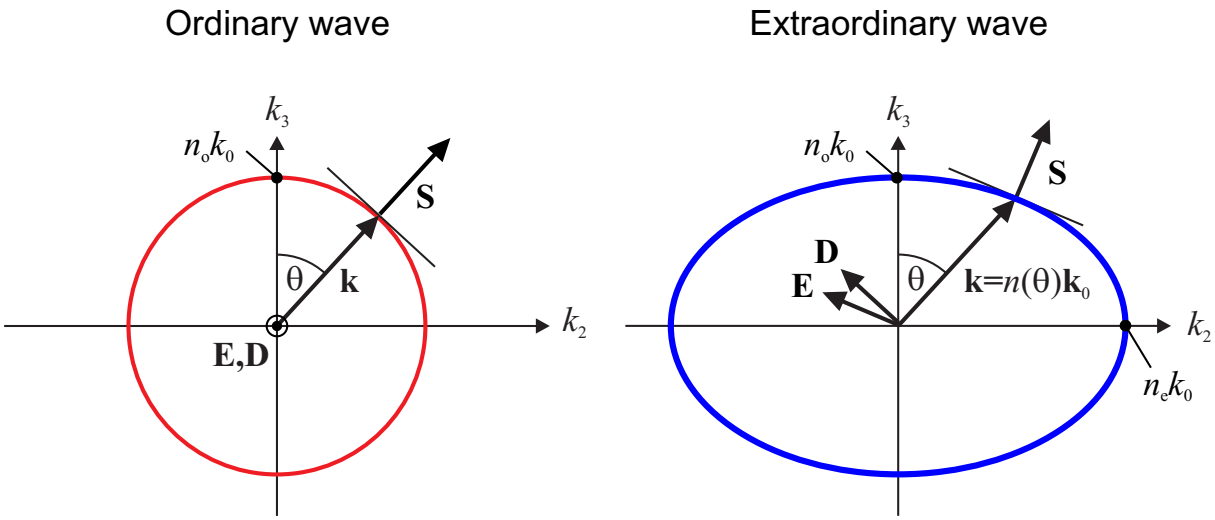


Figure 2.4: Refractive indices of an ordinary (left) and an extraordinary (right) wave in an unaxial crystal.

In order to analyze refraction at the interface between air and an uniaxial crystal, it is important to remember that the parallel component of the wave vector is a conserved quantity. First, we discuss the ordinary wave that is polarized perpendicular to the optical axis of the crystal. The absolute value of the wave vector $|\mathbf{k}_o|$ is independent from the propagation direction and the ordinary law of refraction holds:

$$|\mathbf{k}_i|\sin(\theta_i) = |\mathbf{k}_o|\sin(\theta_o) = k_0 n_o \sin(\theta_o). \tag{2.2.22}$$

In the case of the extraordinary wave, we obtain the relation:

$$|\mathbf{k}_i|\sin(\theta_i) = k_0 n_e(\theta_e+\theta_a)\sin(\theta_e). \tag{2.2.23}$$

The angle θ_e can be easily found by using the scheme depicted in Fig. 2.5. Note that the wave vector $\mathbf{k}$ and the Poynting vector $\mathbf{S}$ of the extraordinary wave are in general not parallel.

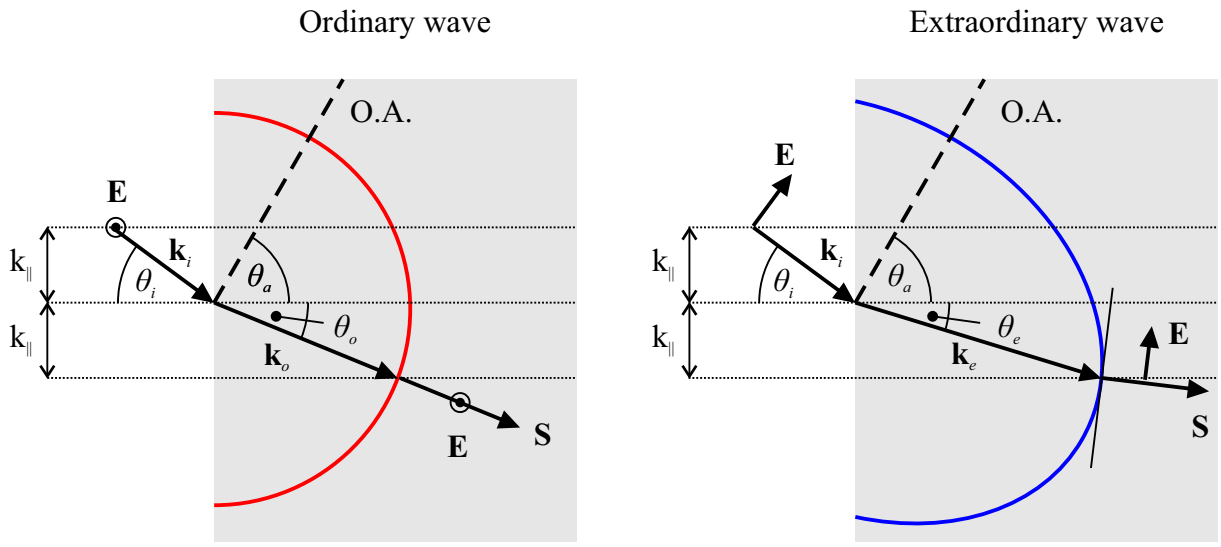


Figure 2.5: Refraction of a wave at the interface between air and an uniaxial crystal. Left: Ordinary wave. Right: Extraordinary wave.

2.3 Optical activity

An optical active material has the ability to continuously rotate the polarization direction of a linearly polarized wave as it propagates through the medium. The rotation angle is independent from the initial orientation of the polarization and the sense of rotation does not change when the wave is propagating in the opposite direction (see Fig. 2.6).

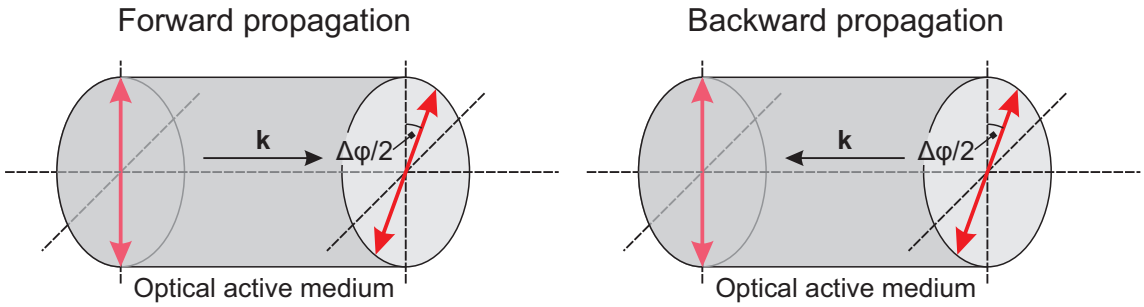


Figure 2.6: Rotation of the direction of polarization of a linearly polarized wave by an optically active medium.

The constitutive equations of an optical active medium have the following form:

$$\mathbf{D}(\mathbf{r},\omega) = \epsilon(\omega)\epsilon_0\mathbf{E}(\mathbf{r},\omega) - \imath\frac{\kappa(\omega)}{c_0}\mathbf{H}(\mathbf{r},\omega), \tag{2.3.1}$$

$$\mathbf{B}(\mathbf{r},\omega) = \imath\frac{\kappa(\omega)}{c_0}\mathbf{E}(\mathbf{r},\omega) + \mu(\omega)\mu_0\mathbf{H}(\mathbf{r},\omega) \quad . \tag{2.3.2}$$

Here, $\kappa(\omega)$ is the so-called chirality parameter that is a measure of the optical activity. It can be shown (Proof: Exercise) that the polarization eigenstates (normal modes) of an optical active material are circularly polarized waves. The refractive indices of left and right circularly polarized waves can be calculated as

$$n_L = \sqrt{\epsilon\mu} - \kappa, \tag{2.3.3}$$
$$n_R = \sqrt{\epsilon\mu} + \kappa, \tag{2.3.4}$$

A linearly polarized wave can be written as a superposition of a left and a right circularly polarized wave:

$$\mathbf{E} = E_0 \hat{\mathbf{e}}_x \, e^{\imath(kz-\omega t)} = \mathbf{E}_L + \mathbf{E}_R, \tag{2.3.5}$$

where

$$\mathbf{E}_L = \frac{E_0}{2} \left(\hat{\mathbf{e}}_x + \imath \hat{\mathbf{e}}_y\right) \, e^{\imath(kz-\omega t)}, \tag{2.3.6}$$
$$\mathbf{E}_R = \frac{E_0}{2} \left(\hat{\mathbf{e}}_x - \imath \hat{\mathbf{e}}_y\right) \, e^{\imath(kz-\omega t)}. \tag{2.3.7}$$

After propagation through an optical active medium of thickness d, the wave is given by:

$$\mathbf{E} = \mathbf{E}_L \, e^{\imath\varphi_L} + \mathbf{E}_R \, e^{\imath\varphi_R} \tag{2.3.8}$$

with

$$\varphi_L = \frac{2\pi}{\lambda_0} d n_L, \tag{2.3.9}$$
$$\varphi_R = \frac{2\pi}{\lambda_0} d n_R. \tag{2.3.10}$$

This can be rewritten as:

$$\mathbf{E} = E_0 \, e^{\imath\left(\frac{\varphi_L+\varphi_R}{2}\right)} \left[\cos\left(\frac{\Delta\varphi}{2}\right) \hat{\mathbf{e}}_x - \sin\left(\frac{\Delta\varphi}{2}\right) \hat{\mathbf{e}}_y\right] e^{\imath(kz-\omega t)} \tag{2.3.11}$$

with

$$\Delta\varphi = \varphi_R - \varphi_L = \frac{2\pi}{\lambda_0} d \left(n_R - n_L\right). \tag{2.3.12}$$

Equation (2.3.11) shows that the wave behind the optical active medium is still linearly polarized. However, the polarization direction is rotated by an angle $\frac{\Delta\varphi}{2}$. The rotary power, i.e, the rotation angle per unit length, of the medium is hence given by

$$\rho = \frac{\pi}{\lambda_0} \left(n_R - n_L\right). \tag{2.3.13}$$

2.4 Faraday effect

Many materials have the ability to rotate the polarization direction of a linearly polarized wave if one applies a static magnetic field $\mathbf{B}_{stat}$ in the propagation direction of the wave. This phenomenon is called the Faraday effect. The rotary power of the Fraday medium is given by

$$\rho = V|\mathbf{B}_{stat}|, \tag{2.4.1}$$

where V is the so-called Verdet constant of the material. An importance difference between the Faraday effect and optical activity is that in the case of the Faraday effect the sense of rotation is different for different directions of propagation (see Fig. 2.7).

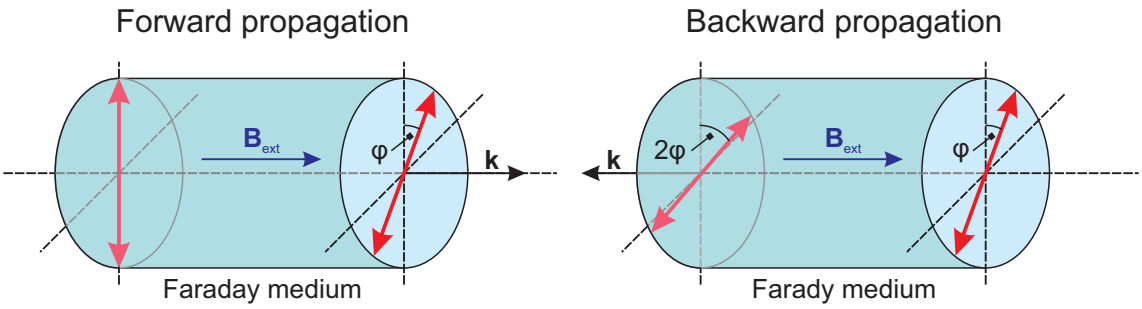


Figure 2.7: Rotation of the direction of polarization of a linearly polarized wave by a Faraday medium.

2.5 Polarization devices

Polarization devices play an important role in many optical setups. In the following, we will discuss several polarization devices that allow to control or modify the polarization state of an incident light wave.

2.5.1 Linear polarizers

A linear polarizer is a device that lets pass only the linear polarization component of an incident light field that is oriented along a specific direction, the so-called transmission axis of the polarizer. The linear polarization component orthogonal to this direction is blocked or deflected. This can be achieved either by selective absorption, selective reflection, or selective refraction.

The extinction ratio ER of a polarizer is defined as the ratio of transmission of the desired polarization T to the transmission of the orthorgonal polarization $T_\perp$. High quality linear polarizer with $ER = 10^5 - 10^6$ are typically composed of two prisms made from an unixaxial material. As an effect of the birefringence, the ordinary and the extraordinary beam are separated and exit the polarizer in different directions. Depending on the relative orientation of the optical axis of the two prisms, different beam configurations can be achieved (see Fig. 2.8)

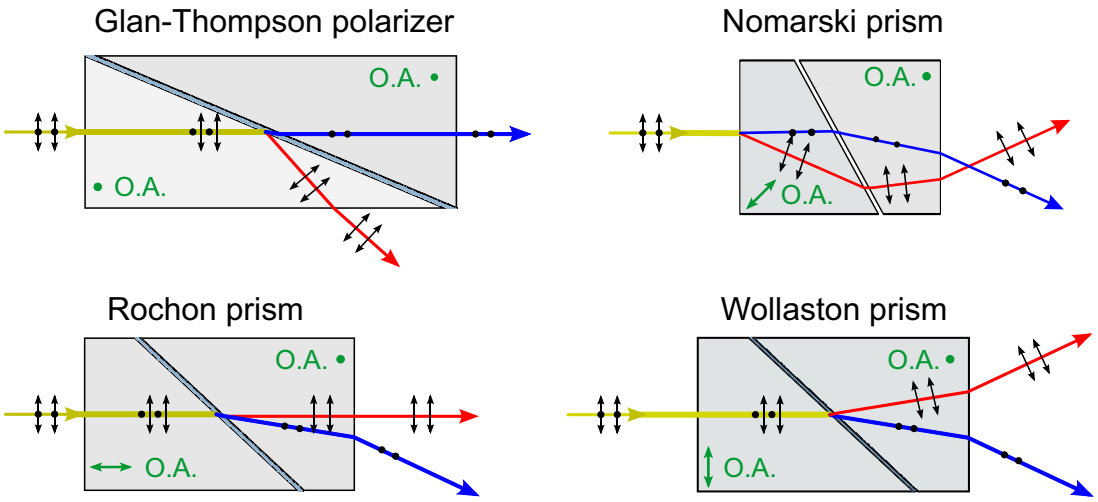


Figure 2.8: Different birefringent linear polarizers. Images adopted from Wikipedia.

2.5.2 Waveplates

Waveplates (or retarders) are devices that can be used to convert the polarization state of an electromagnetic wave. They are often made from uniaxial materials such as quartz, calcite, or mica. The operation principle of a waveplate is to shift the phase between two perpendicular polarization components of the light wave.

Consider a plate of thickness d made from a uniaxial material, whose optical axis is oriented parallel to the interface. We choose our coordinate such that the $\hat{\mathbf{e}}_z$ is normal to the interface (and hence normal to the optical axis), $\hat{\mathbf{e}}_e$ is parallel to the optical axis, and $\hat{\mathbf{e}}_o$ is normal to both $\hat{\mathbf{e}}_z$ and $\hat{\mathbf{e}}_e$. A linearly polarized plane wave impinging on the waveplate can be written as:

$$\mathbf{E}(z,t) = E_o\,\hat{\mathbf{e}}_o\,e^{\imath(k_o z-\omega t)} + E_e\,\hat{\mathbf{e}}_e\,e^{\imath(k_e z-\omega t)} \tag{2.5.1}$$

with

$$k_o = n_o\frac{2\pi}{\lambda_0}, \quad k_e = n_e\frac{2\pi}{\lambda_0}. \tag{2.5.2}$$

Since the ordinary and the extraordinary refractive index are not the same, the two components of the optical wave experience different optical path length during propagation through the wave plate. The resulting phase difference between the two components of the wave at the exit of the wave plate is given by

$$\Delta\varphi = k_o d - k_e d = \frac{2\pi}{\lambda_0} d\,(n_o - n_e)\,. \tag{2.5.3}$$

Quarter wave plate

For a quarter wave plate, the phase difference between the polarization component is $\Delta\varphi = \pm\pi/2$. This is achieved by choosing the material of the plate and its thickness d in such a way that the condition $|d(n_o - n_e)| = \lambda_0/4$ is fulfilled. It is easy to show that an incident linearly polarized wave, whose electric field vector is rotated by 45° with respect to the optical axis ($E_o = E_e$), is converted into a circularly polarized wave by the quarter wave plate.

Half wave plate

A so-called half wave plate is designed such that the condition $d(n_o - n_e) = \lambda_0/2$ holds. The corresponding phase difference is $\Delta\varphi = \pm\pi$. A half wave plate can be used to rotate the polarization direction of a linearly polarized wave (Proof: Exercise).

Babinet-Soleil compensator

A Babinet-Soleil compensator is a variable retarder that consists of a birefringent plate and two birefringent wedges. The optical axis of the plate is oriented perpendicular to that of the two wedges. By moving one wedge in or out of the beam, the phase difference between the two orthogonal polarization components can be continuously varied.

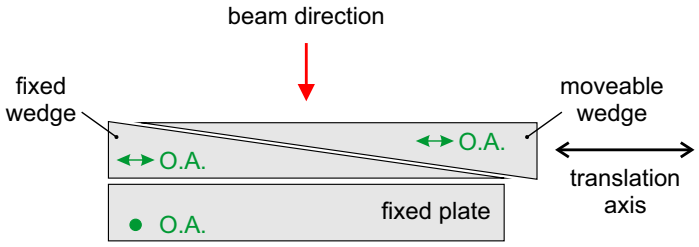


Figure 2.9: Schematic view of a Babinet-Soleil compensator.

2.5.3 Liquid crystal displays

Liquid crystals (LCs) consist of elongated molecules which exhibit an orientational order in the liquid phase. LCs can be found in different phases:

- Nematic phase: The LC molecules have no positional order, but they are self-align with their long axes roughly parallel.
- Smectic phase: The LC-molecules form well-defined layers. Within each layer the molecules are aligned parallel to each other.
- Cholesteric phase: The LC-molecules are arranged in layers. The different layers are twisted with respect to each other.

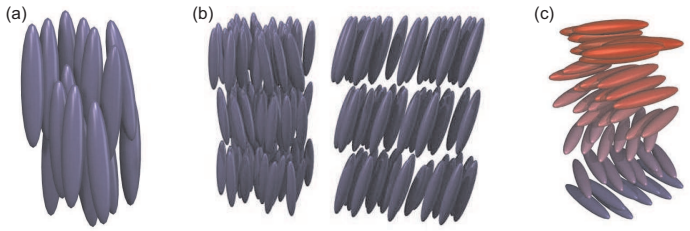


Figure 2.10: (a) Nematic phase. (b) Smectic phase. (c) Cholesteric phase. (Images taken from Wikipedia).

Fig. 2.11 depicts the scheme of a twisted nematic liquid crystal modulator as used in many displays. It consists of a thin cell filled with a nematic LC sandwiched between two transparent electrodes. Without applied voltage, the orientation of the LC molecules is determined by the surface properties of electrodes. To align the LC molecules, the electrodes are covered with a thin layer of a brushed polymer. The electrodes are arranged such that the preferential directions of the two electrodes are perpendicular to each other. This induces in the cell a 90° twist of the LC molecule orientation from one electrode to the other. The front polarizer prepares the polarization of the incident light parallel to LC molecules at the front electrode. As the light propagates through the cell, the polarization of light wave follows the twist of the LC molecules. At the second electrode, the polarization of the outgoing light is hence rotated by 90° with respect to the incident light. If the applied electric field is large enough, the LC molecules untwist and align parallel to the field. Hence, the polarization of the incident light is not rotated as it passes through the liquid crystal layer. A second polarizer behind the cell can be used to translate the polarization modulation into an amplitude modulation.

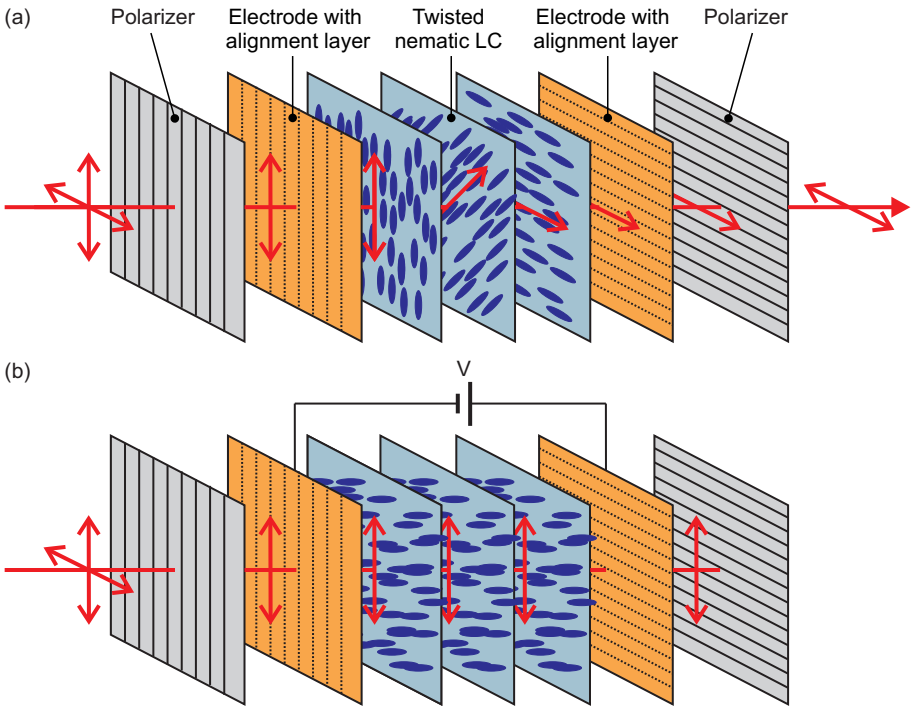


Figure 2.11: Twisted nematic liquid crystal modulator: (a) Without applied voltage. (c) With applied voltage.

2.5.4 Faraday isolator

A Faraday isolator is a device that transmits light in the one direction (forward direction) and blocks light propagating in the opposite direction (backward direction). It consists of a Faraday medium placed between two linear polarizers whose transmission axes make an angle of 45° with respect to each other (see Fig. 2.12). The magnetic field strength applied to the Faraday medium and the thickness of the Faraday medium are chosen such that the polarization of a transmitted wave is rotated by 45°.

2.6 Jones calculus

The Jones calculus is an elegant method to describe the influence of different optical systems on the polarization properties of an optical wave. Within this framework, the polarization state of an optical wave

$$\mathbf{E}(\mathbf{r},t) = (A_x\,\hat{\mathbf{e}}_x + A_y\,\hat{\mathbf{e}}_y)\;e^{\imath(kz-\omega t)} \tag{2.6.1}$$

$$= (a_x\,e^{\imath\varphi_x}\,\hat{\mathbf{e}}_x + a_y\,e^{\imath\varphi_y}\,\hat{\mathbf{e}}_y)\;e^{\imath(kz-\omega t)} \tag{2.6.2}$$

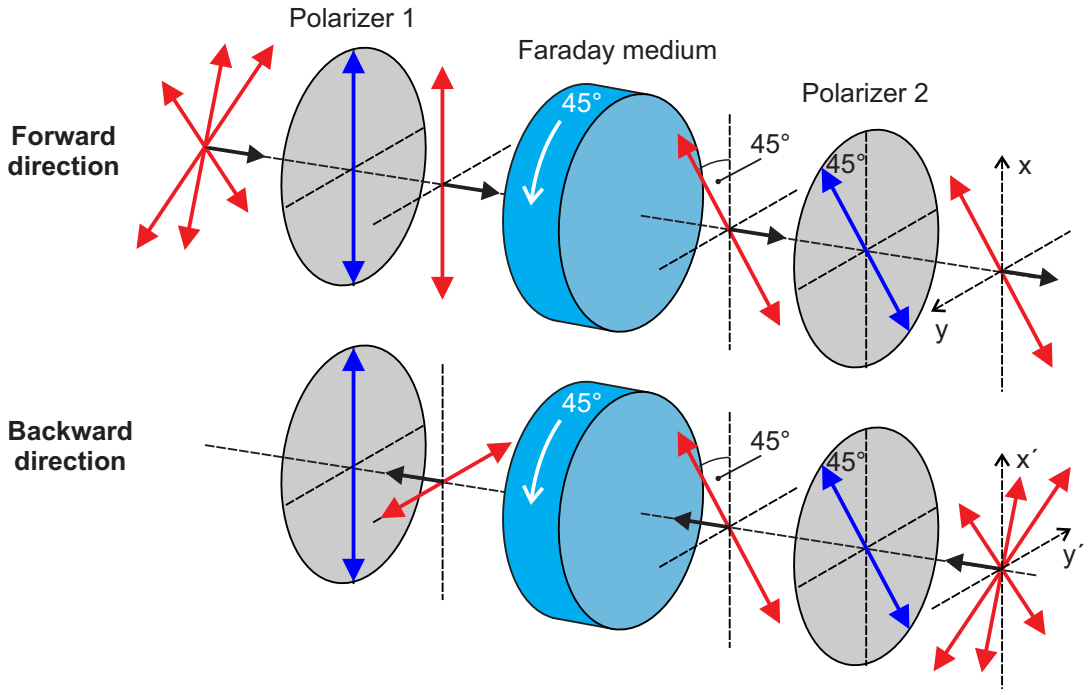


Figure 2.12: Schematic view of a Faraday isolator.

is characterized by the Jones vector

$$\mathbf{J} = \frac{1}{A}\begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{2.6.3}$$

with $A = \sqrt{a_x^2 + a_y^2}$.

Examples of normalized Jones vectors

- Linearly polarized wave (x-polarization):

$$\mathbf{J} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{2.6.4}$$

- Linearly polarized wave (y-polarization) :

$$\mathbf{J} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.6.5}$$

- Right circularly polarized wave:

$$\mathbf{J} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \tag{2.6.6}$$

- Left circularly polarized wave:

$$\mathbf{J} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \tag{2.6.7}$$

Two polarization states with Jones vector $\mathbf{J}_1$ and $\mathbf{J}_2$, respectively, are said to be orthogonal if the inner product between $\mathbf{J}_1$ and $\mathbf{J}_2$ is zero:

$$(\mathbf{J}_1, \mathbf{J}_2) = A_{1x} A_{2x}^* + A_{1y} A_{2y}^*. \tag{2.6.8}$$

An arbitrary Jones vector $\mathbf{J}$ can always be written as a weighted superposition of two orthogonal Jones vectors:

$$\mathbf{J} = \alpha_1 \mathbf{J}_1 + \alpha_2 \mathbf{J}_2 \tag{2.6.9}$$

with

$$\alpha_1 = (\mathbf{J}, \mathbf{J}_1), \quad \alpha_2 = (\mathbf{J}, \mathbf{J}_2). \tag{2.6.10}$$

Next, we consider an optical wave that passes through an optical system. We assume that the polarization of the wave is altered on this occasion from the input state $\mathbf{J}_\mathrm{i}$ to the output state $\mathbf{J}_\mathrm{t}$. If the optical system is linear, i.e., the superposition principle is obeyed, its effect on the polarization of the optical wave can be described by a 2×2 matrix, the so called Jones matrix $\mathbf{T}$, which connects $\mathbf{J}_\mathrm{i}$ with $\mathbf{J}_\mathrm{t}$:

$$\mathbf{J}_\mathrm{t} = \mathbf{T}\, \mathbf{J}_\mathrm{i}, \tag{2.6.11}$$

where

$$\mathbf{T} = \begin{pmatrix} T_{xx} & T_{xy} \\ T_{yx} & T_{yy} \end{pmatrix}. \tag{2.6.12}$$

The effect of a sequence of optical systems on a plane wave can be calculated by multiplication of the corresponding Jones matrices:

$$\mathbf{T} = \mathbf{T}_n \cdots \mathbf{T}_2 \mathbf{T}_1. \tag{2.6.13}$$

Here, $\mathbf{T}_1$ is the Jones matrix of the optical system that is first traversed by the optical wave, $\mathbf{T}_2$ is connected with the second optical system, and so on and so forth.

Jones matrices of selected optical systems:

- Linear polarizer along x-direction:

$$\mathbf{T} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.6.14}$$

- Linear polarizer along y-direction:

$$\mathbf{T} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.6.15}$$

- Wave plate with optical axis aligned in the x-direction:

$$\mathbf{T} = \begin{pmatrix} 1 & 0 \\ 0 & e^{\imath\Delta\varphi} \end{pmatrix} \tag{2.6.16}$$

  with $\Delta\varphi = \frac{2\pi}{\lambda_0} d\,(n_o - n_e)$.

- Polarization rotator (optical active medium or Faraday medium that rotates the polarization of a linearly polarized wave by an angle θ):

$$\mathbf{T} = \begin{pmatrix} cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{2.6.17}$$

The elements of a given Jones vector $\mathbf{J}$ depend on the choice of the coordinate system. With the help of the rotation matrix $\mathbf{R}(\theta)$, we can express the polarization state of the wave in a coordinate system rotated by an angle α as:

$$\mathbf{J}' = \mathbf{R}(\alpha)\mathbf{J} \tag{2.6.18}$$

with

$$\mathbf{R}(\alpha) = \begin{pmatrix} cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}. \tag{2.6.19}$$

We can transform the Jones matrix $\mathbf{T}$ of a given optical system into the new coordinate system by the following matrix relation (Proof: Exercise):

$$\mathbf{T}' = \mathbf{R}(\alpha)\,\mathbf{T}\,\mathbf{R}(-\alpha). \tag{2.6.20}$$

Analogously, one finds that

$$\mathbf{T} = \mathbf{R}(-\alpha)\,\mathbf{T}'\,\mathbf{R}(\alpha). \tag{2.6.21}$$

Example 1: Linear polarizer, whose transmission axis is tilted by an angle α with respect to the x-axis

We first consider the polarizer in a new coordinate system, which is rotated by the angle α with respect to the x-axis. In this new coordinate system, the Jones matrix $\mathbf{T}'$ of the polarizer is given by equation (2.6.14). Using equation (2.6.21), we can then write the Jones matrix $\mathbf{T}$ of the polarizer in the original coordinate system as:

$$\begin{aligned}\mathbf{T} = \mathbf{R}(-\alpha)\,\mathbf{T}'\,\mathbf{R}(\alpha) &= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\alpha) & \cos(\alpha)\sin(\alpha) \\ \cos(\alpha)\sin(\alpha) & \sin^2(\alpha) \end{pmatrix}. \end{aligned} \tag{2.6.22}$$

Example 2: Wave plate whose optical axis is tilted by an angle α with respect to the x-axis

Again, we first consider the wave plate in the rotated coordinate system. In this coordinate system, the Jones matrix $\mathbf{T}'$ of the wave plate is given by equation (2.6.16). Transforming back to the original coordinate system, we can then write the Jones matrix $\mathbf{T}$ of the wave plate as:

$$\begin{aligned}\mathbf{T} &= \mathbf{R}(-\alpha)\,\mathbf{T}'\,\mathbf{R}(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{\imath\Delta\varphi} \end{pmatrix} \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\alpha) + \sin^2(\alpha)e^{\imath\Delta\varphi} & \cos(\alpha)\sin(\alpha)\left[1 - e^{\imath\Delta\varphi}\right] \\ \cos(\alpha)\sin(\alpha)\left[1 - e^{\imath\Delta\varphi}\right] & \sin^2(\alpha) + \cos^2(\alpha)e^{\imath\Delta\varphi} \end{pmatrix}. \end{aligned} \tag{2.6.23}$$

3 Beam optics

3.1 Paraxial Helmholtz equation

In this section we investigate the properties of optical beams. Here we concentrate on so-called paraxial waves whose wavefronts make small angles with the average propagation direction (in our case: the z-axis). A paraxial wave locally resembles a plane wave. Its electric field can be written as the product of a plane wave times a spatial envelope function:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{A}(\mathbf{r})\, e^{\imath(kz-\omega t)}. \tag{3.1.1}$$

We assume that the envelope function $\mathbf{A}(\mathbf{r})$ varies slowly along the average propagation direction z such that the following relations are fulfilled (slowly varying envelope approximation):

$$\partial\mathbf{A}/\partial z \;\ll\; k\mathbf{A}, \tag{3.1.2}$$
$$\partial^2\mathbf{A}/\partial z^2 \;\ll\; k^2\mathbf{A}. \tag{3.1.3}$$

Substituting $\mathbf{E}(\mathbf{r})$ in the Helmholtz equations (1.2.25) and neglecting $\partial^2\mathbf{A}/\partial z^2$ in comparison with $k\partial\mathbf{A}/\partial z$ and $k^2\mathbf{A}$ yields the so-called paraxial Helmholtz equation for the envelope function:

$$\nabla_t^2\mathbf{A}(\mathbf{r}) + \imath 2k\frac{\partial\mathbf{A}(\mathbf{r})}{\partial z} = 0 \quad with \quad \nabla_t^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2. \tag{3.1.4}$$

3.2 Paraboloidal waves

As a first example, we consider a spherical wave that is centered at the origin. For this purpose we start with the solution of the scalar wave equation in spherical coordinates:

$$\Psi(r,t) = \frac{A_0}{r}\, e^{\imath(kr-\omega t)}. \tag{3.2.1}$$

with

$$r^2 = x^2 + y^2 + z^2 = \rho^2 + z^2. \tag{3.2.2}$$

In the following we restrict ourselves to a small region around the z-axis. For $z^2 \gg \rho^2$, we can approximate r by

$$r = \sqrt{\rho^2 + z^2} \approx z\left(1 + \frac{1}{2}\frac{\rho^2}{z^2}\right). \tag{3.2.3}$$

Substituting $r = z\left(1 + \frac{1}{2}\frac{\rho^2}{z^2}\right)$ into the phase and $r = z$ into the denominator of $\Psi(r,t)$, we obtain the scalar paraboloidal wave:

$$\Psi(r,t) \approx \frac{A_0}{z}\,\exp\left(\imath k\frac{\rho^2}{2z}\right) e^{\imath(kz-\omega t)}. \tag{3.2.4}$$

Next, we take the vector nature of the electromagnetic field into account. For this purpose we introduce the polarization direction $\hat{\mathbf{e}}_p$ which fulfills the condition $\hat{\mathbf{e}}_p \perp \hat{\mathbf{e}}_z$. One can easily show (proof: exercise) that the vectorial paraboloidal wave

$$\mathbf{A}\,(\rho,z) = \frac{A_0\hat{\mathbf{e}}_p}{z}\,\exp\left(\imath k\frac{\rho^2}{2z}\right) \tag{3.2.5}$$

satisfies the paraxial Helmholtz equation. The substitution $z \to z - z_0, z_0 \in \Re$ in equation (3.2.5) yields another solution of the paraxial Helmholtz equation:

$$\mathbf{A}\,(\rho,z) = \frac{A_0\hat{\mathbf{e}}_p}{z - z_0}\,\exp\left(\imath k\frac{\rho^2}{2(z-z_0)}\right). \tag{3.2.6}$$

This new solution is a paraboloidal wave whose origin is shifted to z_0. We note that paraboloidal waves serve in the Fresnel diffraction theory as secondary waves that determine the wave front at a later time.

3.3 Gaussian beams

Next, we consider the substitution $z \to q(z) = z - \imath z_0, z_0 \in \Re$ in equation (3.2.5). One can show (proof: exercise) that the so-called Gaussian beam

$$\mathbf{A}\,(\rho,z) = \frac{\tilde{\mathbf{A}}_0}{q(z)}\exp\left(\imath k\frac{\rho^2}{2q(z)}\right) \tag{3.3.1}$$

is also a solution of the paraxial Helmholtz equation[1]. Gaussian beams play an important role in optics because many lasers emit beams that approximate a Gaussian intensity profile (see below).

[1] Note that a Gaussian beam is not a solution of the Helmholtz equation!

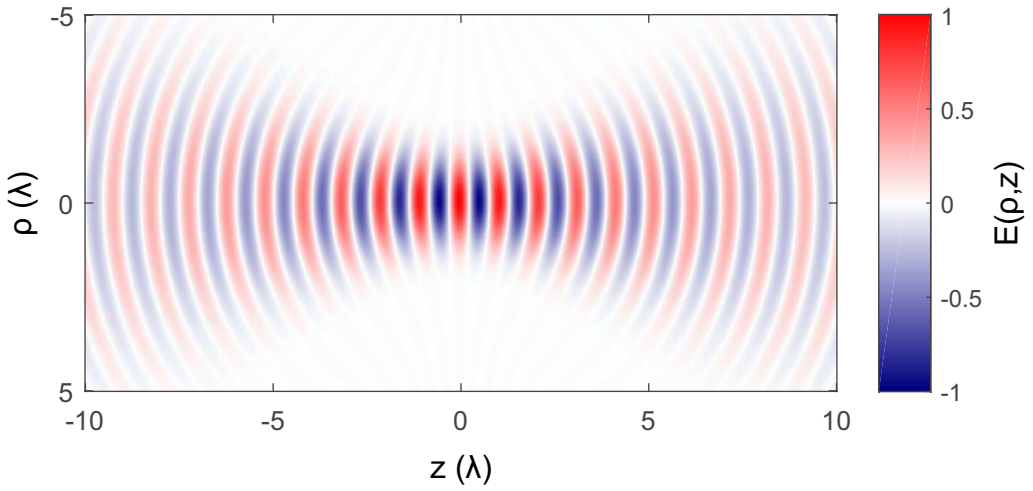


Figure 3.1: Electric field distribution of a Gaussian beam in the vicinity of the beam waist.

The quantity $q(z)$ is the so-called q-parameter of the beam and z_0 is usually denoted as the Rayleigh range. To obtain a better insight into the properties of Gaussian beams, it is useful to define two new real-valued functions $R(z)$ and $W(z)$ that are related to the q-parameter through

$$\frac{1}{q(z)} = \frac{z + \imath z_0}{z^2 + z_0^2} = \frac{1}{R(z)} + \imath \frac{2}{kW^2(z)}. \tag{3.3.2}$$

With

$$\exp\left(\imath k \frac{\rho^2}{2q(z)}\right) = \exp\left(-\frac{\rho^2}{W^2(z)}\right) \exp\left(\imath k \frac{\rho^2}{2R(z)}\right) \tag{3.3.3}$$

and[2]

$$\frac{i}{q(z)} = -\frac{1}{z_0} \frac{W_0}{W(z)} \exp\left(-\imath \eta(z)\right) \tag{3.3.4}$$

we obtain the following alternative representation of a Gaussian beam[3]:

$$\mathbf{E}\left(\rho, z\right) = \mathbf{A}_0 \frac{W_0}{W(z)} \exp\left(-\frac{\rho^2}{W^2(z)}\right) \exp\left(\imath k \frac{\rho^2}{2R(z)}\right) \exp\left(\imath kz - \imath \eta(z)\right). \tag{3.3.5}$$

Here, we have introduced the beam parameters:

$$W_0 = \sqrt{\frac{\lambda z_0}{\pi}}, \tag{3.3.6}$$

[2] $a + ib = \sqrt{a^2 + b^2} \exp\left(\imath \arctan(b/a)\right)$

[3] In contrast to equation(3.3.1), we also include here the plane wave factor $\exp(\imath kz)$.

$$W(z) = W_0\sqrt{1+\left(\frac{z}{z_0}\right)^2}, \tag{3.3.7}$$

$$R(z) = z\left(1+\left(\frac{z_0}{z}\right)^2\right) \tag{3.3.8}$$

$$\eta(z) = \arctan\left(z/z_0\right), \tag{3.3.9}$$

$$\mathbf{A}_0 = \frac{\imath\tilde{\mathbf{A}}_0}{z_0}. \tag{3.3.10}$$

The time-averaged intensity of a Gaussian beam in vacuum as a function of z and ρ is given by:

$$I\left(\rho,z\right) = \frac{c\epsilon_0}{2}|\mathbf{E}\left(\rho,z\right)|^2 = I_0\left(\frac{W_0}{W(z)}\right)^2\exp\left(-2\frac{\rho^2}{W^2(z)}\right) \tag{3.3.11}$$

with

$$I_0 = \frac{c\epsilon_0}{2}|\mathbf{A}_0|^2. \tag{3.3.12}$$

The name Gaussian beam can be traced back to the radial intensity profile of $I\left(\rho,z\right)$. For any value of z, the function $I\left(\rho,z=const\right)$ is a Gaussian function of the radial distance ρ that assumes its maximum at $\rho=0$, i.e., on the optical axis (see Fig. 3.2).

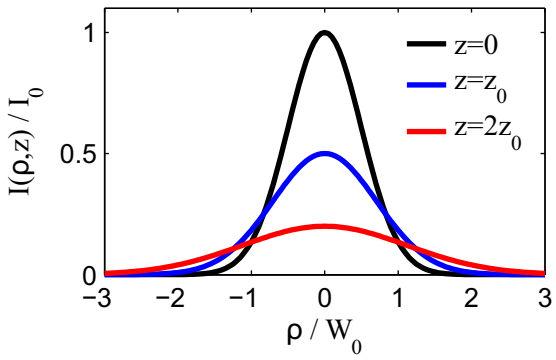


Figure 3.2: Radial intensity profiles of a Gaussian beam for three different .

The axial intensity profile of a Gaussian beam is given by:

$$I\left(\rho=0,z\right) = I_0\left(\frac{W_0}{W(z)}\right)^2 = \frac{I_0}{1+\left(z/z_0\right)^2}. \tag{3.3.13}$$

It has a peak at $z=0$ with $I(0,0)=I_0$ and drops to half the peak value at $z=\pm z_0$, i.e, at the distance of one Rayleigh range from the peak (see 3.3).

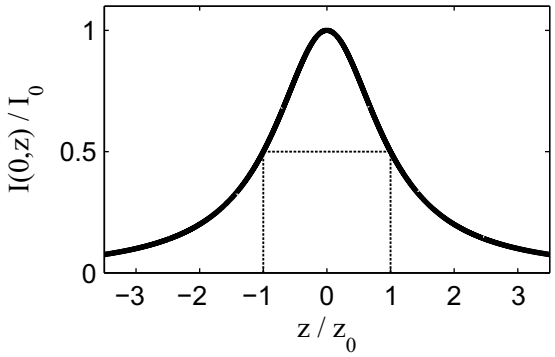


Figure 3.3: Axial intensity profile of a Gaussian beam.

The total power carried by a Gaussian beam can be calculated by integrating equation (3.3.11) over an arbitrary transverse plane:

$$P_{total} = \int_0^\infty I(\rho, z)\, 2\pi\rho d\rho = \frac{1}{2} I_0 \left(\pi W_0^2\right). \tag{3.3.14}$$

A circle of radius ρ_0 centered around the optical axis contains the fraction

$$\frac{P(\rho_0)}{P_{total}} = \frac{\int_0^{\rho_0} I(\rho, z)\, 2\pi\rho d\rho}{\int_0^\infty I(\rho, z)\, 2\pi\rho d\rho} = 1 - \exp\left(-2\frac{\rho_0^2}{W^2(z)}\right) \tag{3.3.15}$$

of the total power carried by the beam.

The function $W(z)$ is often taken as a measure of the width of a Gaussian beam. Thus, $W(z)$ is usually referred to as the beam width. At a distance $W(z)$, the intensity is by a factor of $1/e^2$ smaller than the on-axis intensity (see Fig.3.4). A circle with radius $\rho_0 = W(z)$ contains $\sim 87\%$ of the total power.

According to equation (3.3.7), $W(z)$ is defined by:

$$W(z) = W_0 \sqrt{1 + \left(\frac{z}{z_0}\right)^2}, \tag{3.3.16}$$

where

$$W_0 = \sqrt{\frac{\lambda z_0}{\pi}} \tag{3.3.17}$$

is the so-called waist radius. For $z \gg z_0$, the beam width increases linearly with z:

$$W(z) \approx W_0 \frac{z}{z_0} = \theta_{div} z. \tag{3.3.18}$$

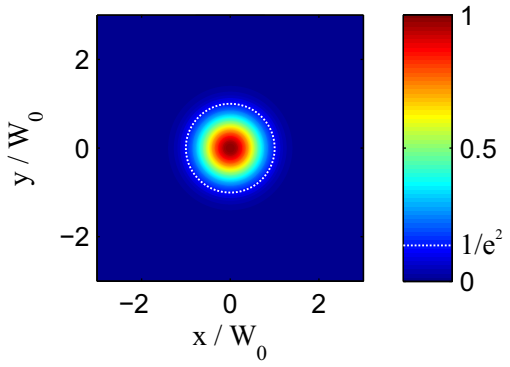


Figure 3.4: Transverse intensity distribution of a Gaussian beam at the beam waist ($z = 0$). The dotted white line is a circle with radius $W(z = 0) = W_0$.

Here, we have introduced in the last step the beam divergence angle (see Fig.3.5)

$$\theta_{div} = \frac{W_0}{z_0} = \frac{\lambda}{\pi W_0}. \tag{3.3.19}$$

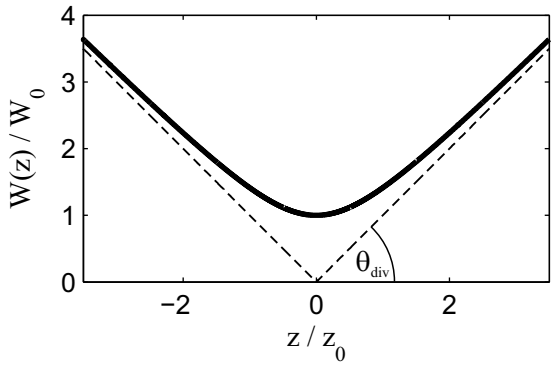


Figure 3.5: Beam width $W(z)$ and divergence angle θ_{div} of a Gaussian beam.

The phase of the electric field of a Gaussian beam is given by

$$\varphi(\rho, z) = kz - \eta(z) + \frac{k\rho^2}{2R(z)}, \tag{3.3.20}$$

where

$$\eta(z) = \arctan\left(z/z_0\right) \tag{3.3.21}$$

is the Gouy phase. On the optical axis ($\rho = 0$) equation (3.3.20) simplifies to

$$\varphi(0, z) = kz - \eta(z). \tag{3.3.22}$$

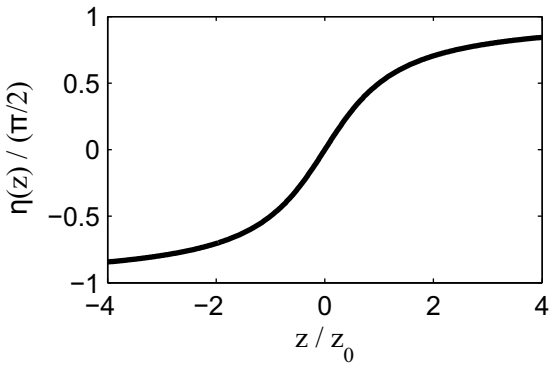


Figure 3.6: Gouy phase versus axial position z.

The Gouy phase $\eta(z)$ describes an additional retardation of the Gaussian beam with respect to a plane wave (kz). The retardation becomes $-\pi/2$ for $z = -\infty$ and $\pi/2$ for $z = \infty$.

The wave fronts of a Gaussian beam fulfill the condition

$$kz - \eta(z) + \frac{k\rho^2}{2R(z)} = 2\pi p, \tag{3.3.23}$$

where p is a constant.

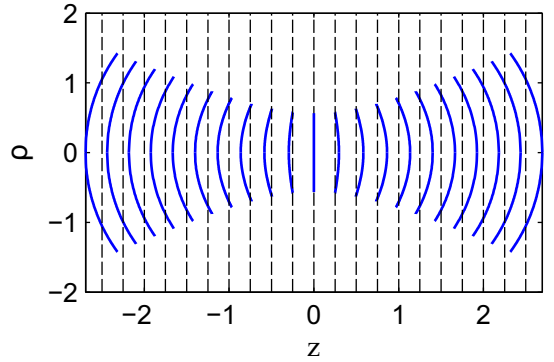


Figure 3.7: Wave fronts of a Gaussian beam (blue curves) and a plane wave (black dottet lines). The phase changes by $\pi/2$ between adjacent wave fronts.

Equation (3.3.23) can be rewritten as

$$z + \frac{\rho^2}{2R(z)} = \lambda p + \frac{\lambda\eta(z)}{2\pi}. \tag{3.3.24}$$

Both $\eta(z)$ and $R(z)$ are relatively slowly varying functions so that these two quantities are effectively constant within the beam width on each wavefront. For constant R and η, equation (3.3.24) represents a paraboloidal surface with radius of curvature R. Near the beam waist ($z = 0$) the phase fronts are (almost) planar and R tends to infinity. For large z, the radius of curvature increases almost linearly and the wavefronts are approximately those of a spherical wave.

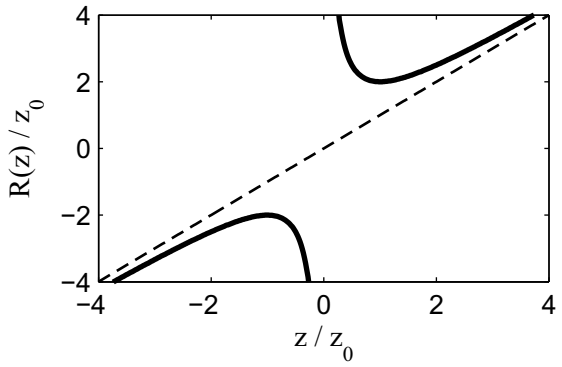


Figure 3.8: Radius of curvature of a Gaussian beam (solid curve). The dotted curve represents the radius of curvatute of a spherical wave.

3.3.1 Beam quality

Optical beams emitted by real lasers deviate from the ideal Gaussian beam profile. To characterize the quality of a laser beam, we define the so-called $\mathbb{M}^2$-factor as the ratio of the waist-diameter-divergence product of the laser to that of a Gaussian beam:

$$\mathbb{M}^2 = \frac{2W_l 2\theta_l}{4\lambda/\pi}. \tag{3.3.25}$$

Here, $2W_l$ is the diameter of the laser beam and $2\theta_l$ is its angular divergence. For the Gaussian beam, we have used that $2W_0 2\theta_0 = 4\lambda/\pi$.

One can show that a Gaussian beam exhibits the smallest possible divergence angle for a given waist diameter $\Rightarrow \mathbb{M}^2 \geq 1$.

Laser type	Typical $\mathbb{M}^2$-value
He-Ne laser	1.1
Ion laser	1.1 - 1.3
High-power multimode laser diodes	3 -4

3.3.2 Beam focusing

The electric field of a Gaussian beam directly behind the lens with focal length f is given by the product of its electric field directly in front of the lens times the transfer function of the lens:

$$\mathbf{E}_{out}(z,\rho) = h(\rho)\mathbf{E}_{in}(z,\rho), \tag{3.3.26}$$

where

$$h(\rho) = \exp\left[-\imath\frac{k_0\rho^2}{2f}\right]. \tag{3.3.27}$$

Transfer function of a lens

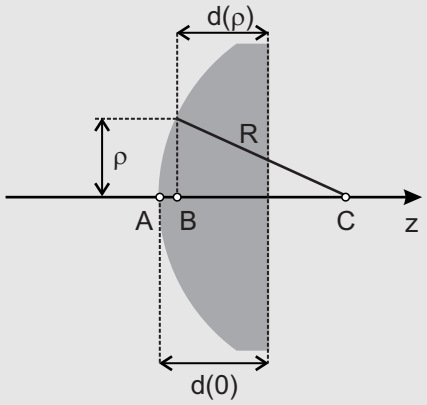
The transfer function of a lossless dielectric medium with refractive index n and thickness profile $d(x,y)$ is given by:

$$h(x,y) = \exp\left[\imath(n-1)k_0 d(x,y)\right]. \tag{3.3.28}$$

In the case of planoconfex lens, the thickness profile changes with the distance $\rho = \sqrt{x^2+y^2}$ from the optical axis according to

$$d(\rho) = d(0) - \overline{AB} = d(0) - \left(R - \overline{BC}\right) = d(0) - \left(R - \sqrt{R^2-\rho^2}\right), \tag{3.3.29}$$

where R is the radius of curvature of the spherical interface (see figure below).



For $\rho^2 \ll R^2$, we obtain

$$d(\rho) \approx d(0) - \frac{\rho^2}{2R}. \tag{3.3.30}$$

This allows us to write the transfer function of the lens as

$$h(x,y) = \exp\left[-\imath\frac{k_0\rho^2}{2f}\right] \tag{3.3.31}$$

with

$$f = \frac{R}{n-1} \tag{3.3.32}$$

Since the lens only modifies the phase of the beam, the beam width W directly behind the lens is the same as the beam width directly in front of the lens.

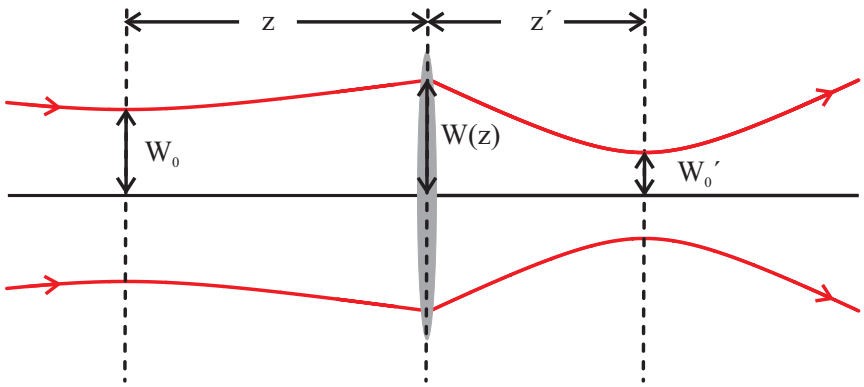


Figure 3.9: Focusing of a Gaussian beam with a thin lens.

In the following, we assume that the Gaussian beam is centered at $z = 0$ and has a beam waist W_0. The lens is located at the position z. The phase of the beam directly in front of the lens is thus given by:

$$\varphi(\rho,z) = kz - \eta(z) + \frac{k\rho^2}{2R(z)}. \tag{3.3.33}$$

After transversing the lens, the phase of the beam becomes:

$$\varphi'(\rho,z) = kz - \eta(z) + \frac{k\rho^2}{2R(z)} - \frac{k\rho^2}{2f} = kz - \eta(z) + \frac{k\rho^2}{2R'(z)}, \tag{3.3.34}$$

with

$$\frac{1}{R'(z)} = \frac{1}{R(z)} - \frac{1}{f}. \tag{3.3.35}$$

The parameters of a Gaussian beam can be determined if the radius of curvature R and the beam width W are known for a given point of the beam axis. It can be shown (proof: exercise), that the distance of this point to the beam waist, z, and the waist radius, W_0, are given by

$$W_0 = \frac{W}{\sqrt{1 + \left(\pi W^2/\lambda R\right)^2}}, \tag{3.3.36}$$

$$z = \frac{R}{1 + \left(\lambda R/\pi W^2\right)^2}. \tag{3.3.37}$$

We can use equations (3.3.35)-(3.3.37) to determine the position of the beam waist, z', and the waist radius, W_0', of the Gaussian beam behind the lens:

$$W_0' = \frac{W}{\sqrt{1 + \left(\pi W^2/\lambda R'\right)^2}}, \tag{3.3.38}$$

$$-z' = \frac{R'}{1 + \left(\lambda R'/\pi W^2\right)^2}. \tag{3.3.39}$$

The minus sign in the last equation indicates that the beam waist is located right of the lens.

In the following, we assume that the lens is located at the waist of the incident beam ($z = 0$, $W = W_0$, $R = \infty$). In this case, the radius of curvature of the Gaussian beam directly behind the lens is $R' = -f$. The waist of the beam is located at the distance

$$|z'| = \frac{f}{1 + \left(\lambda f/\pi W_0^2\right)^2} \tag{3.3.40}$$

behind the lens and the corresponding waist radius is given by

$$W_0' = \frac{W_0}{\sqrt{1 + \left(\pi W_0^2/\lambda f\right)^2}}. \tag{3.3.41}$$

For an incident collimated beam ($z_0 \gg f$), the last equation simplifies to

$$W_0' \approx \frac{\lambda}{\pi W_0} f. \tag{3.3.42}$$

To achieve the smallest possible spot size in the focus, the waist of the incident beam must be as large as possible, i.e., it must fill the aperture (diameter: D) of the lens. With $D = 2W_0$, we obtain:

$$2W_{0,min}' \approx \frac{4}{\pi} \lambda F_{\#}, \tag{3.3.43}$$

where the F-number is defined by

$$F_{\#} = \frac{f}{D}. \tag{3.3.44}$$

3.4 Optical tweezers

Light does not only carry energy but also momentum. Thus, if a beam of light impinges on an object and is absorbed or scattered it will exert a force on the object. Under everyday conditions, this optical force is much to small to be detected. However, in the case of a strongly focused laser beam and small objects, the optical force becomes relevant. Applications of the optical force include the trapping of ultracold neutral atoms and the manipulation of biological systems with so-called optical tweezers.

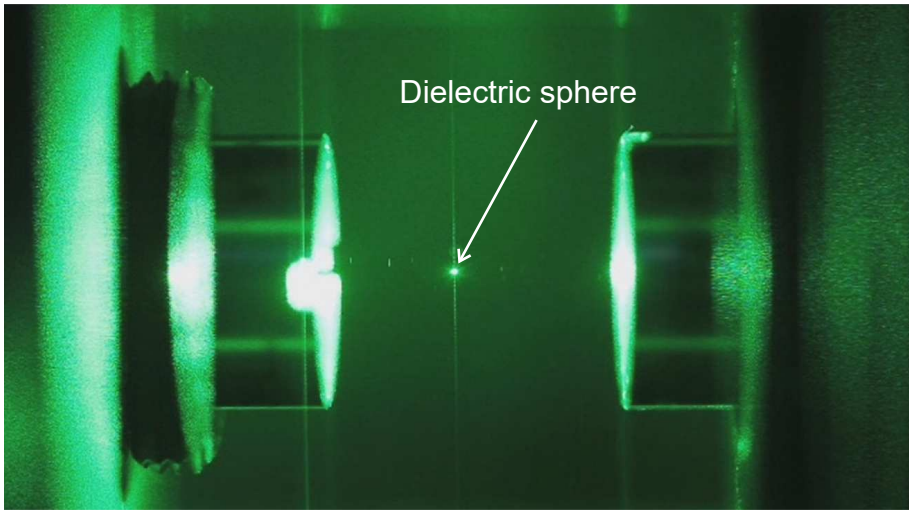


Figure 3.10: Dielectric sphere trapped by a strongly focused laser beam (source: Wikipedia).

In the following we consider a small dielectric sphere with radius a and dielectric constant ϵ_s that is illuminated by a focused laser beam with intensity $I(\mathbf{r})$. The dielectric constant of the surrounding medium is ϵ_m. The time average of the total optical force acting on the sphere has two contributions:

$$\langle \mathbf{F}_{total}(\mathbf{r}) \rangle = \langle \mathbf{F}_{scat}(\mathbf{r}) \rangle + \langle \mathbf{F}_{grad}(\mathbf{r}) \rangle. \tag{3.4.1}$$

The first term is the so-called scattering force and the second term the so-called gradient force. We will show below that the scattering force tends to 'push' the sphere away from the focus while the gradient forces tends to 'pull' the sphere into the high intensity region of the beam.

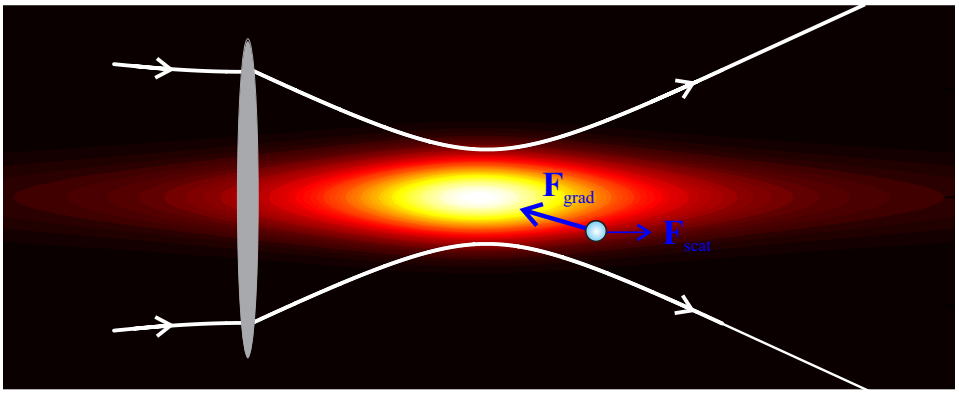


Figure 3.11: Illustration of the forces acting on a small spherical particle in a strongly focused laser beam. The intensity distribution is represented by the color plot. The white lines indicate the beam radius.

3.4.1 Scattering force

The scattering force results from scattering of the incident laser beam from the dielectric sphere. The total scattered power is given by

$$P_{scat}(\mathbf{r}) = \sigma_{scat} I(\mathbf{r}), \tag{3.4.2}$$

where σ_{scat} is the total scattering cross-section of the sphere. If the radius a of the sphere is much smaller than the wavelength λ, the scattering characteristics of the sphere is that of an electric dipole. Its scattering cross section can then be written as

$$\sigma_{scat} = \frac{k^4}{6\pi}|\alpha|^2, \tag{3.4.3}$$

with the polarizability

$$\alpha = 4\pi a^3 \frac{\epsilon_s - \epsilon_m}{\epsilon_s + 2\epsilon_m}. \tag{3.4.4}$$

The occurrence of the scattering force can be best understood in the photon picture. A single scattering event can be considered as the absorption of a photon from the incident laser beam followed by the subsequent reemission of a photon with the same energy. Each absorbed photon transfers the momentum $\hbar k \hat{\mathbf{e}}_{\mathbf{k}}$ to the sphere. Accordingly, if the photon is reemitted in the direction $\hat{\mathbf{e}}'_{\mathbf{k}}$ it will transfer the momentum $-\hbar k \hat{\mathbf{e}}'_{\mathbf{k}}$ to the sphere. However, since photons are reemitted with the same probability in opposite direction, the time-averaged contribution of all the remission processes to the momentum transfer is zero. The resulting the time-averaged scattering force experienced by the sphere is given by

$$\langle \mathbf{F}_{scat}(\mathbf{r}) \rangle = \frac{n_m P_{scat}(\mathbf{r})}{c_0} \hat{\mathbf{e}}_{\mathbf{k}} = \frac{128\pi^5 a^6 n_m}{3\lambda^4 c_0} \left(\frac{\epsilon_s - \epsilon_m}{\epsilon_s + 2\epsilon_m} \right)^2 I(\mathbf{r})\, \hat{\mathbf{e}}_{\mathbf{k}}. \tag{3.4.5}$$

3.4.2 Gradient force

The incident laser beam polarizes the sphere and induces the dipole moment

$$\mathbf{p}(\mathbf{r},t) = \epsilon_0\epsilon_m\alpha\mathbf{E}(\mathbf{r},t). \tag{3.4.6}$$

The potential energy of the sphere in the laser beam is given by

$$U(\mathbf{r},t) = -\frac{1}{2}\mathbf{p}(\mathbf{r},t)\cdot\mathbf{E}(\mathbf{r},t) = -\frac{1}{2}\epsilon_0\epsilon_m\alpha\mathbf{E}(\mathbf{r},t)\cdot\mathbf{E}(\mathbf{r},t). \tag{3.4.7}$$

The factor 1/2 accounts for the fact that the dipole is not permanent but induced by the laser beam. The time average of the potential energy can be calculated as

$$\langle U(\mathbf{r})\rangle = -\frac{1}{4}\epsilon_0\epsilon_m\alpha|\mathbf{E}(\mathbf{r})|^2, \tag{3.4.8}$$

From this follows the gradient force

$$\langle\mathbf{F}_{grad}(\mathbf{r})\rangle = -\nabla\langle U(\mathbf{r})\rangle = \frac{1}{4}\epsilon_0\epsilon_m\alpha\nabla|\mathbf{E}(\mathbf{r})|^2 = \frac{2\pi a^3 n_m}{c_0}\frac{\epsilon_s-\epsilon_m}{\epsilon_s+2\epsilon_m}\nabla I(r). \tag{3.4.9}$$

3.4.3 Trapping condition

The axial scattering force and the axial gradient force acting on a small dielectric sphere exerted by Gaussian beam with beam waist at $z = 0$ are given by

$$\langle F_{scat}(z)\rangle = \frac{128\pi^5 a^6 n_m}{3\lambda^4 c_0}\left(\frac{\epsilon_s-\epsilon_m}{\epsilon_s+2\epsilon_m}\right)^2\frac{1}{\left[1+(z/z_0)^2\right]}I_0, \tag{3.4.10}$$

$$\langle F_{grad}(z)\rangle = -\frac{2\pi a^3 n_m}{c_0}\left(\frac{\epsilon_s-\epsilon_m}{\epsilon_s+2\epsilon_m}\right)\frac{2z/z_0^2}{\left[1+(z/z_0)^2\right]^2}I_0, \tag{3.4.11}$$

respectively. The ratio of the magnitude of the two forces can be calculated as

$$R = \frac{|\langle F_{grad}(z)\rangle|}{|\langle F_{scat}(z)\rangle|} = \frac{3\lambda^4}{64\pi^4 a^3}\frac{\epsilon_s+2\epsilon_m}{\epsilon_s-\epsilon_m}\frac{2z/z_0^2}{1+(z/z_0)^2}. \tag{3.4.12}$$

It reaches its maximum at $z = z_0$. The corresponding maximum value is

$$R_{max} = \frac{3\lambda^4}{64\pi^4 a^3}\frac{\epsilon_s+2\epsilon_m}{\epsilon_s-\epsilon_m}\frac{1}{z_0}. \tag{3.4.13}$$

The dielectric sphere can be trapped by the Gaussian beam if the gradient force is larger than the scattering force for some axial position. A necessary condition for this is that $R_{max} > 1$.

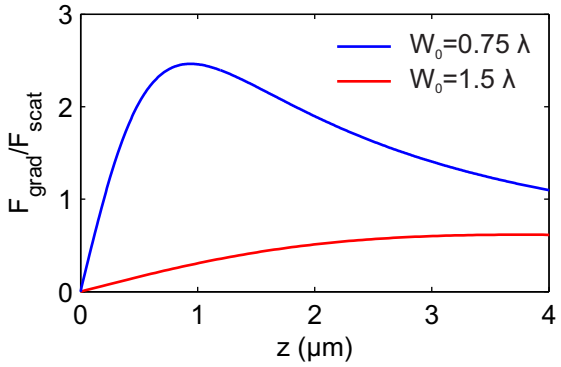


Figure 3.12: Ratio of the axial gradient force and the axial scattering force acting on a small dielectric sphere for two different Gaussian beams. Parameters: $\lambda_0 = 532\,\mathrm{n}m$, $\epsilon_s = 2.25$ (glass), $\epsilon_m = 1.769$ (water), $a = 40\,\mathrm{n}m$.

3.5 Angular momentum of paraxial beams

In the previous section we have seen that a beam of light can exert a force on an object. The electromagnetic momentum density connected with an electromagnetic wave is given by

$$\mathbf{g} = \frac{\mathbf{S}}{c^2} = \epsilon_0 \left(\mathbf{E} \times \mathbf{B}\right). \tag{3.5.1}$$

In the following, we discuss the angular momentum carried by a paraxial beam. For this purpose, we calculate the ratio of the z-component of the angular momentum of the beam per unit length J_z and the energy density per unit length U:

$$\frac{\langle J_z \rangle}{\langle U \rangle} = \frac{\Re\left(\int dxdy \left[\mathbf{r} \times (\mathbf{E}^* \times \mathbf{B})\right] \cdot \hat{\mathbf{e}}_z\right)}{\frac{1}{2}\Re\left(\int dxdy \left[\mathbf{E}^* \cdot \mathbf{E} + c^2 \mathbf{B}^* \cdot \mathbf{B}\right]\right)} \tag{3.5.2}$$

For a paraxial beam, the following approximations hold:

$$\frac{\partial \mathbf{E}}{\partial z} \approx \imath k \mathbf{E}, \tag{3.5.3}$$

$$\mathbf{E}^* \cdot \mathbf{E} \approx \mathbf{E}_\perp^* \cdot \mathbf{E}\perp, \tag{3.5.4}$$

where $\mathbf{E}_\perp$ is the component of $\mathbf{E}$ perpendicular to the optical axis $\hat{\mathbf{e}}_z$. After some algebra, we obtain:

$$\frac{\langle J_z \rangle}{\langle U \rangle} = \frac{\Re\left(\int dxdy \mathbf{E}_\perp^* \cdot \left[\mathbf{r} \times (-\imath\nabla)\right]_z \mathbf{E}_\perp - \imath \left[\mathbf{E}_\perp^* \times \mathbf{E}_\perp\right]_z\right)}{\omega \int dxdy |\mathbf{E}_\perp|^2} \tag{3.5.5}$$

Next, we express the polarization state of the beam as a superposition of circularly polarized waves:

$$\mathbf{E} = \begin{bmatrix} E_+ \\ E_- \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} E_x - \imath E_y \\ E_x + \imath E_y \end{bmatrix}. \tag{3.5.6}$$

In this basis, $\mathbf{E}_- = 0$ corresponds to a left circularly polarized beam and $\mathbf{E}_+ = 0$ characterizes a right circularly polarized beam. We assume that the intensity of the beam is normalized such that

$$\langle \mathbf{E}|\mathbf{E}\rangle = \int dxdy |\mathbf{E}_\perp|^2 = 1. \tag{3.5.7}$$

The z components of the orbital angular momentum operator

$$l_z = \left[\mathbf{r} \times (-\imath\hbar\nabla)\right]_z = -\imath\hbar\frac{\partial}{\partial\phi} \tag{3.5.8}$$

and the spin-1 angular momentum operator

$$s_z = \hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.5.9}$$

allow us to rewrite equation (3.5.5) as

$$\frac{\langle J_z\rangle}{\langle U\rangle} = \frac{\langle \mathbf{E}|l_z|\mathbf{E}\rangle + \langle \mathbf{E}|s_z|\mathbf{E}\rangle}{\hbar\omega}. \tag{3.5.10}$$

Writing the ratio $\langle J_z\rangle/\langle U\rangle$ in this form suggests that the angular momentum of the beam has two contributions: The orbital angular momentum $\langle \mathbf{E}|l_z|\mathbf{E}\rangle$ and the the spin angular momentum $\langle \mathbf{E}|s_z|\mathbf{E}\rangle$.

The spin angular momentum only depends on the polarization state of the beam:

$$\langle \mathbf{E}|s_z|\mathbf{E}\rangle = \hbar \int dxdy \left(|E_+|^2 - |E_-|^2\right). \tag{3.5.11}$$

It vanishes for a linearly polarized beam ($|E_+| = |E_-|$) and takes the values $+\hbar$ and $-\hbar$ for a left circularly polarized beam and a right circularly polarized beam, respectively.

The orbital angular momentum of a beam is determined by the ϕ dependence of $\mathbf{E}_\perp(\rho, \phi, z, t)$. Obviously, a simple Gaussian beam does not carry orbital angular momentum since its electric field does not depend on ϕ. However, by sending a Gaussian beam through an appropriate phase mask we can generate a beam with non-vanishing orbital angular momentum.

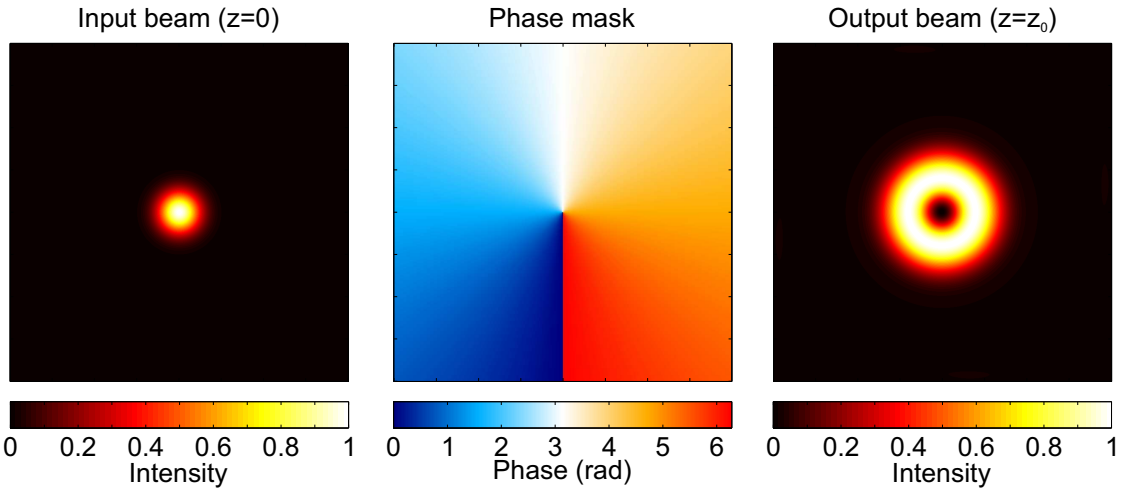
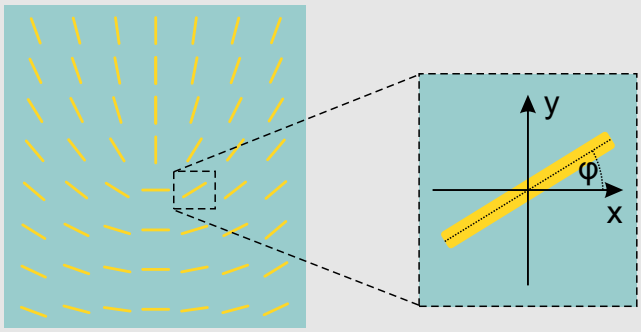


Figure 3.13: Transmission of a Gaussian beam (left) through a phase mask (middle) generates a beam with non-vanishing orbital angular momentum (right).

Reasearch topic: Metasurfaces
A metasurface is an artificial sheet material with sub-wavelength thickness. Its electromagnetic properties can be spatially tailored by nanostructuring of the constituents of the metasurface. This approach opens interesting new possibilities to locally modify the wave front of a light beam that is either transmitted through or reflected from the metasurface. Metasurfaces can be used to create optical elements like lenses or phase plates with sub-wavelength thickness.

In the following, we discuss a class of metasurfaces that utilizes simple metallic rod antennas to modify the phase of a transmitted beam. To explain the underlying operation principle, we first consider the excitation of a single metallic rod antenna. We assume that the antenna is in the xy-plane and that is axis is tilted by an angle φ relative to the $\hat{\mathbf{e}}_x$ axis.



Let the incident beam be a circularly polarized wave

$$\mathbf{E}_{\text{in}} = \mathbf{E}^{\sigma} = E_x\hat{\mathbf{e}}_x + E_y\hat{\mathbf{e}}_y = \frac{E}{\sqrt{2}}\left(\hat{\mathbf{e}}_x + \sigma\imath\hat{\mathbf{e}}_y\right) = E\hat{\mathbf{e}}_\sigma, \tag{3.5.12}$$

where $\sigma = 1$ corresponds to right-circularly polarized light and $\sigma = -1$ corresponds to left-circularly polarized light. The incident wave excites a current oscillation in the metallic rod antenna. The resulting dipole moment can be written as

$$\begin{bmatrix} p_x \\ p_y \end{bmatrix} = \epsilon_0\alpha \begin{pmatrix} \cos^2(\varphi) & \sin(\varphi)\cos(\varphi) \\ \sin(\varphi)\cos(\varphi) & \sin^2(\varphi) \end{pmatrix} \begin{bmatrix} E_x \\ E_y \end{bmatrix}. \tag{3.5.13}$$

By substituting equation (3.5.12) into (3.5.13) we obtain after a short calculation:

$$\mathbf{p} = \frac{\alpha\epsilon_0 E}{2}\left[\hat{\mathbf{e}}_\sigma + e^{\imath 2\varphi}\hat{\mathbf{e}}_{-\sigma}\right]. \tag{3.5.14}$$

The far-field radiation of a dipole can be expressed as:

$$\mathbf{E}_{rad} = \frac{e^{\imath kr}}{-\imath kr}\frac{\imath k^3}{4\pi}\hat{\mathbf{e}}_r \times \left(\hat{\mathbf{e}}_r \times \mathbf{p}\right), \tag{3.5.15}$$

where k is the wavenumber and r is the distance from the dipole and $\hat{\mathbf{e}}_r$ is a unit vector in the observation direction. Thus, in the forward direction ($\hat{\mathbf{e}}_r = \hat{\mathbf{e}}_z$), the radiated light field of the rod antenna has two circularly polarized light components:

$$\mathbf{E}_{\text{rad}} = \frac{e^{\imath kr}k^2}{4\pi r}\frac{\alpha\epsilon_0 E}{2}\left[\hat{\mathbf{e}}_\sigma + e^{\imath 2\varphi}\hat{\mathbf{e}}_{-\sigma}\right]. \tag{3.5.16}$$

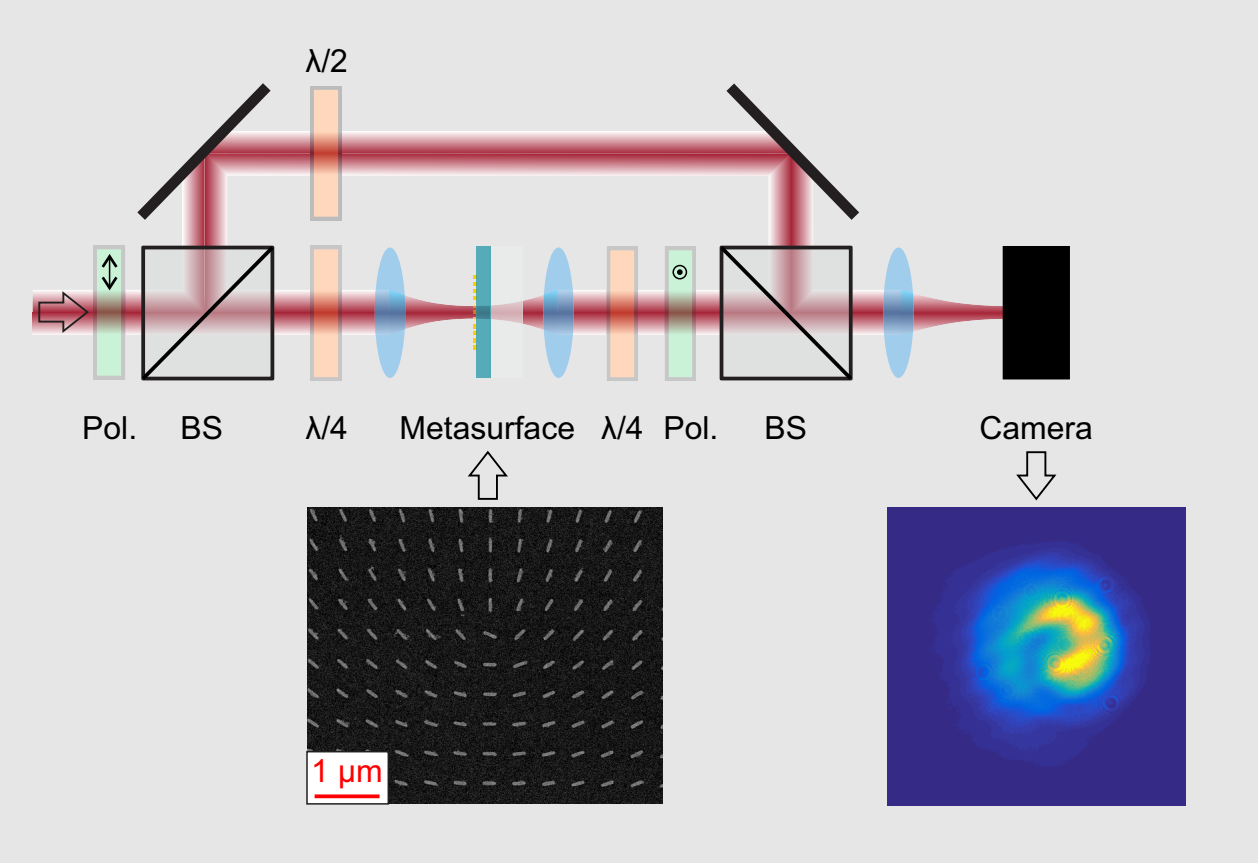
The total electric field is given by the superposition of the incident wave and the radiated light field of the antenna:

$$\mathbf{E}_{total} = \mathbf{E}_{\text{in}} + \mathbf{E}_{rad}. \tag{3.5.17}$$

If a circular polarizer is placed behind the sample that filters out the incident circular polarization, the transmitted field is given by:

$$\mathbf{E}_{trans} = \left(\mathbf{E}_{total}\cdot\hat{\mathbf{e}}_{-\sigma}\right)\hat{\mathbf{e}}_{-\sigma} = \frac{e^{\imath kr}k^2}{4\pi r}\frac{\alpha\epsilon_0 E}{2}e^{\imath 2\varphi}\hat{\mathbf{e}}_{-\sigma}. \tag{3.5.18}$$

This shows that the interaction of the incident wave with the rod antenna results in the creation of a wave with opposite circular polarization state. Interestingly, its phase can be simply controlled by the orientation of the rod antenna. Thus, a metasurface consisting of a sub-wavelength array of rod antennas with different orientations allows us to imprint a specific phase profile on the transmitted beam. This can be used, for instance, to create a beam with non-vanishing orbital angular momentum (Metasurface & interference image: Alexander Fassbender, Nanophotonics group, University of Bonn).


λ/2
Pol.
BS
λ/4
Metasurface
λ/4
Pol.
BS
Camera
1 µm

4 Light pulses

4.1 Wave packets of light and the time-bandwidth product

So far, we have focused on monochromatic waves with a constant amplitude vector. In contrast, the amplitude vector of a light pulse at a fixed position z_0 changes with time:

$$\mathbf{E}(z_0, t) = \mathbf{A}(z_0, t)\, e^{-\imath\omega_0 t}. \tag{4.1.1}$$

Here, $\mathbf{A}(z_0, t)$ is the so-called envelope of the pulse and ω_0 is its carrier frequency. Alternatively, a light pulse can be also described as a wave packet. This can be seen by switching from the time domain to the frequency domain and back via the Fourier transform:

$$\mathbf{E}(z_0, \omega) = \int_{-\infty}^{\infty} \mathbf{A}(z_0, t)\, e^{\imath(\omega-\omega_0)t} dt. \tag{4.1.2}$$

$$\mathbf{E}(z_0, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{E}(z_0, \omega)\, e^{-\imath\omega t} d\omega. \tag{4.1.3}$$

Here, $\mathbf{E}(z_0, \omega)$ characterizes the contribution (amplitude and phase) of the frequency component ω to the wave packet.

Example: Gaussian pulse

We consider in this example a light pulse with Gaussian envelope propagating in vacuum along the z-direction. At a fixed position, e.g. $z_0 = 0$, its electric field is given by

$$\mathbf{E}(t) = \mathbf{E}_0\, e^{-\frac{2\ln(2)t^2}{\tau_p^2}}\, e^{-\imath\omega_0 t}. \tag{4.1.4}$$

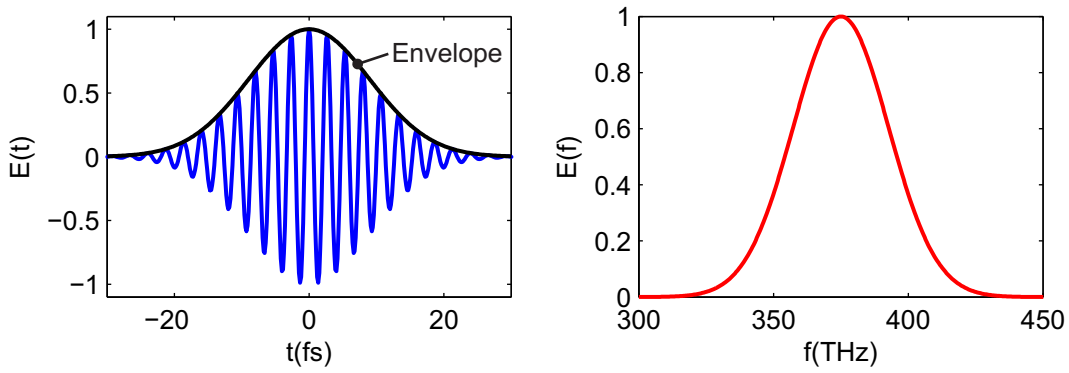


Figure 4.1: Light pulse with Gaussian profile. Left: Time domain. Right: Frequency domain.

Here, ω_0 is the carrier frequency and τ_p is a measure for the pulse duration. The intensity of the light pulse can be calculated as

$$I(t) = \frac{1}{2}\epsilon_0 c_0 |\mathbf{E}_0|^2 \, e^{-\frac{4\ln(2)t^2}{\tau_p^2}}. \tag{4.1.5}$$

For $t = \pm\tau_p/2$, the intensity takes the value

$$I(\pm\tau_p/2) = \frac{I(0)}{2}. \tag{4.1.6}$$

Hence, we can identify τ_p as the full width at half maximum (FWHM) of the intensity of the pulse.

The Fourier transform of $E(t)$ is given by

$$\mathbf{E}(\omega) = \sqrt{\frac{\pi}{2\ln(2)}}\tau_p \, \mathbf{E}_0 \, e^{-\frac{(\omega-\omega_0)^2\tau_p^2}{8\ln(2)}}. \tag{4.1.7}$$

The corresponding spectral intensity (or spectrum) is

$$S(\omega) = |\mathbf{E}_0|^2 \, e^{-\frac{(\omega-\omega_0)^2\tau_p^2}{4\ln(2)}}. \tag{4.1.8}$$

The spectral intensity falls to half of its peak value for

$$S(\omega_0 \pm \Delta\omega/2) = \frac{S(\omega_0)}{2}, \tag{4.1.9}$$

where

$$\Delta\omega = \frac{4\ln(2)}{\tau_p} \tag{4.1.10}$$

is the FWHM of the spectral intensity. Obviously, $\Delta\omega$ and τ_p are related to each other via the time-bandwidth product

$$\tau_p \, \Delta\omega = 4 \ln(2). \tag{4.1.11}$$

A corresponding time-bandwidth product can be also derived for other pulse shapes. The general rule is that short optical pulses require a broad spectrum!

4.2 Light pulses in dispersive media

Next, we consider the propagation of a light pulse in a dispersive medium. For each frequency component ω of the pulse, the electric field strength after the propagation distance z is given by:

$$\mathbf{E}\,(z,\omega) = \mathbf{E}\,(0,\omega)\, e^{\imath k(\omega) z} \tag{4.2.1}$$

with

$$k(\omega) = n(\omega)\, \omega / c_0. \tag{4.2.2}$$

Describing the pulse as a wave packet, we find:

$$\mathbf{E}\,(z,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{E}\,(0,\omega)\, h\,(\omega, z)\, e^{-\imath\omega t} d\omega. \tag{4.2.3}$$

Here, we have introduced the transfer function of the medium:

$$h\,(\omega, z) = e^{\imath k(\omega) z}. \tag{4.2.4}$$

Next, we expand the wave vector $k(\omega)$ in a Taylor series around ω_0:

$$k(\omega) = \beta_0 + \beta'\,(\omega - \omega_0) + \frac{1}{2}\beta''\,(\omega - \omega_0)^2 + \cdots \tag{4.2.5}$$

where

$$\beta_0 = k(\omega_0),\ \beta' = \left(\frac{dk}{d\omega}\right)_{\omega_0},\ \beta'' = \left(\frac{d^2k}{d\omega^2}\right)_{\omega_0}. \tag{4.2.6}$$

With these abbreviations, the transfer function can be written as:

$$h\,(\omega, z) = e^{\imath\beta_0 z}\, e^{\imath\beta'(\omega-\omega_0)z}\, e^{\imath\frac{1}{2}\beta''(\omega-\omega_0)^2 z}. \tag{4.2.7}$$

We first consider the case of a material with weak dispersion, i.e., $\beta'' = 0$. By substituting $h(\omega, z) = e^{\imath\beta_0 z}\, e^{\imath\beta'(\omega-\omega_0)z}$ into equation (4.2.3), we obtain:

$$\begin{aligned}\mathbf{E}(z,t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{E}(0,\omega)\; e^{\imath\beta_0 z}\, e^{\imath\beta'(\omega-\omega_0)z} e^{-\imath\omega t} d\omega \\ &= \frac{e^{\imath(\beta_0-\beta'\omega_0)z}}{2\pi}\int_{-\infty}^{\infty}\mathbf{E}(0,\omega)\; e^{-\imath\omega(t-\beta' z)} d\omega. \end{aligned} \tag{4.2.8}$$

Next, we define the group index

$$n_g(\omega) = n(\omega) + \omega\frac{dn(\omega)}{d\omega}, \tag{4.2.9}$$

the group velocity

$$v_g = \frac{1}{\beta'} = \frac{d\omega}{dk} = \frac{c_0}{n_g(\omega)}, \tag{4.2.10}$$

and the group delay

$$\tau_g = \frac{z}{v_g}. \tag{4.2.11}$$

Using these quantities, equation (4.2.8) can be rewritten as

$$\mathbf{E}(z,t) = e^{\imath(\beta_0-\omega_0/v_g)z}\,\mathbf{E}(0, t-\tau_g). \tag{4.2.12}$$

This shows that the pulse propagates with the group velocity v_g through the weakly dispersive medium without changing its temporal profile.

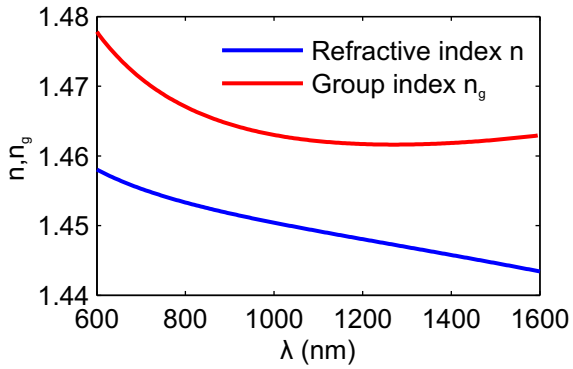


Figure 4.2: Refractive index and group index of fused silica glass.

In dispersive media with $\beta'' \neq 0$, the envelope changes as the pulse propagates. This can be seen by considering the difference of the group delay of two groups of waves centered at the frequencies ω_0 and $\omega_1 = \omega_0 + \Delta\omega$, respectively:

$$\begin{aligned} \Delta\tau_g &= [\beta'(\omega_1) - \beta'(\omega_0)]\, z \\ &\approx \left[\beta'(\omega_0) + \frac{d\beta'}{d\omega}\Delta\omega - \beta'(\omega_0)\right] z \\ &= z\beta''\Delta\omega. \end{aligned} \tag{4.2.13}$$

For $\beta'' \neq 0$, the different spectral components of a pulse propagate with different velocities through the medium. This effect is called group velocity dispersion (GVD). The dispersion coefficient is defined by:

$$D(\lambda) = -\frac{2\pi c_0}{\lambda^2}\beta''. \tag{4.2.14}$$

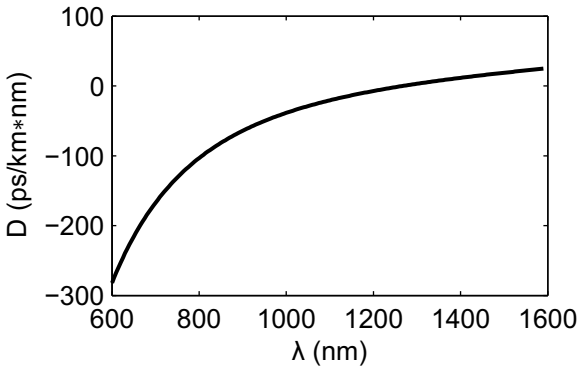


Figure 4.3: Dispersion coefficient of fused silica glass.

With the dispersion coefficient, we can rewrite equation (4.2.13) as:

$$\Delta\tau_g = zD(\lambda)\Delta\lambda, \tag{4.2.15}$$

where

$$\Delta\lambda = -\frac{\Delta\omega\lambda^2}{2\pi c_0}. \tag{4.2.16}$$

Equation (4.2.15) can be interpreted as follows: A pulse with spectral bandwidth $\Delta\lambda$ is stretched by a time $\Delta\tau_g$ after the propagation distance z in a medium with dispersion coefficient $D(\lambda)$. The effect of GVD on an optical impulse is exemplified in Fig. 4.4.

GVD is an important issue in the context of telecommunication based on fiber-optic networks. Here, the information is transmitted by a stream of short optical pulses. In the

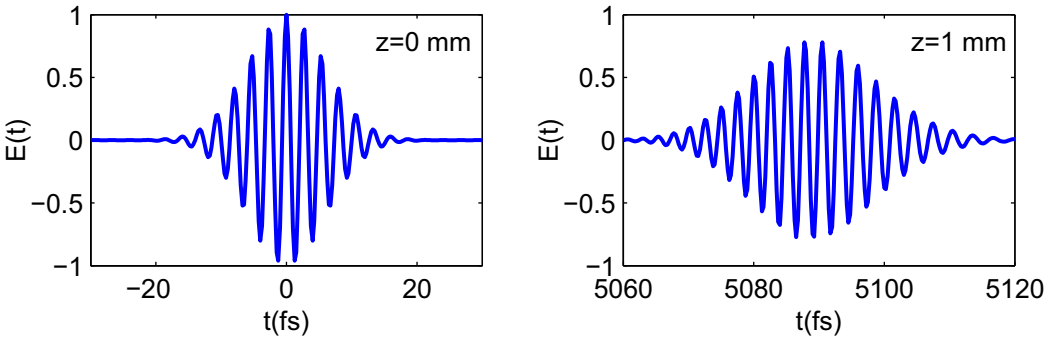


Figure 4.4: Left: Gaussian pulse with $\tau_p = 10\,\mathrm{f}s$ and $\omega_0 = 2\pi \times 375\,\mathrm{T}Hz$. Right: The same pulse after propagation through 1 mm of BK7-glass.

so-called ON-OFF keying scheme, the presence or absence of an optical pulse represent the bits "1" and "0", respectively. To avoid errors during the data transfer, subsequent optical pulses must not overlap at the end of the fiber. Pulse broadening due to GVD in the optical fiber thus sets a limit to the maximum achievable data rate.

Example: Propagation of a Gaussian pulse in BK7-Glas.

- Carrier frequency: $\omega_0 = 2\pi \times 375\,\mathrm{T}Hz$.

  $\Rightarrow$ Center wavelength: $\lambda_0 = 800\,\mathrm{n}m$.

- Propagation distance: $z = 1\,\mathrm{m}m$.

- Refractive index: $n(800\,\mathrm{n}m) = 1.51$.

- Group index: $n_g(800\,\mathrm{n}m) = 1.527$.

  $\Rightarrow$ Group delay $\tau_g = 5090\,\mathrm{f}s$.

- Dispersions coefficient of BK7: $D(800\,\mathrm{n}m) = -128\,\mathrm{p}s/km \times nm$

- Pulse length: $\tau_p = 10\,\mathrm{f}s$.

  $\Rightarrow$ Spectral width: $\Delta\lambda = 94\,\mathrm{n}m$.

  $\Rightarrow \Delta\tau_g \approx 12\,\mathrm{f}s$.

- Pulse length: $\tau_p = 100\,\mathrm{f}s$.

  $\Rightarrow$ Spectral width: $\Delta\lambda = 9.4\,\mathrm{n}m$.

$\Rightarrow \Delta\tau_g \approx 1.2\,\mathrm{f}s.$

The 10-fs-pulse is strongly distorted after 1 mm of BK7 Glas! In contrast, the 100-fs-pulse experiences almost no temporal broadening.

Outlook:
In media with anomalous dispersion ($\frac{dn(\omega)}{d\omega} < 0$), the group velocity v_g can exceed c_0 and can even become negative. A negative group velocity corresponds to the apparently paradoxical situation that the maximum of the transmitted pulse appears at the back side of the medium before the peak of the incident pulse has entered the medium. The reason why this situation is not in conflict with causality of relativity is explained in the following paper: D. J. Gauthier, R. Boyd, *F*ast Light, Slow Light and Optical Precursors: What Does it All Mean? Photonics Spectra, January 2007, Seite 82-90.

4.3 Gaussian pulses in weakly dispersive media

In the previous section, we analyzed the propagation of an optical pulse in a dispersive media using a frequency domain approach. Here, we want to consider the evolution of the pulse envelope $\mathbf{A}(z,t)$ in the time domain. In doing so, we restrict ourselves to weakly dispersive media, for which the approximation $k(\omega) = \beta_0 + \beta'(\omega-\omega_0) + \frac{1}{2}\beta''(\omega-\omega_0)^2$ holds. All higher terms in the expansion are neglected.

We start by writing the pulse envelope $\mathbf{A}(z,t)$ as:

$$\begin{aligned}\mathbf{A}(z,t) &= \mathbf{E}(z,t)\, e^{-\imath(\beta_0 z-\omega_0 t)} \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{E}(0,\omega)\, e^{\imath\left[\beta'(\omega-\omega_0)z+\frac{1}{2}\beta''(\omega-\omega_0)^2 z\right]} e^{-\imath(\omega-\omega_0)t} d\omega. \end{aligned} \tag{4.3.1}$$

Next, we calculate the spatial derivative

$$\begin{aligned}\frac{d\mathbf{A}}{dz} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{E}(0,\omega)\left[\imath\beta'(\omega-\omega_0)\right] e^{\imath\left[\beta'(\omega-\omega_0)z+\frac{1}{2}\beta''(\omega-\omega_0)^2 z\right]} e^{-\imath(\omega-\omega_0)t} d\omega + \\ &\quad \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{E}(0,\omega)\left[\imath\frac{1}{2}\beta''(\omega-\omega_0)^2\right] e^{\imath\left[\beta'(\omega-\omega_0)z+\frac{1}{2}\beta''(\omega-\omega_0)^2 z\right]} e^{-\imath(\omega-\omega_0)t} d\omega. \end{aligned} \tag{4.3.2}$$

The two integrals can be easily identified as $-\beta' \frac{d\mathbf{A}}{dt}$ and $-\frac{\imath\beta''}{2}\frac{d^2\mathbf{A}}{dt^2}$, respectively. Thus, the evolution of the pulse envelope is governed by the following equation:

$$\frac{d\mathbf{A}\,(z,t)}{dz} + \beta' \frac{d\mathbf{A}\,(z,t)}{dt} + \imath\frac{1}{2}\beta'' \frac{d^2\mathbf{A}\,(z,t)}{dt^2} = 0. \tag{4.3.3}$$

In the remainder of this section, we consider a Gaussian pulse:

$$\mathbf{A}_{Gauss}\,(z,t) = \hat{\mathbf{e}}_p \sqrt{\frac{\tau_0^2}{\tau_0^2 - 2\imath\beta'' z}} \exp\left[-\frac{(t-\beta' z)^2}{\tau_0^2 - 2\imath\beta'' z}\right], \tag{4.3.4}$$

where $\hat{\mathbf{e}}_p$ is a unit vector that defines the polarization state of the pulse and τ_0 characterizes the pulse width. By substituting this envelope function in equation (4.3.3), we can directly confirm that $\mathbf{A}_{Gauss}$ is a valid solution. We can easily verify that the pulse propagates with the group velocity v_g through the medium. The absolute value of the envelope,

$$|\mathbf{A}_{Gauss}\,(z,t)| = \left(\frac{1}{1+(2\beta'' z/\tau_0^2)^2}\right)^{1/4} \exp\left[-\frac{(t-\beta' z)^2}{\tau_0^2 + (2\beta'' z/\tau_0)^2}\right], \tag{4.3.5}$$

remains a Gaussian function of time.

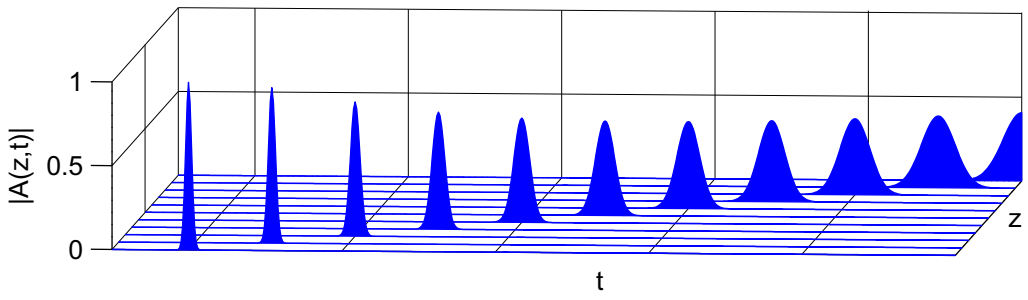


Figure 4.5: Evolution of the envelope of a Gaussian pulse propagating in a dispersive medium.

The pulse width at the entrance of the medium ($z = 0$) is

$$\tau_p(0) = \tau_0\sqrt{2\ln(2)}. \tag{4.3.6}$$

As the pulse propagates through the medium, $\tau_p(z)$ increases as:

$$\tau_p(z) = \tau_p(0)\sqrt{1+\left(\frac{2\beta'' z}{\tau_0^2}\right)^2}. \tag{4.3.7}$$

After the distance $z_0 = \frac{\sqrt{3}}{2}\frac{\tau_0^2}{\beta''}$, the pulse width has doubled compared to the initial value. The pulse spreading is a consequence of the fact that the different spectral components

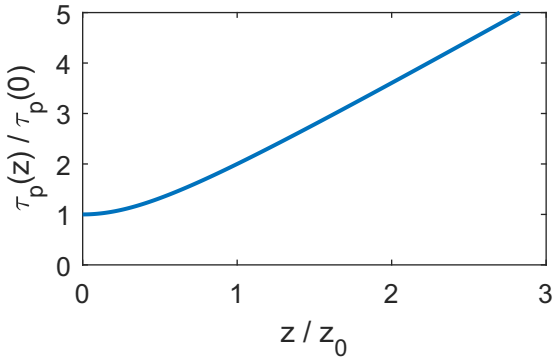


Figure 4.6: Pulse width of a Gaussian pulse as a function of the propagating distance in a dispersive medium.

travel with a different velocity through the medium. This also results in a "chirping" of the pulse, i.e., the instantaneous frequency of the pulse varies with time:

$$\omega(z,t) = \frac{d\Phi(z,t)}{dt} = \omega_0 + \frac{4\beta'' z}{\tau_0^4 + (2\beta'' z)^2}\,(t - \beta' z)\,, \tag{4.3.8}$$

where ω_0 is the carrier frequency and $\Phi(z,t)$ is the phase of the envelope function.

5 Optical resonators

An optical resonator is a device that is used to confine and concentrate light by means of multiple refelctions. In its simplest form, an optical resonator consists of a set of mirrors that creates a standing wave pattern. Optical resonators have found wide application as frequency selective optical filters and as "containers" for light in lasers.

5.1 Planar mirror resonator

5.1.1 Planar resonator with perfect mirrors

We start with a resonator formed by two perfect mirrors ($R_m = 1$) separated by a distance L. The refractive index of the medium inside of the resonator is assumed to be n. Reflections at the two mirrors result in the formation of standing waves. The tangential component of the electric field of the corresponding resonator modes vanishes at the mirror surfaces.

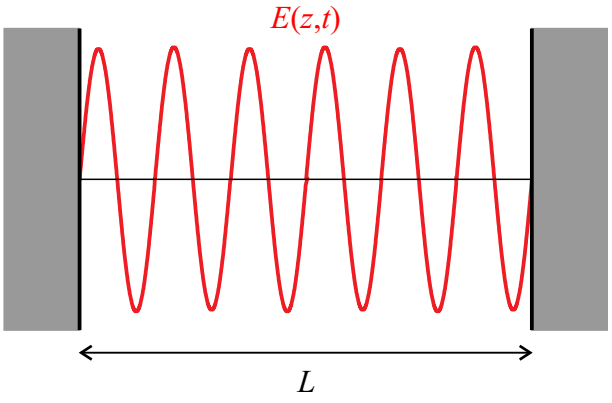


Figure 5.1: Standing wave in a planar mirror resonator.

The electric field strength of a resonator mode is given by:

$$E(z,t) = E_0 \sin(kz) \cos(2\pi\nu t). \tag{5.1.1}$$

The boundary conditions $E(0,t) = E(L,t) = 0$ imply that the allowed values of k of a resonator mode are restricted to

$$k_m = \frac{m\pi}{L}, m = 1, 2, 3, \ldots. \tag{5.1.2}$$

The corresponding allowed frequencies of the resonator modes are given by:

$$\nu_m = \frac{mc_0}{2nL}, m = 1, 2, 3, \ldots. \tag{5.1.3}$$

The spacing between two adjacent modes is:

$$\Delta\nu = \frac{c_0}{2nL}. \tag{5.1.4}$$

5.1.2 Fabry-Perot etalon

A Fabry-Perot etalon is a planar resonator with partially reflective mirrors. To derive the spectral properties of a Fabry-Perot etalon, we consider the situation depicted in Fig. 5.2. For simplicity, we assume that the two mirrors have identical properties. The incident wave E_0 is partially reflected ($E_o r$) and partially transmitted ($E_0 t$) at the front interface of the Fabry-Perot etalon. The light coupled into the Fabry-Perot etalon travels back and forth. At each reflection, a fraction of the light leaks out of the resonator.

Transmission and reflection characteristics

The total transmitted wave results from the coherent superposition of the transmitted partial waves:

$$E_t = E_0 t t' \underbrace{\left[1 + r'^2 e^{i\delta} + \left(r'^2 e^{i\delta}\right)^2 + \left(r'^2 e^{i\delta}\right)^3 + \cdots\right]}_{Geometrical\,series} = E_0 t t' \frac{1}{1 - r'^2 e^{i\delta}}, \tag{5.1.5}$$

with

- t: Amplitude transmission coefficient of the front mirror for light imping from outside on the Fabry-Perot etalon.
- r: Amplitude reflection coefficient of the front mirror for light imping from outside on the Fabry-Perot etalon.
- t': Amplitude transmission coefficient for light imping from inside on a mirror.

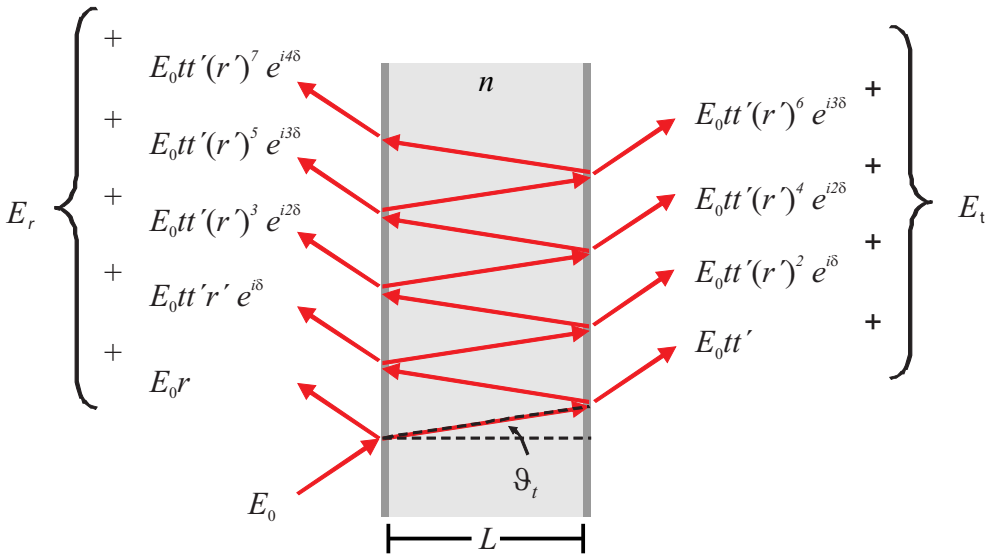


Figure 5.2: The total reflected and the transmitted field result from the coherent superposition of the reflected and transmitted partial waves.

- r': Amplitude reflection coefficient for light imping from inside on a mirror.
- Phase difference between two adjacent partial waves (Proof: Exercise):

$$\delta = 2k_0 n \cos(\theta_t) L. \tag{5.1.6}$$

.

For simplicity reasons, we assume that losses in the Fabry-Perot etalon can be neglected. In that case, the amplitude coefficients are related via:

$$r' = -r \quad and \quad r^2 + tt' = 1. \tag{5.1.7}$$

With this assumption, the transmittance of the Fabry-Perot etalon becomes:

$$T = \frac{|E_t|^2}{|E_0|^2} = \frac{1}{1 + F\sin^2(\delta/2)}, \tag{5.1.8}$$

where the coefficient of finesse of the resonator is given by

$$F = \left(\frac{2r}{1-r^2}\right)^2. \tag{5.1.9}$$

The reflectance of the Fabry-Perot etalon is given by

$$R = \frac{F\sin^2(\delta/2)}{1 + F\sin^2(\delta/2)}. \tag{5.1.10}$$

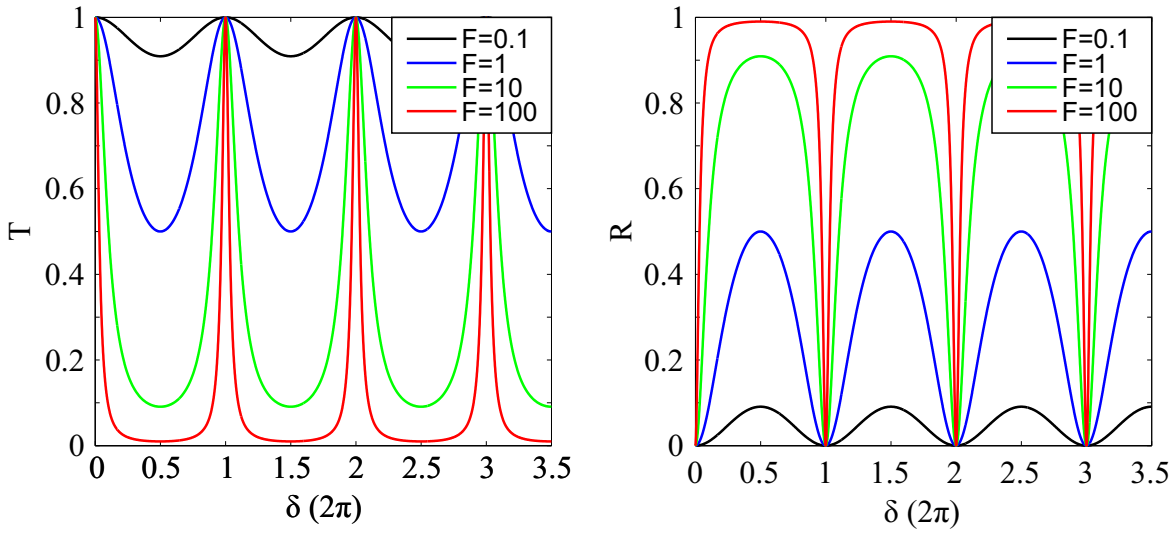


Figure 5.3: Transmittance (left) and reflectance (right) of a Fabry-Perot etalon for different values of the coefficient of finesse.

Here, we have used the condition $T + R = 1$ for loss-free media

According to equation (5.1.8), the transmittance T exhibits a maximum if the phase difference δ is a multiple of 2π:

$$\delta = 2k_0 n \cos(\theta_t) L = 2m\pi, m = 1, 2, 3, \ldots. \tag{5.1.11}$$

This corresponds to a situation, in which all transmitted partial waves are in phase.

The condition $\delta = 2m\pi$ is fulfilled if the frequency of the mode takes the value

$$\nu_m = \frac{mc_0}{2n\cos(\theta_t)L}, m = 1, 2, 3, \ldots. \tag{5.1.12}$$

The spectral separation of two adjacent transmittance maxima is the so-called free spectral range:

$$\Delta\nu_{\mathrm{FSR}} = \nu_{m+1} - \nu_m = \frac{c_0}{2n\cos(\theta_t)L}. \tag{5.1.13}$$

For $F \gg 1$, the full width at half maximum (FWHM) of a transmittance peak is given by

$$\Delta\nu_{\mathrm{FWHM}} = \frac{c_0}{\sqrt{F}\pi n\cos(\theta_t)L}. \tag{5.1.14}$$

The finesse $\mathcal{F}$ of the resonator is defined by

$$\mathcal{F} = \frac{\Delta\nu_{\mathrm{FSR}}}{\Delta\nu_{\mathrm{FWHM}}}. \tag{5.1.15}$$

It is easy to show that the finesse $\mathcal{F}$ and the coefficient of finesse F are related by

$$\mathcal{F} = \frac{\pi}{2}\sqrt{F}. \tag{5.1.16}$$

Intracavity intensity

In what follows, we consider the intracavity intensity of a lossless Fabry Perot etalon. Let I_0 be the intensity of the incident wave and I_t be the intensity of the transmitted wave. The mirror reflectivity is assumed to be R_m. The corresponding mirror transmittance is given by $T_m = 1 - R_m$. Inside of the resonator, the electric field is given by the superposition of a right-traveling wave with intensity I_+ and a left-traveling wave with intensity I_-. The intensities are related by

$$\begin{aligned} I_t &= I_0 \frac{(1-R_m)^2}{(1-R_m)^2 + 4R_m \sin^2(\delta/2)} && (5.1.17)\\ I_t &= T_m I_+ = (1-R_m)\, I_+ && (5.1.18)\\ I_- &= R_m I_+. && (5.1.19) \end{aligned}$$

On resonance ($\delta = 2m\pi$), the intracavity intensities are given by:

$$\begin{aligned} I_+ &= \frac{1}{(1-R_m)} I_0, && (5.1.20)\\ I_- &= \frac{R_m}{(1-R_m)} I_0. && (5.1.21) \end{aligned}$$

The superposition of the right-traveling wave and the left-traveling wave results in an interference pattern with the intensity distribution:

$$I(z) = I_+ + I_- + 2\sqrt{I_+ I_-}\cos(2kz + \phi_0) \tag{5.1.22}$$

where k is the wave vector in the resonator and ϕ_0 is a constant phase that takes phase shifts caused by mirror reflections into account. It follows from equation (5.1.22) that the average intensity inside of the cavity is given by

$$I_{avg} = \frac{1+R_m}{1-R_m} I_0. \tag{5.1.23}$$

Example: Intensity of a Fabry Perot etalon with $R_m = 0.99$.

$\Rightarrow I_{avg} = \frac{1+R_m}{1-R_m} I_0 = \frac{1.99}{0.01} I_0 = 199 I_0$.

The average intensity inside of the resonator is more than two orders of magnitude larger than the intensity of the incident wave!

Fabry-Perot etalon with material absorption

So far, we have considered a lossless medium in the Fabry-Perot etalon. We want to discuss now the effect of material absorption on the transmission characteristics. We assume that the amplitude of the electric field is attenuated by a factor $\sqrt{A} = e^{-\alpha L}$ after a single pass througth the medium. The amplitude of the total transmitted field then becomes:

$$\begin{aligned} E_t &= E_0 t t' e^{-\alpha L} \underbrace{\left[1 + r'^2 e^{i\delta} e^{-2\alpha L} + \left(r'^2 e^{i\delta} e^{-2\alpha L}\right)^2 + \left(r'^2 e^{i\delta} e^{-2\alpha L}\right)^3 + \cdots\right]}_{Geometrical\,series} \\ &= E_0 t t' e^{-\alpha L} \frac{1}{1 - r'^2 e^{i\delta} e^{-2\alpha L}}, \end{aligned} \tag{5.1.24}$$

If we assume that the mirrors are lossless, i.e., $r'^2 + tt' = 1$, we find after a short calculation (proof: exerecise) that the intensity of the transmitted wave is given by

$$T = \frac{|E_t|^2}{|E_0|^2} = \frac{(1-R)^2 A}{(1-RA)^2 + 4RA\sin^2(\delta/2)}. \tag{5.1.25}$$

The maximum transmission is given by:

$$T_{max} = \frac{(1-R)^2 A}{(1-RA)^2}. \tag{5.1.26}$$

5.2 Spherical mirror resonators

Planar mirror resonators are very sensitive to misalignment of the mirrors. In contrast, properly designed spherical mirror resonators are less susceptible to misalignment. For that reason, resonators comprising spherical mirrors have found widespread application.

5.2.1 Modes of a spherical mirror resonator

We proceed to show that the modes of spherical mirror resonators are Gaussian beams. For this purpose we start with examining the reflection of a Gaussian beam from a spherical mirror (see Fig. 5.4). Special care must be paid to the sign conventions employed for the radius of curvature. Here and in the following we define that the radius of curvature of a spherical mirror is negative (positive) if the mirror is concave (convex). For the radius of curvature of a Gaussian beam, we employ the sign convention introduced in section 3.3, i.e., R is positive (negative) for positions right (left) of the beam waist.

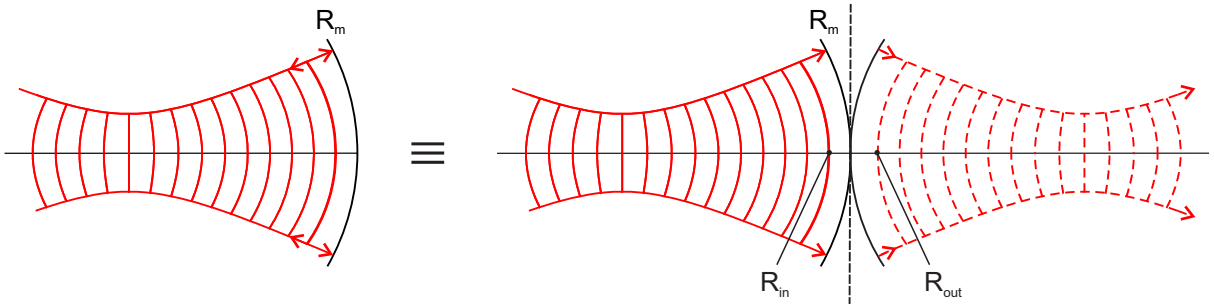


Figure 5.4: Reflection of a Gaussian beam from a spherical mirror.

We first consider the reflection of a Gaussian beam from a single spherical mirror. The beam width is not altered upon reflection. However, the radius of curvature R_{out} of the beam is changed. After reflection it is given by[1]:

$$\frac{1}{R_{out}} = \frac{1}{R_{in}} + \frac{2}{R_{\mathrm{m}}}, \tag{5.2.1}$$

where R_{in} is the radius of curvature of the incident beam at the mirror and R_{m} is the radius of curvature of the spherical mirror. The reflected beam retraces the incident beam ($R_{out} = -R_{in}$) if the condition $R_{in} = -R_{\mathrm{m}}$ is fulfilled, i.e, the electric field of the beam has on every point of the surface of the concave mirror the same phase.

Next, we consider a spherical-mirror resonator consisting of two spherical mirrors of radii R_1 and R_2 separated by a distance d. Let the positions of the two mirrors be z_1 and $z_2 = z_1 + d$ (see Fig. 5.5). A Gaussian beam is a mode of this resonator if it retraces itself as it is reflected back and forth between the two mirrors. As shown above, this requires that the radius of curvature of the Gaussian beam $R(z)$ fulfills the following conditions[2]:

$$R(z_1) = R_1, \tag{5.2.2}$$

$$R(z_2) = -R_2. \tag{5.2.3}$$

[1] This situation is equivalent to focusing of a Gaussian beam by a thin lens (see equation (3.3.35)).
[2] Here, we consider the forward propagating Gaussian beam.

To determine the Gaussian beam parameters, we use these two conditions together with equation (3.3.8) to obtain

$$R_1 = z_1 + \frac{z_0^2}{z_1}, \tag{5.2.4}$$

$$-R_2 = z_2 + \frac{z_0^2}{z_2}. \tag{5.2.5}$$

Here, z_0 is the Rayleigh length of the beam. Solving for z_1, z_2, and z_0^2, we find:

$$z_1 = \frac{-d\,(R_2 + d)}{R_1 + R_2 + 2d} \tag{5.2.6}$$

$$z_2 = d + z1 \tag{5.2.7}$$

$$z_0^2 = \frac{-d\,(R_1 + d)\,(R_2 + d)\,(R_2 + R_1 + d)}{(R_2 + R_1 + 2d)^2}. \tag{5.2.8}$$

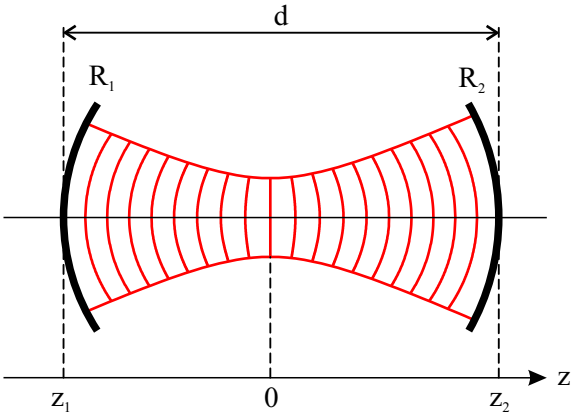


Figure 5.5: Gaussian beam as a mode of a spherical mirror resonator.

In the case of a symmetric resonator with concave mirrors ($R_1 = R_2 = -|R|$), equation (5.2.8) simplifies to:

$$z_0^2 = \frac{(2|R| - d)d}{4}. \tag{5.2.9}$$

The corresponding waist radius W_0 is given by:

$$W_0 = \left(\frac{\lambda z_0}{\pi}\right)^{1/2} = \left(\frac{\lambda}{\pi}\right)^{1/2} \left(\frac{2|R|d - d^2}{4}\right)^{1/4}. \tag{5.2.10}$$

5.2.2 Allowed Frequencies of a spherical mirror resonator

Next, we determine the allowed frequencies of the Gaussian mode inside of the spherical mirror resonator. On the optical axis ($\rho = 0$), the phase of a Gaussian beam is given by:

$$\varphi(0, z) = kz - \eta(z). \tag{5.2.11}$$

As the Gaussian beam propagates from the first mirror to the second mirror, its phase changes by:

$$\varphi(0, z_2) - \varphi(0, z_1) = k(z_2 - z_1) - [\eta(z_2) - \eta(z_1)] = kd - \Delta\eta \tag{5.2.12}$$

with

$$\Delta\eta = \eta(z_2) - \eta(z_1). \tag{5.2.13}$$

After one complete round trip, the phase of a mode must change by a multiple of 2π:

$$2kd - 2\Delta\eta = 2\pi m, m = 0, \pm 1, \pm 2, \ldots. \tag{5.2.14}$$

The corresponding frequencies are given by:

$$\nu_m = m\frac{c_0}{2d} + \frac{\Delta\eta}{\pi}\frac{c_0}{2d}. \tag{5.2.15}$$

Like in the case of a planar resonator, the allowed frequencies of a spherical mirror resonator are separated by $\delta\nu = c_0/2d$. However, the Gouy phase of the Gaussian beam adds a constant offset to the resonance frequencies.

5.2.3 Stability criterion for spherical mirror resonators

Not all spherical mirror resonator configurations give rise to a stable mode confinement. In order to derive a mode stability criterion, it is convenient to introduce the so-called g-parameters of the resonator:

$$g_1 = \left(1 + \frac{d}{R_1}\right), \tag{5.2.16}$$

$$g_2 = \left(1 + \frac{d}{R_2}\right). \tag{5.2.17}$$

With the g-parameters, equation (5.2.8) can be expressed as:

$$z_0^2 = \frac{d^2 g_1 g_2 \left(1 - g_1 g_2\right)}{\left(2 g_1 g_2 - g_1 - g_2\right)^2}. \tag{5.2.18}$$

For a Gaussian beam, the Rayleigh length z_0 must be real and hence $z_0^2 \geq 0$. Inspection of equation (5.2.18) immediately shows that this condition can be only fulfilled, if

$$0 \leq g_1 g_2 \leq 1. \tag{5.2.19}$$

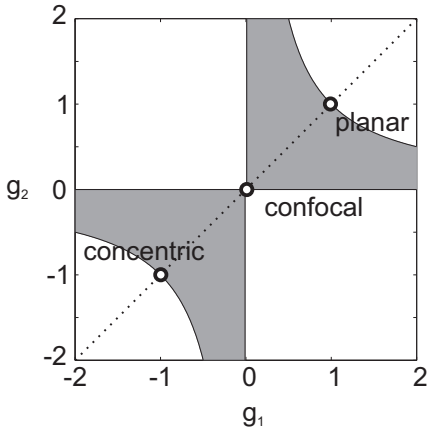


Figure 5.6: Stability diagram for a spherical mirror resonator. The gray-shaded areas correspond to stable configurations.

Examples of stable resonators:

- Symmetric concentric resonator: $R_1 = R_2 = -d/2, g_1 = g_2 = -1$.
- Symmetric confocal resonator: $R_1 = R_2 = -d, g_1 = g_2 = 0$.
- Planar resonator: $R_1 = R_2 = \infty, g_1 = g_2 = 1$.

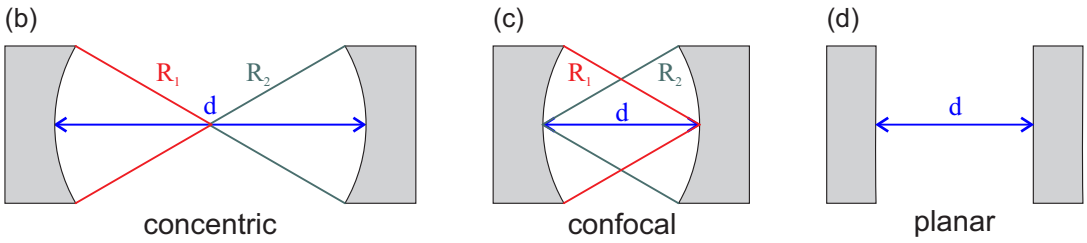


Figure 5.7: Examples of stable resonators.

6 Optical waveguides

An optical waveguide is a device that confines electromagnetic waves and guides them along its structure. Optical waveguides are widely used as "cables" for light in integrated optical circuits and optical communication systems. Common types of optical waveguides include optical fibers, slab waveguides, and ridge waveguides. In the majority of cases, optical waveguides are made from transparent dielectric materials. However, waveguiding can be also achieved at the surface of metals.

6.1 Planar-mirror waveguides

We start our discussion of optical waveguides with the planar-mirror waveguide. It consists of two parallel planar mirrors that are separated by a distance d (see Fig. 6.1). This type of waveguide is not of practical relevance for photonic applications. However, because of its simplicity it can be used as a model system to discuss general properties of optical waveguides.

The principle of operation of the planar-mirror waveguide can be understood in terms of ray-optics. A ray of light is guided by repeated reflections between the two mirrors and describes a "zigzag"-path as it propagates along the waveguide. This simple ray-picture

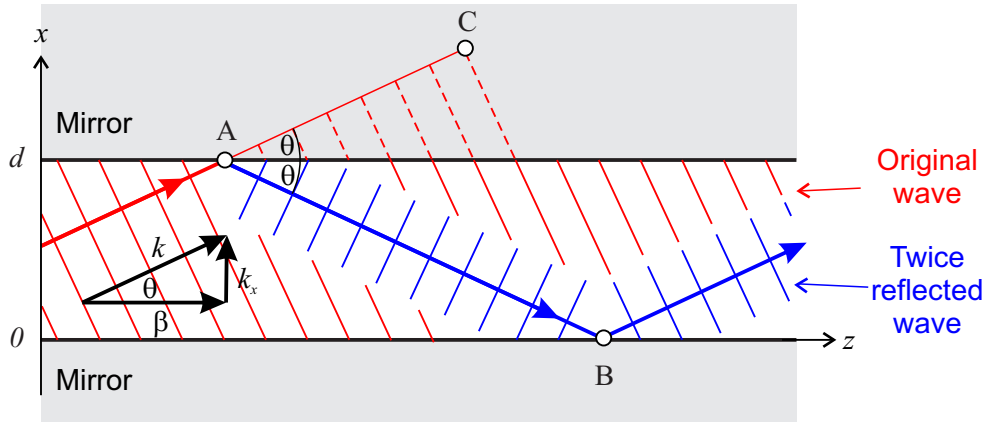


Figure 6.1: Scheme of a planar mirror waveguide.

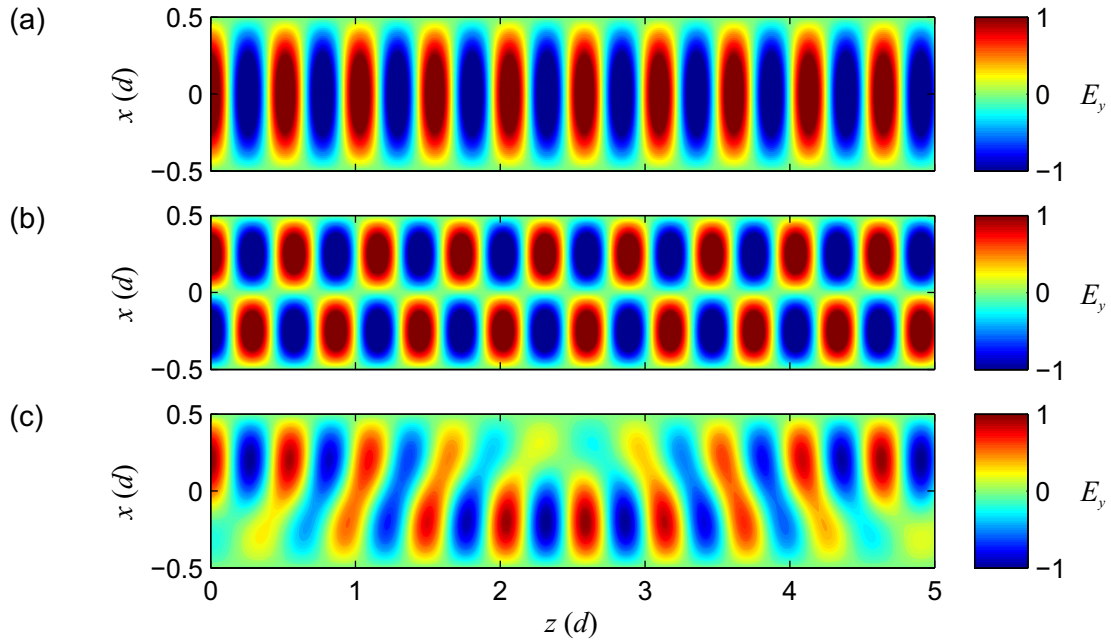


Figure 6.2: Field distritribution of (a) the first mode (m=1), (b) the second mode (m=2), (c) superposition of the first two modes.

is sufficient if the separation of the mirrors d is much larger than the wavelength of light λ. However, if d and λ are of comparable size, the wave nature of light has to be taken into account. In this case, the electromagnetic field in the region between the two mirrors can be written as a superposition of two plane waves, an upward traveling wave ("zig wave") with wavevector $\mathbf{k} = k_x\hat{\mathbf{e}}_x + \beta\hat{\mathbf{e}}_z$ and a downward traveling wave ("zag wave") with wavevector $\mathbf{k} = -k_x\hat{\mathbf{e}}_x + \beta\hat{\mathbf{e}}_z$.

For the following discussion, we assume that the electric field is oriented parallel to the interfaces (TE-polarization). The modes of the waveguide are electromagnetic waves with constant transverse profile as they propagate along the waveguide. For example, Figs. 6.2 (a) and (b) depict the electric field distributions of the first two TE modes of a planar-mirror waveguide. In contrast, the electric field shown in Fig. 6.2 (c) is not a mode of the waveguide.

A waveguide mode reproduces itself after two reflections, i.e., the original wave and the twice reflected wave must be in phase (see Fig. 6.1):

$$\Delta\varphi = \frac{2\pi}{\lambda}\overline{AB} - \pi - \frac{2\pi}{\lambda}\overline{AC} - \pi = 2\pi q, \tag{6.1.1}$$

where $q = 0, 1, 2, \ldots$. Note that the two reflections give rise to a phase change of 2π. For the given geometry, we find that

$$\overline{AB} - \overline{AC} = 2d\sin(\theta). \tag{6.1.2}$$

Thus, the allowed bouncing angles θ_m satisfy the condition:

$$\sin(\theta_m) = m\frac{\lambda}{2d},\ m = (1+q) = 1, 2, 3\ldots \tag{6.1.3}$$

Each bouncing angle θ_m corresponds to a different mode. The $\hat{\mathbf{e}}_x$-component of the wave vector $\mathbf{k}$ of the m-th mode can be expressed as

$$k_{x,m} = k\sin(\theta_m) = m\frac{\pi}{d},\ m =, 1, 2, 3\ldots. \tag{6.1.4}$$

The corresponding propagation constant ($\hat{\mathbf{e}}_z$-component of the wave vector $\mathbf{k}$), β_m, reads

$$\beta_m = \sqrt{k^2 - \frac{m^2\pi^2}{d^2}}. \tag{6.1.5}$$

With $k = \omega/c$, equation (6.1.5) can be rewritten as:

$$\omega_m = c\sqrt{\beta_m^2 + \frac{m^2\pi^2}{d^2}}. \tag{6.1.6}$$

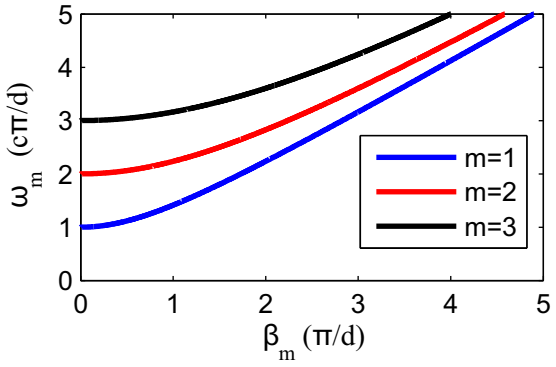


Figure 6.3: Dispersion relations of the first three TE-modes of a planar mirror waveguide.

The dispersion relation $\omega_m(\beta_m)$ of the mirror waveguide is a monotonic function. It approaches the linear relation $\omega_m \approx \beta_m c$ for large values of β_m. For $\beta_m \to 0$, the frequency of the m-th mode reaches the finite value

$$\omega_{m,\text{cut}} = \frac{cm\pi}{d}. \tag{6.1.7}$$

This frequency is the so-called cutoff frequency of the m-th mode, i.e., the mode becomes evanescent (β_m becomes an imaginary quantity) for frequencies below $\omega_{m,cut}$.

The effective index of the m-th mode is defined by:

$$n_{\text{eff}}(\omega) = \frac{\beta_m}{k} = \sqrt{1 - \frac{m^2\pi^2c^2}{\omega^2 d^2}}. \tag{6.1.8}$$

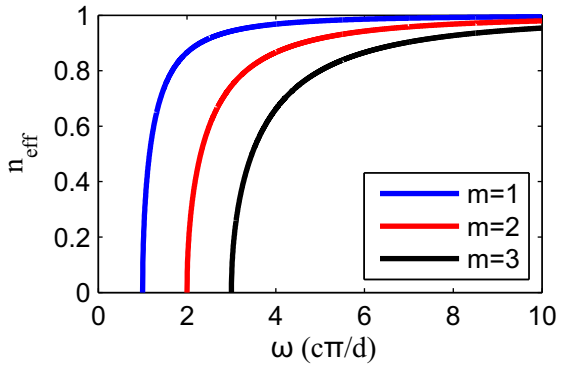


Figure 6.4: Effective refractive index of the first three modes of a planar mirror waveguide.

A waveguide mode can be written as a superposition of an upward and a downward propagating plane wave:

$$\mathbf{E}_m(x,z,t) = A\,\hat{\mathbf{e}}_y\, e^{\imath(\frac{m\pi}{d}x+\beta_m z-\omega_m t)} + B\,\hat{\mathbf{e}}_y\, e^{\imath(-\frac{m\pi}{d}x+\beta_m z-\omega_m t)}. \tag{6.1.9}$$

Here, we have used equation (6.1.4) to express the $\hat{\mathbf{e}}_x$-component of the wave vector. The boundary conditions $\mathbf{E}(x=0,z,t)=0$ and $\mathbf{E}(x=d,z,t)=0$ imply that $A=-B$. We thus obtain:

$$\mathbf{E}_m(x,z,t) = A\,\hat{\mathbf{e}}_y\, \sin\left(\frac{m\pi}{d}x\right) e^{\imath(\beta_m z-\omega_m t)}. \tag{6.1.10}$$

6.2 Total internal reflection revisited

In the following we consider a plane wave that impinges on a planar interface between two dielectric media with $n_i > n_t$. The electric fields of the incident wave, the reflected wave, and the transmitted wave can be written as

$$\mathbf{E}_i(\mathbf{r},t) = \mathbf{E}_i\, e^{\imath(\mathbf{k}_i\cdot\mathbf{r}-\omega t)}, \quad (\text{medium } 1, n=n_i), \tag{6.2.1}$$
$$\mathbf{E}_r(\mathbf{r},t) = r\,\mathbf{E}_i\, e^{\imath(\mathbf{k}_r\cdot\mathbf{r}-\omega t)}, \quad (\text{medium } 1, n=n_i), \tag{6.2.2}$$
$$\mathbf{E}_t(\mathbf{r},t) = t\,\mathbf{E}_i\, e^{\imath(\mathbf{k}_t\cdot\mathbf{r}-\omega t)}, \quad (\text{medium } 2, n=n_t), \tag{6.2.3}$$

respectively.

Substitution of (6.2.3) in the wave equation yields:

$$\mathbf{k}_t\cdot\mathbf{k}_t = \left(k_{t,\parallel}\,\mathbf{e}_\parallel + k_{t,\perp}\,\mathbf{e}_\perp\right)\cdot\left(k_{t,\parallel}\,\mathbf{e}_\parallel + k_{t,\perp}\,\mathbf{e}_\perp\right) = k_{t,\parallel}^2 + k_{t,\perp}^2 = \frac{\omega^2 n_t^2}{c_0^2}. \tag{6.2.4}$$

Here, $k_{t,\parallel}\,\mathbf{e}_\parallel$ and $k_{t,\perp}\,\mathbf{e}_\perp$ are the components of the wave vector $\mathbf{k}_t$ which are oriented parallel and normal to the interface, respectively.

The continuity conditions imply that the parallel components of $\mathbf{k}_i$, $\mathbf{k}_t$, and $\mathbf{k}_r$ are equal:

$$k_{i,\parallel} = k_{r,\parallel} = k_0 n_i \sin(\theta_i) = k_0 n_t \sin(\theta_t) = k_{t,\parallel}. \tag{6.2.5}$$

The normal component of $\mathbf{k}_t$ can be thus written as:

$$k_{t,\perp} = \sqrt{\frac{\omega^2 n_t^2}{c_0^2} - k_0^2 n_t^2 \sin^2(\theta_t)} = k_0 n_t \sqrt{1 - \frac{n_i^2}{n_t^2}\sin^2(\theta_i)}. \tag{6.2.6}$$

For $\theta_i > \theta_c = \arcsin(n_t/n_i)$, $k_{t,\perp}$ becomes a complex quantity. Hence, the electric field is evanescent in medium 2, i.e., it decays exponentially as the distance to the interface increases. There is on average no energy transport across the interface because the time averaged Poynting vector of the evanescent wave vanishes (Proof: Exercise). This effect is called totat internal reflection (TIR).

Figure 6.5: Interface between two media with refractive indices n_i and n_t, respectively.

Figure 6.6: Optical waveguide based on TIR (Image: Wikipedia).
Modern version: http://www.youtube.com/watch?v=hBQ8fh_Fp04 .

6.3 Dielectric waveguides

The basic structure of a dielectric waveguide consists of a high-index material (core) which is surrounded by low-index media (cladding). Typically, the cross-section of an optical waveguide is constant along the direction of propagation.

Depending of the application, different waveguide geometries are employed. Slab waveguides consist of a high index layer sandwiched between two media with smaller refractive index (see Fig. 6.7, left). They are used, for example, to guide light in the active region of semiconductor lasers. Ridge waveguides (see Fig. 6.7, center) serve as "wires" for light in integrated optics and optical fibers (see Fig. 6.7, right) are employed to transport optical signals in telecommunication systems. In the following, we assume that the refractive index profile of the waveguides is step-wise constant.

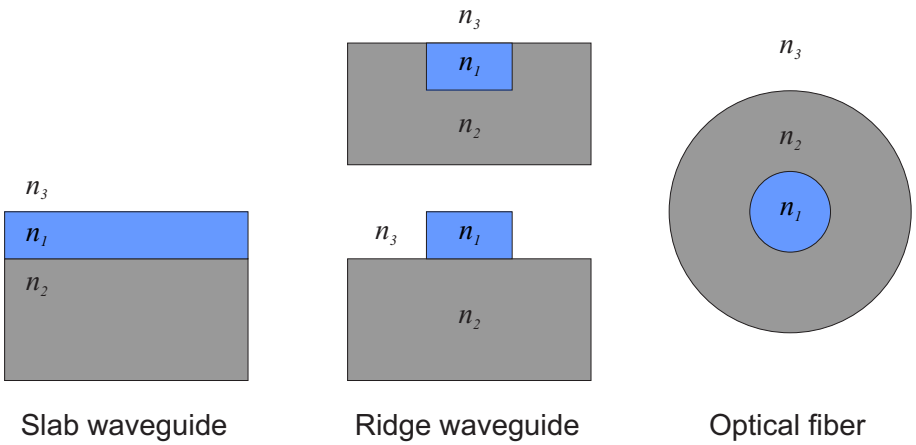


Figure 6.7: Cross-sections of different types of step-index waveguides. Without loss of generality, we assume that $n_1 > n_2 \geq n_3$.

6.4 Slab waveguides

The refractive index profile of a simple slab waveguide is given by:

$$n(x) = \begin{cases} n_c & 0 < x & \text{cladding} \\ n_f & -d \leq x \leq 0 & \text{film} \\ n_s & x < -d & \text{substrate} \end{cases} \tag{6.4.1}$$

The modes of a slab waveguide can be classified according to their polarization state as either transverse electric (TE) or transverse magnetic (TM) modes. TE modes are characterized by an electric field that is oriented perpendicular to the plane of propagation

(xz-plane). Thus, its non-vanishing field components are E_y, H_x, and H_z. For a TM mode, the magnetic field is oriented perpendicular to the plane of propagation and its field components are H_y, E_x, and E_z.

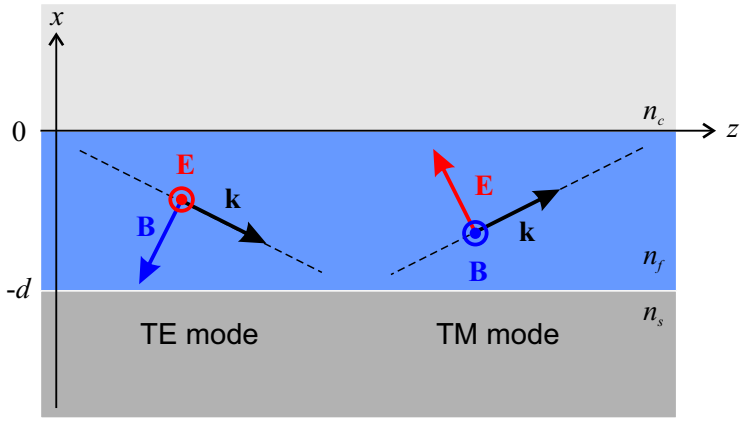


Figure 6.8: Slab waveguide: TE and TM modes.

6.4.1 TE modes

The TE modes of a slab waveguide have the following properties:

- The modes are solutions of the Helmholtz equation in all three media:

$$\left[\nabla^2 + \frac{\omega^2 n_{\text{med}}^2}{c_0^2}\right] \mathbf{E}(\mathbf{r}) = 0, \tag{6.4.2}$$

  where n_{med} is the refractive index of the respective medium.

- In the film layer, the electromagnetic field is given by the superposition of an upward and a downward propagating plane wave ("zigzag wave").

- The fields are evanescent in the substrate and in the cover, i.e., the normal components of the wave vector in the substrate and in the cover are complex quantities.

- E_y and H_z are continuous at the interfaces.

- The tangential component of the wave vector β is the same in all three media.

- The possible values of β are restricted to the range $k_0 n_s < \beta < k_0 n_f$ (see Fig. 6.9). The lower bound follows from the fact that the mode must not become radiative in the substrate. The upper bound must be fulfilled to obtain a propagating-wave solution in the film layer.

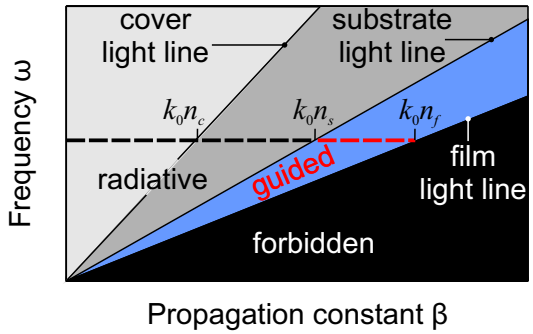


Figure 6.9: Possible values of the propagation constant for radiative modes and guides modes.

We choose the following ansatz for the electric field distribution:

$$\mathbf{E}(x,y,z,t) = \hat{\mathbf{e}}_y E_y(x)\, e^{\imath(\beta z - \omega t)} \tag{6.4.3}$$

with

$$E_y(x) = \begin{cases} A\, e^{-qx} & 0 \le x & \text{(cover)} \\ B\, e^{\imath hx} + C\, e^{-\imath hx} & -d \le x \le 0 & \text{(film)} \\ D\, e^{px} & x \le -d & \text{(substrate)} \end{cases} \tag{6.4.4}$$

Here, h, q, and p are the transverse components of the wave vector in the film, the cover, and the substrate, respectively. The constants A, B, C, and D follow from the continuity conditions at the interfaces (see below).

Upon inserting $\mathbf{E}(x,y,z,t)$ in the Helmholtz equation, we obtain:

$$h = \sqrt{k_0^2 n_f^2 - \beta^2} \quad \text{(film)}, \tag{6.4.5}$$

$$q = \sqrt{\beta^2 - k_0^2 n_c^2} \quad \text{(cover)}, \tag{6.4.6}$$

$$p = \sqrt{\beta^2 - k_0^2 n_s^2} \quad \text{(substrate)}. \tag{6.4.7}$$

The magnetic field strength of the mode follows from $\nabla \times \mathbf{E} = \imath\omega\mathbf{B}$ and $\mathbf{B} = \mu\mu_0\mathbf{H}$. The transverse component $H_z(x)$ reads:

$$H_z(x) = \begin{cases} \frac{-qA}{\imath\omega\mu_0}\, e^{-qx} & 0 \le x & \text{(cover)} \\ \frac{1}{\imath\omega\mu_0}\left[\imath hB\, e^{\imath hx} - \imath hC\, e^{-\imath hx}\right] & -d \le x \le 0 & \text{(film)} \\ \frac{pD}{\imath\omega\mu_0}\, e^{px} & x \le -d & \text{(substrate)} \end{cases} \tag{6.4.8}$$

$E_y(x)$ and $H_z(x)$ are continuous at the two interfaces. By imposing the continuity conditions at $x = 0$, we find:

$$\begin{aligned} A &= B + C, && (6.4.9)\\ -qA &= \imath hB - \imath hC. && (6.4.10) \end{aligned}$$

Solving for B and C yields:

$$\begin{aligned} B &= \frac{\imath h - q}{2\imath h} A, && (6.4.11)\\ C &= \frac{\imath h + q}{2\imath h} A. && (6.4.12) \end{aligned}$$

From the continuity condition for $E_y(x)$ at $x = -d$ it follows that

$$B\, e^{-\imath hd} + C\, e^{\imath hd} = D\, e^{-pd}. \tag{6.4.13}$$

With these conditions, we can rewrite $E_y(x)$ and $H_z(x)$ as

$$E_y(x) = \begin{cases} Ae^{-qx} & 0 \le x \quad \text{(cover)} \\ A\left[\cos(hx) - \frac{q}{h}\sin(hx)\right] & -d \le x \le 0 \quad \text{(film)} \\ A\left[\cos(hd) + \frac{q}{h}\sin(hd)\right] e^{[p(x+d)]} & x \le -d \quad \text{(substrate)} \end{cases} \tag{6.4.14}$$

and

$$H_z(x) = \begin{cases} \frac{-qA}{\imath\omega\mu_0} \exp(-qx) & 0 \le x \quad \text{(cover)} \\ \frac{-hA}{\imath\omega\mu_0}\left[\sin(hx) + \frac{q}{h}\cos(hx)\right] & -d \le x \le 0 \quad \text{(film)} \\ \frac{pA}{\imath\omega\mu_0}\left[\cos(hd) + \frac{q}{h}\sin(hd)\right] \exp\left[p(x+d)\right] & x \le -d \quad \text{(substrate)} \end{cases} \tag{6.4.15}$$

Dispersion relation

By imposing the continuity condition on $H_z(x)$ at $x = -d$, we get:

$$h\sin(hd) - q\cos(hd) = p\left[\cos(hd) + \frac{q}{h}\sin(hd)\right]. \tag{6.4.16}$$

This equation can be rewritten as

$$\tan(hd) = \frac{p+q}{h\left(1 - pq/h^2\right)}. \tag{6.4.17}$$

Using the identity $\arctan\left(\frac{x+y}{1-xy}\right) = \arctan(x) + \arctan(y) + m\pi$, we obtain

$$hd = \arctan\left(\frac{p}{h}\right) + \arctan\left(\frac{q}{h}\right) + m\pi. \tag{6.4.18}$$

By substituting h, p, and q in this equation, we finally arrive at the dispersion relation of the TE-waveguide modes in implicit form:

$$d\sqrt{k_0^2 n_f^2 - \beta^2} = \arctan\left(\sqrt{\frac{k_0^2(n_f^2 - n_s^2)}{k_0^2 n_f^2 - \beta^2} - 1}\right) + \arctan\left(\sqrt{\frac{k_0^2(n_f^2 - n_c^2)}{k_0^2 n_f^2 - \beta^2} - 1}\right) + m\pi. \tag{6.4.19}$$

By numerically solving the dispersion relation (6.4.19), we obtain the propagation constant β_m of the TE_m mode for a given frequency ω. The normal components of the wavevectors in the three media, h, p, q, follow from equations (6.4.5) to (6.4.7).

The ansatz (6.4.3) describes a guided wave, if both p and q are real quantities. For $n_s > n_c$, we can use equation (6.4.7) and the condition $p = 0$ to determine the cutoff frequency of the TE_m mode. Inserting $\beta_{cut} = \frac{\omega_{cut}}{c_0} n_s$ into equation (6.4.19) yields the cut-off frequency of the TE_m mode:

$$\omega_{\mathrm{cut}} = \frac{c_0\left(\arctan\left(\sqrt{\frac{n_s^2 - n_c^2}{n_f^2 - n_s^2}}\right) + m\pi\right)}{d\sqrt{n_f^2 - n_s^2}} \tag{6.4.20}$$

Note that the TE_0 mode of a symmetric slab waveguide ($n_s = n_c$) is always guided, i.e., $\omega_{cut,\mathrm{TE}_0} = 0$.

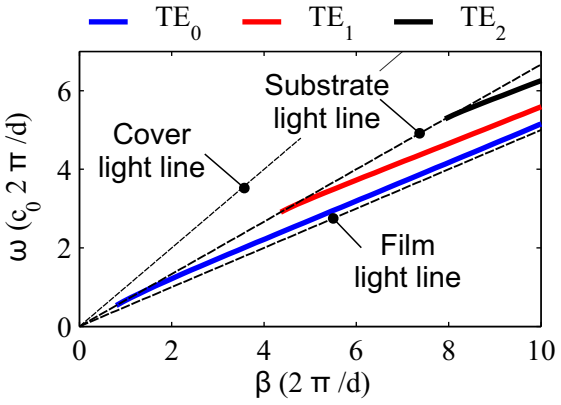


Figure 6.10: Dispersion relations of the TE_0, TE_1, and TE_2 mode of a slab waveguide with $n_f = 2.0$, $n_s = 1.5$, and $n_c = 1.0$.

Mode profile

Fig. 6.11 exemplifies the electric field profiles of the first three modes. The TE_m mode has m zeros of the electric field in the film layer.

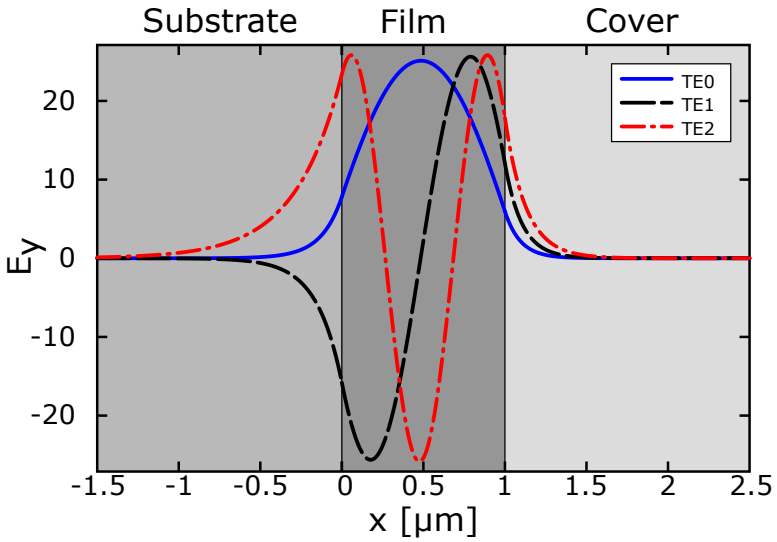


Figure 6.11: Electric field profiles of the first three TE modes.

An online multilayer slab waveguide mode solver can be found here:
`http://www.computational-photonics.eu/oms.html`

Effective refractive index

The effective index of the TE_m mode is defined as:

$$n_{\mathrm{eff,TE}_m} = \frac{\beta_m}{k_0}. \tag{6.4.21}$$

The effective index changes with frequency even if we assume that the refractive indices of the waveguide materials exhibit no dispersion. One finds for all modes the same trends (see also Fig. 6.12):

- $n_{\mathrm{eff}} \to n_s$ for $\omega \to \omega_{\mathrm{cut,TE}}$,
- $n_{\mathrm{eff}} \to n_f$ for $\omega \to \infty$.

This behavior is a result of the changing mode profile (see Fig. 6.13). In the following discussion, the TE_0 mode serves as the example. However, a similar behavior is also found

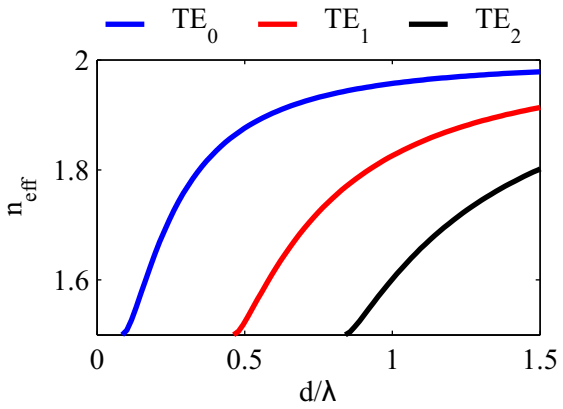


Figure 6.12: Effective refractive index of the TE_0, the TE_1, and the TE_2 mode of a slab waveguide with $n_f = 2.0$, $n_s = 1.5$, and $n_c = 1.0$.

for the higher order modes. For frequencies slightly above the cutoff (blue curve), the electric field penetrates deeply into the substrate. Hence, the field "sees" predominantly the refractive index of the substrate. As the frequency increases, the electric field is more concentrated in the film layer so that the mode primarily experiences the refractive index of the film.

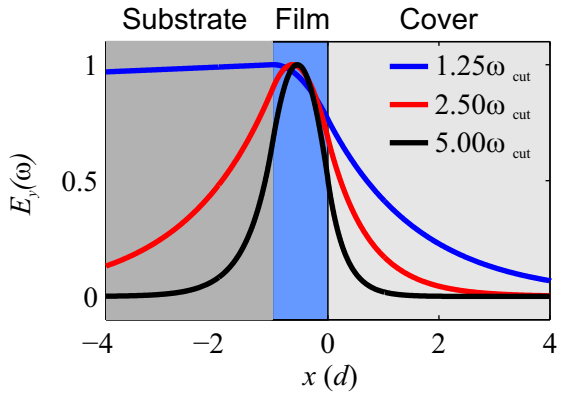


Figure 6.13: Electric field profile of the TE_0 mode for three different fequencies.

Orthogonality relation

In loss free waveguides, the waveguide modes fulfill the following orthogonality relation:

$$\frac{1}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\mathbf{E}_m \times \mathbf{H}_l^* + \mathbf{E}_l^* \times \mathbf{H}_m)\cdot \mathbf{e}_z\, dxdy = P_m \delta_{ml}. \tag{6.4.22}$$

Here, $(\mathbf{E}_m, \mathbf{H}_m)$ and $(\mathbf{E}_l, \mathbf{H}_l)$ are the electromagnetic fields of the TE_m and the TE_l mode, respectively, P_m is the power transported by the TE_m mode, and δ_{ml} is the Kronecker

delta. With $H_{x,m} = \frac{-\beta_m}{\omega\mu_0} E_{y,m}$, equation (6.4.22) can be rewritten for TE modes as:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_{y,m}E_{y,l}^{*}\,dxdy = 2\frac{\omega\mu_0}{\beta_m}P_m\delta_{ml}. \tag{6.4.23}$$

The orthogonality relation shows that the total power flow in a loss free waveguide is given by the sum of the power transported by the individual modes.

Proof of equation (6.4.22)

Let $(\mathbf{E}_m, \mathbf{H}_m)$ and $(\mathbf{E}_l, \mathbf{H}_l)$ be the electromagnetic fields of the TE_m and the TE_l mode. With the vector identity $\nabla\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot(\nabla\times\mathbf{A}) - \mathbf{A}\cdot(\nabla\times\mathbf{B})$, we obtain from Maxwell's equations for loss free media ($\epsilon = \epsilon^*$):

$$\nabla\cdot(\mathbf{E}_m\times\mathbf{H}_l^{*} + \mathbf{E}_l^{*}\times\mathbf{H}_m) = 0. \tag{6.4.24}$$

Next, we replace ∇ by $\nabla_t + \hat{\mathbf{e}}_z\partial/\partial z$, where ∇_t is the transverse gradient operator and $\hat{\mathbf{e}}_z$ is a unit vector along the propagation direction of the modes. With that, equation (6.4.24) can be rewritten as

$$\nabla_t\cdot(\mathbf{E}_m\times\mathbf{H}_l^{*} + \mathbf{E}_l^{*}\times\mathbf{H}_m) + \imath\,(\beta_m-\beta_l)\,\hat{\mathbf{e}}_z\cdot(\mathbf{E}_m\times\mathbf{H}_l^{*} + \mathbf{E}_l^{*}\times\mathbf{H}_m) = 0. \tag{6.4.25}$$

Here, β_m and β_l are the respective propagation constants of the two modes.

Integration over an arbitrary surface S in the xy plane and applying the two-dimensional form of the divergence theorem yields:

$$\begin{aligned}\int\int_S \nabla_t\cdot(\mathbf{E}_m\times\mathbf{H}_l^{*} + \mathbf{E}_l^{*}\times\mathbf{H}_m)\,da &= \oint_C (\mathbf{E}_m\times\mathbf{H}_l^{*} + \mathbf{E}_l^{*}\times\mathbf{H}_m)\cdot\hat{\mathbf{e}}_n dl \\ &= -\imath\,(\beta_m-\beta_l)\int\int_S \hat{\mathbf{e}}_z\cdot(\mathbf{E}_m\times\mathbf{H}_l^{*} + \mathbf{E}_l^{*}\times\mathbf{H}_m)\,da.\end{aligned} \tag{6.4.26}$$

Here, C is the contour of S and $\mathbf{n}$ is a unit vector normal to C and $\hat{\mathbf{e}}_z$. If we take S as the entire xy plane, the contour integral vanishes and we obtain:

$$(\beta_l-\beta_m)\int\int_S \hat{\mathbf{e}}_z\cdot(\mathbf{E}_m\times\mathbf{H}_l^{*} + \mathbf{E}_l^{*}\times\mathbf{H}_m)\,da = 0. \tag{6.4.27}$$

With the definition of the time averaged Poynting vector, this finally proofs equation (6.4.22).

6.4.2 TM-Moden

Analogous!

6.5 Excitation of guided modes

6.5.1 End-fire coupling

The simplest method to excite guided modes is end-fire coupling (see Fig. 6.14, left). Here, light is directly focused on a facet of the waveguide. The polarization and the transverse distribution of the light in the focus should match those of the desired waveguide mode. Because of the small dimensions of the waveguide film, end-fire coupling is usually an inefficient method. Moreover, in the multi-mode regime, this method usually results in the excitation of several different guided modes.

We can use ray optics to obtain an estimate for the optimal focusing condition. The largest possible bouncing angle of a ray guided by TIR is given by $\theta_{f,max} = \pi/2 - \theta_c$, where θ_c is the critical angle for TIR at the substrate-film interface. To calculate the corresponding external angle, θ_a, we have to take refraction at the input-facet of the waveguide into account. With the help of Snell's law, we find

$$\sin(\theta_a) = n_f \sin(\theta_{f,max}) = n_f \cos(\theta_c) = \sqrt{n_f^2 - n_s^2} = NA, \tag{6.5.1}$$

where NA is the numerical aperture of the waveguide.

6.5.2 Prism coupling

Prism coupling is an efficient excitation method for guided waves in slab waveguides. Here, the base of a prism made from a high index material is placed at a short distance, d_p, from the film layer (see Fig. 6.14, center). An incident light beam creates an evanescent field at the base of the prism, if the angle of incidence, θ_p, is larger than the critical angle of the base-interface. The evanescent wave can efficiently couple to a specific waveguide mode if (i) d_p is not to large and (ii) if the tangential component of the incident wave vector matches the propagation constant of the targeted waveguide mode:

$$\beta_m = k_0 n_p \sin(\theta_p). \tag{6.5.2}$$

Different waveguide modes can be selectively excited by setting the appropriate angle of incidence.

6.5.3 Grating coupling

A diffraction grating can be used to split an incident beam into several diffracted beams. The component of the wave vector parallel to the grating period of a diffracted beam is given by

$$k_{j,||} = k_{0,||} + jg, j = \pm 1, \pm 2, \ldots, \tag{6.5.3}$$

where, $\mathbf{k}_0$ is the wave vector of the incident beam, and $\mathbf{g}$ is the primitive reciprocal lattice vector of the grating with $g = \frac{2\pi}{a}$. A grating deposited on top of the film layer can be used to couple a plane wave to a waveguide mode (see Fig. 6.14, right). The phase matching condition reads:

$$\beta_m = k_0 \sin(\theta_i) \pm j\frac{2\pi}{a}, j = 1, 2, 3, \cdots \tag{6.5.4}$$

Different waveguide modes can be selectively excited by setting the appropriate angle of incidence.

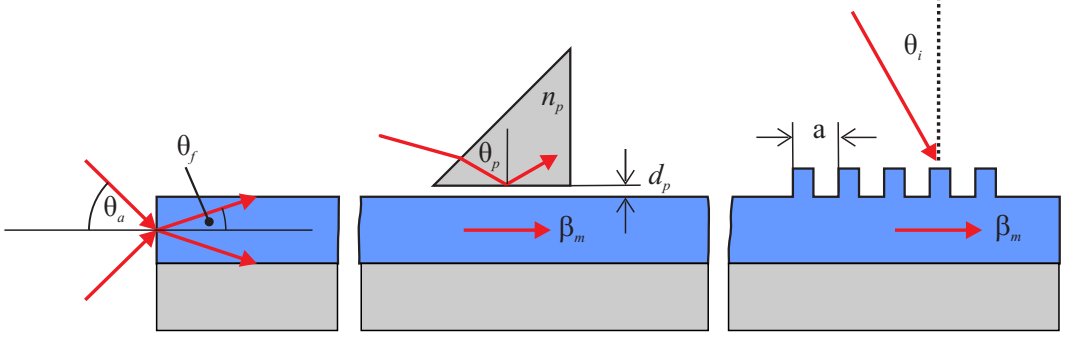


Figure 6.14: Excitation of guided modes. Left: End-fire coupling. Middle: Prism coupling. Right: Grating coupling.

6.6 Two-dimensional dielectric waveguides

In a slab waveguide, TIR confines the modes in only one dimension (in our case the x-direction). By structuring the film layer, we can confine the modes also in a second direction, e.g., in the y-direction. These 2D waveguides can be used as "wires for light" in which light propagates without spreading like in the case of a free beam.

In general, the electromagnetic field distribution and the dispersion relation of a 2D waveguide can not be calculated analytically but require either numerical methods or

approximation methods. Here, we briefly discuss the effective index method. The basic idea of this approximation method is to replace the 2D waveguide by an equivalent 1D slab waveguide. A prerequisite is that the electric (magnetic) field of the 2D waveguide mode is dominated by one vector component. In this case, we can adopt a scalar treatment of the problem.

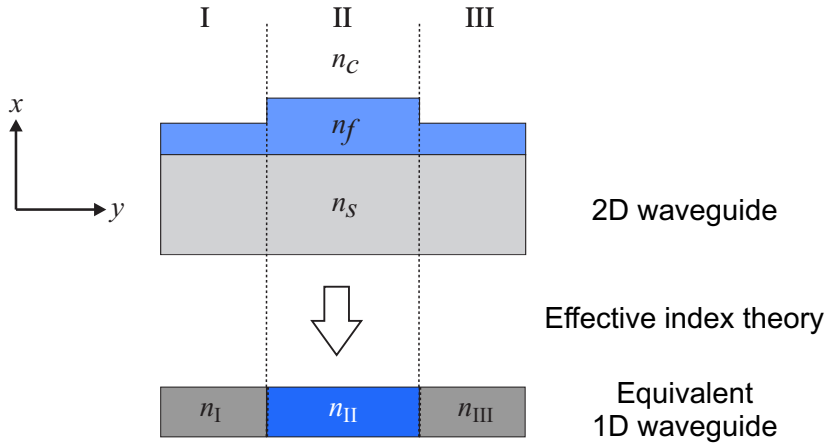


Figure 6.15: Principle of the effective index method.

Consider a TE mode of the 2D waveguide:

$$E(x,y,z,t) = E_y(x,y)e^{\imath(\beta z-\omega t)}. \tag{6.6.1}$$

The mode profile $E_y(x,y)$ satisfies the Helmholtz equation:

$$\left(\partial_x^2+\partial_y^2+\frac{\omega^2}{c_0^2}n^2(x,y)\right)E_y(x,y)=\beta^2E_y(x,y). \tag{6.6.2}$$

We choose the product ansatz:

$$E_y(x,y) = F(x,y)G(y). \tag{6.6.3}$$

Here, we assume that $F(x,y)$ depends only weakly on y such that $\partial_y^2F(x,y)=0$.

Substitution of equation (6.6.3) in (6.6.2) yields a set of two coupled equations:

$$\left(\partial_x^2+\frac{\omega^2}{c_0^2}n^2(x,y)\right)F(x,y)=\frac{\omega^2}{c_0^2}n_{eff}^2(y)F(x,y), \tag{6.6.4}$$

$$\left(\partial_y^2+\frac{\omega^2}{c_0^2}n_{eff}^2(y)\right)G(y)=\beta^2G(y). \tag{6.6.5}$$

Here, $F(x,y)$ and $n_{eff}(y)$ are only weakly dependent on y. Hence, equation (6.6.4) describes a 1D slab waveguide (in x-direction) in which y acts as a parameter. Its solution defines an effective index profile of a 1D slab waveguide - this time in the y-direction. The solution of equation (6.6.5) finally yields an approximate solution of the propagation constant β of the 2D waveguide mode.

6.7 Coupled mode theory

In what follows, we investigate the coupling between two parallel single-mode waveguides (see Fig. 6.16). Let the index profile of the composite structure be given by

$$\epsilon(x,y) = \epsilon_s + \Delta\epsilon_a(x,y) + \Delta\epsilon_b(x,y), \tag{6.7.1}$$

where $\Delta\epsilon_a(x,y)$ and $\Delta\epsilon_a(x,y)$ are zero outside of the core of waveguide a and b, respectively.

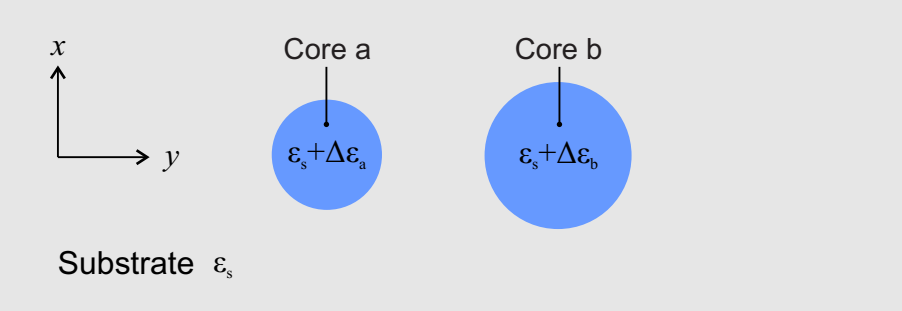


Figure 6.16: Index profile of a structure consisting of two waveguides.

We first consider the waveguide modes $E_\alpha(x,y,z,t) = E_\alpha(x,y)e^{\imath(\beta_\alpha z - \omega t)}$ $(with\, \alpha = a,b)$ of the two individual waveguides[1]. These modes satisfy the equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\omega^2}{c_0^2}\left[\epsilon_s + \Delta\epsilon_\alpha(x,y)\right]\right) E_\alpha(x,y) = \beta_\alpha^2\, E_\alpha(x,y). \tag{6.7.2}$$

The effect of coupling can be ignored if the separation between the two waveguides is much larger than the decay lengths of the evanescent fields outside of the waveguide cores. In this case, $E_a(x,y,z,t)$ and $E_b(x,y,z,t)$ are also modes of the composite structure and thus obey the orthogonality relation

$$\frac{\beta_a}{2\omega\mu_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_a E_b^*\, dxdy = \delta_{ab}. \tag{6.7.3}$$

[1]We assume for reasons of simplicity that the electric field of the modes has one dominant vector component such that a scalar treatment is possible.

Here, we assume that the power carried by the modes has been normalized to 1 Watt. The total electric field in the composite structure can be written as

$$E_{total}(x,y,z,t) = A\,E_a(x,y,z,t) + B\,E_b(x,y,z,t), \tag{6.7.4}$$

where, A and B are constant mode amplitudes.

For intermediate separations, the modes of the composite structure can be still approximated by $E_a(x,y,z,t)$ and $E_b(x,y,z,t)$, respectively. However, the coupling modifies the amplitudes such that the total field becomes:

$$E_{\text{total}}(x,y,z,t) = A(z)\,E_a(x,y,z,t) + B(z)\,E_b(x,y,z,t). \tag{6.7.5}$$

$\mathbf{E}_{\text{total}}(\mathbf{r},t)$ is a solution of the Helmholtz equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + \frac{\omega^2}{c_0^2}\left[\epsilon_s + \Delta\epsilon_a(x,y) + \Delta\epsilon_b(x,y)\right]\right) E_{\text{total}}(x,y,z,t) = 0. \tag{6.7.6}$$

Substitution of the ansatz (6.7.5) in equation (6.7.6) yields

$$\begin{aligned} &2\imath\beta_a\frac{dA}{dz}E_a(x,y)e^{\imath(\beta_a z-\omega t)} + 2\imath\beta_b\frac{dB}{dz}E_b(x,y)e^{\imath(\beta_b z-\omega t)} \\ &\quad = -\frac{\omega^2}{c_0^2}\Delta\epsilon_b(x,y)AE_a(x,y)e^{\imath(\beta_a z-\omega t)} - \frac{\omega^2}{c_0^2}\Delta\epsilon_a(x,y)BE_b(x,y)e^{\imath(\beta_b z-\omega t)}. \end{aligned} \tag{6.7.7}$$

Here we have assumed that $A(z)$ and $B(z)$ are slowly varying functions of z with $\frac{d^2A}{dz^2} = \frac{d^2B}{dz^2} = 0$.

Next, we multiply equation (6.7.7) with $E_a^*(x,y)$ and $E_b^*(x,y)$ and integrate over the entire xy plane. With the normalization condition (6.7.3), we obtain

$$\frac{dA}{dz} = \imath\kappa_{ab}Be^{-\imath(\beta_a-\beta_b)z} + \imath\kappa_{aa}A \tag{6.7.8}$$

$$\frac{dB}{dz} = \imath\kappa_{ba}Ae^{\imath(\beta_a-\beta_b)z} + \imath\kappa_{bb}B \tag{6.7.9}$$

with

$$\kappa_{ab} = \frac{\omega}{4}\epsilon_0\int\int E_a^*(x,y)\Delta\epsilon_a(x,y)E_b(x,y)dxdy \tag{6.7.10}$$

$$\kappa_{ba} = \frac{\omega}{4}\epsilon_0\int\int E_b^*(x,y)\Delta\epsilon_b(x,y)E_a(x,y)dxdy \tag{6.7.11}$$

$$\kappa_{aa} = \frac{\omega}{4}\epsilon_0\int\int E_a^*(x,y)\Delta\epsilon_b(x,y)E_a(x,y)dxdy \tag{6.7.12}$$

$$\kappa_{bb} = \frac{\omega}{4}\epsilon_0\int\int E_b^*(x,y)\Delta\epsilon_a(x,y)E_b(x,y)dxdy. \tag{6.7.13}$$

$$\tag{6.7.14}$$

The terms with κ_{aa} and κ_{bb} lead to small corrections of the propagation constants due to the presence of the other waveguide:

$$\beta_a \rightarrow \beta_a + \kappa_{aa} \tag{6.7.15}$$
$$\beta_b \rightarrow \beta_a + \kappa_{bb}. \tag{6.7.16}$$

If we take the total electric field as

$$E_{\text{total}}(x,y,z,t) = A(z)\,E_a(x,y)e^{\imath[(\beta_a+\kappa_{aa})z-\omega t]} + B(z)\,E_b(x,y)e^{\imath[(\beta_b+\kappa_{bb})z-\omega t]}, \tag{6.7.17}$$

the coupled equations (6.7.8) and (6.7.9) simplify to

$$\frac{dA}{dz} = \imath\kappa_{ab}Be^{-\imath\Delta\beta z} \tag{6.7.18}$$
$$\frac{dB}{dz} = \imath\kappa_{ba}Ae^{\imath\Delta\beta z}. \tag{6.7.19}$$

Here, we have introduced the phase mismatch

$$\Delta\beta = (\beta_a + \kappa_{aa}) - (\beta_b + \kappa_{bb}). \tag{6.7.20}$$

In what follows we consider the case that at $z = 0$ the field in one of the waveguides is zero, e.g., $A(z = 0) = A_0$ and $B(z = 0) = 0$. In this case, solutions of equations (6.7.18) and (6.7.19) are given by (proof: exercise)

$$A(z) = A_0e^{-\imath\frac{\Delta\beta}{2}z}\left[\cos(\gamma z) + \imath\frac{\Delta\beta}{2\gamma}\sin(\gamma z)\right] \tag{6.7.21}$$
$$B(z) = \imath\frac{\kappa}{\gamma}A_0e^{\imath\frac{\Delta\beta}{2}z}\sin(\gamma z). \tag{6.7.22}$$

Here, we have assumed $\kappa_{ab} = \kappa_{ba} = \kappa$ and γ is defined by

$$\gamma^2 = \left(\frac{\Delta\beta}{2}\right)^2 + \kappa^2. \tag{6.7.23}$$

The power flow in the two waveguides is given by:

$$P_a(z) = |A(z)|^2 = P_a(0)\left[\cos^2(\gamma z) + \left(\frac{\Delta\beta}{2\gamma}\right)^2\sin^2(\gamma z)\right] \tag{6.7.24}$$
$$P_b(z) = |B(z)|^2 = P_a(0)\left(\frac{\kappa}{\gamma}\right)^2\sin^2(\gamma z). \tag{6.7.25}$$

The power is periodically exchanged between the two waveguides. For identical waveguides ($\Delta\beta = 0$), the power is completely transferred from waveguide a to waveguide b at the coupling length

$$L_0 = \frac{\pi}{2\kappa}. \tag{6.7.26}$$

At the distance $L_0/2$, half of the power is coupled to the other waveguide. Hence, the two waveguides act at this distance as a 50/50 beamsplitter. By varying the length of the coupling region we can transfer any desired fraction of the optical power from one waveguide to the other.

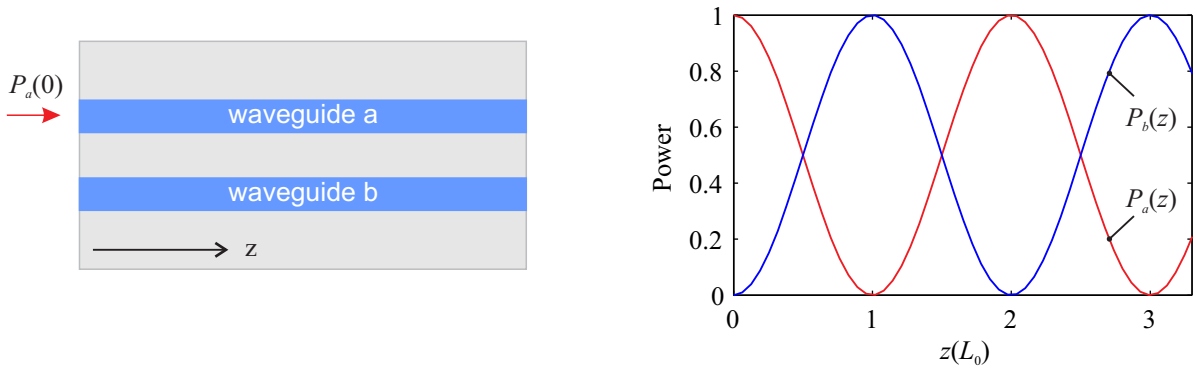


Figure 6.17: Power exchange between two identical coupled waveguides.

6.8 Optical fibers

An optical fiber is a cylindrical waveguide made from low loss materials, e.g., silica glass. In the following, we consider step-index optical fibers in which a uniform core of diameter $2a$ is embedded in a uniform cladding material with a lower index of refraction.

Optical fibers are widely used in telecommunications and are essential for the Internet. Charles Kuen Kao was awarded half of the 2009 Nobel Prize in Physics for "groundbreaking achievements concerning the transmission of light in fibers for optical communication".

Typical parameters of a telecommunication fiber

- Operation wavelength: $\lambda = 1.55\mu m$ (low losses).
- Cladding: Ultra pure fused silica glass ($n_2 = 1.46$).

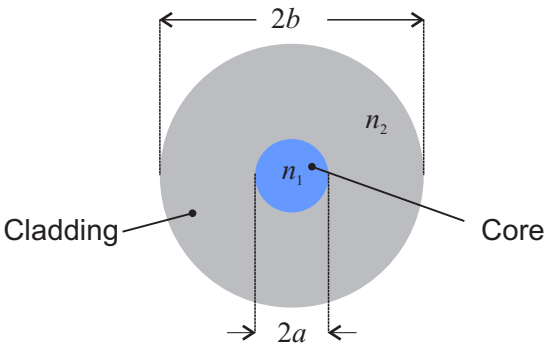


Figure 6.18: Cross section of a step-index optical fiber.

- Core: Ge doped silica glass.
- Relative refractive index difference for typical doping concentrations:

$$\Delta = \frac{n_1 - n_2}{n_1} \approx 0.01 - 0.02. \tag{6.8.1}$$

  $\Rightarrow$ Most optical fibers used for practical applications are weakly guiding fibers.
- Geometric dimensions of a typical single mode optical fiber:
  - Core diameter: $2a = 8\mu m$.
  - Cladding diameter: $2b = 125\mu m$.

6.8.1 Weakly guiding fibers — LP modes

In a weakly guiding fiber, the electric (magnetic) field is almost transversal, i.e., the field components in the propagation direction can be neglected[2]. In the following, we will thus look for guided modes of the form:

$$\mathbf{E}(r,\phi,z,t) = \mathbf{e}_{tr}\, U(r,\phi) e^{\imath(\beta z - \omega t)} \quad with \quad \mathbf{e}_{tr} \cdot \mathbf{e}_z \approx 0. \tag{6.8.2}$$

These linearly polarized modes are the so-called LP modes.

[2]This aspect can be easily understood in terms of ray-optics. A small relative refractive index difference implies that the bouncing angles of the guided rays are small. By replacing the guided rays with plane waves, we immediately see that the expected field distributions have only small lateral field components.

The function $U(r,\phi)$ satisfies the Helmholtz equation in cylindrical coordinates:

$$\left[\Delta_r + \Delta_\phi + k_0^2 n^2(r)\right] U(r,\phi) = \beta^2 U(r,\phi) \quad with \quad \begin{cases} \Delta_r = \frac{1}{r}\partial_r r \partial_r \\ \Delta_\phi = \frac{1}{r^2}\partial_\phi^2 \end{cases} \tag{6.8.3}$$

Substituting the ansatz

$$U(r,\phi) = u(r)e^{\imath l\phi} \tag{6.8.4}$$

in the Helmholtz equation leads to the following equations for the radial profile $u(r)$:

- Core ($r \le a$):

$$\left[\Delta_r + h^2 - \frac{l^2}{r^2}\right] u(r) = 0 \quad with \quad h^2 = n_1^2 k_0^2 - \beta^2, \tag{6.8.5}$$

- Cladding: ($r > a$):

$$\left[\Delta_r - q^2 - \frac{l^2}{r^2}\right] u(r) = 0 \quad with \quad q^2 = \beta^2 - n_2^2 k_0^2. \tag{6.8.6}$$

The solutions of equations (6.8.5) and (6.8.6) are the family of of Bessel functions. For physical reasons, we have to exclude those solutions which diverge at $r = 0$ in the core or $r \to \infty$ in the cladding. With that, we obtain the bound solutions:

$$u(r) \propto \begin{cases} J_l(hr) & r < a \quad \text{(Core)}, \\ K_l(qr) & r > a \quad \text{(Cladding)}. \end{cases} \tag{6.8.7}$$

Here, $J_l(x)$ is the Bessel function of the first kind and order l, and $K_l(x)$ is the modified Bessel function of the second kind and order l.

For reasons of convenience, we introduce normalized coordinates

$$X = ha, \tag{6.8.8}$$
$$Y = qa, \tag{6.8.9}$$

and the V-parameter (normalized frequency) of the fiber:

$$V = \sqrt{X^2 + Y^2} = \frac{2\pi a}{\lambda_0}\sqrt{n_1^2 - n_2^2}. \tag{6.8.10}$$

The continuity of $u(r)$ and $u'(r)$ at $r = a$ requires that

$$X\frac{J_{l\pm1}(X)}{J_l(X)} = \pm Y\frac{K_{l\pm1}(Y)}{K_l(Y)}. \tag{6.8.11}$$

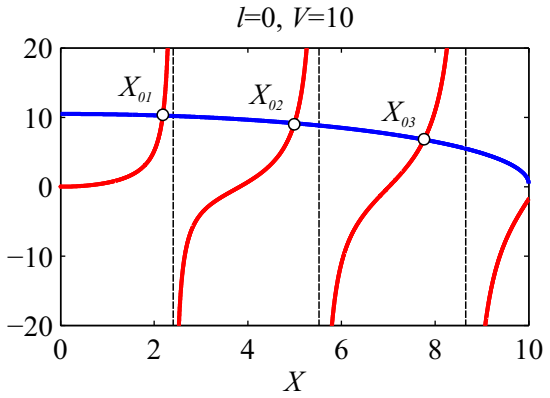


Figure 6.19: Graphical solution of the dispersion relation of the LP modes for $l = 0$ and $V = 10$. The red and blue curve correspond to the LHS and RHS, respectively, of equation (6.8.11). The black dashed lines are the roots of $J_0(X)$.

l	m=1	m=2	m=3
0	0	3.832	7.016
1	2.405	5.520	8.654

Table 6.1: Cutoff frequencies of several LP modes.

This equation is an implicit representation of the dispersion relation of the LP modes. Each of the m intersection points X_{lm} of the curves $X\frac{J_{l\pm1}(X)}{J_l(X)}$ and $\pm Y\frac{K_{l\pm1}(Y)}{K_l(Y)}$ gives rise to a different LP mode. The propagation constant, β_{lm}, of the LP_{lm} mode can be calculated as

$$\beta_{lm} = \sqrt{n_1^2 k_0^2 - \frac{X_{lm}^2}{a^2}}. \tag{6.8.12}$$

The dispersion relation (6.8.11) has at least one solution so that the LP_{01} mode is always guided. All other LP modes are only guided for frequencies above the respective cutoff frequency. In the frequency interval $0 \leq V \leq 2.405$, an optical fiber operates as a single-mode waveguide.

The intensity profile of selected LP modes is depicted in Fig. 6.20.

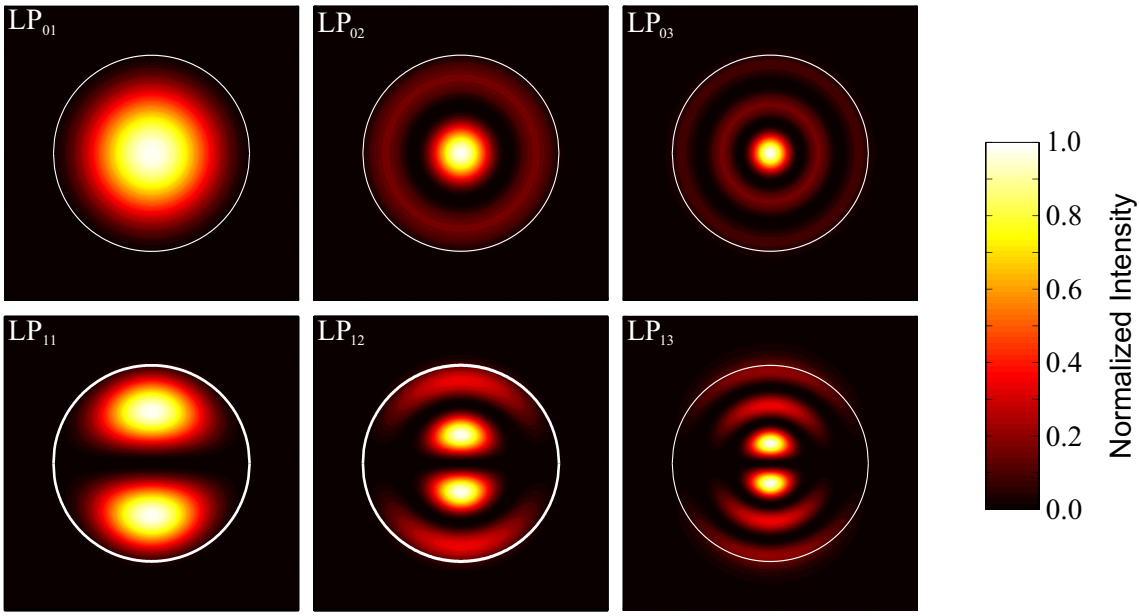


Figure 6.20: Intensityprofile of selected LP modes for $V = 10$. The white line marks the interface between the core and the cladding.

6.8.2 Dispersion and absorption in optical fibers

Dispersion

Dispersion in optical fibers plays an important role in the context of telecommunications. There are several sources of dispersion in optical fibers:

- **Material dispersion**: See section 1.4.1.
- **Waveguide dispersion**: The effective refractive index of a mode depends on frequency because of the changing mode profile.
- **Modal dispersion**: Different modes travel at different velocities.
- **Nonlinear dispersion**: Nonlinear optical effects lead to an intensity dependent refractive index.

In the following example, we discuss the influence of the first three contributions. We consider a step-index optical fiber with core radius $a = 3.7\,\mu\text{m}$. The cladding is assumed to be made of fused silica (see black curve in Fig. 6.21 (a)) and the relative refractive index difference between core and cladding is $\Delta = 0.005$.

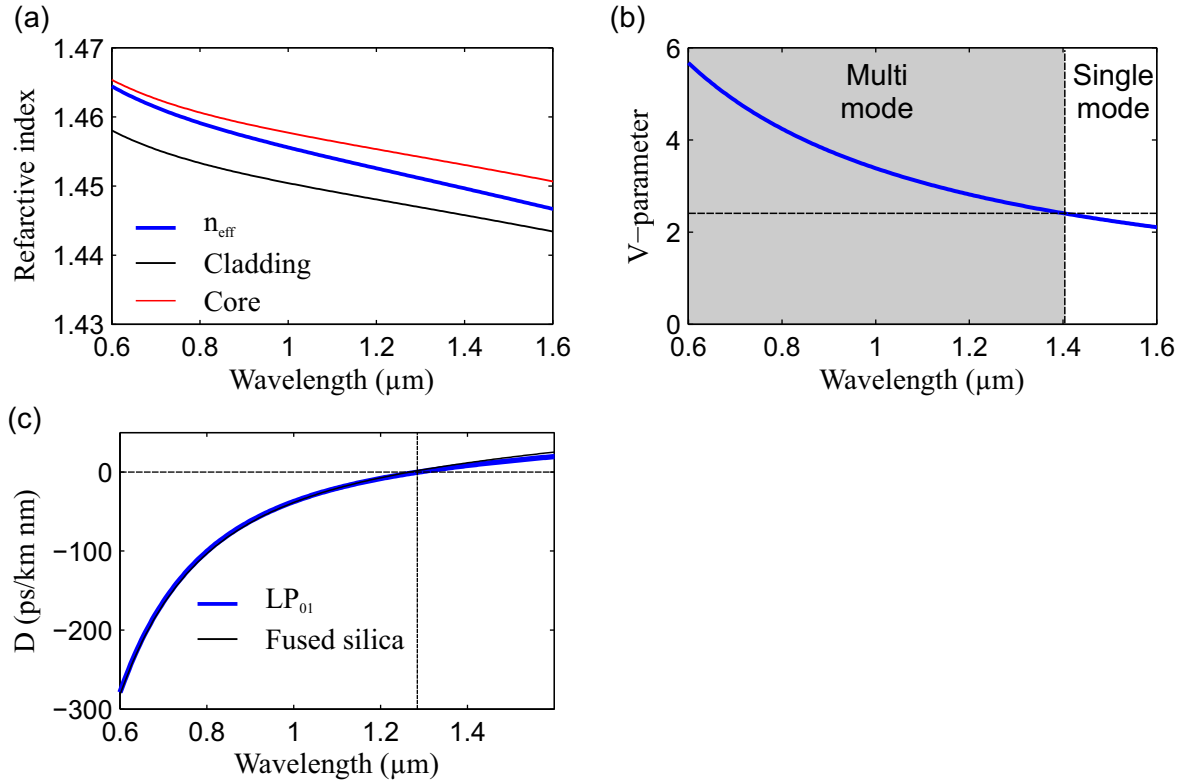


Figure 6.21: (a) Calculated effective refractive index of the LP_{01} mode (blue) and refractive index of the cladding (black) and the core (red). (b) V-parameter. (c) Dispersion coefficient of the LP_{01} mode (blue) and fused silica (black).

The effective refractive index $n_{eff} = \beta/k_0$ of the LP_{01} mode (see blue curve in Fig. 6.21 (a)) exhibits the expected behavior : For small wavelength (large frequencies), n_{eff} approaches the refractive index of the core while for large wavelength (small frequencies) n_{eff} tends to the refractive index of the cladding.

Fig. 6.21 (b) depicts the V-parameter of the optical fiber for the given geometry and materials. The optical fiber supports single mode operation for wavelength $\lambda \geq 1.4\,\mu m$. For smaller wavelength, the fiber has several guided modes and modal dispersion becomes important.

The dispersion coefficient $D(\lambda)$ of the LP_{01} mode (blue curve) is shown together with the dispersion coefficient of fused silica (black curve) in Fig. 6.21 (c). The two curves lie almost on top of each other. Hence, the dispersion of the LP_{01} mode is dominated by material dispersion. The waveguide dispersion plays for the given parameters only a minor role.

Absorption

Absorption and scattering cause an exponential decay of the optical power of a guided wave, $P(L)$, with the propagation distance, L. The corresponding attenuation coefficient, α is usually given in units of dB/km:

$$\alpha = \frac{1}{L} 10 \log_{10} \left(\frac{1}{T(L)} \right), \tag{6.8.13}$$

where $T(L) = P(L)/P(0)$ is the transmittance of the optical fiber.

Example: Conversion of dB/km in $T(1km)$

- 0 dB/km $\leftrightarrow T = 1$
- 3 dB/km $\leftrightarrow T \approx 0.5$
- 10 dB/km $\leftrightarrow T = 0.1$
- 20 dB/km $\leftrightarrow T = 0.01$

Different mechanisms contribute to the fiber losses:

- UV-absorption caused by the electronic polarizability of SiO_2.
- Infrared-absorption in SiO_2 due to molecular vibrations.
- Rayleigh-scattering caused by inhomogeneities in the glass.
- OH-absorption.

The respective weights of these loss mechanisms depend on the operation wavelength. The loss coefficients of high quality single mode fibers can be as low as $\alpha \approx 0.15\,\mathrm{d}B/km$ at the operation wavelength $\lambda = 1.55\,\mu\mathrm{m}$.

6.9 Surface plasmon polaritons

Light can not propagate in bulk metal for frequencies below the plasma frequency, ω_p. However, the interface between a metal and a dielectric medium supports guided waves in this frequency range. These so-called surface plasmon polaritons (SPPs) are longitudinal charge density oscillations coupled to an TM polarized optical wave. The electromagnetic

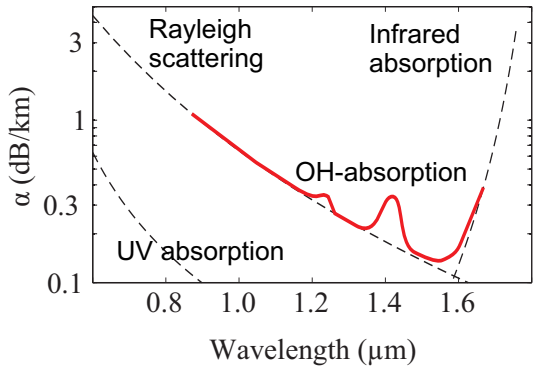


Figure 6.22: Scheme of the attenuation coefficient of a high quality single mode fiber.

field of the SPP has its maximum field strength at the interface and falls off exponentially in both media.

6.9.1 SPP dispersion relation

To derive the SPP dispersion relation, we start with the ansatz:

- Dielectric medium $(x > 0, \epsilon(x) = \epsilon_d)$:

$$\mathbf{H}^+(\mathbf{r},t) = (0,A,0)\, e^{-k_x^+ x}\, e^{\imath(\beta_{SPP} z - \omega t)}, \tag{6.9.1}$$

$$\Rightarrow \mathbf{E}^+(\mathbf{r},t) = \left(\frac{\beta_{SPP} A}{\omega\epsilon_0\epsilon_d}, 0, \frac{-\imath k_x^+ A}{\omega\epsilon_0\epsilon_d}\right) e^{-k_x^+ x}\, e^{\imath(\beta_{SPP} z - \omega t)}. \tag{6.9.2}$$

- Metal $(x < 0, \epsilon(x) = \epsilon_m(\omega))$:

$$\mathbf{H}^-(\mathbf{r},t) = (0,B,0)\, e^{k_x^- x}\, e^{\imath(\beta_{SPP} z - \omega t)}, \tag{6.9.3}$$

$$\Rightarrow \mathbf{E}^-(\mathbf{r},t) = \left(\frac{\beta_{SPP} B}{\omega\epsilon_0\epsilon_m(\omega)}, 0, \frac{\imath k_x^- B}{\omega\epsilon_0\epsilon_m(\omega)}\right) e^{k_x^- x}\, e^{\imath(\beta_{SPP} z - \omega t)} \tag{6.9.4}$$

From the continuity of the tangential component of the electric and the magnetic field strength, respectievly, at the interface $(x = 0)$ follows that

$$H_y^-(x=0,t) = H_y^+(x=0,t) \Rightarrow A = B, \tag{6.9.5}$$

$$E_z^-(x=0,t) = E_z^+(x=0,t) \Rightarrow \frac{k_x^+}{\epsilon_d} = -\frac{k_x^-}{\epsilon_m(\omega)}. \tag{6.9.6}$$

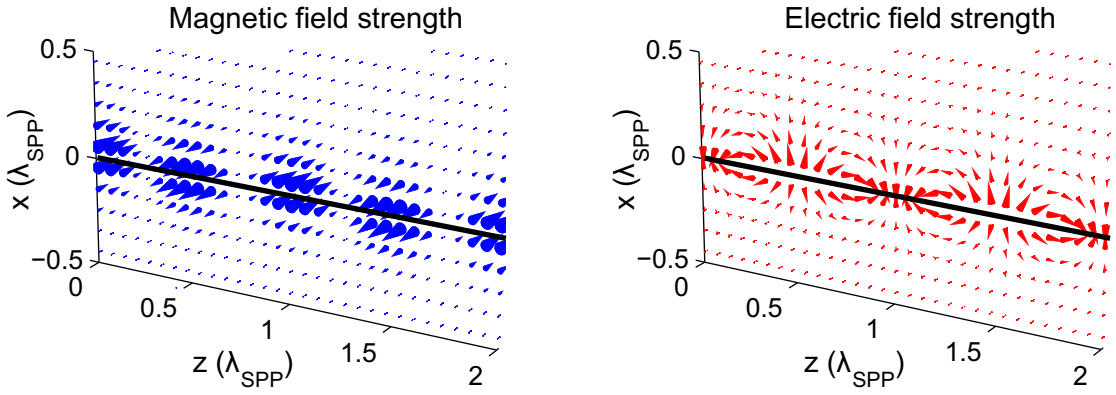


Figure 6.23: Snapshot of the magnetic (left) and the electric (right) field strength of a SPP.

The latter condition implies that ϵ_d and $\epsilon_m(\omega)$ have opposite algebraic signs.

Substituting the ansatz in the wave equation yields:

$$\beta_{SPP}^2 - (k_x^+)^2 = \epsilon_d \frac{\omega^2}{c_0^2}, \tag{6.9.7}$$

$$\beta_{SPP}^2 - (k_x^-)^2 = \epsilon_m(\omega) \frac{\omega^2}{c_0^2}. \tag{6.9.8}$$

After a short calculation (Proof: Exercise), we obtain the SPP dispersion relation:

$$\beta_{SPP} = \frac{\omega}{c_0} \sqrt{\frac{\epsilon_d \epsilon_m(\omega)}{\epsilon_d + \epsilon_m(\omega)}}. \tag{6.9.9}$$

Small propagation losses require that the real part of the function $[\epsilon_d \epsilon_m(\omega)] / (\epsilon_d + \epsilon_m(\omega)]$ is positive. This implies that the dielectric functions have to fulfill the conditions:

$$\epsilon_d \epsilon_m(\omega) < 0, \tag{6.9.10}$$

$$\epsilon_d + \epsilon_m(\omega) < 0. \tag{6.9.11}$$

Fig. 6.24 depicts the SPP dispersion relation for the case of an interface between a dielectric medium and a Drude metal characterized by the plasma frequency ω_p. As expected, the SPP dispersion curve (blue curve) lies on the right side of the light line of the dielectric medium (black curve). Furthermore, it is easy to show that for $\beta_{SPP} \to \infty$, the frequency of the SPP approaches the limiting frequency

$$\omega_{SPP} = \frac{\omega_p}{\sqrt{1 + \epsilon_d}}. \tag{6.9.12}$$

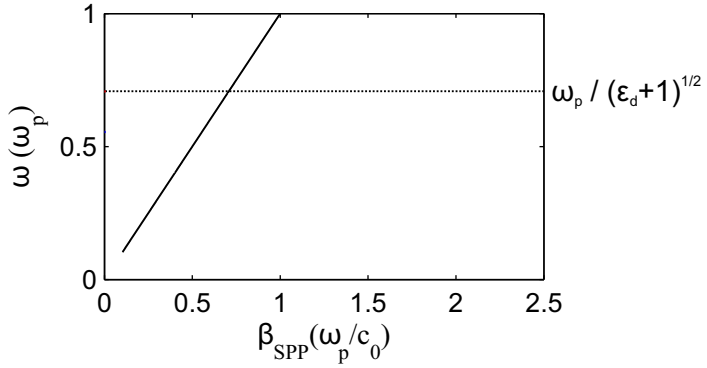


Figure 6.24: SPP dispersions relation (blue curve) for the case of an interface between a dielectric medium with dielectric constant ϵ_d and a Drude metal characterized by the plasma frequency ω_p.

In "real" metals, interband transitions result in a strong damping of the SPP and limit β'_{SPP} to finite values. This effect is shown in Fig 6.25 for the case of a silver-air interface.

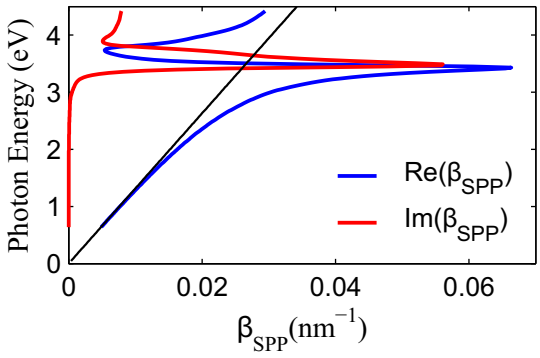


Figure 6.25: SPP dispersion relation for an Ag-air interface.

For "good" metals ($\epsilon'_m \gg \epsilon''_m$), the following approximations are valid:

$$\beta'_{SPP} \approx \frac{\omega}{c_0}\left(\frac{\epsilon_d \epsilon'_m(\omega)}{\epsilon_d + \epsilon'_m(\omega)}\right)^{1/2} \tag{6.9.13}$$

$$\beta''_{SPP} \approx \frac{\omega}{c_0}\left(\frac{\epsilon_d \epsilon'_m(\omega)}{\epsilon_d + \epsilon'_m(\omega)}\right)^{3/2} \frac{\epsilon''_m(\omega)}{2(\epsilon'_m(\omega))^2} \tag{6.9.14}$$

The real and the imaginary part of β_{SPP} are related to the SPP wavelength, λ_{SPP}, and the SPP propagation length, L_{SPP}, respectively:

$$\lambda_{SPP} = \frac{2\pi}{\beta'_{SPP}}, \tag{6.9.15}$$

$$L_{SPP} = \frac{1}{2\beta''_{SPP}}. \tag{6.9.16}$$

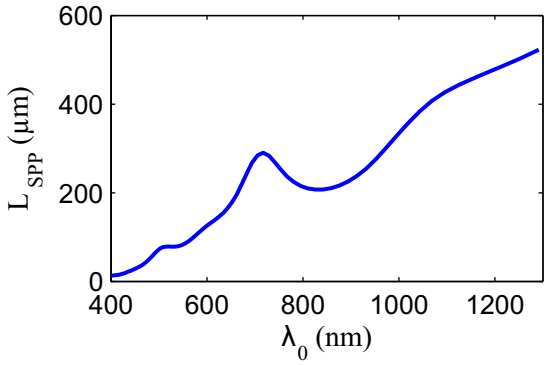


Figure 6.26: Propagation length of the SPP at the Ag-air interface as a function of the vacuum wavelength.

6.9.2 Excitation of SPPs

The wave number of the SPP is larger than that of light in the dielectric medium. As a consequence, a SPP cannot be directly excited by an incident light beam impinging on a flat metallic. In the following, we briefly discuss three common techniques to excite SPPs.

Prism coupling - Kretschmann configuration

In the Kretschmann configuration, a thin metal film (typically the thickness is less than 100 nm) is evaporated onto the base of a prism. The metal film is illuminated through the prism with a p-polarized light beam. If the phase matching condition

$$\beta_{SPP} = \sqrt{\epsilon_p} k_0 \sin(\theta_p) \tag{6.9.17}$$

is fulfilled, light can evanescently couple through the metal film to the SPP at the metal-air interface. This results in a pronounced dip (surface plasmon resonance) in the reflectance of the metal film (see Fig. 6.28).

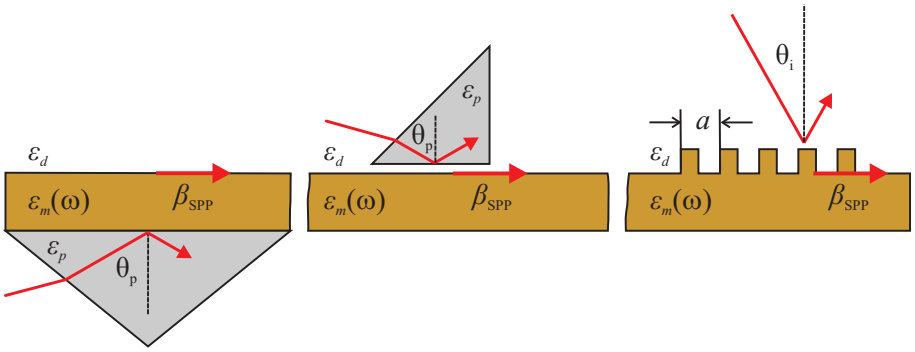


Figure 6.27: Excitation of SPPs: Prism coupling in Kretschmann configuration (left), prism coupling in Otto configuration (center), and grating coupling (right).

Prism coupling - Otto configuration

Prism coupling in the Otto configuration is analogous to prism coupling in the case of a dielectric slab waveguide (see section 6.5.2). The phase matching condition reads:

$$\beta_{SPP} = \sqrt{\epsilon_p} k_0 \sin(\theta_p). \tag{6.9.18}$$

Grating coupling

A grating with period a deposited on top of the metal film can be used to couple a p-polarized plane wave to the SPP. The phase matching condition reads:

$$\beta_{SPP} = k_0 \sin(\theta_i) \pm j\frac{2\pi}{a}, j = 1, 2, 3, \cdots . \tag{6.9.19}$$

6.9.3 Application: Surface plasmon resonance biosensor

SPPs are very sensitive to small changes of the dielectric environment in the vicinity of the metal-dielectric interface. Surface plasmon resonance (SPR) biosensors use this property to detect small amounts of a specific analyte, e.g., an antibody, in a test solution.

Most SPR biosensors are based on a prism coupler in Kretschmann configuration (see Fig.6.30). Here, the thin metal film is coated with the respective conjugated antigen of the analyte. In a first step, the SPP resonance condition (angle of incidence or resonance wavelength) is measured with a buffer solution without the analyte. Then the measurement is repeated with the test solution. Binding of the analyte to the conjugated antigen

changes the dielectric function of the functional layer $\epsilon_{\mathrm{ml}} \rightarrow \epsilon_{\mathrm{ml}} + \delta\epsilon$. This in turn modifies the SPP resonance condition. Fig. 6.31 depicts the calculated shift of the SPR of a 50 nm thick Ag film due to the presence of 2 nm thick layer of a protein. (Parameters: $\epsilon_d = \epsilon_{ml} = 1.76, \delta\epsilon = 0.04$).

Figure 6.28: Calculated reflectance of a 50 nm thick Ag film illuminated in the Kretschmann configuration.

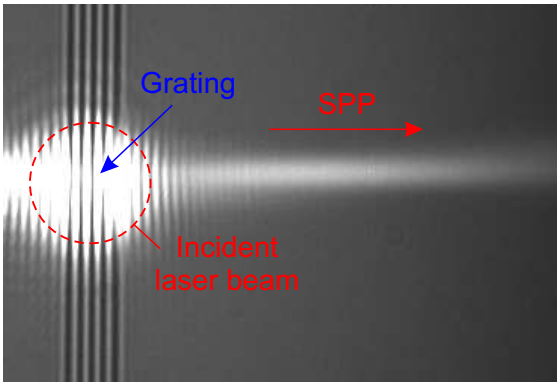


Figure 6.29: Leakage radiation microscopy image of a SPP excited by grating coupling.

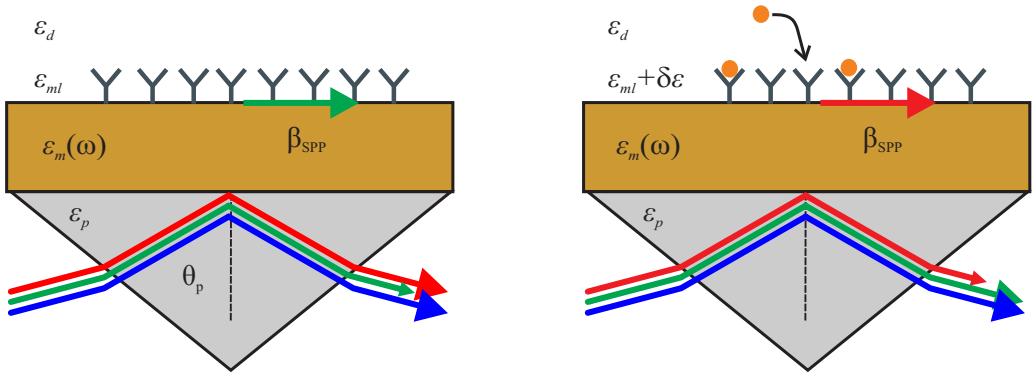


Figure 6.30: SPP based biosensor.

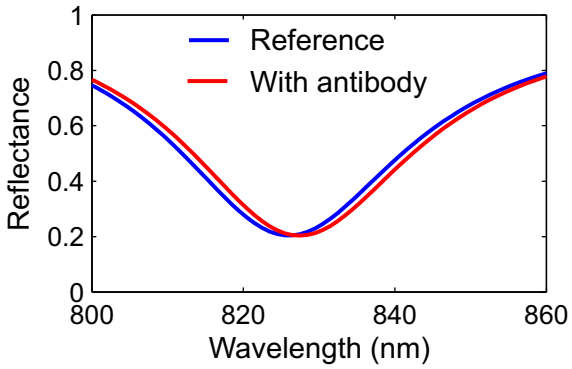


Figure 6.31: Calculated shift of the SPR.

7 Laser

The word laser has been originally introduced as an acronym for **l**ight **a**mplification by **s**timulated **e**mission of **r**adiation. However, nowadays it is mostly used to denote a device that emits coherent light generated by stimulated emission. The first laser was constructed in 1960 by Theodore Maiman. This invention triggered a phenomenal boom of research activities that still lasts today. By now, lasers have found widespread application in almost any branch of the natural sciences and engineering.

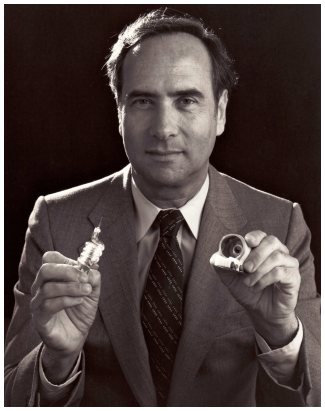


Figure 7.1: Theodore Maiman constructed the first laser. (Image source: Wikipedia).

7.1 Emission and absorption of electromagnetic radiation

In this section, we dicuss the interaction a monochromatic light field with an ensemble of identical atoms. We assume that the radiation is resonant to an optical allowed transition between two atomic states, $|1\rangle$ and $|2\rangle$, with energies E_1 and E_2, respectively, i.e., the light frequency ν matches the transition frequency

$$\nu_0 = \frac{E_2 - E_1}{h}. \tag{7.1.1}$$

All other atomic transitions are assumed to be sufficiently detuned such that the atoms can be treated as two-level systems.

The emission and absorption spectra of the atoms are not δ-functions, i.e., the optical transition is smeared out over a frequency range $\Delta\nu$ around the transition frequency ν_0. Spectral broadening is the result of the finite lifetimes, τ_1 and τ_2, of the levels $|1\rangle$ and $|2\rangle$, respectively. To account for this fact, we introduce the normalized line shape function $\hat{g}(\nu)$ with

$$\int_0^\infty \hat{g}(\nu)\, d\nu = 1. \tag{7.1.2}$$

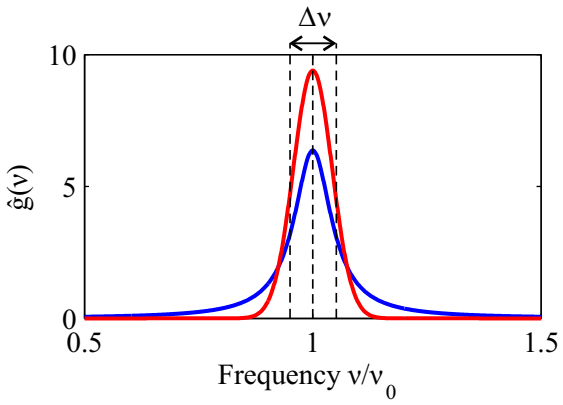


Figure 7.2: Spectral line shape functions: Homogeneous broadening (blue) and inhomogeneous broadening (red).

7.1.1 Transition rates

We will now discuss the transition rates of the elementary processes relevant for the interaction of a two-level system with a (near) resonant monochromatic light field (see Figure 7.3).

Absorption rate

We assume that the two-level system is initially in the lower energy state $|1\rangle$. By absorption of an incident photon, the two-level system can make a transition to the upper energy state $|2\rangle$. Since this transition is induced by the light field, it can be plausibly expected that the absorption rate is proportional to the energy density $u(\nu)$ of the light field, i.e., the number of photons, and to the spectral overlap of the light field with the atomic transition charcterized by $\hat{g}(\nu)$. The rate of transition from $|1\rangle$ to $|2\rangle$ due to absorption of one photon with frequency in the range between ν and $\nu + d\nu$ can thus be written as

$$W_{12}(\nu)\, d\nu = B_{12} u(\nu) \hat{g}(\nu)\, d\nu, \tag{7.1.3}$$

where B_{12} is a constant.

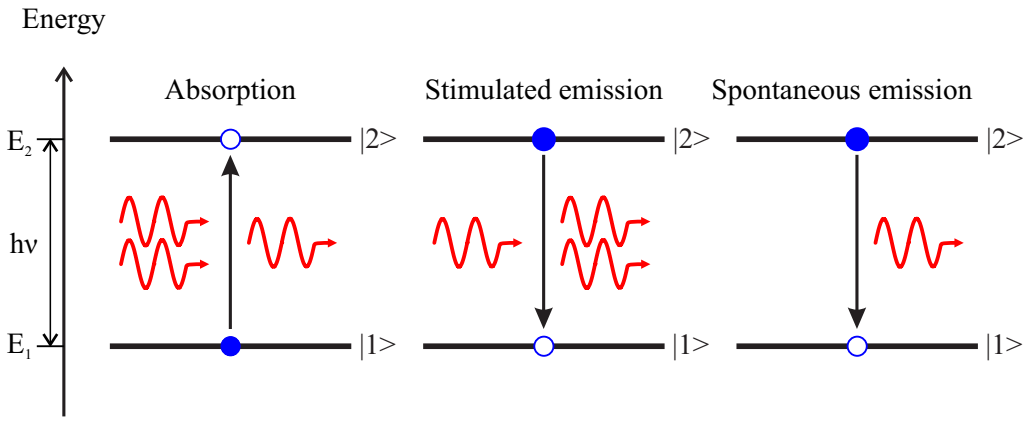


Figure 7.3: Interaction of a two-level system with a (near) resonant monochromatic light field.

Rate of stimulated emission

If the two-level system is initially in the upper energy state $|2\rangle$, it can be stimulated by an incident photon to make a transition to the lower energy state $|1\rangle$. In this process, a second photon is emitted that is a cloned version of the incident photon, i.e., the emitted photon inherits the frequency, the wave vector, and the polarization of the incident photon. Stimulated emission can be considered as the reverse process of absorption. The rate of stimulated emission for photons in the frequency range between ν and $\nu + d\nu$ is thus given by

$$W_{21}(\nu)\, d\nu = B_{21} u(\nu) \hat{g}(\nu)\, d\nu, \tag{7.1.4}$$

where B_{21} is a constant.

Rate of spontaneous emission

Finally, if the two-level system is initially in the upper energy state $|2\rangle$, it can make a spontaneous (without incident photon) transition to the lower energy state $|1\rangle$. The energy is released in the form of one photon. The rate of spontaneous emission for the frequency range between ν and $\nu + d\nu$ is given by

$$W_{sp}(\nu)\, d\nu = A_{21} \hat{g}(\nu)\, d\nu. \tag{7.1.5}$$

Here, A_{21} is a constant. Note that $W_{sp}(\nu)$ is independent of the energy density $u(\nu)$ of the light field. The total rate of spontaneous emission can be calculated as

$$W_{sp} = \int_0^\infty W_{sp}(\nu)\, d\nu = A_{21}. \tag{7.1.6}$$

Einstein coefficients

The three constants A_{21}, B_{12}, B_{21} that characterize the three processes are the so-called Einstein coefficients. In the following, we show that the Einstein coefficients are not independent of each other but directly proportional to another. For this purpose, we consider an ensemble of identical atomic systems in thermal equilibrium that interacts with the blackbody radiation.

The energy density of the blackbody radiation is given by the Planck-formula:

$$u(\nu) = \frac{8\pi}{c^3} \frac{h\nu^3}{\exp(h\nu/k_B T) - 1}. \tag{7.1.7}$$

In thermal equilibrium, the population densities N_1 and N_2 of the energy levels $|1\rangle$ and $|2\rangle$, respectively, are governed by the Boltzmann distribution:

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \exp\left(-h\nu/k_B T\right). \tag{7.1.8}$$

Here, g_1 and g_2 are the degeneracy factors of the two levels, k_B is the Boltzmann constant, and T is the temperature of the ensemble. The rate of decrease of the population density N_1 due to absorption of a photon with frequency in the range between ν and $\nu + \Delta\nu$ is given by:

$$\frac{dN_1}{dt} = -N_1 W_{12}(\nu)\, d\nu = -N_1 B_{12} u(\nu) \hat{g}(\nu)\, d\nu. \tag{7.1.9}$$

The corresponding rate of decrease of the population density N_2 due to emission of a photon with frequency in the range between ν and $\nu + \Delta\nu$ reads:

$$\frac{dN_2}{dt} = -N_2 \left[W_{21}(\nu) + W_{sp}(\nu)\right] d\nu = -N_2 \left[B_{21} u(\nu) \hat{g}(\nu) + A_{21} \hat{g}(\nu)\right] d\nu. \tag{7.1.10}$$

The principle of detailed balance demands that these two rates are equal:

$$N_2 \left[W_{21}(\nu) + W_{sp}(\nu)\right] = N_1 W_{12}(\nu). \tag{7.1.11}$$

Rearranging the last equation yields:

$$\frac{N_2}{N_1} = \frac{W_{12}(\nu)}{W_{21}(\nu) + W_{sp}(\nu)} = \frac{B_{12}\, u(\nu)}{B_{21}\, u(\nu) + A_{21}}. \tag{7.1.12}$$

Solving for $u(\nu)$, we obtain:

$$u(\nu) = \frac{A_{21}/B_{21}}{(g_1 B_{12}/g_2 B_{21}) \exp\left(h\nu/k_B T\right) - 1}. \tag{7.1.13}$$

The comparison of equations (7.1.7) and (7.1.13) yields:

$$\frac{A_{21}}{B_{21}} = \frac{8\pi h\nu^3}{c^3} \tag{7.1.14}$$

and

$$g_1 B_{12} = g_2 B_{21}. \tag{7.1.15}$$

Equations (7.1.14) and (7.1.15) have been derived for an ensemble of two-level systems in thermal equilibrium with the blackbody radiation. Since the properties of the two-level system do not depend on the thermodynamic state, these two equations are generally valid, e.g., also in the case of a laser.

Next, we consider a two-level system in its excited state $|2\rangle$ in the absence of radiation ($u(\nu) = 0$). Spontaneous emission leads to transitions from state $|2\rangle$ to state $|1\rangle$. The so-called spontaneous radiative lifetime of state $|2\rangle$ is given by

$$\tau_{sp} = \frac{1}{W_{sp}} = \frac{1}{A_{21}}. \tag{7.1.16}$$

We thus find:

$$W_{sp}(\nu) = \frac{1}{\tau_{sp}}\hat{g}(\nu), \tag{7.1.17}$$

$$W_{21}(\nu) = \frac{c^3}{8\pi h\nu^3\tau_{sp}}u(\nu)\hat{g}(\nu), \tag{7.1.18}$$

and

$$W_{12}(\nu) = \frac{g_2}{g_1}W_{21}(\nu). \tag{7.1.19}$$

Finally, we note that the Einstein coefficients do not solely characterize intrinsic properties of the two-level system. They are also influenced by the environment. For example, the rate of spontaneous emission can be increased by placing the two-level system in a resonant cavity (Purcell effect).

7.1.2 Transition cross sections

The transition rates can be conveniently expressed in terms of transition cross sections. We define the stimulated emission cross section, $\sigma_e(\nu)$, through

$$W_{21}(\nu) = \frac{I(\nu)}{h\nu}\sigma_{\mathrm{e}}(\nu) \tag{7.1.20}$$

with

$$I(\nu) = cu(\nu). \tag{7.1.21}$$

A comparison with equation (7.1.18) yields:

$$\sigma_{\mathrm{e}}(\nu) = \frac{c^2}{8\pi\nu^2\tau_{sp}}\hat{g}(\nu). \tag{7.1.22}$$

The absorption cross-section is defined accordingly through

$$W_{12}(\nu) = \frac{I(\nu)}{h\nu}\sigma_{\mathrm{a}}(\nu) \tag{7.1.23}$$

with (see equation (7.1.19))

$$\sigma_{\mathrm{a}}(\nu) = \frac{g_2}{g_1}\sigma_e(\nu). \tag{7.1.24}$$

7.2 Laser amplifier

In this section, we investigate under which conditions an electromagnetic wave can be coherently amplified by a laser medium. We assume that the laser medium can be described by an ensemble of identical two-level systems. As shown before, both stimulated emission and spontaneous emission increase the number of photons. However, only stimulated emission is a desired process for a laser amplifier as it leads to "photon cloning". In contrast, the spontaneously emitted photons and the incident photons are not correlated. Hence, spontaneous emission acts as an unwanted source of noise. It is neglected in the following discussion.

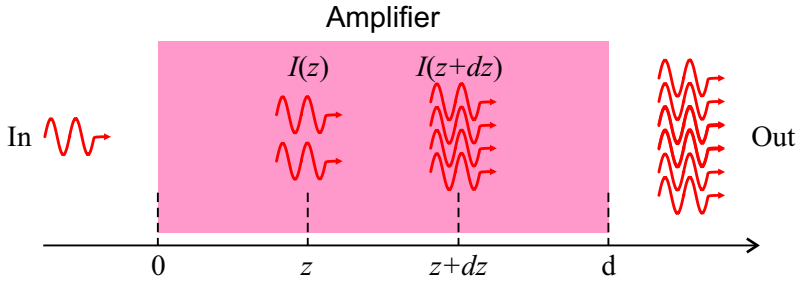


Figure 7.4: Amplification of an optical wave by a laser amplifier.

Let $I(z)$ be the intensity of a monochromatic wave at the position z inside of the laser medium. After propagation by a distance dz, the intensity changes by an amount

$$dI = I(z+dz) - I(z) = h\nu \left(N_2 W_{21} - N_1 W_{12}\right) dz. \tag{7.2.1}$$

Here, the first term on the right hand-side accounts for the photons gained between z and $z+dz$ by stimulated emission while the second term describes the corresponding absorption losses.

With equation (7.1.19), we obtain:

$$\frac{dI}{dz} = \left[N_2 - \frac{g_2}{g_1} N_1\right] \sigma_e(\nu) I(z) = \gamma(\nu) I(z). \tag{7.2.2}$$

Here, we have introduced in the last step the gain coefficient

$$\gamma(\nu) = \left[N_2 - \frac{g_2}{g_1} N_1\right] \sigma_e(\nu) = \left[N_2 - \frac{g_2}{g_1} N_1\right] \frac{c^2}{8\pi\nu^2 \tau_{sp}} \hat{g}(\nu). \tag{7.2.3}$$

Integration of equation (7.2.2) yields:

$$I(z) = I(0) \, \exp\left(\gamma(\nu) z\right). \tag{7.2.4}$$

Here, we have tacitly assumed that $\left[N_2 - \frac{g_2}{g_1} N_1\right]$ is constant. The intensity of the wave increases if the gain coefficient is positive. Obviously, this requires a positive population difference

$$N = \left[N_2 - \frac{g_2}{g_1} N_1\right]. \tag{7.2.5}$$

Example: Gain coefficient of a ruby laser amplifier

Ruby (Al_2O_3 doped with Cr^{3+})
$\nu_0 = 4.326 \cdot 10^{14}\mathrm{H}z \Rightarrow \lambda = 694.3\mathrm{n}m$
$\tau_{sp} = 3\mathrm{m}s$
n=1.78
$g(\nu_0) \approx 1/\Delta\nu = 5 \cdot 10^{-12}\mathrm{s}$
$\left(N_2 - \frac{g_2}{g_1} N_1\right) = 5 \cdot 10^{17}\mathrm{c}m^{-3}$ (typical value for flashlamp pumping).

$\Rightarrow \gamma(\nu_0) = N \frac{c^2}{8\pi\nu^2\tau_{sp}} \hat{g}(\nu_0) = 0.05\,\mathrm{c}m^{-1}$

7.2.1 Amplifier pumping

Amplification of a wave in a laser medium requires population inversion, i.e., the population difference, $N = \left[N_2 - \frac{g_2}{g_1} N_1\right]$, must be positive. This requires an external energy source to "pump" the two-level systems from state $|1\rangle$ to $|2\rangle$. Depending on the laser medium, population inversion may be achieved either by optical pumping (flashlamp, laser) or by electrically pumping (gas discharge, charge carrier injection).

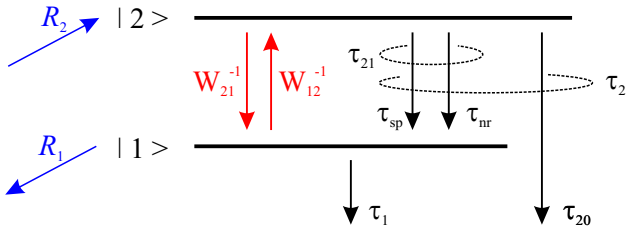


Figure 7.5: Energy levels $|1\rangle$ and $|2\rangle$ and their decay times. Pumping increases (decreases) the population density of $|2\rangle$ ($|1\rangle$) at the rate $\mathcal{R}_2$ ($\mathcal{R}_1$).

In the following, we derive the so-called rate equations of a laser amplifier for continuous-wave (CW) operation. For this purpose, we analyze the population densities, N_1 and N_2, for a steady-state situation (see Fig. 7.5). The degeneracies of these two levels are assumed to be equal ($g_1 = g_2$). The population densities of additional higher or lower energy states are not explicitly considered. However, these level play an important role for the operation of a laser amplifier. Energy states with energies $E_h > E_2$ are required for the pumping process. Furthermore, the levels $|1\rangle$ and $|2\rangle$ may radiatively or nonradiatively decay to states with energies $E_l < E_1$.

The total decay rate of state $|2\rangle$ in the absence of a light field is given by:

$$\frac{1}{\tau_2} = \frac{1}{\tau_{21}} + \frac{1}{\tau_{20}}, \tag{7.2.6}$$

where τ_{21} is the lifetime due to transitions $|2\rangle \rightarrow |1\rangle$, and τ_{20} is associated with all transitions from $|2\rangle$ to states with energies $E_l < E_1$.

The decay rate $1/\tau_{21}$ can be written as the sum of a radiative and a nonradiative decay rate:

$$\frac{1}{\tau_{21}} = \frac{1}{\tau_{sp}} + \frac{1}{\tau_{nr}}. \tag{7.2.7}$$

Rate equations in the absence of amplifier radiation

First, we consider the amplifier in the absence of radiation ($u(\nu) = 0$). The temporal variation of the population densities follows from

$$\frac{dN_2}{dt} = \mathcal{R}_2 - \frac{N_2}{\tau_2}, \tag{7.2.8}$$

$$\frac{dN_1}{dt} = -\mathcal{R}_1 - \frac{N_1}{\tau_1} + \frac{N_2}{\tau_{21}}. \tag{7.2.9}$$

Here, $\mathcal{R}_2$ is the pumping rate of state $|2\rangle$ and $\mathcal{R}_1$ is the depumping rate of state $|1\rangle$. Under steady state conditions ($\frac{dN_2}{dt} = \frac{dN_1}{dt} = 0$), we obtain the following population difference in the absence of amplifier radiation:

$$N_0 = (N_2 - N_1) = \mathcal{R}_2\tau_2 \left(1 - \frac{\tau_1}{\tau_{21}}\right) + \mathcal{R}_1\tau_1. \tag{7.2.10}$$

Rate equations in the presence of amplifier radiation

In the presence of a radiation field $u(\nu)$, the rate equations become:

$$\frac{dN_2}{dt} = \mathcal{R}_2 - \frac{N_2}{\tau_2} - N_2W_e + N_1W_e, \tag{7.2.11}$$

$$\frac{dN_1}{dt} = -\mathcal{R}_1 - \frac{N_1}{\tau_1} + \frac{N_2}{\tau_{21}} + N_2W_e - N_1W_e. \tag{7.2.12}$$

Here, we have have set $W_{12} = W_{21} = W_\mathrm{e}$. Under steady state conditions ($\frac{dN_2}{dt} = \frac{dN_1}{dt} = 0$), we find:

$$N = N_2 - N_1 = \frac{N_0}{1 + \tau_s W_\mathrm{e}}, \tag{7.2.13}$$

with the saturation time constant

$$\tau_\mathrm{s} = \tau_2 + \tau_1 \left(1 - \frac{\tau_2}{\tau_{21}}\right). \tag{7.2.14}$$

We find that the radiation field depletes the population difference!

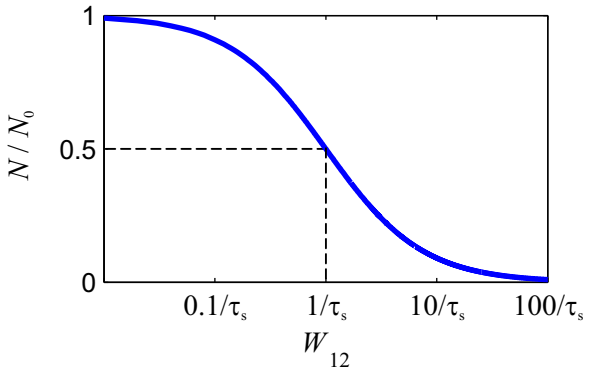


Figure 7.6: Depletion of the population difference N as a function of the rate of stimulated emission W_e.

Four-level pumping

The four-level pumping scheme (see Fig. 7.7) is a good approximation for many real-world laser materials, e.g. Nd:YAG crystals. Atoms are excited at a rate R from the ground state $|0\rangle$ to a high energy state $|3\rangle$. From this state, the atoms undergo a quick nonradiative transition to the long-lived upper laser state $|2\rangle$. Because of the short lifetime of state $|3\rangle$, we can set the pumping rate of level $|2\rangle$ to $R_2 = R$. We neglect pumping in or out of level $|1\rangle$, i.e., we set $\mathcal{R}_1 = 0$.

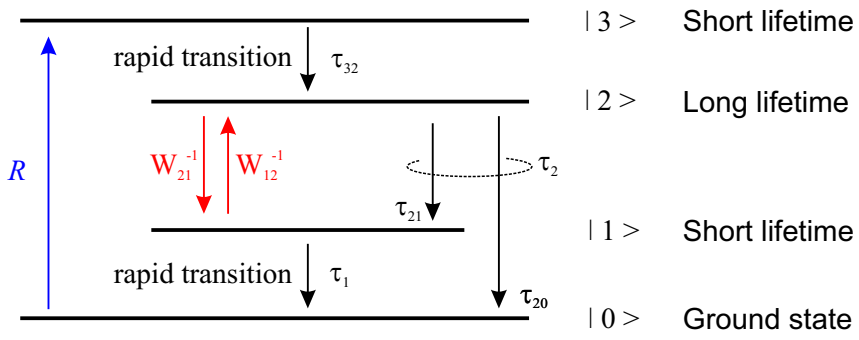


Figure 7.7: Level scheme for a four-level system.

Applying our previous considerations to the four-level pumping scheme, we obtain the following steady state population difference in the absence of amplifier radiation:

$$N_0 = (N_2 - N_1) = \mathcal{R}\tau_2 \left(1 - \frac{\tau_1}{\tau_{21}}\right). \tag{7.2.15}$$

In a typical four-level laser, the following conditions hold: $\tau_{sp} \ll \tau_{nr}$, and $\tau_{20} \gg \tau_{sp} \gg \tau_1$.

We thus obtain:

$$N_0 \approx \mathcal{R}\tau_{sp}, \tag{7.2.16}$$

and

$$\tau_{\mathrm{s}} \approx \tau_{sp}. \tag{7.2.17}$$

The corresponding steady state population difference in the presence of amplifier radiation becomes:

$$N \approx \frac{\mathcal{R}\tau_{sp}}{1 + \tau_{sp}W_{\mathrm{e}}}. \tag{7.2.18}$$

So far we have implicitly assumed that the pumping rate $\mathcal{R}$ is independent of the population difference N. The following discussion shows that this assumption breaks down for large N. The total density of atoms is given by:

$$N_a = N_g + N_1 + N_2 + N_3. \tag{7.2.19}$$

In the case of optical pumping, the pumping rate $\mathcal{R}$ is given by

$$\mathcal{R} = (N_g - N_3)W, \tag{7.2.20}$$

where W is the corresponding transition probability. Since the states $|1\rangle$ and $|3\rangle$ are short-lived, we can set

$$N_1 \approx N_3 = 0. \tag{7.2.21}$$

The pump rate can thus be written as

$$\mathcal{R} \approx (N_a - N)W \tag{7.2.22}$$

and equation (7.2.18) becomes

$$N \approx \frac{\tau_{sp}N_aW}{1 + \tau_{sp}W + \tau_{sp}W_{\mathrm{e}}}. \tag{7.2.23}$$

This equation shows that the population difference N saturates for strong pumping ($W \gg 1/\tau_{sp}$).

7.2.2 Amplifier nonlinearity

We have seen in the last section that stimulated emission leads to a depletion of the population difference:

$$N = \frac{N_0}{1 + \tau_s W_e}. \tag{7.2.24}$$

In terms of the radiation intensity $I(z,\nu)$, this equation can be rewritten as

$$N = \frac{N_0}{1 + I(z,\nu)/I_{\mathrm{s}}(\nu)}, \tag{7.2.25}$$

with the saturation intensity

$$I_{\mathrm{s}}(\nu) = \frac{h\nu}{\tau_{\mathrm{s}}\sigma_e(\nu)}. \tag{7.2.26}$$

Substituting equation (7.2.25) in equation (7.2.3) yields the the saturated gain coefficient

$$\gamma(\nu) = \frac{\gamma_0(\nu)}{1 + I(z,\nu)/I_s(\nu)}, \tag{7.2.27}$$

where

$$\gamma_0(\nu) = N_0\sigma_e(\nu) = N_0\frac{c^2}{8\pi\nu^2\tau_{\mathrm{sp}}}\hat{g}(\nu) \tag{7.2.28}$$

is the so-called small-signal gain coefficient. The variation of the intensity inside of the laser amplifier is governed by

$$\frac{dI}{dz} = \gamma(\nu)I(z,\nu) = \frac{\gamma_0(\nu)I(z,\nu)}{1 + I(z,\nu)/I_s(\nu)}. \tag{7.2.29}$$

In general, this nonlinear equation can only be solved numerically. However, for very small or very large input intensities $I(0,\nu)$, we can give the following approximate solutions:

- Case 1 - Small input signal

  For $I(z,\nu)/I_s(\nu) \ll 1$, equation (7.2.29) can be approximated by

  $$\frac{dI}{dz} \approx \gamma_0(\nu)I(z,\nu). \tag{7.2.30}$$

  The corresponding solution is

  $$I(d,\nu) = I(0,\nu)\,\exp\left[\gamma_0(\nu)d\right]. \tag{7.2.31}$$

- Case 2 - Large input signal

  For $I(0,\nu)/I_s(\nu) \gg 1$, equation (7.2.29) becomes

  $$\frac{dI}{dz} \approx \gamma_0(\nu)I_s(\nu). \tag{7.2.32}$$

  In this case, $I(d,\nu)$ is given by

  $$I(d,\nu) = I(0,\nu) + \gamma_0(\nu)I_s(\nu)d. \tag{7.2.33}$$

7.3 Laser oscillator

In the previous section we have seen that a laser amplifier operated under population inversion conditions can be used to amplify an incident optical wave. By adding an optical resonator as a feedback element to the laser amplifier, we obtain a laser oscillator. In the following, we investigate the conditions under which "lasing" occurs.

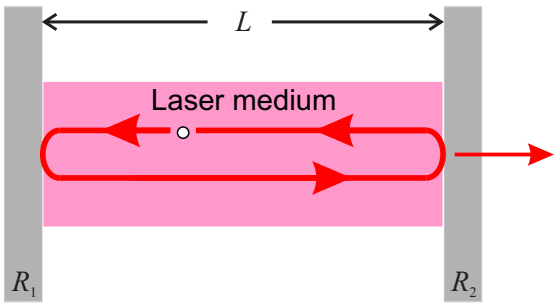


Figure 7.8: Scheme of a simple laser oscillator.

7.3.1 Threshold condition

For reasons of simplicity, we assume that the laser cavity is formed by two planar mirrors (Fabry-Perot resonator) with reflectivities R_1 and R_2 which are separated by a distance L. The laser cavity is filled with the laser medium. Losses due to scattering or absorption in the laser medium are characterized by the attenuation coefficient $\alpha(\nu)$ (loss per unit length).

Let $I(0,\nu)$ be the intensity of a very weak, monochromatic light field at some point in the laser medium. After one complete round-trip in the resonator, the intensity becomes:

$$I(2L,\nu) = R_1 R_2 I(0,\nu) \exp\left([\gamma_0(\nu) - \alpha(\nu)]2L\right), \tag{7.3.1}$$

where γ_0 is the small-signal gain coefficient of the laser medium.

For $I(2L,\nu) > I(0,\nu)$, the intensity will initially grow exponentially in time and laser oscillations set in. The threshold condition for lasing is thus given by:

$$R_1 R_2 \exp\left([\gamma_0(\nu) - \alpha(\nu)]2L\right) \geq 1. \tag{7.3.2}$$

The threshold gain coefficient $\gamma_{thr}(\nu)$ is the smallest gain coefficient that satisfies the threshold condition. Its value is given by

$$\gamma_{thr}(\nu) = \alpha(\nu) - \frac{1}{2L}\ln(R_1 R_2). \tag{7.3.3}$$

The corresponding threshold population inversion follows from equation (7.2.28):

$$N_{thr} = \frac{8\pi\tau_{sp}}{c^2}\frac{\nu^2}{\hat{g}(\nu)}\gamma_{thr}(\nu) = \frac{8\pi\tau_{sp}}{c^2}\frac{\nu^2}{\hat{g}(\nu)}\left(\alpha(\nu) - \frac{1}{2L}\ln(R_1R_2)\right). \tag{7.3.4}$$

For a typical four-level laser, the required pump rate is given by (see equation (7.2.16)):

$$\mathcal{R}_{thr} = \frac{N_{thr}}{\tau_{sp}}. \tag{7.3.5}$$

7.3.2 Pump-power dependence

In the following, we consider the characteristics of a typical four-level laser for different pumping rates $\mathcal{R}$.

- For $\mathcal{R} < \mathcal{R}_{thr}$, the gain is not sufficient to overcome the losses. Hence, laser oscillations do not occur ($I(\nu) = 0$) and the output of the laser is dominated by spontaneous emission. In this regime, the population inversion increases linearly with the pumping rate (see equation (7.2.16)):

  $$N_0 = R\tau_{sp}. \tag{7.3.6}$$

- At the threshold ($\mathcal{R} = \mathcal{R}_{thr}$), the gain just compensates the losses. According to equation (7.3.4), the population inversion is given by:

  $$N_{thr} = \mathcal{R}_{thr}\tau_{sp} = \frac{8\pi\tau_{sp}}{c^2}\frac{\nu^2}{\hat{g}(\nu)}\gamma_{thr}(\nu). \tag{7.3.7}$$

- Above threshold ($\mathcal{R} > \mathcal{R}_{thr}$), the laser sustains oscillations and its output is dominated by stimulated emission.

  When switching on the laser, the population difference N_0 is initially larger than the threshold inversion and the intensity grows exponentially. As the intensity increases, saturation effects become important and the gain coefficient decreases. The steady-state condition is reached ($I(2L,\nu) = I(0,\nu)$), when the saturated gain becomes equal to the distributed loss of the laser oscillator:

  $$\exp\left[\gamma(\nu)2L\right] = R_1R_2\exp\left[\alpha(\nu)2L\right] \tag{7.3.8}$$

  or

  $$\gamma(\nu) = \gamma_{thr}(\nu). \tag{7.3.9}$$

The corresponding light intensity in the laser medium is given by:

$$I(\nu) = I_s \left(\frac{\gamma_0(\nu)}{\gamma_{thr}(\nu)} - 1 \right) = I_s \left(\frac{N_0}{N_{thr}} - 1 \right) = I_s \left(\frac{\mathcal{R}}{\mathcal{R}_{thr}} - 1 \right). \tag{7.3.10}$$

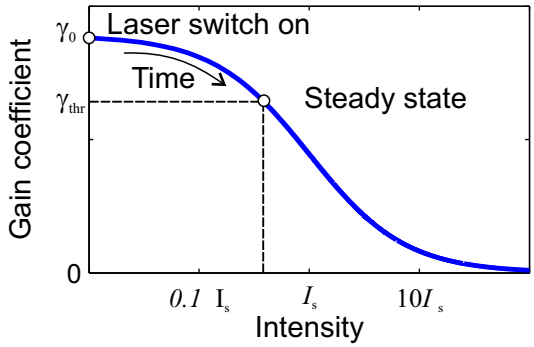


Figure 7.9: Gain coefficient of a laser oscillator.

Note that stronger pumping does not increase the steady-state population difference N. Above threshold, N is "clamped" to the corresponding threshold value:

$$N = N_{thr}. \tag{7.3.11}$$

Figure (7.10) summarizes the results of this section.

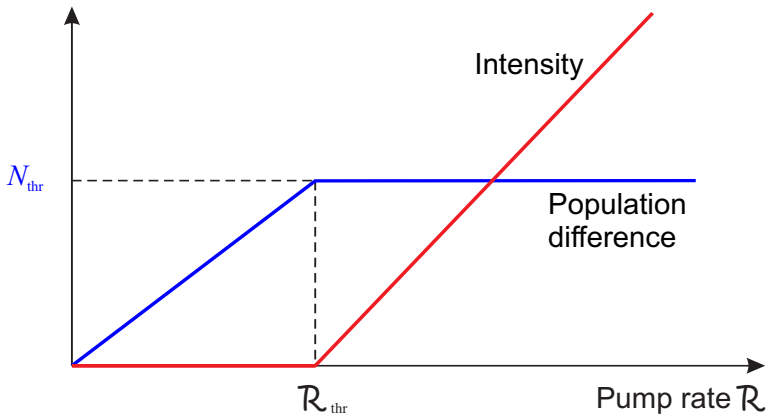


Figure 7.10: Population difference (blue curve) and laser intensity (red curve) as a function of the pumping rate.

7.4 Some types of lasers

7.4.1 Helium-neon laser

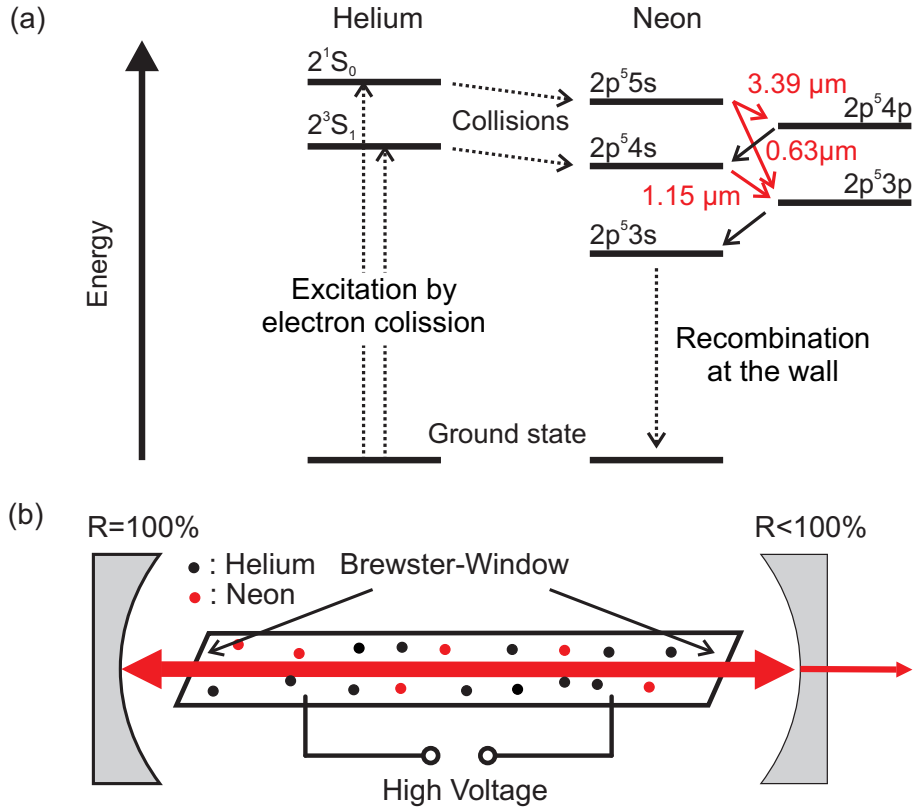


Figure 7.11: (a) Energy level diagram of a HeNe laser. (b) Schematic diagram of a HeNe laser.

- First publication: A. Javan *et* al., Phys. Rev. Lett. **6**, 106 (1961).
- The HeNe laser is a gas laser whose gain medium consists of a mixture of He and Ne. The He atoms are excited by inelastic electron collisions from the ground state to long-lived metastable states. The energy is then efficiently transfered to the Ne atoms by collisions.
- Typical parameters:
  - Gas pressure: 100 Pa
  - Partialpressure He/Ne: 10/1
  - Ignition Voltage: 10-15 kV

– Discharge Voltage: 1-2 kV

- Applications: Laser interferometer, alignment laser

7.4.2 Ti:sapphir laser

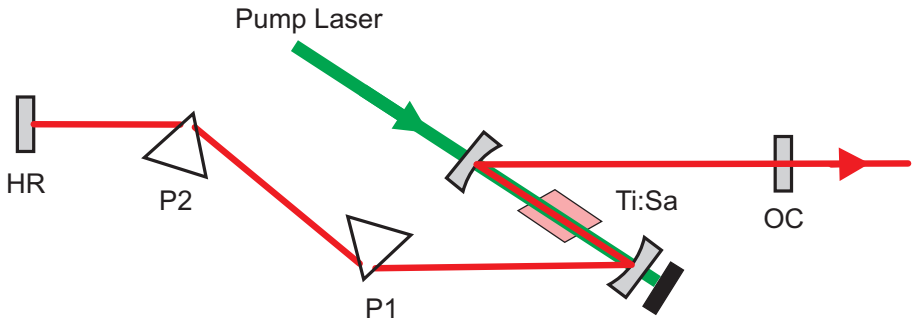


Figure 7.12: Schematic diagram of a fs-Ti:Sa-Laser. HR: High reflector ($R \approx 100\%$), OC: Output coupler ($R \approx 98\%$). The prisms P1 and P2 are used to control the dispersion of the cavity.

- First publication: P. F. Moulton, J. Opt. Soc. Am. B **3**, 125 (1986).

- The Ti:Sa laser is a solid state laser which uses a crystal of sapphire (Al2O3) that is doped with Ti^{3+}-ions as the laser medium. The energy levels of the Ti^{3+}-ions are spectrally broadened by crystal field effects and coupling to phonons. This results in a very large gain bandwidth ranging from 670 nm to 1100 nm. Ti:Sa laser are normally pumped with a continuous wave Argon laser or a frequency-doubled Nd:YVO_4 laser.

- Applications: Generation of ultrashort laser pulses for time-resolved spectroscopy.

7.4.3 Semiconductor laser diodes

- First publication: R. N. Hall *et* al., Phys. Rev. Lett. **9**, 366 (1962).

- A homojunction laser diode consists of a highly doped p-n junction. Applying a forward electrical bias causes the injection of holes and electrons from opposite sides of the p-n junction into the depletion region. Here, the electrons and holes recombine and generate photons. The corresponding emission frequency is determined by the size of the electronic band gap. Often, the semiconductor crystal is cleaved in such a way that two parallel facets of the crystal can be used as the mirrors of a simple Fabry-Perot cavity. Homojunction laser diodes are very inefficient as they require

large current densities because of the large volume of the active region. Hence, they can be only operated in pulsed mode.

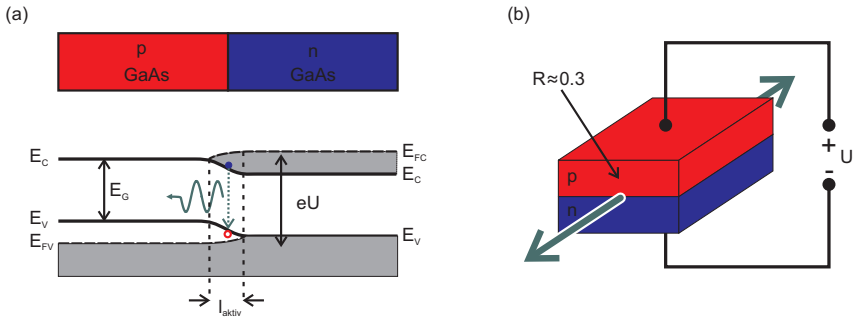


Figure 7.13: (a) Band structure of a homojunction laser diode. (b) Edge emitter: The mirrors of the cavity are defined by the facets of the semiconductor crystal.

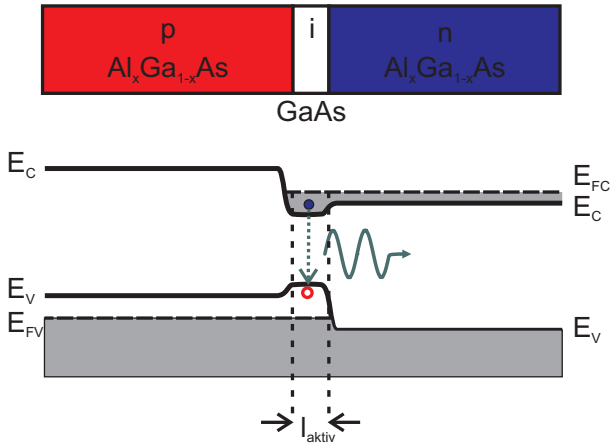


Figure 7.14: Band structure of a double-heterojunction laser diode.

- In a double-heterojunction laser diode, a thin layer of a low band gap material (e.g. galium arsenid) is sandwiched between two high band gap layers (e.g. aluminum galium arsenid). This results in a better confinement of the carriers in the active region. Furthermore, the refractive index profile of the double-heterojunction causes a confinement of the optical field to the active region. Both effects reduce the required current density such that continuous wave operation of double-heterojunction laser diodes at room temperature is possible.

- Applications: Telecommunications, pump laser, DVD player, ...

7.5 Generation of laser pulses

7.5.1 Gain Switching

Idea: Generate laser pulses by temporal variation of the pump rate.

- For $t < 0$, the pump rate $\mathcal{R}(t) = \mathcal{R}_a$ is smaller than the threshold pump rate $\mathcal{R}_{thr}$. The population difference $N(t) = N_a$ is not sufficient to start lasing ($I(t) = 0$).
- At $t = 0$, the pump rate is increased above the threshold pump rate ($\mathcal{R}(t) = \mathcal{R}_b > \mathcal{R}_{thr}$). As a consequence, the population difference $N(t)$ increases with time.
- As soon as the population difference $N(t)$ becomes larger as the threshold population difference N_{thr}, laser emission starts and the light intensity $I(t)$ increases.
- Stimulated emission reduces the population difference until the steady state condition $N(t) = N_{thr}$ is met.
- At $t = t_2$, the pump rate is reduced to its initial vale $\mathcal{R}_a$. The population difference decreases and the laser action stops.

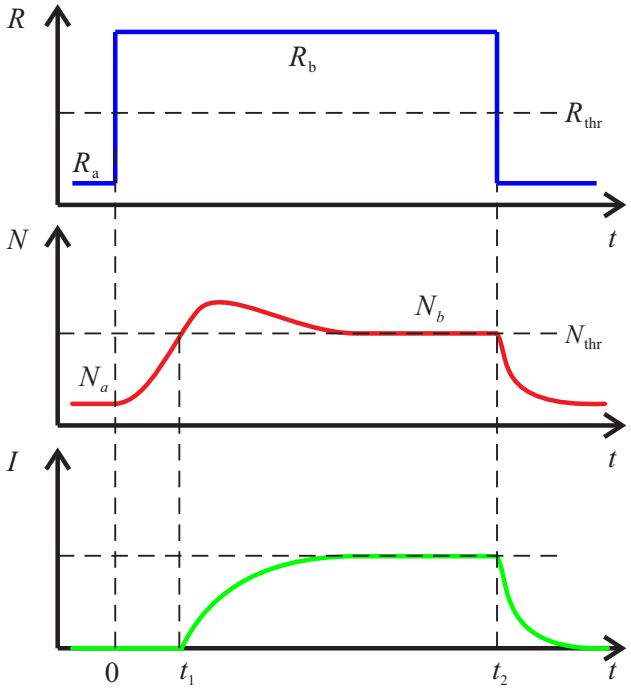


Figure 7.15: Gain modulation: Population difference $N(t)$ and light intensity $I(t)$ as a function of a variable pump rate $\mathcal{R}(t)$.

7.5.2 Q-switching

Idea: Generate laser pulses by modulating the absorption losses of the cavity.

- For $t < t_1$, the loss coefficient of the cavity $\alpha(t) = \alpha_a$ is so large that the pump rate $\mathcal{R}$ is not sufficient to reach the threshold ($N(t) = N_a < N_{thr,a} \Rightarrow I(t) = 0$).
- At $t = t_1$, the losses of the cavity are reduced. The population inversion $N(t_1) = N_a$ is now larger than the new threshold population difference $N_{thr,b}$ and the laser emission starts.
- The light intensity $I(t)$ increases quickly and, as a consequence, the population difference $N(t)$ decreases.
- As soon as $N(t)$ becomes smaller as the threshold population difference $N_{thr,b}$, lasing stops and the light intensity $I(t)$ falls off to zero.
- At $t = t_2$, the losses of the cavity are increased to the initial value ($\alpha(t) = \alpha_a$). The population difference recovers and tends towards the initial value ($N(t) \to N_a$).

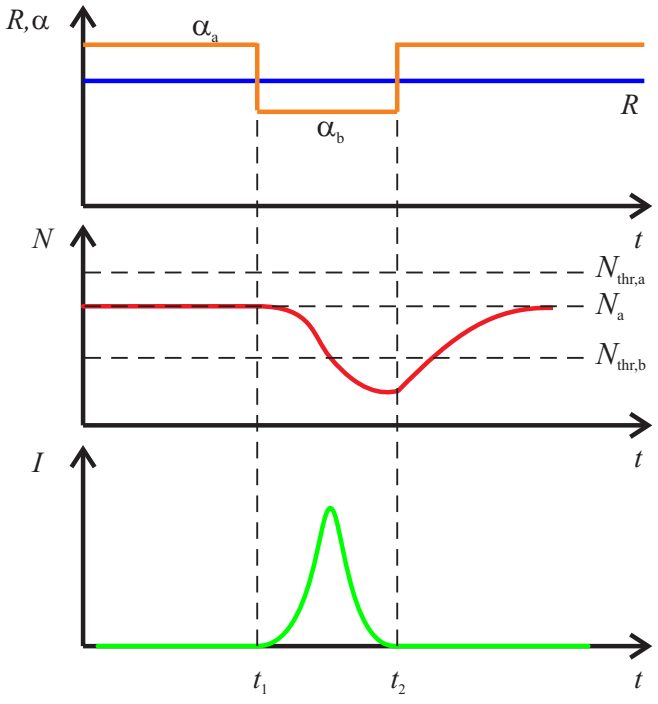


Figure 7.16: Q-switching: Population difference $N(t)$ and light intensity $I(t)$ as a function of a variable cavity absorption coefficient $\alpha(t)$.

7.5.3 Mode locking

In section (7.3), we have derived the threshold condition of the laser for a single mode of the cavity. For gain media with large spectral band width, several cavity modes might fulfill the threshold condition (7.3.2) and start to lase. The characteristics of the laser output depends on the phase relationship between the different modes:

- **Continuous wave (CW)**: If the different modes have no fixed phase relationship, constructive and destructive interference maintain balance and the output intensity of the laser exhibits only small temporal fluctuations around the average value.
- **Mode-locking**: If the the phase of the different modes are locked, interference of the modes creates a train of laser pulses.

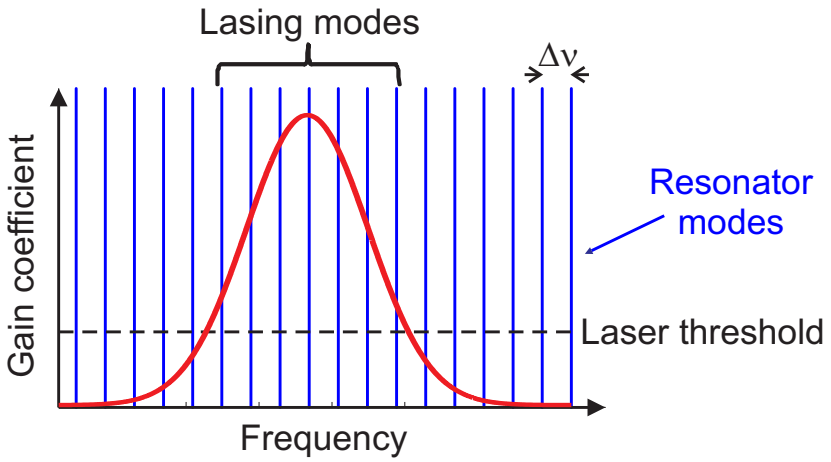


Figure 7.17: Gain medium with large gain bandwidth.

In what follows, we discuss a simple model for mode-locking. We consider $2N+1$-modes with frequency $\nu_m = \nu_0 + m\Delta\nu, m = -N, -N+1, \cdots, N-1, N$, phase difference $\Delta\varphi = 0$, and equal amplitude E_0.

The total electric field strength is given by the superposition of the electric field strengths of the different modes:

$$\begin{aligned} E(t) &= \sum_{m=-N}^{N} E_0\, e^{i2\pi(\nu_0+m\Delta\nu)t} \\ &= E_0\, e^{i2\pi\nu_0 t}\, e^{-i2\pi N\Delta\nu t} \sum_{m=0}^{2N} e^{i2\pi m\Delta\nu t} \\ &= E_0\, e^{i2\pi\nu_0 t}\, e^{-i2\pi N\Delta\nu t}\, \frac{1-e^{i2\pi(2N+1)\Delta\nu t}}{1-e^{i2\pi\Delta\nu t}} \\ &= E_0\, e^{i2\pi\nu_0 t}\, \frac{e^{-i\pi(2N+1)\Delta\nu t}}{e^{-i\pi\Delta\nu t}}\, \frac{e^{i\pi(2N+1)\Delta\nu t}}{e^{i\pi\Delta\nu t}}\, \frac{e^{-i\pi(2N+1)\Delta\nu t}-e^{i\pi(2N+1)\Delta\nu t}}{e^{-i\pi\Delta\nu t}-e^{i\pi\Delta\nu t}} \\ &= E_0\, e^{i2\pi\nu_0 t}\, \frac{\sin\left[(2N+1)\pi\Delta\nu t\right]}{\sin\left[\pi\Delta\nu t\right]}. \end{aligned} \tag{7.5.1}$$

The corresponding intensity is:

$$I(t) = I_0 \frac{\sin^2\left[(2N+1)\pi\Delta\nu t\right]}{\sin^2\left[\pi\Delta\nu t\right]}, \tag{7.5.2}$$

where I_0 is the intensity of a single mode. The time lag τ_r between two consecutive pulses is given by:

$$\tau_r = \frac{1}{\Delta\nu} = \frac{2L}{c}. \tag{7.5.3}$$

Vividly, τ_r is the time that a pulse needs to complete one round trip in the cavity. The pulse length τ_p is reciprocal to the spectral width of the pulse:

$$\tau_p = \frac{1}{(2N+1)\Delta\nu}. \tag{7.5.4}$$

The peak intensity of the pulse is given by:

$$I_{max} = (2N+1)^2\, I_0. \tag{7.5.5}$$

For comparison, the intensity of a comparable CW-laser is $I_{CW} = (2N+1)\, I_0$.

Methods for mode-locking:

- Active mode-locking: The cavity absorption coefficient is periodically modulated with the modulation frequency $\nu_{mod} = \Delta\nu$. This can be achieved by adding, e.g., an acousto-optic modulator to the cavity. The periodic variation of the amplitude

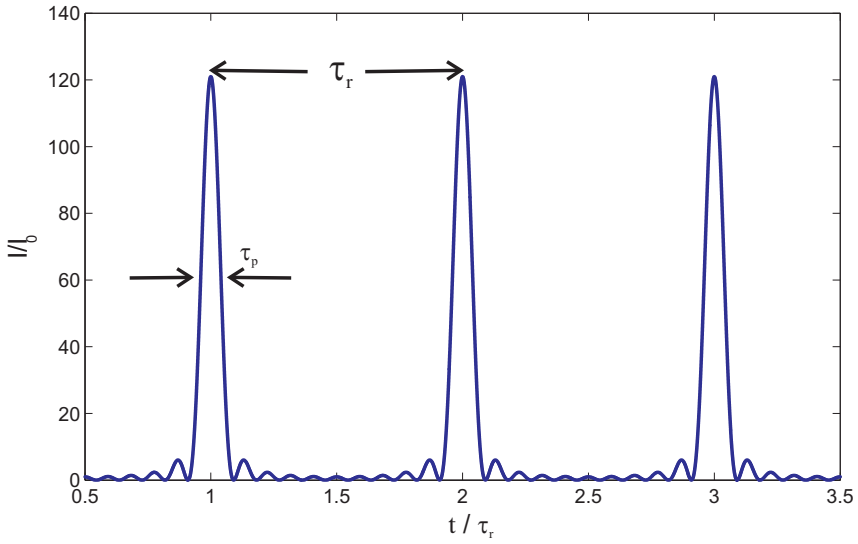


Figure 7.18: Intensity of a periodic pulse train for 11 phase-locked modes with equal amplitude.

of the mode with frequency ν leads to side bands with frequencies $\nu \pm \Delta\nu$. These sidebands are phase-locked with the center frequency. Modulation of the side bands creates new phase-locked side-bands with frequencies $\nu \pm 2\Delta\nu$, and so on.

Active-mode locking can be intuitively understood in the time domain. A short pulses travels in the cavity with the round-trip time $\tau_r = 1/\Delta\nu$ back and forth. The modulator "unlocks" the cavity whenever the pulse arrives at the position of the modulator. The pulse "sees" only small losses can be hence amplified in the laser.

- Passive mode-locking: The absorption of the cavity is modulated with a passive element, e.g., a saturable absorber.

8 Optical modulators

8.1 Electro-optic modulators

8.1.1 Electro-optic media

In a linear anisotropic medium, the electric displacement field $\mathbf{D}$ and the electric field strength $\mathbf{E}$ are related to each other through the electric permeability tensor $\boldsymbol{\epsilon}$:

$$\mathbf{D} = \epsilon_0 \boldsymbol{\epsilon} \mathbf{E}. \tag{8.1.1}$$

In the following, we assume that the electric permittivity tensor is symmetric: $\epsilon_{ij} = \epsilon_{ji}$. We use for our analysis the principal coordinate system, in which the electric permittivity tensor becomes diagonal:

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_{11} & 0 & 0 \\ 0 & \epsilon_{22} & 0 \\ 0 & 0 & \epsilon_{33} \end{pmatrix}. \tag{8.1.2}$$

With the electric impermeability tensor

$$\boldsymbol{\eta} = \boldsymbol{\epsilon}^{-1} \tag{8.1.3}$$

we can rewrite equation (8.1.1) as

$$\mathbf{E} = \frac{\boldsymbol{\eta}}{\epsilon_0} \mathbf{D}. \tag{8.1.4}$$

In the principal coordinate system, $\boldsymbol{\eta}$ is also diagonal and symmetric:

$$\eta_{ii} = \frac{1}{\epsilon_{ii}} = \frac{1}{n_i^2}. \tag{8.1.5}$$

The index ellipsoid is the quadratic representation of the electric impermeability tensor:

$$\sum_{i,j} \eta_{ij} \, x_i' x_j' = 1. \tag{8.1.6}$$

This surface does not depend on the particular choice of the coordinate system (x_1', x_2', x_3'). In the the principal coordinate system (x_1, x_2, x_3), the index ellipsoid can be written as:

$$\sum_{ij} \eta_{ij}\, x_i x_j = \frac{x_1^2}{n_1^2} + \frac{x_2^2}{n_2^2} + \frac{x_3^2}{n_3^2} = 1. \tag{8.1.7}$$

The optical properties of so-called electro-optic media can be changed by applying an external electric field $\mathbb{E}$:

$$\eta_{ij}(\mathbb{E}) = \eta_{ij}(0) + \Delta\eta_{ij}(\mathbb{E}). \tag{8.1.8}$$

Here, $\eta_{ij}(0)$ is the element of the electric impermeability tensor without applied field and $\Delta\eta_{ij}(\mathbb{E})$ is a small change caused by $\mathbb{E}$.

Next, we expand $\Delta\eta_{ij}(\mathbb{E})$ as a power series:

$$\Delta\eta_{ij}(\mathbb{E}) = \sum_k \mathfrak{r}_{ijk}\mathbb{E}_k + \sum_{kl} \mathfrak{s}_{ijkl}\mathbb{E}_k\mathbb{E}_l. \tag{8.1.9}$$

The first term describes a linear change of the electric impermeability due to $\mathbb{E}$. This linear electro-optic effect is known as the Pockels effect. The coefficients $\{\mathfrak{r}_{ijk}\}$ are the so-called Pockels coefficients. The quadratic variation of the impermeabilty is usually referred to as the Kerr effect and the coefficients $\{\mathfrak{s}_{ijkl}\}$ are called Kerr coefficients.

8.1.2 Pockels effect

In the following, we concentrate on the Pockels effect:

$$\Delta\eta_{ij}(\mathbb{E}) = \sum_k \mathfrak{r}_{ijk}\mathbb{E}_k. \tag{8.1.10}$$

The index ellipsoid of the crystal is modified by the Pockels effect and reads:

$$\begin{aligned}\sum_{ij} \eta_{ij}\, x_i x_j \;\; = \;\; & \sum_k \left(\frac{1}{n_1^2} + \mathfrak{r}_{11k}\mathbb{E}_k\right) x_1^2 + \left(\frac{1}{n_2^2} + \mathfrak{r}_{22k}\mathbb{E}_k\right) x_2^2 + \left(\frac{1}{n_3^2} + \mathfrak{r}_{33k}\mathbb{E}_k\right) x_3^2 \\ & + 2x_1x_2\mathfrak{r}_{12k}\mathbb{E}_k + 2x_2x_3\mathfrak{r}_{23k}\mathbb{E}_k + 2x_1x_3\mathfrak{r}_{13k}\mathbb{E}_k = 1.\end{aligned} \tag{8.1.11}$$

Because of the symmetry properties of the impermeability tensor ($\eta_{ij} = \eta_{ji}$), the Pockels coefficients have to fulfill the condition

$$\mathfrak{r}_{ijk} = \mathfrak{r}_{jik}. \tag{8.1.12}$$

Hence, the Pockels tensor $\mathfrak{r}$ has at most 6×3 independent elements. The number of independent elements can be further decreased through symmetry properties of the crystal. In particular, all elements vanish for materials with inversion symmetry.

It is convenient to use the following contracted notation:

$$\mathfrak{r}_{hk} = \mathfrak{r}_{ijk} \tag{8.1.13}$$

with

$$h = \begin{cases} 1 & for \quad ij = 11, \\ 2 & for \quad ij = 22, \\ 3 & for \quad ij = 33, \\ 4 & for \quad ij = 23 \quad oder \quad 32, \\ 5 & for \quad ij = 13 \quad oder \quad 31, \\ 6 & for \quad ij = 12 \quad oder \quad 21. \end{cases} \tag{8.1.14}$$

With this new notation, equation (8.1.10) can be rewritten as:

$$\Delta\eta_h = \sum_k \mathfrak{r}_{hk}\mathbb{E}_k. \tag{8.1.15}$$

The corresponding matrix notation reads:

$$\begin{bmatrix} \Delta\eta_1 \\ \Delta\eta_2 \\ \Delta\eta_3 \\ \Delta\eta_4 \\ \Delta\eta_5 \\ \Delta\eta_6 \end{bmatrix} = \begin{bmatrix} \mathfrak{r}_{11} & \mathfrak{r}_{12} & \mathfrak{r}_{13} \\ \mathfrak{r}_{21} & \mathfrak{r}_{22} & \mathfrak{r}_{23} \\ \mathfrak{r}_{31} & \mathfrak{r}_{32} & \mathfrak{r}_{33} \\ \mathfrak{r}_{41} & \mathfrak{r}_{42} & \mathfrak{r}_{43} \\ \mathfrak{r}_{51} & \mathfrak{r}_{52} & \mathfrak{r}_{53} \\ \mathfrak{r}_{61} & \mathfrak{r}_{62} & \mathfrak{r}_{63} \end{bmatrix} \begin{bmatrix} \mathbb{E}_1 \\ \mathbb{E}_2 \\ \mathbb{E}_3 \end{bmatrix} \tag{8.1.16}$$

Example I: $LiNbO_3$

$LiNbO_3$ is a negative uniaxial crystal ($n_e < n_o$) which belongs to the crystallographic point group 3m. We assume that the optical axis is oriented parallel to the x_3-direction. The corresponding Pockels tensor then reads:

$$\mathfrak{r} = \begin{bmatrix} 0 & -\mathfrak{r}_{22} & \mathfrak{r}_{13} \\ 0 & \mathfrak{r}_{22} & \mathfrak{r}_{13} \\ 0 & 0 & \mathfrak{r}_{33} \\ 0 & \mathfrak{r}_{51} & 0 \\ \mathfrak{r}_{51} & 0 & 0 \\ -\mathfrak{r}_{22} & 0 & 0 \end{bmatrix} \tag{8.1.17}$$

If the static field $\mathbb{E}$ is applied along the optical axis (x_3 axis) of the $LiNbO_3$ crystal, the modified index ellipsoid is given by

$$\left(\frac{1}{n_o^2} + \mathfrak{r}_{13}\mathbb{E}\right)\left(x_1^2 + x_2^2\right) + \left(\frac{1}{n_e^2} + \mathfrak{r}_{33}\mathbb{E}\right)x_3^2 = 1. \tag{8.1.18}$$

For this configuration, the Pockels effects does not change the orientation of the prinicpal axes. The oridinary and the extraordinary indices, respectively, can be deduced from

$$\frac{1}{n_o^2(\mathbb{E})} = \frac{1}{n_o^2} + \mathfrak{r}_{13}\mathbb{E}, \tag{8.1.19}$$

$$\frac{1}{n_e^2(\mathbb{E})} = \frac{1}{n_e^2} + \mathfrak{r}_{33}\mathbb{E}. \tag{8.1.20}$$

Because $\mathfrak{r}_{13}\mathbb{E}$ and $\mathfrak{r}_{33}\mathbb{E}$ are small, we obtain:

$$n_o(\mathbb{E}) \approx n_o - \frac{1}{2}n_o^3\mathfrak{r}_{13}\mathbb{E}, \tag{8.1.21}$$

$$n_e(\mathbb{E}) \approx n_e - \frac{1}{2}n_e^3\mathfrak{r}_{33}\mathbb{E}. \tag{8.1.22}$$

For $\lambda = 500\,\mathrm{n}m$, the Pockels coefficients of $LiNbO_3$ have the following values:

- $\mathfrak{r}_{13} = 9.6\,\mathrm{p}m/V$
- $\mathfrak{r}_{22} = 6.8\,\mathrm{p}m/V$
- $\mathfrak{r}_{33} = 30.9\,\mathrm{p}m/V$
- $\mathfrak{r}_{51} = 32.6\,\mathrm{p}m/V$
- $n_o = 2.341$
- $n_e = 2.246$

To get a feel for the magnitude of the Pockels effect, let us consider an example. For $\mathbb{E} = 10^6\,\mathrm{V}/m$ (10 kV applied across a 1 cm thick crystal), the extraordinary refractive index changes by $\Delta n_e = -\frac{1}{2}n_e^3\mathfrak{r}_{33}\mathbb{E} = -1.75 \times 10^{-4}$. This shows that the refractive index change induced by the Pockels effect is typically quite small.

Example II: KDP

KDP is a negative uniaxial crystal which belongs to the crystallographic point group $\bar{4}2m$. We assume that the optical axis is oriented parallel to the x_3-direction. The corresponding Pockels tensor reads:

$$\mathfrak{r} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \mathfrak{r}_{41} & 0 & 0 \\ 0 & \mathfrak{r}_{41} & 0 \\ 0 & 0 & \mathfrak{r}_{63} \end{bmatrix} \tag{8.1.23}$$

If the static field $\mathbb{E}$ is applied along the optical axis (x_3 axis) of the KDP crystal, the modified index ellipsoid reads:

$$\frac{x_1^2 + x_2^2}{n_o^2} + \frac{x_3^2}{n_e^2} + 2\mathfrak{r}_{63}\mathbb{E}x_1x_2 = 1. \tag{8.1.24}$$

The new principal axis are obtained by a 45°-rotation about the x_3-axis:

$$x_1' = \frac{x_1 + x_2}{\sqrt{2}}, \tag{8.1.25}$$

$$x_2' = \frac{x_1 - x_2}{\sqrt{2}}, \tag{8.1.26}$$

$$x_3' = x_3. \tag{8.1.27}$$

The index ellipsoid reads in the new coordinate system:

$$\frac{x_1'^2}{n_1^2(\mathbb{E})} + \frac{x_2'^2}{n_2^2(\mathbb{E})} + \frac{x_3'^2}{n_3^2(\mathbb{E})} = 1 \tag{8.1.28}$$

with

$$\frac{1}{n_1^2(\mathbb{E})} = \frac{1}{n_o^2} + \mathfrak{r}_{63}\mathbb{E}, \tag{8.1.29}$$

$$\frac{1}{n_2^2(\mathbb{E})} = \frac{1}{n_o^2} - \mathfrak{r}_{63}\mathbb{E}, \tag{8.1.30}$$

$$n_3(\mathbb{E}) = n_e. \tag{8.1.31}$$

After a short calculation, we obtain:

$$n_1(\mathbb{E}) \approx n_o - \frac{1}{2} n_o^3 \mathfrak{r}_{63} \mathbb{E}, \tag{8.1.32}$$

$$n_2(\mathbb{E}) \approx n_o + \frac{1}{2} n_o^3 \mathfrak{r}_{63} \mathbb{E}, \tag{8.1.33}$$

$$n_3(\mathbb{E}) = n_e. \tag{8.1.34}$$

If the external field is applied in the direction of the optical axis, the KDP behaves as a result of the Pockels effect like a biaxial crystal.

8.1.3 Phase modulators

The Pockels effect can be used to modulate the phase of an optical wave. Depending on to the orientation of the external field $\mathbb{E}$ relative to the propagation direction $\mathbf{k}$ of the optical wave, one distinguishes two configurations:

- Transverse phase modulators: $\mathbb{E} \perp \mathbf{k}$,
- Longitudinal phase modulators: $\mathbb{E} || \mathbf{k}$.

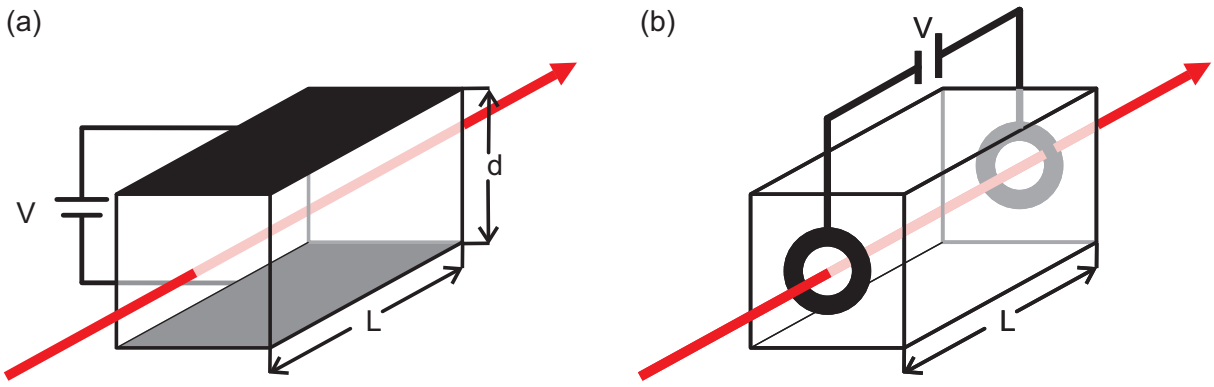


Figure 8.1: (a) Scheme of a transverse phase modulator. (b) Scheme of a longitudinal phase modulator.

The operation frequency of electro-optic modulators is limited by electrical capacitive effects and by the transition time of the light through the material. Commercially available electro optic modulators for free space optics can be typically operated at several hundred MHz. The operation frequency can be considerably increased by integrating the electro-optic modulator in a waveguide structure. In this case, modulation speeds up to 100 GHz have been achieved.

Transverse phase modulators

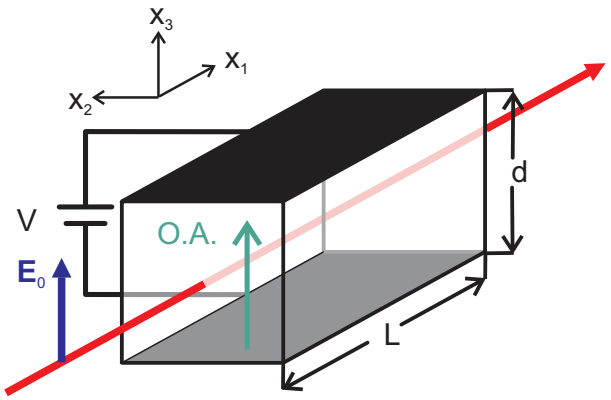


Figure 8.2: Transverse phase modulator.

In the following, we discuss the example of a transverse phase modulator based on a $LiNbO_3$ crystal of thickness L. We assume that the optical axis is oriented parallel to the x_3-direction. A voltage $V = \mathbb{E}d$ is applied along the x_3-direction, where d is the separation of the two electrodes. The optical wave propagates in the x_1-direction and is polarized in the x_3-direction (extraordinary wave):

$$\mathbf{E}(x_1, t) = E_0 \hat{\mathbf{e}}_3 \exp(ik_e x_1 - i\omega t). \tag{8.1.35}$$

The phase shift behind the modulator due to the Pockels effect is given by:

$$\Delta\varphi(\mathbb{E}) = \varphi(\mathbb{E}) - \varphi(0) = k_0 \left[n_e(\mathbb{E}) - n_e(0)\right] L = -\frac{\pi n_e^3 \mathfrak{r}_{33} L}{\lambda_0}\mathbb{E}. \tag{8.1.36}$$

The last equation can be rewritten as

$$\Delta\varphi(V) = -\pi\frac{V}{V_\pi}, \tag{8.1.37}$$

where

$$V_\pi = \frac{d}{L}\frac{\lambda_0}{n_e^3 \mathfrak{r}_{33}} \tag{8.1.38}$$

is the so-called half-wave voltage of the modulator. This is the voltage which is required to induce a phase retardation of π .

Next, we consider a harmonic modulation of the applied voltage with frequency Ω:

$$V(t) = V_0 \sin(\Omega t). \tag{8.1.39}$$

The electric field of the optical wave behind the modulator is then given by:

$$\mathbf{E}(L,t) = E_0\hat{\mathbf{e}}_3 \exp i\varphi_0 \exp\left[-i\left(\omega t + \pi\frac{V_0}{V_\pi}\sin(\Omega t)\right)\right]. \tag{8.1.40}$$

With the identity for Bessel-functions

$$\exp\left[-i\delta\sin(\Omega t)\right] = \sum_{n=-\infty}^{\infty} J_n(\delta)\exp(-in\Omega t), \tag{8.1.41}$$

we obtain:

$$\mathbf{E}(L,t) = E_0\hat{\mathbf{e}}_3 \exp i\varphi_0 \sum_{n=-\infty}^{\infty} J_n(\pi V_0/V_\pi)\exp(-i(\omega + n\Omega)t). \tag{8.1.42}$$

The modulation creates side-bands with frequency $\omega\pm\Omega, \omega\pm 2\Omega, \omega\pm 3\Omega, \cdots$, respectively. A possible application of this property is the use of a periodically driven electro-optic phase modulator as an active mode locking device in a pulsed laser.

Longitudinal phase modulators

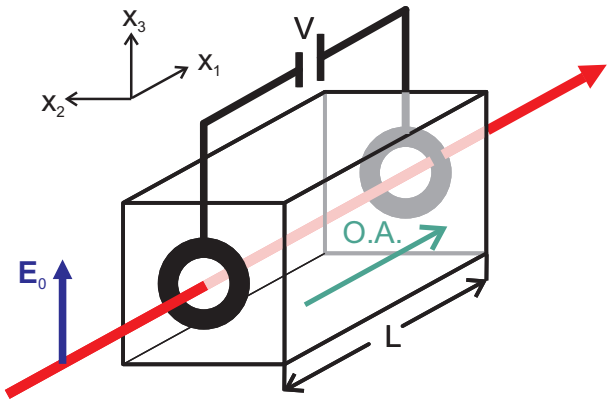


Figure 8.3: Longitudinal phase modulator.

Here, we consider a $LiNbO_3$-crystal whose optical axis is aligned parallel to the x_1-axis. The external static field $\mathbb{E}$ is applied along the x_1-axis. The optical wave propagates in the x_1-direction and is polarized in the x_3-direction (ordinary wave):

$$\mathbf{E}(x_1,t) = E_0\hat{\mathbf{e}}_3 \exp(ik_o x_1 - i\omega t). \tag{8.1.43}$$

The phase shift behind the modulator due to the Pockels effect is given by:

$$\Delta\varphi(\mathbb{E}) = \varphi(\mathbb{E}) - \varphi(0) = k_0\left[n_o(\mathbb{E}) - n_o(0)\right]L = -\pi\frac{n_o^3\mathfrak{r}_{13}}{\lambda_0}\mathbb{E}L. \tag{8.1.44}$$

With $V = \mathbb{E}L$, the last equation can be rewritten as:

$$\Delta\varphi(V) = -\pi\frac{V}{V_\pi}, \tag{8.1.45}$$

where the half-wave voltage V_π of a longitudinal phase modulator is given by

$$V_\pi = \frac{\lambda_0}{n_o^3 \mathfrak{r}_{13}}. \tag{8.1.46}$$

Note that V_π does not depend on the crystal length L in the case of a longitudinal phase modulator.

8.1.4 Polarization modulators

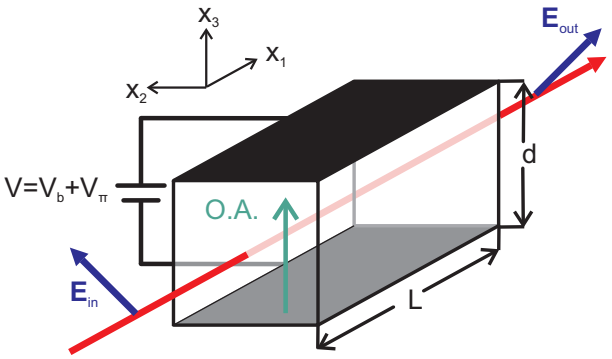


Figure 8.4: Scheme of a polarization modulator.

The Pockels effect can be also used to change the polarization of an optical wave if the electric field of the incident wave is not parallel to one of the three principal axes of the Pockels medium. The optical wave can then be written as the superposition of two polarization eigenmodes whose relative phase can be controlled by applying an external electric field $\mathbb{E}$.

In the following, we consider again a $LiNbO_3$-crystal whose optical axis is parallel to the x_3-direction. The external electric field $\mathbb{E}$ is applied along the x_3-direction. The incident optical wave propagates in x_1-direction and is polarized along $(\hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3)/\sqrt{2}$.

Behind the crystal, the electric field of the wave is given by

$$\mathbf{E}(L,t) = E_0\frac{\hat{\mathbf{e}}_2}{\sqrt{2}}\exp(ik_oL - i\omega t) + E_0\frac{\hat{\mathbf{e}}_3}{\sqrt{2}}\exp(ik_eL - i\omega t). \tag{8.1.47}$$

The phase difference $\Delta\varphi$ between the ordinary and the extraordinary component of the wave is

$$\Delta\varphi = \frac{\pi}{\lambda_0}\left[2\left(n_o - n_e\right)L + \left(n_e^3\mathfrak{r}_{33} - n_o^3\mathfrak{r}_{13}\right)\frac{VL}{d}\right]. \tag{8.1.48}$$

Without applied electric field, the phase difference is given by

$$\Delta\varphi_0 = \frac{2\pi}{\lambda_0}\left(n_o - n_e\right)L. \tag{8.1.49}$$

The last equation can be written as

$$\Delta\varphi_0 = 2m\pi + \frac{2\pi}{\lambda_0}\left(n_o - n_e\right)\Delta L, \tag{8.1.50}$$

where m is an integer with

$$-\pi < \Delta\varphi_0 - 2m\pi = \frac{2\pi}{\lambda_0}\left(n_o - n_e\right)\Delta L \leq \pi. \tag{8.1.51}$$

$\Delta\varphi_0$ can be compensated by applying a background voltage

$$V_b = \frac{2\left(n_e - n_o\right)d}{\left(n_e^3\mathfrak{r}_{33} - n_o^3\mathfrak{r}_{13}\right)L}\Delta L \tag{8.1.52}$$

to the modulator.

By changing the voltage V one can vary the phase difference between the two polarization components and hence use the polarization modulator as a variable wave plate. For instance, the polarization modulator acts as a $\lambda/2$-plate if $\Delta\varphi = \pi$. The necessary voltage is given by

$$V = V_b + V_\pi, \tag{8.1.53}$$

with the half-wave voltage

$$V_\pi = \frac{\lambda_0}{n_e^3\mathfrak{r}_{33} - n_o^3\mathfrak{r}_{13}}\frac{d}{L}. \tag{8.1.54}$$

In our example, the polarization of the optical wave is rotated by 90°.

8.1.5 Intensity modulators

In the following, we discuss how to modulate the intensity of an optical wave by means of an electro-optic modulator.

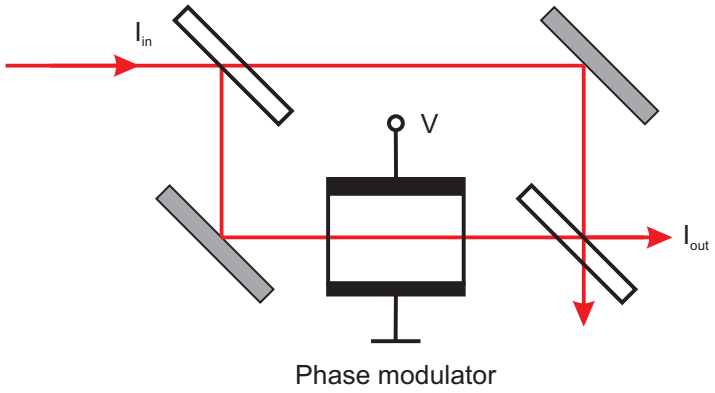


Figure 8.5: Scheme of a Mach-Zehnder modulator

Mach-Zehnder modulators

Consider a Mach-Zehnder interferometer, which contains a phase modulator in one of its arms, e.g., arm 1. If the beam splitters divide the power equally (50:50-beam splitters), the intensity at the output of the interferometer is given by

$$I_{out} = \frac{1}{2} I_{in} + \frac{1}{2} I_{in} \cos(\varphi) = I_{in} \cos^2(\varphi/2), \tag{8.1.55}$$

where $\varphi = \varphi_1(V) - \varphi_2$ is the phase difference between the two waves in the two arms of the interferometer.

The phase of the optical wave in arm 1 depends on the applied voltage:

$$\varphi_1(V) = \varphi_1(V = 0) - \pi \frac{V}{V_\pi}. \tag{8.1.56}$$

With $\varphi_0 = \varphi_1(V = 0) - \varphi_2$, the intensity at the output of the Mach-Zehnder modulator can be written as

$$I_{out} = I_{in} \cos^2 \left(\frac{\varphi_0}{2} - \frac{\pi}{2} \frac{V}{V_\pi} \right). \tag{8.1.57}$$

Intensity modulation by a combination of a polarization modulator and a polarizer

Exercise!

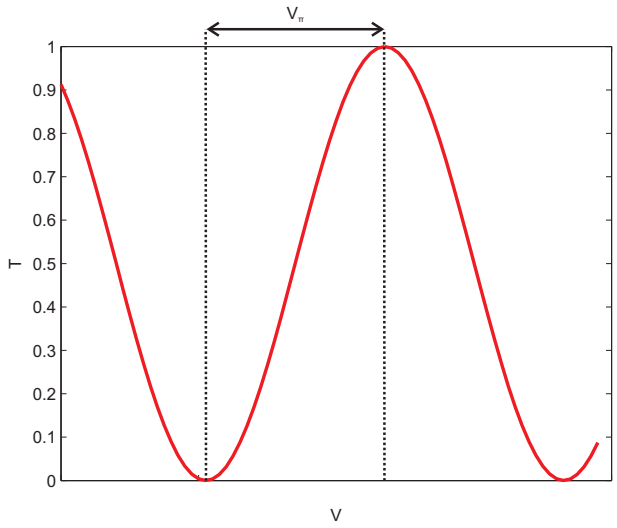


Figure 8.6: Transmittance $T(V) = I_{out}/I_{in}$ of Mach-Zehnder modulator as a function of the applied voltage.

8.2 Acousto-opticcal modulators

8.2.1 Photoelastic effect

A sound wave that travels through a medium leads to an elastic deformation of the medium. This deformation can be characterized by the strain tensor

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{8.2.1}$$

Here, $x_i (i = 1, 2, 3)$ is the coordinate of a small volume element of the medium and $u_i (i = 1, 2, 3)$ is its displacement from the equilibrium position. The diagonal elements $S_{ii} (i = 1, 2, 3,)$ characterize the linear expansion or contraction of the medium along the respective coordinate axis. The off-diagonal elements represent a shear strain.

The mechanical deformation of the medium by the sound wave leads to a modification of its optical properties. The strain-optic tensor $\mathbf{p}$ connects the variation of the impermeability $\boldsymbol{\Delta\eta}$ with the strain tensor $\mathbf{S}$:

$$\Delta\eta_{ij} = p_{ijkl} S_{kl}. \tag{8.2.2}$$

Since both $\boldsymbol{\Delta\eta}$ and $\mathbf{S}$ are symmetric tensors, the elements of the strain-optic tensor show the following permutation symmetry:

$$p_{ijkl} = p_{jikl} = p_{ijlk} = p_{jilk}. \tag{8.2.3}$$

It is convenient to use the contracted notation introduced in section 8.1.2. Equation (8.2.4) then reads:

$$\Delta\eta_g = p_{gh} S_h \quad (g, h = 1, 2, 3, 4, 5, 6). \tag{8.2.4}$$

In the following, we discuss the acousto-optic effect in an isotropic medium (e.g., fused silica) with refractive index n. The strain-optic tensor of the medium reads:

$$\mathfrak{p} = \begin{bmatrix} \mathfrak{p}_{11} & \mathfrak{p}_{12} & \mathfrak{p}_{12} & 0 & 0 & 0 \\ \mathfrak{p}_{12} & \mathfrak{p}_{11} & \mathfrak{p}_{12} & 0 & 0 & 0 \\ \mathfrak{p}_{12} & \mathfrak{p}_{12} & \mathfrak{p}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathfrak{p}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathfrak{p}_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathfrak{p}_{44} \end{bmatrix}, \quad \mathfrak{p}_{44} = \frac{1}{2}(\mathfrak{p}_{11} - \mathfrak{p}_{12}). \tag{8.2.5}$$

Consider a longitudinal sound wave that propagates in the x_3-direction. The resulting displacement of a small volume element of the medium is given by

$$\mathbf{u}(z, t) = -a\hat{\mathbf{e}}_z \cos(Kz - \Omega t). \tag{8.2.6}$$

Here, a is the oscillation amplitude, Ω is the angular frequency of the sound wave, and $\mathbf{K}$ is the wave vector. The only non-vanishing element of the strain tensor is given in this case by

$$S_3(z, t) = Ka \sin(Kz - \Omega t) = S_0 \sin(Kz - \Omega t). \tag{8.2.7}$$

The corresponding change in the impermeability tensor can be calculated as

$$\begin{bmatrix} \Delta\eta_1 \\ \Delta\eta_2 \\ \Delta\eta_3 \\ \Delta\eta_4 \\ \Delta\eta_5 \\ \Delta\eta_6 \end{bmatrix} = \begin{bmatrix} \mathfrak{p}_{11} & \mathfrak{p}_{12} & \mathfrak{p}_{12} & 0 & 0 & 0 \\ \mathfrak{p}_{12} & \mathfrak{p}_{11} & \mathfrak{p}_{12} & 0 & 0 & 0 \\ \mathfrak{p}_{12} & \mathfrak{p}_{12} & \mathfrak{p}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathfrak{p}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathfrak{p}_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathfrak{p}_{44} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ S_3 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{8.2.8}$$

or equivalently,

$$\Delta\eta_1 = p_{12} S_0 \sin(Kz - \Omega t), \tag{8.2.9}$$

$$\Delta\eta_2 = p_{12} S_0 \sin(Kz - \Omega t), \tag{8.2.10}$$

$$\Delta\eta_3 = p_{11} S_0 \sin(Kz - \Omega t), \tag{8.2.11}$$

$$\Delta\eta_4 = \Delta\eta_5 = \Delta\eta_6 = 0. \tag{8.2.12}$$

From the modified index ellipsoid

$$\left(\frac{1}{n^2} + p_{12}S_0 \sin\left(Kz - \Omega t\right)\right) x^2 + \left(\frac{1}{n^2} + p_{12}S_0 \sin\left(Kz - \Omega t\right)\right) y^2 \tag{8.2.13}$$

$$+ \left(\frac{1}{n^2} + p_{11}S_0 \sin\left(Kz - \Omega t\right)\right) z^2 = 1 \tag{8.2.14}$$

we obtain the principal refractive indices

$$n_x(z,t) = n - \frac{1}{2}n^3 p_{12} S_0 \sin\left(Kz - \Omega t\right), \tag{8.2.15}$$

$$n_y(z,t) = n - \frac{1}{2}n^3 p_{12} S_0 \sin\left(Kz - \Omega t\right), \tag{8.2.16}$$

$$n_z(z,t) = n - \frac{1}{2}n^3 p_{11} S_0 \sin\left(Kz - \Omega t\right). \tag{8.2.17}$$

We note that the longitudinal sound wave causes a periodic variation of the refractive indices with a grating period given by

$$\Lambda = \frac{2\pi}{K}. \tag{8.2.18}$$

8.2.2 Bragg scattering

In the following, we consider the effect of the periodic refractive index variation caused by a sound wave on an optical wave. We note that the speed of sound is much smaller than the speed of light in the medium. Hence, we can assume in a first step that the refractive index variation is "frozen" in time.

The local variation of the refractive index gives rise to light scattering. The scattered waves interfere constructively, if the Bragg condition

$$\sin(\theta) = \frac{\lambda}{2\Lambda} \tag{8.2.19}$$

is fulfilled. Here, λ is the wavelength of light in the medium, and θ is the angle of incidence (see Fig. 8.7).

Example: Bragg scattering of a HeNe laser from a longitudinal sound wave with frequency $f = 1\,\mathrm{G}Hz$ in Flint glass.

- Speed of sound: $v_s = 3\,\mathrm{k}m/s$

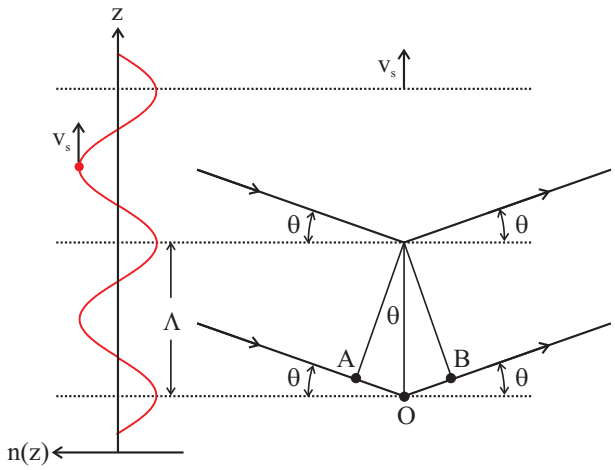


Figure 8.7: Bragg scattering of an optical wave from a periodic variation of the refractive index caused by a sound wave.

- Wavelength of the sound wave: $\Lambda = v_s/f = 3 \times 10^3 \mathrm{m} s^{-1} / 1 \times 10^9 \mathrm{s}^{-1} = 3\,\mu\mathrm{m}$
- Optical wavelength in vacuum: $\lambda_0 = 632.8\,\mathrm{n}m$ (HeNe-Laser)
- Refractive index of Flint glass: $n = 1.95$
- Optical wavelength in the medium: $\lambda = \lambda_0/n = 324.5\,\mathrm{n}m$
- Internal Bragg angle (in the medium): $\theta = 3.1°$
- External Bragg angle (refraction at the glass/air interface) $\theta' \approx n\theta = 6°$

Next, we take the movement of the sound wave and hence also the refractive index grating through the medium into account. The velocity component in the direction of the incident optical wave of the sound wave is given by

$$v = v_s \sin(\theta). \tag{8.2.20}$$

The moving grating can be considered as a moving source such that the Doppler shifted angular frequency of the diffracted optical wave is given by

$$\omega' = \omega \left(1 + 2\frac{v_s \sin(\theta)}{c}\right). \tag{8.2.21}$$

Using the relations $\sin(\theta) = \frac{\lambda}{2\Lambda}$, $c = \lambda\omega/2\pi$, and $v_s = \Lambda\Omega/2\pi$, we can rewrite equation (8.2.21) as

$$\omega' = \omega + \Omega. \tag{8.2.22}$$

The interaction of sound and light can be also understood in the framework of quantum theory. Here, the acoustic wave is considered as a stream of quanta, so called phonons, which each exhibit the energy $\hbar\Omega$ and the momentum $\hbar\mathbf{K}$. The diffracted light signal is created by a series of scattering events. In each of these scattering events, an incident photon (momentum $\hbar\mathbf{k}$, energy: $\hbar\omega$) and a phonon (momentum $\hbar\mathbf{K}$, energy: $\hbar\Omega$) are annihilated and a new photon with momentum $\hbar\mathbf{k}'$ and energy $\hbar\omega'$ is created. Conservation of momentum and energy requires that:

$$\hbar\mathbf{k} + \hbar\mathbf{K} = \hbar\mathbf{k}', \tag{8.2.23}$$

$$\hbar\omega + \hbar\Omega = \hbar\omega'. \tag{8.2.24}$$

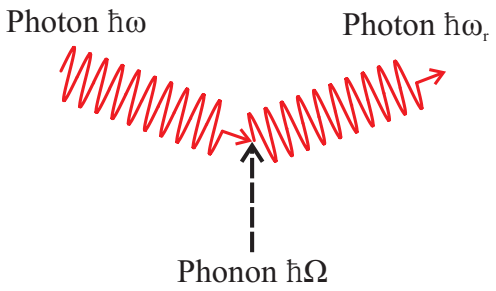


Figure 8.8: Bragg scattering in the framework of quantum theory.

8.3 Bragg cells

A Bragg cell consists of a transparent crystal to which a piezoelectric transducer is attached. The transducer is caused to vibrate by an applied oscillating electric signal. These vibrations in turn create a sound wave in the crystal. For weak sound intensities, the intensity of the diffracted light wave is proportional to the intensity of the sound wave[1]. In this regime, the Bragg cell works as a linear analog modulator [see Fig. 8.9 (a)]. If the Bragg cell is operated with large acoustic power, the diffraction efficiency can become close to unity and the Bragg cell acts as a switch [see Fig. 8.9 (b)].

[1]See, e.g., *P*hotonics - Optical electronics in modern communications by A. Yariv and P. Yeh, Oxford University Press

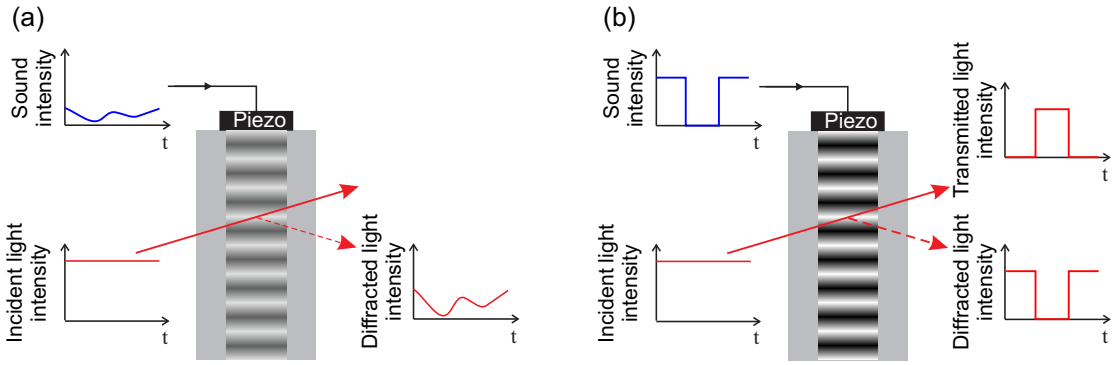


Figure 8.9: Bragg cell: (a) Linear analog modulator, (b) switch.

9 Nonlinear optics

9.1 Anharmonic oscillator model

The Lorentz-oscillator model (see section 1.4.1) describes the linear optical properties of dielectric materials. For small displacements x of the electrons from their equilibrium positions, we can approximate the potential energy function of the electrons by a harmonic potential:

$$U_{\mathrm{LO}}(x) = \frac{1}{2} m_e \omega_e^2 x^2. \tag{9.1.1}$$

For strong electric fields and, hence, for large displacements, deviations from the harmonic potential become important. In the following, we discuss a potential of the form

$$U_{\mathrm{AO}}(x) = \frac{1}{2} m_e \omega_e^2 x^2 + \frac{1}{3} m_e a x^3, \tag{9.1.2}$$

which can be used to describe media without inversion symmetry.

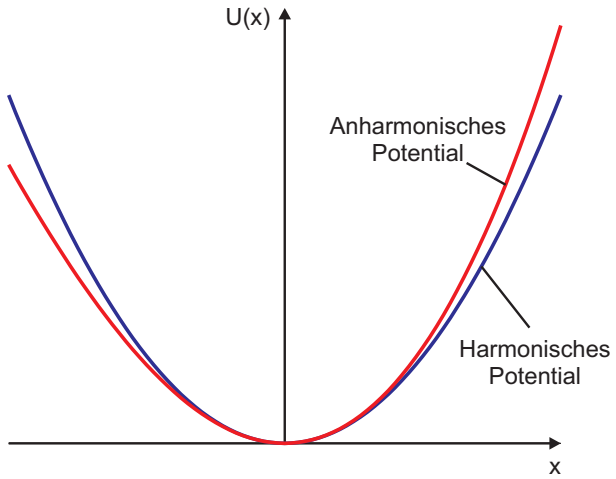


Figure 9.1: Comparison between a harmonic potential (blue curve) and an anharmonic potential (red curve).

The equation of motion of an anharmonic oscillator reads:

$$\ddot{x} + 2\gamma\dot{x} + \omega_e^2 x + a x^2 = -qE(t)/m_e. \tag{9.1.3}$$

In the following, we consider that the incident electric field[1] has two frequency components which oscillate at the frequency ω_1 and ω_2, respectively:

$$E(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + c.c. \tag{9.1.4}$$

We assume that $\omega_e^2 x \gg a x^2$. Under this condition, we can use perturbation theory to solve equation (9.1.3) . We formally replace $E(t)$ by $\xi E(t)$, where $\xi \in [0, 1]$ is the perturbation parameter. The equation of motion (9.1.3) thus reads:

$$\ddot{x} + 2\gamma\dot{x} + \omega_e^2 x + a x^2 = -\xi q E(t)/m_e. \tag{9.1.5}$$

Next, we are looking for a solution which can be written as a power series in ξ:

$$x = \xi x^{(1)} + \xi^2 x^{(2)} + \xi^3 x^{(3)} + \cdots \tag{9.1.6}$$

Substituting this ansatz in the equation of motion and sorting by powers of ξ yields:

$$\begin{aligned}
\mathcal{O}(\xi) &: \ \ddot{x}^{(1)} + 2\gamma\dot{x}^{(1)} + \omega_e^2 x^{(1)} = -qE(t)/m_e && (9.1.7)\\
\mathcal{O}(\xi^2) &: \ \ddot{x}^{(2)} + 2\gamma\dot{x}^{(2)} + \omega_e^2 x^{(2)} + a\left[x^{(1)}\right]^2 = 0 && (9.1.8)\\
\mathcal{O}(\xi^3) &: \ \ddot{x}^{(3)} + 2\gamma\dot{x}^{(3)} + \omega_e^2 x^{(3)} + 2a x^{(1)} x^{(2)} = 0 && (9.1.9)\\
&\cdots
\end{aligned}$$

In linear order $\mathcal{O}(\xi)$, the equation of motion is that of a harmonic oscillator with the well known steady-state solution:

$$x^{(1)}(t) = x^{(1)}(\omega_1) e^{-i\omega_1 t} + x^{(1)}(\omega_2) e^{-i\omega_2 t} + c.c., \tag{9.1.10}$$

where

$$x^{(1)}(\omega_j) = \frac{-q}{m_e} \frac{E_j}{D(\omega_j)} \tag{9.1.11}$$

and

$$D(\omega_j) = \omega_e^2 - \omega_j^2 - 2i\omega_j\gamma. \tag{9.1.12}$$

The corresponding linear susceptibility can be calculated as (see section 1.4.1):

$$\chi^{(1)}(\omega_j) = \frac{N q^2/m_e}{\epsilon_0 D(\omega_j)}. \tag{9.1.13}$$

[1] It is very important to use real-valued fields in this calculation!

Next, we consider the second order response $\mathcal{O}(\xi^2)$. $\left[x^{(1)}(t)\right]^2$ acts in equation (9.1.8) as a "driving force" with frequency components $\pm 2\omega_1$, $\pm 2\omega_2$, $\pm(\omega_1+\omega_2)$, $\pm(\omega_1-\omega_2)$, and 0.

We start with the component that oscillates at frequency $2\omega_1$:

$$\ddot{x}^{(2)} + 2\gamma\dot{x}^{(2)} + \omega_e^2 x^{(2)} = -\frac{a\left(qE_1/m_e\right)^2 e^{-i2\omega_1 t}}{D^2(\omega_1)}. \tag{9.1.14}$$

This is the equation of motion of a harmonic oscillator with driving force $-\frac{a(qE_1/m_e)^2 e^{-i2\omega_1 t}}{D^2(\omega_1)}$. The steady-state solution is given by

$$x^{(2)}(t) = x^{(2)}(2\omega_1)e^{-i2\omega_1 t} + c.c. \tag{9.1.15}$$

with

$$x^{(2)}(2\omega_1) = \frac{-a(q/m_e)^2 E_1^2}{D(2\omega_1)D^2(\omega_1)}. \tag{9.1.16}$$

The corresponding second order susceptibility[2] can be calculated as

$$\begin{aligned}
\chi^{(2)}(2\omega_1,\omega_1,\omega_1) &= \frac{Na(q^3/m_e^2)}{\epsilon_0 D(2\omega_1)D^2(\omega_1)} && (9.1.17)\\
&= \frac{\epsilon_0^2 m_e a}{N^2 q^3}\chi^{(1)}(2\omega_1)\left[\chi^{(1)}(\omega_1)\right]^2. && (9.1.18)
\end{aligned}$$

Analogously, we obtain after a short calculation the second order susceptibility for the other frequency components:

$$\begin{aligned}
\chi^{(2)}(2\omega_2,\omega_2,\omega_2) &= \frac{\epsilon_0^2 m_e a}{N^2 q^3}\chi^{(1)}(2\omega_2)\left[\chi^{(1)}(\omega_2)\right]^2, && (9.1.19)\\
\chi^{(2)}(\omega_1+\omega_2,\omega_1,\omega_2) &= 2\frac{\epsilon_0^2 m_e a}{N^2 q^3}\chi^{(1)}(\omega_1+\omega_2)\chi^{(1)}(\omega_1)\chi^{(1)}(\omega_2), && (9.1.20)\\
\chi^{(2)}(\omega_1-\omega_2,\omega_1,-\omega_2) &= 2\frac{\epsilon_0^2 m_e a}{N^2 q^3}\chi^{(1)}(\omega_1-\omega_2)\chi^{(1)}(\omega_1)\chi^{(1)}(-\omega_2), && (9.1.21)\\
\chi^{(2)}(0,\omega_1,-\omega_1) &= 2\frac{\epsilon_0^2 m_e a}{N^2 q^3}\chi^{(1)}(0)\chi^{(1)}(\omega_1)\chi^{(1)}(-\omega_1), && (9.1.22)\\
\chi^{(2)}(0,\omega_2,-\omega_2) &= 2\frac{\epsilon_0^2 m_e a}{N^2 q^3}\chi^{(1)}(0)\chi^{(1)}(\omega_2)\chi^{(1)}(-\omega_2). && (9.1.23)
\end{aligned}$$

[2]Here, we use the following notation. The first frequency argument in the bracket corresponds to the oscillation frequency of the second order polarization. The second and third frequency argument are two frequency components of the incident electric field that generate the polarization.

The third order displacement $x^{(3)}(t)$ can be calculated by substitution of $x^{(1)}(t)$ and $x^{(2)}(t)$ in equation(9.1.9) $[\mathcal{O}(\xi^3)]$, and so on and so forth.

For materials with inversion symmetry, we can choose the following ansatz for the the anharmonic potential :

$$U_{\mathrm{NO}}(x) = \frac{1}{2} m_e \omega_e^2 x^2 - \frac{1}{4} m_e b x^4. \tag{9.1.24}$$

It is easy to show that all the even orders of the susceptibility are identical to zero.

Second order susceptibility of a medium with inversion symmetry.

In general, $\chi^{(2)}$ is a tensor which "connects" the second order polarization with the incident electric field:

$$P_i^{(2)}(\omega_1 + \omega_2) = \epsilon_0 \chi_{ijk}^{(2)}(\omega_1 + \omega_2, \omega_1, \omega_2) E_j(\omega_1) E_k(\omega_2). \tag{9.1.25}$$

In the case of the inversion of the coordinates ($\mathbf{r} \to -\mathbf{r}$), both, the electric field and the polarization change their sign: $\mathbf{E} \to -\mathbf{E}, \mathbf{P}^{(2)} \to -\mathbf{P}^{(2)}$.

Thus, we obtain:

$$\chi_{ijk}^{(2)}(\omega_1 + \omega_2, \omega_1, \omega_2) = -\chi_{ijk}^{(2)}(\omega_1 + \omega_2, \omega_1, \omega_2). \tag{9.1.26}$$

This condition can be only fulfilled if $\chi_{ijk}^{(2)} = 0$!

9.2 Wave propagation in nonlinear media

In the following, we assume that the polarization P can be expanded as a power series[3] in E :

$$P(t) = \epsilon_0 \left[\chi^{(1)} E(t) + \chi^{(2)} E^2(t) + \chi^{(3)} E^3(t) + \cdots \right] \tag{9.2.1}$$

The first term is the well known linear polarization:

$$P_L(t) = \epsilon_0 \chi^{(1)} E(t). \tag{9.2.2}$$

[3]For reasons of simplicity, we neglect the vector character of the fields.

All higher order terms are subsumed under the nonlinear polarization

$$P_{NL}(t) = \epsilon_0 \left[\chi^{(2)} E^2(t) + \chi^{(3)} E^3(t) + \cdots\right]. \tag{9.2.3}$$

With these abbreviations, the wave equation can be written as ($M = 0, j = 0$)

$$\nabla^2 E - \frac{1}{c_0^2}\frac{\partial^2 E}{\partial t^2} = \mu_0 \frac{\partial^2 P_L}{\partial t^2} + \mu_0 \frac{\partial^2 P_{NL}}{\partial t^2} \tag{9.2.4}$$

$$\Rightarrow \nabla^2 E - \frac{n^2}{c_0^2}\frac{\partial^2 E}{\partial t^2} = \mu_0 \frac{\partial^2 P_{NL}}{\partial t^2} \tag{9.2.5}$$

The nonlinear polarization P_{NL} acts as a source term on the right-hand side of the wave equation. We thus expect the generation of new frequency components no present in the incident light field.

9.3 Second-order nonlinear processes

In this section, we consider effects which arise from the second-order nonlinear polarization:

$$P_{NL}(t) = \epsilon_0 \chi^{(2)} E^2(t). \tag{9.3.1}$$

We assume that the incident electric field has two components which oscillate at frequency ω_1 and ω_2, respectively:

$$E(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + c.c. \tag{9.3.2}$$

The corresponding nonlinear polarization can be written as:

$$\begin{aligned} P_{NL}(t) = \quad & P_{NL}(0) + [P_{NL}(2\omega_1)e^{-i2\omega_1 t} + P_{NL}(2\omega_2)e^{-i2\omega_2 t} \\ + \quad & P_{NL}(\omega_1+\omega_2)e^{-i(\omega_1+\omega_2)t} + P_{NL}(\omega_1-\omega_2)e^{-i(\omega_1-\omega_2)t} + c.c.] \end{aligned} \tag{9.3.3}$$

with

$$P_{NL}(0) = 2\epsilon_0\chi^{(2)}(0,\omega_1,-\omega_1)|E_1|^2 + 2\epsilon_0\chi^{(2)}(0,\omega_2,-\omega_2)|E_2|^2 \tag{9.3.4}$$

$$P_{NL}(2\omega_1) = \epsilon_0\chi^{(2)}(2\omega_1,\omega_1,\omega_1)E_1^2 \tag{9.3.5}$$

$$P_{NL}(2\omega_2) = \epsilon_0\chi^{(2)}(2\omega_2,\omega_2,\omega_2)E_2^2 \tag{9.3.6}$$

$$P_{NL}(\omega_1+\omega_2) = 2\epsilon_0\chi^{(2)}(\omega_1+\omega_2,\omega_1,\omega_2)E_1E_2 \tag{9.3.7}$$

$$P_{NL}(\omega_1-\omega_2) = 2\epsilon_0\chi^{(2)}(\omega_1-\omega_2,\omega_1,-\omega_2)E_1E_2^* \tag{9.3.8}$$

P_{NL} has components which oscillate at the frequency $\omega = 2\omega_1$, $\omega = 2\omega_2$, $\omega = \omega_1 + \omega_2$ and $\omega = \omega_1 + \omega 2$, respectively:

- $P_{NL}(2\omega_1)$ generates light with the frequency $\omega = 2\omega_1$. This process is called second harmonic generation[4] (SHG) .
- $P_{NL}(2\omega_2)$ generates light with the frequency $\omega = 2\omega_2$.
- $P_{NL}(\omega_1 + \omega_2)$ generates light with the frequency $\omega = \omega_1 + \omega_2$. This process is called sum frequency generation (SFG) .
- $P_{NL}(\omega_1 - \omega_2)$ generates light with the frequency $\omega = \omega_1 - \omega_2$. This process is called difference frequency generation(DFG) .

Additionally, there is a DC component ($\omega = 0$). The corresponding process is called optical rectification (OR).

SHG, SFG and DFG can be interpreted as the interaction of three photons mediated by the nonlinear medium.

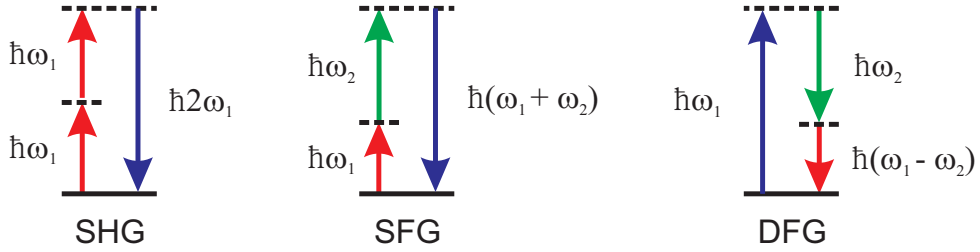


Figure 9.2: SHG, SFG and DFG can be interpreted as the interaction of three photons mediated by the nonlinear medium.

9.3.1 Second harmonic generation

In the following, we discuss SHG in more detail. The electric fields of the fundamental wave and the second harmonic wave are given by:

$$E(z,t) = \sum_{j=1}^{2} E_j(z) e^{i(k_j z - \omega_j t)} + c.c. \tag{9.3.9}$$

with

$$\omega_2 = 2\omega_1, k_1 = \frac{\omega_1}{c_0} n(\omega_1) \quad and \quad k_2 = \frac{2\omega_1}{c_0} n(2\omega_1). \tag{9.3.10}$$

[4]First experimental observation: P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, Phys. Rev. Lett. **7**, 118 (1961).

We assume that the amplitudes $E_j(z), j = 1, 2$ of the two waves vary slowly with position $\left(\left|\frac{\partial^2}{\partial z^2}E_j(z)\right| \ll \left|\frac{\partial}{\partial z}E_j(z)\right|\right)$ such that the following relation is fulfilled

$$\frac{\partial^2}{\partial z^2}E_j(z)e^{i(k_j z)} \approx e^{i(k_j z)}\left[2ik_j\frac{\partial}{\partial z} - k_j^2\right]E_j(z) \tag{9.3.11}$$

The wave equation can be thus written as

$$\sum_{j=1}^{2} e^{ik_j z}\left[2ik_j\frac{\partial}{\partial z}\underbrace{-k_j^2 + \frac{n^2(\omega_j)\omega_j^2}{c_0^2}}_{0}\right]E_j(z)e^{-i\omega_j t} + c.c. = -\sum_{j=1}^{2}\frac{\omega_j^2}{\epsilon_0 c_0^2}P^{(2)}(\omega_j)e^{-i\omega_j t} + c.c., \tag{9.3.12}$$

with

$$P^{(2)}(\omega_1) = 2\epsilon_0\chi^{(2)}E_2(z)e^{ik_2 z}E_1^*(z)e^{-ik_1 z}, \tag{9.3.13}$$
$$P^{(2)}(2\omega_1) = \epsilon_0\chi^{(2)}E_1(z)^2e^{2ik_1 z}. \tag{9.3.14}$$

Rearranging leads to the following coupled-amplitude equations:

$$\frac{\partial}{\partial z}E_1(z) = \frac{i\,\omega_1\,\chi^{(2)}}{c_0\,n(\omega_1)}E_2(z)E_1^*(z)e^{i\Delta k z}, \tag{9.3.15}$$
$$\frac{\partial}{\partial z}E_2(z) = \frac{i\,\omega_1\,\chi^{(2)}}{c_0\,n(2\omega_1)}E_1(z)^2e^{-i\Delta k z}, \tag{9.3.16}$$

with the phase mismatch

$$\Delta\mathbf{k} = \mathbf{k}_2 - 2\mathbf{k}_1. \tag{9.3.17}$$

For collinear propagation of the two waves, the phase mismatch Δk is given by:

$$\Delta k = \frac{2\omega_1}{c_0}\left[n(2\omega_1) - n(\omega_1)\right]. \tag{9.3.18}$$

In the following, we assume that the fundamental wave has a constant amplitude (undepleted-pump approximation) and that no second-harmonic light is incident on the nonlinear medium. In that case, the intensity of the SHG wave after the propagation distance z is given by

$$I_2(z) = C^2\,z^2\,I_1^2\,\frac{\sin^2(\Delta k z/2)}{(\Delta k z/2)^2}. \tag{9.3.19}$$

Here, C is a constant which is proportional to the nonlinear susceptibility and I_1 is the intensity of the fundamental wave. For the following discussion it is convenient to introduce the coherence length

$$l_{coh} = \frac{\pi}{\Delta k}. \tag{9.3.20}$$

In the case of phase matching ($\Delta k = 0$), the intensity of the second harmonic wave grows quadratically with the propagation distance in the nonlinear medium. In contrast, I_2 shows an oscillatory behaviour for $\Delta k \neq 0$. The second harmonic intensity grows within the first coherence length l_{coh}, falls of again, vanishes after $2l_{coh}$, rises again, and so on and so forth. This oscillatory behaviour is an interference effect. For $\Delta k \neq 0$, the fundamental wave and the second harmonic wave propagate with different velocities through the nonlinear medium. The second harmonic waves generated at two different points separated by the distance l_{coh} are π out of phase and, hence, interfere destructively.

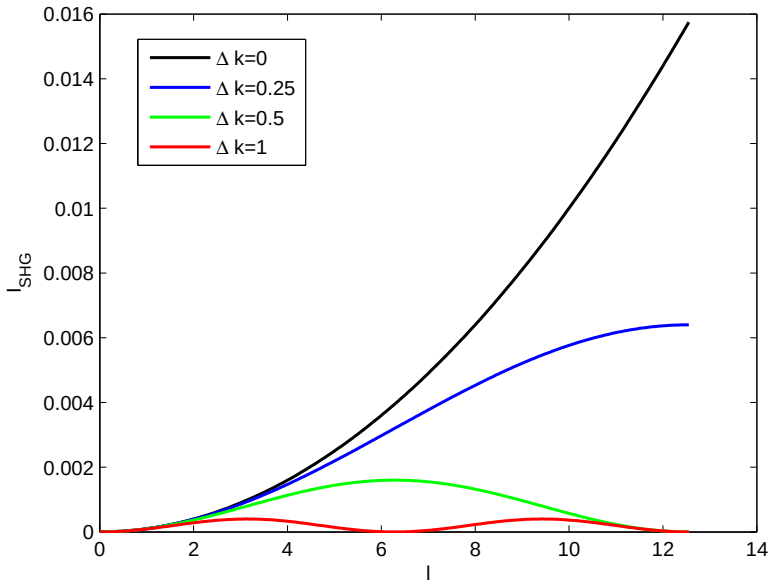


Figure 9.3: Second harmonic intensity as a function of propagation distance for different phase mismatches.

9.3.2 Phase matching

The previous section has shown that phase matching ($\Delta k = 0$) is essential for efficient SHG. Because of linear dispersion ($n(2\omega_1) > n(\omega_1)$), this condition can not be fulfilled in an isotropic material. However, in unaxial materials, we can use the different refractive indices of the ordinary and the extraordinary wave to achieve phase matching.

Type I Phase matching

We initially consider a positive unaxial crystal with $n_o(2\omega_1) < n_e(\omega_1)$. For type I phase matching, the polarization of the fundamental wave is chosen such that it propagates as an extraordinary wave through the crystal. The second harmonic wave is polarized perpendicular to the optical axis (ordinary wave). Phase matching is achieved if $n_o(2\omega_1) = n_e(\omega_1, \theta)$ (see Fig. 9.4) .

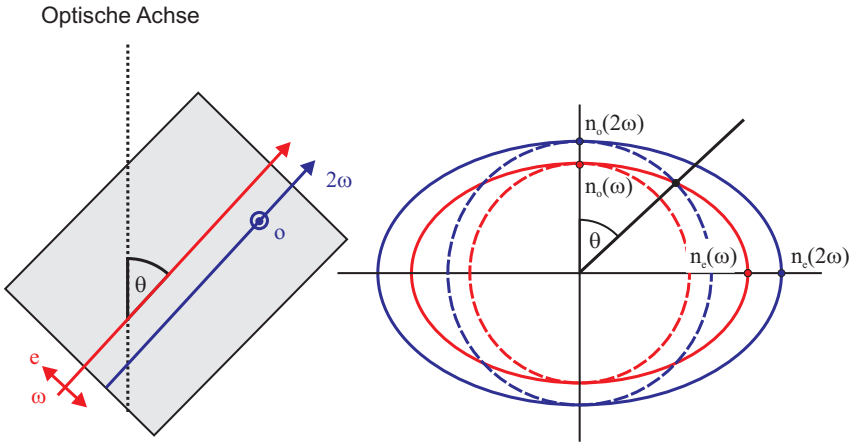


Figure 9.4: Type I phase matching for a positive unaxial crystal.

In the case of a negative unaxial crystal, the polarizations of the fundamental and the second harmonic wave change their roles.

Type II phase matching

For type II phase matching, the polarization of the fundamental wave is chosen such that is has simultaneously an ordinary and an extraordinary component. Phase matching is achieved, if

- Positive uniaxial crystal: $2n_o(2\omega_1) = n_o(\omega_1) + n_e(\omega_1, \theta)$.
- Negative uniaxial crystal: $2n_e(2\omega_1, , \theta) = n_o(\omega_1) + n_e(\omega_1, \theta)$.

Quasi phase matching

Type I and type II phase matching can not be achieved in materials with strong linear dispersion, e.g., $LiNbO_3$. In this situation, one can (sometimes) adopt quasi phase matching as an alternative approach. The idea behind quasi phase matching is to periodically

change the sign of the nonlinear coefficient. In the case of $LiNbO_3$, this can be achieved by electric poling of the ferroelectric domains with periodically structured electrodes.

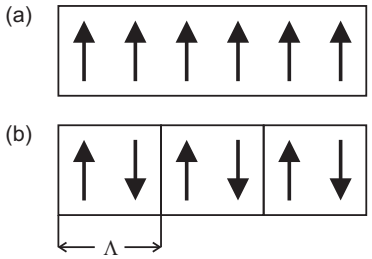


Figure 9.5: Schematic representation of a nonlinear medium. The arrows represent the orientation of the domains and, hence, the sign of the nonlinear coefficient. (a) Homogeneous crystal. (b) Periodically poled crystal.

In the following, we assume that the sign of the nonlinear coefficient $\chi^{(2)}$ changes along the propagation direction of the fundamental wave with period Λ. The Fourier series of $\chi^{(2)}(z)$ is then given by:

$$\chi^{(2)}(z) = \chi^{(2)}_{\mathrm{B}ulk} \left[\sum_{m=-\infty}^{\infty} c_m \exp\left(im \frac{2\pi}{\Lambda} z \right) \right]. \tag{9.3.21}$$

with

$$c_m = \frac{2}{m\pi} \sin(m\pi/2). \tag{9.3.22}$$

Substitution of equation (9.3.21) in equation (9.3.16) yields:

$$\frac{\partial}{\partial z} E_2(z) = i \frac{\omega_1}{c_0 n(2\omega_1)} \chi^{(2)}_{\mathrm{B}ulk} E_1^2(z) \sum_{m=-\infty}^{\infty} c_m e^{i\left(m\frac{2\pi}{\Lambda} - k_2 + 2k_1\right)z}. \tag{9.3.23}$$

Quasi phase matching is achieved, if the condition

$$m\frac{2\pi}{\Lambda} - k_2 + 2k_1 = m\frac{2\pi}{\Lambda} - \Delta k = 0 \tag{9.3.24}$$

is fulfilled for a specific value of m. In the following, we consider that the last equation is met for $m = 1$. The corresponding poling period is given by

$$\Lambda = \frac{2\pi}{\Delta k} = 2l_{coh}. \tag{9.3.25}$$

If we neglect the contributions of the other Fourier components with $m \neq 1$, we obtain the following effective nonlinear coefficient:

$$\chi^{(2)}_{eff} = c_1 \chi^{(2)}_{\mathrm{B}ulk} = \frac{2}{\pi} \chi^{(2)}_{\mathrm{B}ulk}. \tag{9.3.26}$$

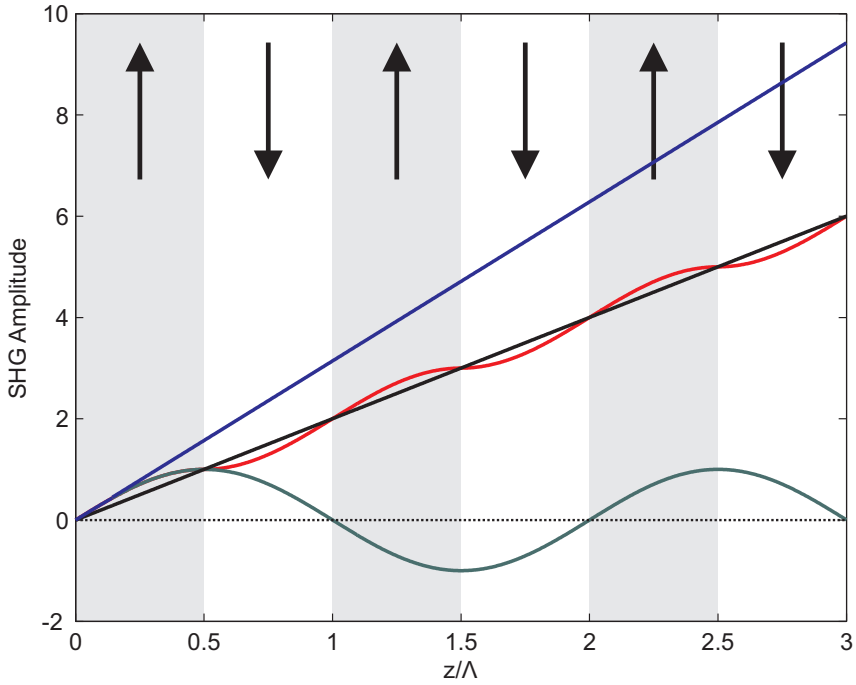


Figure 9.6: Amplitude of the second harmonic wave as a function of the propagation distance. The arrows represent the orientation of the domains. Blue curve: Perfect phasematching ($\Delta k = 0$). Green curve: No phase matching ($\Delta k \neq 0$). Red curve: Quasi phase matching, all Fourier components are taken into account. Black curve: Quasi phase matching, Fourier components with $m \neq 1$ are neglected.

9.4 Third-order nonlinear processes

In this section, we discuss effects related to the third-order nonlinear polarization:

$$P_{NL}(t) = \epsilon_0 \chi^{(3)} E^3(t). \tag{9.4.1}$$

For an incident wave $E(t) = [E(\omega_1)e^{-i\omega_1 t} + c.c.]$, the nonlinear polarization has components which oscillate with the frequency ω_1 and $3\omega_1$, respectively:

$$P_{NL}(t) = (P_{NL}(\omega_1)e^{-i\omega_1 t} + P_{NL}(3\omega_1)e^{-i3\omega_1 t} + c.c.), \tag{9.4.2}$$

where

$$P_{NL}(\omega_1) = 3\epsilon_0 \chi^{(3)} |E(\omega_1)|^2 E(\omega_1), \tag{9.4.3}$$
$$P_{NL}(3\omega_1) = \epsilon_0 \chi^{(3)} E^3(\omega_1). \tag{9.4.4}$$

9.4.1 Third-harmonic generation

The polarization $P_{NL}(3\omega_1)$ is the source of radiation with the frequency $3\omega_1$. This effect is called third-harmonic generation (THG) and can be treated analogously to SHG. The corresponding phase matching condition reads:

$$\mathbf{k}_3 - 3\mathbf{k}_1 = 0. \tag{9.4.5}$$

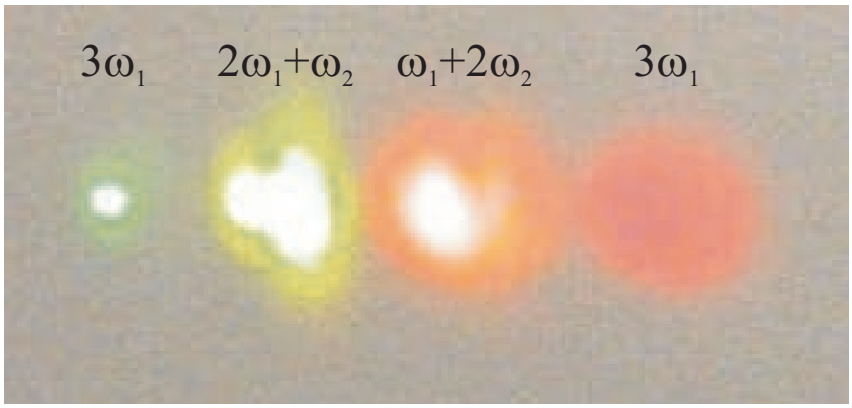


Figure 9.7: Generation of new frequency components from two near-infrared laser pulses by third-order nonlinear processes in glas. Image provided by Christiane Becker.

9.4.2 Optical Kerr effect

The third-order polarization also has a component which oscillates at the frequency ω_1 of the fundamental wave. The total polarization at frequency ω_1 is thus given by:

$$P(\omega_1) = \epsilon \left[\chi^{(1)} + \underbrace{3\chi^{(3)}|E(\omega_1)|^2}_{\Delta\chi} \right] E(\omega_1). \tag{9.4.6}$$

The incremental change of the susceptibility $\Delta\chi$ is responsible for an intensity dependence of the refractive index:

$$n \to n + n_2 I. \tag{9.4.7}$$

With $n^2 = 1 + \chi$ and $I = n_0\epsilon_0 c_0|E|^2/2$, we obtain:

$$n_2 = \frac{3}{n_0^2 c_0 \epsilon_0}\chi^{(3)}. \tag{9.4.8}$$

The intensity dependence of the refractive index is known as the optical Kerr effect. Third-order linear media are hence often denoted as Kerr media and n_2 is called the Kerr coefficient.

In the following, we discuss one consequence of the optical Kerr effect. Consider an ultrashort laser pulse with pulse length τwhich propagates through a Kerr Medium:

$$E(t) = A(t)e^{-i\omega_0 t}, \tag{9.4.9}$$

where $A(t)$ is the envelope function and ω_0 is the carrier frequency.

We assume that the medium exhibits an instantaneous response. The refractive index is thus a function of time:

$$n(t) = n_0 + n_2 I(t). \tag{9.4.10}$$

The electric field of the pulse behind the Kerr medium is given by:

$$E_{out}(0,t) = A(t)e^{-i\omega_0 t}e^{-i\Phi_{\mathrm{N}L}(t)} \tag{9.4.11}$$

with

$$\Phi_{\mathrm{N}L}(t) = -\frac{\omega_0}{c_0} n_2 I(t) d. \tag{9.4.12}$$

The nonlinear refractive index cause a temporal variation of the phase that is proportional to the instantaneous intensity. This effect is called self phase modulation (SPM).

The spectrum of the pulse behind the Kerr medium is calculated by Fourier transformation:

$$S(\omega) = \left| \int_{-\infty}^{\infty} \mathbf{E}_{in}(t)\, e^{-\omega_0 t} e^{-i\Phi(t)} e^{\imath\omega t} dt \right|^2 . \tag{9.4.13}$$

For further analysis, we define the instantaneous frequency $\omega(t)$ through

$$\omega(t) = \omega_0 + \delta\omega(t) \tag{9.4.14}$$

with

$$\delta\omega(t) = \frac{d}{dt}\Phi_{\mathrm{N}L}(t). \tag{9.4.15}$$

The temporal variation of the nonlinear phase shows that SPM creates new frequency components. An estimate of the maximum frequency shift is given by:

$$\delta\omega_{max} \approx \frac{\Delta\Phi_{\mathrm{N}L}^{(max)}}{\tau_p/2} \tag{9.4.16}$$

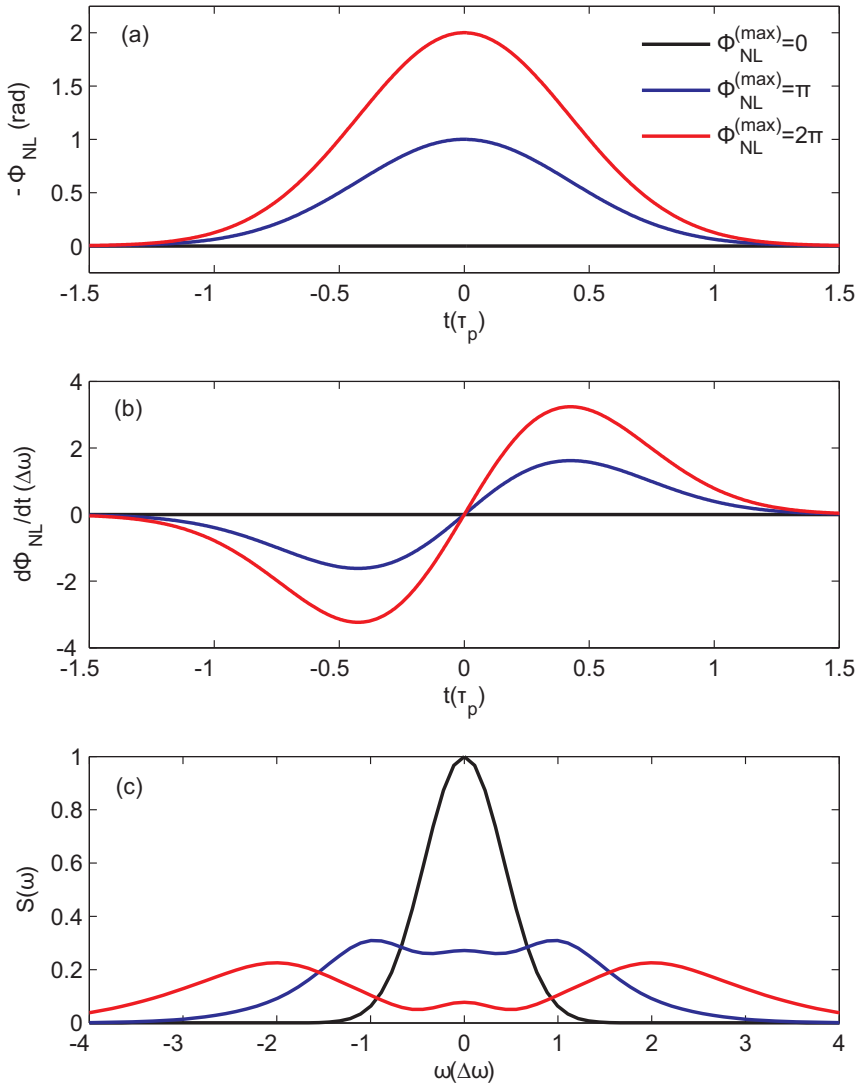


Figure 9.8: SPM of a Gaussian pulse.

with

$$\Delta\Phi_{\mathrm{NL}}^{(max)} \approx \frac{\omega_0}{c_0} n_2 I_0 d. \tag{9.4.17}$$

This estimate illustrates that spectral broadening due to SPM is best achieved with short, intense pulses.

9.5 Nonlinear frequency conversion devices

9.5.1 Green laser pointer

Green laser pointers usually are DPSSFD laser (DPSSFD:diode pumped solid state frequency-doubled). Fig. 9.9 depicts the scheme of such a device. A high-power AlGaAs-Laser diode (typical power: 100 mW - 300 mW) operated at $\lambda = 808\,\mathrm{n}m$ is used to pump a Nd:YVO_4 crystal. This crystal efficiently emits light at the wavelength $\lambda = 1064\,\mathrm{n}m$. The cavity of the Nd:$YVO_4$ laser is formed by the Nd:YVO_4-crystal and a nonlinear KTP crystal. The laser mirrors are directly coated on the front surface of the Nd:YVO_4 crystal and on the back surface of the KTP crystal, respectively. The near infrared radiation is frequency doubled in the cavity by the KDP crystal. The emitted green light has a wavelength of $\lambda = 532\,\mathrm{n}m$.

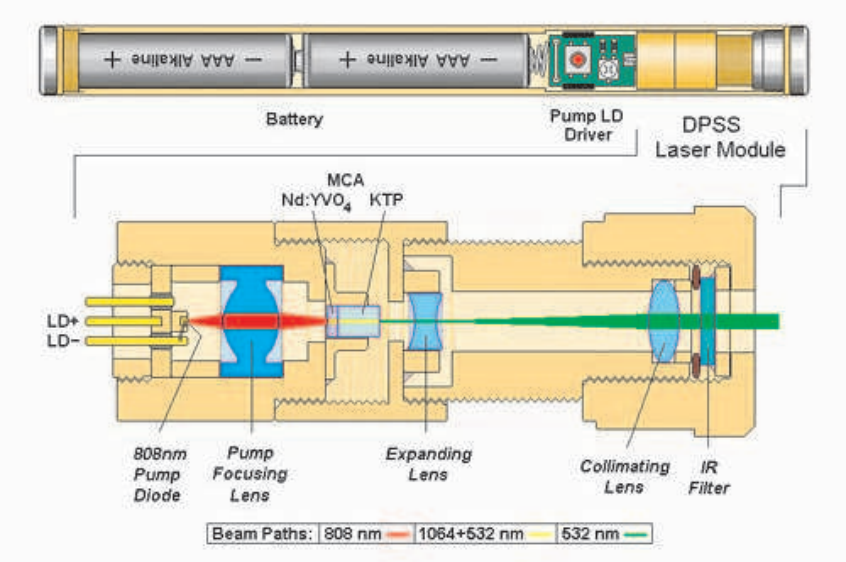


Figure 9.9: Scheme of a green laser pointer. Image taken from Wikipedia.

9.5.2 Optical parametric devices

As the tunability of lasers is limited by the gain bandwidth of the respective active material, broadband tunable coherent light sources mainly rely on frequency conversion processes. In particular $\chi^{(2)}$-based nonlinear processes play an important role in this context.

Optical parametric amplifier

In an optical parametric amplifier (OPA) difference frequency generation between a strong pump wave with frequency ω_3 and a weak signal wave with frequency ω_1 is used to amplify the weak signal. The nonlinear interaction of the two waves in the $\chi^{(2)}$-medium creates a third wave, the so-called idler, with frequency

$$\omega_2 = \omega_3 - \omega_1. \tag{9.5.1}$$

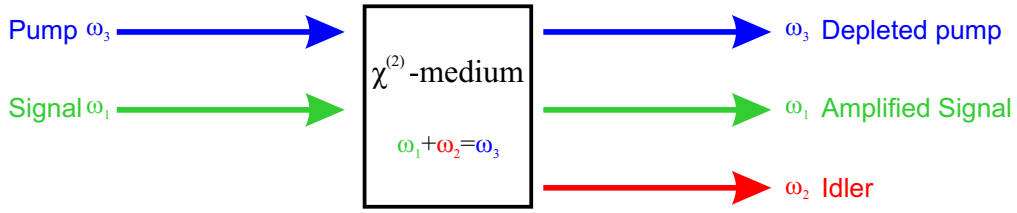


Figure 9.10: Scheme of the optical parametric amplification process.

Like in the case of SHG, one can derive a set of coupled-amplitude equations that governs the spatial evoulution of the amplitudes E_1, E_2, E_3 of the signal, the idler, and the pump, respectively. Assuming a constant amplitude of the pump wave $E_3(z) = \text{const}$ (undepleted pump approximation) and co-linear propagation of the three waves, the coupled-amplitude equations are given by:

$$\frac{dE_1(z)}{dz} = \frac{4\pi\imath\omega_1^2\chi^{(2)}}{k_1 c_0^2} E_3 E_2^*(z) e^{\imath\Delta k z}, \tag{9.5.2}$$

$$\frac{dE_2(z)}{dz} = \frac{4\pi\imath\omega_2^2\chi^{(2)}}{k_2 c_0^2} E_3 E_1^*(z) e^{\imath\Delta k z}, \tag{9.5.3}$$

with the phase mismatch

$$\Delta k = k_3 - k_1 - k_2. \tag{9.5.4}$$

In the following, we consider the case $\Delta k = 0$, which requires that

$$n(\omega_3)\omega_3 = n(\omega_1)\omega_1 + n(\omega_2)\omega_2. \tag{9.5.5}$$

Differentiation of equation (9.5.2) with respect to z and introducing the complex conjugate of equation (9.5.3) yields

$$\frac{d^2E_1(z)}{dz^2} = \kappa^2 E_1(z), \tag{9.5.6}$$

where the coupling constant κ is given by

$$\kappa^2 = \frac{16\pi\omega_1^2\omega_2^2\left(\chi^{(2)}\right)^2}{k_1 k_2 c_0^4}|E_3|^2. \tag{9.5.7}$$

Equation (9.5.6) has the general solution

$$E_1(z) = C\sinh(\kappa z) + D\cosh(\kappa z), \tag{9.5.8}$$

where C and D are integration constants. If the idler wave has zero amplitude at the entrance of the nonlinear medium ($E_2(0) = 0$), the amplitudes of the signal and idler are given by

$$E_1(z) = E_1(0)\cosh(\kappa z) \tag{9.5.9}$$

$$E_2(z) = i\sqrt{\frac{n(\omega_1)\omega_2}{n(\omega_2)\omega_1}}\frac{E_3}{|E_3|}E_1^*(0)\sinh(\kappa z). \tag{9.5.10}$$

Inspection of equations (9.5.9) and (9.5.10) shows that both waves experience monotonic growth. Importantly, the phase of the signal wave is conserved in this amplification process.

Fig. 9.11 depicts the scheme of a two-pass OPA which can be used as a tunable femtosecond light source for near-infrared pulses. The OPA is pumped by an amplified femtosecond Ti:Sapphire laser. A part of the pump is used to generate a white light continuum in a sapphire crystal. The white light continuum acts as the seed for the optical paramateric amplification process in the nonlinear crystal. Typical nonlinear crystal materials are beta barium borate (BBO), potassium dihydrogen phosphate (KDP), and lithium triborate (LBO). The frequencies of the signal and idler are determined by the phase matching condition which can be changed by tilting or heating the nonlinear crystal.

Optical parametric generator

The classical wave analysis given above requires a signal wave $E_1(0) \neq 0$ (or a idler wave $E_2(0) \neq 0$) to seed the optical parametric amplification process. For $E_1(0) = E_2(0) = 0$, we expect that no signal and idler wave are created in the nonlinear medium. In contrast to our expectation, one finds experimentally that an intense pump beam can give rise to the generation of a signal and idler wave even if no signal or idler input are present. The nonlinear process responsible for this effect is called spontaneous parametric down conversion. It can be understood as the time reversed process of sum frequency generation. In the photon picture, spontaneous parametric down conversion is the spontaneous

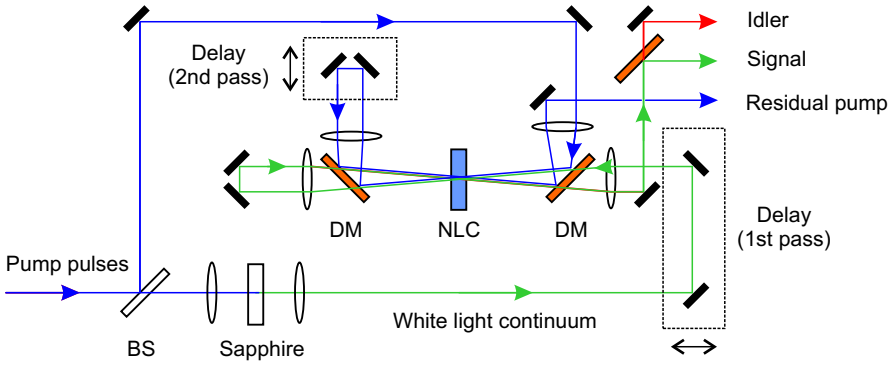


Figure 9.11: Scheme of of a tuneable fs-OPA. BS: Beam splitter. DM: Dichroic mirror. NLC: Nonlinear crystal.

splitting of a pump photon into a signal photon and an idler photon triggered by vacuum fluctuations.

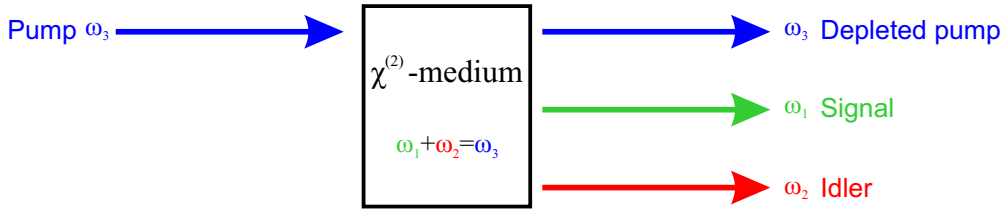


Figure 9.12: Scheme of the optical parametric generation process.

An Optical parametric generator (OPG) uses spontaneous parametric down conversion to seed the optical parametric amplification process. Since spontaneous parametric down conversion is a very inefficient process, OPGs typically requires pump pulses with high peak intensities. Fig. 9.13 depicts the scheme of a two-pass OPG that can be used as a tunable femtosecond light source for near-infrared pulses.

Optical parametric oscillator

An optical parametric oscillator (OPO) consists of a nonlinear medium that is enclosed by an optical resonator. A pump wave serves as the input that generates in the nonlinear medium a signal and a idler wave via spontaneous parametric down conversion. The optical resonator is designed such that it provides feedback for either the signal wave, the idler wave, or both waves. This feedback significantly reduces the required intensity of

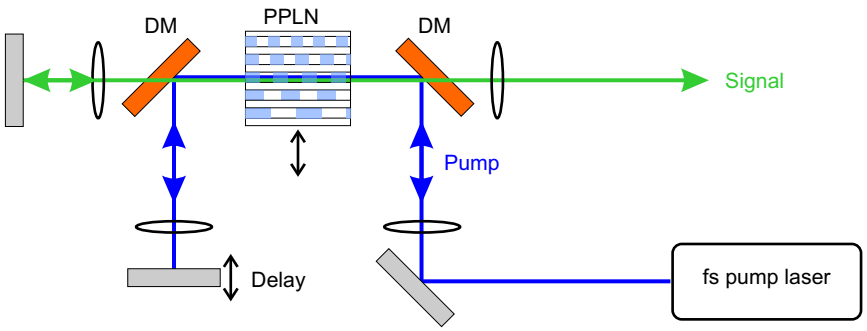


Figure 9.13: Scheme of of a tuneable fs-OPG. DM: Dichroic mirror. PPLN: Periodically poled $LiNbO_3$ crystal.

the pump wave to produce an appreciable signal and idler intensity. Hence, OPOs can be used in combination with fs-oscillator systems whereas OPAs and OPGs typically use amplified laser systems as pump sources. Frequency tuning is achieved in an OPO by varying the phase matching condition.

Figure 9.14: Scheme of the optical parametric oscillation process.

10 Photo detectors

A photo detector is a device that converts the incident optical energy in a measurable signal, e.g., a voltage signal or an electric current. Photo detectors can be subdivided into two main classes:

- **P**hoton detectors are based on the photoelectric effect. The absorption of a photon by the detector material leads to the emission of a free electron or the creation of an electron-hole pair. The measurement signal is proportional to the number of absorbed photons. Since the photoelectric effect requires that the incident photon has a material specific minimum energy, photon detectors exhibit a photon energy cutoff below which its response quickly vanishes.
  Examples: Photo multilplier tubes, photo diodes, photo conductors.

- In a **t**hermal detector the photon energy is converted into heat. The resulting temperature change is measured and converted into an electric signal which is proportional to the absorbed optical power. Thermal detectors usually have a broad spectral response.
  Examples: Bolometer, pyroelectric detectors.

10.1 Detector characteristics

10.1.1 Responsivity $\mathcal{R}$

The responsivity $\mathcal{R}$ relates the electric output of the photo detector (electric current or voltage) to the optical power P_L incident on the device. If the output of the photo detector is a voltage signal v, the responsivity is given by

$$\mathcal{R} = \frac{v}{P_L}, \quad with \quad [\mathcal{R}] = \mathrm{V}\,W^{-1}. \tag{10.1.1}$$

For a photo detector that delivers an electric current i as output signal, the responsivity is defined as

$$\mathcal{R} = \frac{i}{P_L}, \quad with \quad [R] = \mathrm{A}\,W^{-1}. \tag{10.1.2}$$

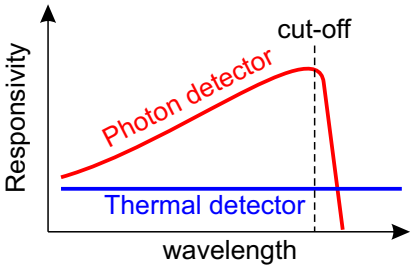


Figure 10.1: Typical spectral responsivity of a photon detector (red curve) and a thermal detector (blue curve).

10.1.2 Quantum efficiency η

The photon flux Φ (photons per second) of a monochromatic light beam can be calculated as:

$$\Phi = \frac{P_L}{\hbar\omega}, \tag{10.1.3}$$

where ω is the angular frequency and P_L is the optical power of the beam.

Let us first consider an ideal photon detector. Here, we assume that every photon impinging on the detector releases one free electron (or creates one electron-hole pair). The corresponding electric current i at the output of an ideal photon detector is hence given by

$$i_{ideal} = e\Phi. \tag{10.1.4}$$

In a real photon detector, the probability that an incident photon releases an electron is smaller than one. The ratio of the release rate of electrons r_e and incident photon flux Φ defines the quantum efficiency η of the detector:

$$\eta = \frac{r_e}{\Phi}. \tag{10.1.5}$$

With this definition, the responsivity of the photon detector can thus be expressed as:

$$\mathcal{R} = \frac{i}{P_L} = \frac{er_e}{\Phi\hbar\omega} = \eta\frac{e}{\hbar\omega}. \tag{10.1.6}$$

Example - Typical Si photo diode (Thorlabs FDS02):
$\mathcal{R}_{800\,\mathrm{nm}} = 0.45\,\mathrm{A\,W^{-1}} \Rightarrow \eta_{800\,\mathrm{nm}} \approx 0.7$

10.1.3 Signal-to-noise ratio SNR and noise equivalent power NEP

The measurement signal $s(t)$ at the output of any photon detector is always superimposed by random noise. For the following discussion, it is convenient to introduce the generalized power of a signal as the square of the measurement signal:

$$P(t) = s^2(t). \tag{10.1.7}$$

The generalized power can be interpreted as the power dissipated by a load resistance of 1Ω attached to the detector.

The signal-to-noise ratio SNR of the measurment signal is defined as

$$SNR = \frac{\mathrm{P}_{signal}}{\mathrm{P}_{noise}}, \tag{10.1.8}$$

where P_{signal} is the generalized power of the desired signal and P_{noise} is the generalized power of the noise.

The noise equivalent power NEP is the optical power P_L required to achieve a signal-to-noise ratio of one in a one hertz output bandwidth. It characterizes the sensitivity of photo detector or detector system. A smaller NEP corresponds to a more sensitive detector. In an alternative definition used by some manufacturers the noise equivalent power is specified in units of $\mathrm{W}/\sqrt{\mathrm{Hz}}$.

Example- Typical Si-photo diode (Thorlabs FDS02): $NEP = 9.3 \times 10^{-15}\mathrm{W}/\sqrt{\mathrm{Hz}}$.

10.1.4 Linearity and dynamic range

In the linear detection range, the responsivity of the detector does not depend on the optical power P_L. However, for optical powers larger than the saturation power P_L^{sat} the response of the detector saturates and the responsivity of the detector is no longer a constant. Thus, the useful measurement range of a given photo detector is within the limits $NEP \leq P_L \leq P_L^{sat}$.

The dynamic range of the detector (in dB) is given by

$$DR = 10\ \log_{10} \frac{P_L^{sat}}{NEP}. \tag{10.1.9}$$

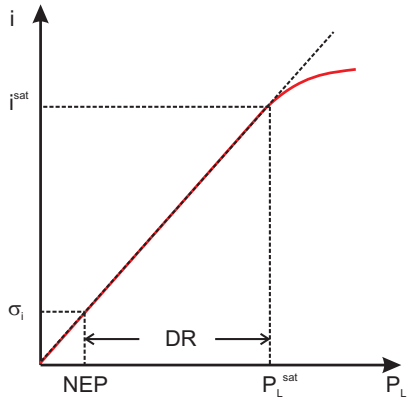


Figure 10.2: Typical response of a phuto detector as a function of the incident optical power.

10.2 Some types of photon detectors

10.2.1 Photomultiplier tubes

A photomultiplier tube (PMT) consists of an evacuated glass housing that contains a photocathode followed by several dynodes and an anode. The electrodes are interconnected such that each electrode is held at larger positive potential than the previous one. An incident photon can release an electron from the photocathode if the photon energy $\hbar\omega$ is larger than the material specific work function W of the photocathode material. The emitted electron is accelerated towards the first dynode. There, it leads to the emission of several secondary electrons. These electrons are accelerated towards the second dynode where they create for their part further secondary electrons, and so on. Hence, the number of electrons increases from dynode stage to next.

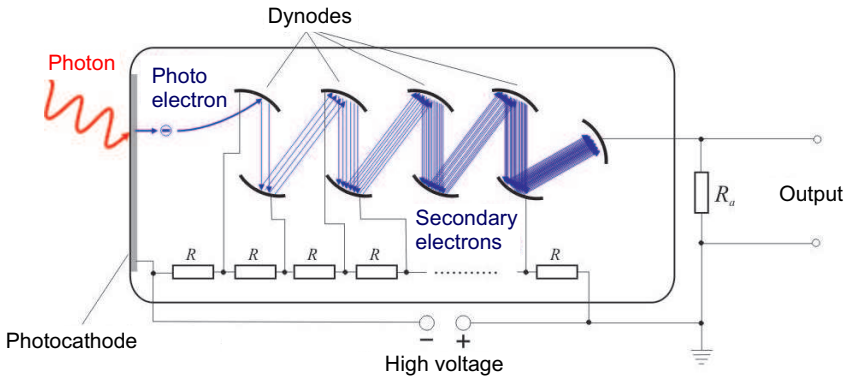


Figure 10.3: Scheme of a PMT (Wikipedia, modified).

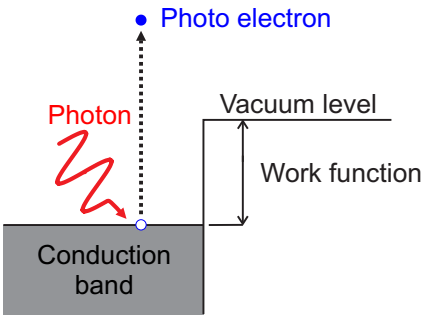


Figure 10.4: Photoelectric effect.

The spectral response of a given PMT is determined by the material of the photocathode and that of the window. Some typical photocathode materials are:

- Solar blind photocathodes (UV only): Cs-Te, Cs-I
- UV-Vis photocathodes: Cs-Sb
- Vis-NIR photocathodes: Ag-O-Cs, GaAs (Cs)

The quantum efficiency of the photocathode depends on the chosen material and on the photon energy. Typical values range in the order of $\eta = 0.01 \cdots 0.45$.

The total current gain of the PMT is defined as the ratio of the anode current I_a and the cathode current I_k

$$G = \frac{I_\mathrm{a}}{I_\mathrm{k}}. \tag{10.2.1}$$

For a PMT with N dynodes and an amplification factor δ per dynode, G can be calculated as

$$G = \delta^N. \tag{10.2.2}$$

A typical PMT has $N = 9$ to 12 dynodes and the amplification factor ranges between $\delta = 3$ to 8.

Example - PMT with $N = 9$ dynodes and amplification factor $\delta = 5$
$\Rightarrow G \approx 2 \times 10^6$.

The responsivity $\mathcal{R}$ of a PMT is given by

$$\mathcal{R} = \frac{\eta G e}{\hbar \omega}. \tag{10.2.3}$$

PMTs can be operated in different modes:

- For low light intensities, PMTs are usually operated in the so-called photon counting or Geiger mode. Here, the voltage between the dynodes is chosen such that a large current gain G is achieved. Incident photons give rise to individual current pulses which are processed by a counting electronics.

- The so-called current mode is used for moderate light intensities. In this case, the PMT is operated with a small current gain G and a large load resistance R_l.

10.2.2 Photoconductors

The density of mobile charge carriers (electrons in the conduction band and holes in the valence band) in a semiconductor can be increased by the absorption of light. The minimum photon energy required for this process depends on the type of semiconductor. In the case of an intrinsic semiconductor, the photon energy required to generate an electron-hole pair must be larger than the band-gap energy. For doped semiconductors, the photon energy can be significantly smaller as carriers are promoted from impurity levels within the band gap to the conduction band (n-doped semiconductor) and the valence band (p-doped semiconductor), respectively. The photogenerated carriers alter the conductivity of the semiconductor (see below). This effect can be used to measure the corresponding photon flux Φ.

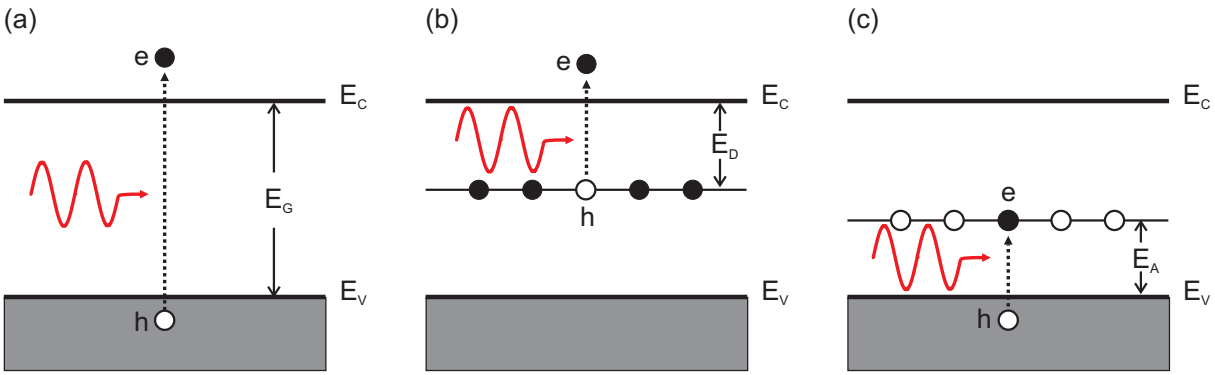


Figure 10.5: Generation of mobile charge carriers in a semiconductor due to light absorption. (a) Intrinsic semiconductor, (b) n-doped semiconductor,(c) p-doped semiconductor.

When an external electric field $\boldsymbol{\mathcal{E}}$ is applied to a semiconductor crystal, the free conduction band electrons are accelerated by this field. After a typical time τ_e, the electrons scatter with a phonon or a lattice impurity and are abruptly decelerated. The interplay of acceleration and deceleration results in a drift motion of the electrons. To calculate the

drift velocity, we start with the classical equation of motion of an electron:

$$m_e^* \dot{\mathbf{v}} + \frac{m_e^*}{\tau_e}\mathbf{v} = -q\boldsymbol{\mathcal{E}}. \tag{10.2.4}$$

Here, m_e^* is the effective mass of the electron. In steady state ($\dot{\mathbf{v}} = 0$), the average electron velocity is the drift velocity

$$\mathbf{v}_{d,e} = \frac{-q\tau_e \boldsymbol{\mathcal{E}}}{m_e^*}. \tag{10.2.5}$$

The electron mobility μ_e is defined by

$$\mu_e = \frac{q\tau_e}{m_e^*}. \tag{10.2.6}$$

Equation (10.2.5) can thus be written as

$$\mathbf{v}_{d,e} = -\mu_e \boldsymbol{\mathcal{E}}. \tag{10.2.7}$$

So far, we have considered only the free electrons. The transport of the free holes can be treated analogously. The total current density results from the motion of both electrons and holes. It is given by:

$$\mathbf{j} = q\left(n_e \mathbf{v}_{d,e} + n_h \mathbf{v}_{d,h}\right) = q\left(n_e \mu_e + \mu_h b_h\right)\boldsymbol{\mathcal{E}}. \tag{10.2.8}$$

Here, n_e and n_h are the densities of free electrons and holes, respectively. According to Ohm's law, the current density j and the electric field $\boldsymbol{\mathcal{E}}$ are related through

$$j = \sigma \boldsymbol{\mathcal{E}}, \tag{10.2.9}$$

where σ is the conductivity. A comparison of equation (10.2.8) and (10.2.9) yields

$$\sigma = q\, n_e\, \mu_e + q\, n_h\, \mu_h. \tag{10.2.10}$$

The total charge carrier density of electrons n_e (holes n_h) is given by the sum of the charge carrier density of electrons $n_{e,0}$ (holes $n_{h,0}$) in the absence of optical illumination and the density of photogenerated excess electrons Δn_e (holes Δn_h):

$$n_e = n_{e,0} + \Delta n_e, \tag{10.2.11}$$
$$n_h = n_{h,0} + \Delta n_h. \tag{10.2.12}$$

The photogenerated free charge carriers change the conductivity of the semiconductor by

$$\Delta\sigma = q(\Delta n_e \mu_e + \Delta n_h \mu_h). \tag{10.2.13}$$

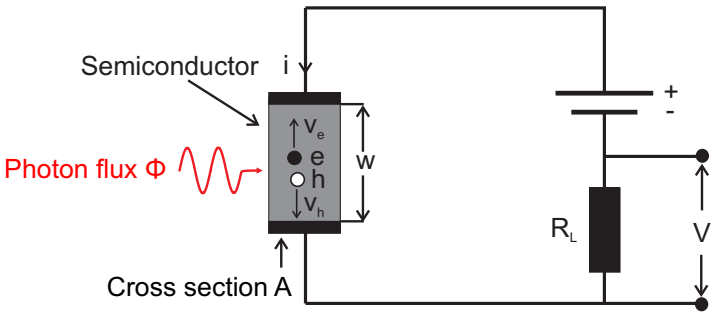


Figure 10.6: Photoconductive detector.

In the following, we consider an intrinsic semiconductor, for which the condition $\Delta n_e = \Delta n_h$ holds. The generation rate G of electron-hole pairs in the crystal with volume $V = wA$ (see fig.10.6) is given by

$$G = \eta\Phi/wA, \tag{10.2.14}$$

where η is the fraction of the incident photons that generate an electron hole-pair and Φ is the photon flux.

The recombination rate of excess charge carries is given by

$$R = \frac{\Delta n}{\tau_R}, \tag{10.2.15}$$

where τ_R is the excess-carrier recombination lifetime.

Under steady-state conditions, the two rates balance each other ($G = R$) and we obtain:

$$\Delta n = \frac{\eta\Phi\tau_R}{wA}. \tag{10.2.16}$$

Inserting Δn in equation (10.2.13) yields

$$\Delta\sigma = q\Delta n(\mu_e + \mu_h) = \frac{q\eta\tau_R(\mu_e + \mu_h)}{wA}\Phi. \tag{10.2.17}$$

The total current i_t can be written as the sum of the current i_0 in the absence of optical illumination and the photo current i_p. The latter is given by

$$i_p = \frac{q\eta\tau_R(v_{d,e} + v_{d,h})}{w}\Phi. \tag{10.2.18}$$

The electron transit time across the sample can be calculated as

$$\tau_{t,e} = w/v_{d,e}. \tag{10.2.19}$$

For many semiconductors, the condition $v_{d,e} \gg v_{d,h}$ holds. The photo current can thus be written as

$$i_p \approx q\eta \frac{\tau_R}{\tau_{t,e}} \Phi. \tag{10.2.20}$$

10.2.3 Photodiodes

Reminder: The p-n junction

In what follows, we consider a p-n junction, i.e., a semiconductor crystal which is doped on the left half with acceptors (p-semiconductor) and on the right half with donors (n-semiconductor). In a thought experiment, we start with the two separate halves. In the n-doped halve, the electrons are the majority carriers and the holes are the minority carriers, i.e., the density of electrons is much larger than the density of holes. The electrons and holes take on the opposite roles in the p-doped halve.

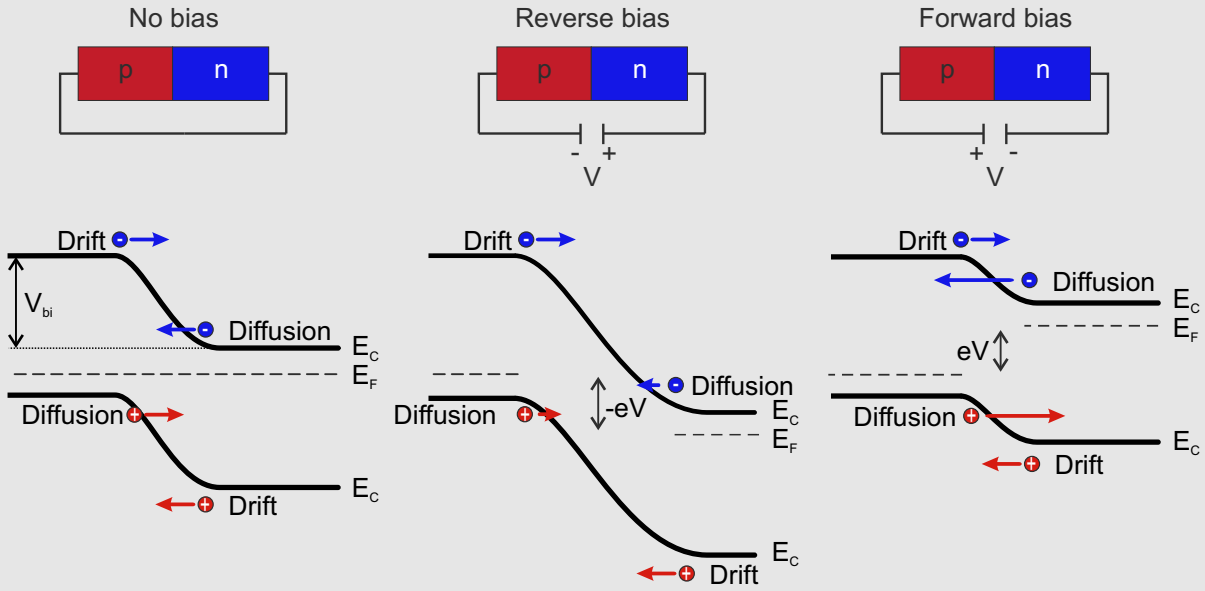
If we combine the two halves, the large carrier concentration gradients of the majority carriers lead to diffusion. The electrons diffuse from the n-side to the p-side and recombine there with holes. Likewise, holes diffuse from the p-side to the n-side and recombine with electrons. As a result, a double layer of uncompensated negative acceptors and positive donors forms near the junction. This space-charge zone gives rise to the so-called built-in voltage V_{bi} that counteracts the diffusion.

In thermal equilibrium, the net current flow of electrons and holes across the junction vanishes. Both for electrons and holes, the drift current caused by V_{bi} is compensated by the diffusion current that results from the respective carrier concentration gradient.

Next, we consider the effect of an external voltage V on the p-n junction. Because of the depletion of free carriers, the space-charge zone has a considerable larger resistance than the rest of the semiconductor crystal. Hence, we can assume that the potential drop across the space-charge zone is equal to the externally applied voltage.

The electron drift currents results from the minority carriers, i.e, electrons in the p-doped semiconductor and holes in the n-doped semiconductor, which have been thermally generated within or near the space-charge zone and which move under the influence of V_{bi} to the other side. Both of these currents are largely independent of the externally applied voltage. The situation is different for the diffusion currents which stem from the majority carriers. Since the majority carriers have to move against the potential, only the fraction $\exp -e(V_{bi} - V)/k_B T$ (Boltzmann-factor) of the majority carriers can overcome

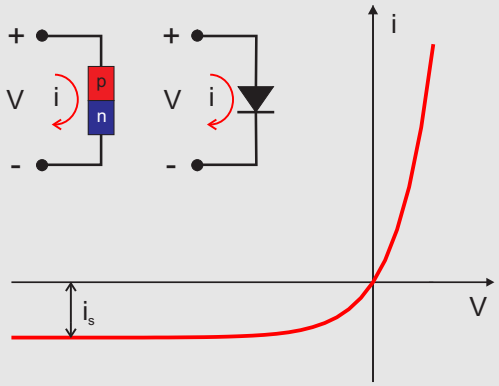
the potential and reach the other side. There they recombine with the corresponding majority carriers.



The combination of these effects gives rise to a current-voltage relation of the form

$$i(V) = i_s \left(e^{\frac{eV}{k_B T}} - 1 \right). \tag{10.2.21}$$

Here, i_s is the so-called saturation current.



Principle of operation

The absorption of light by a semiconductor can lead to the creation of electron-hole pairs. In doped semiconductors, this process usually hardly changes the concentration of the majority carriers. However, the concentration of the minority carriers can be significantly increased through optical illumination. This effect is used in photodiodes to measure the incident photon flux.

Consider the p-n junction depicted in Fig. 10.7. If an electron-hole pair is generated in the depletion region of the p-n junction, or one diffusion length away from it, the respective minority carrier is transported by the built-in electric field of the depletion region to the other side. Hence, holes generated in the n-doped region move to the p-doped halve, and electrons generated in the p-doped region to the n-doped halve. As a result, the drift currents increase under optical illumination. Since the concentrations of the majority carriers are hardly affected by the optical illumination, the diffusion currents are not significantly altered.

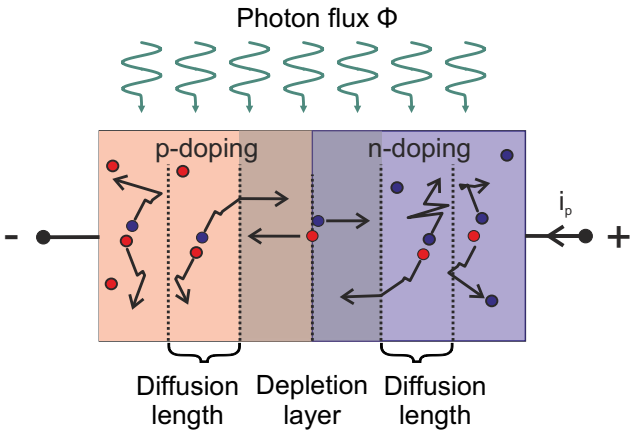


Figure 10.7: Illuminated p-n junction operated under reverse bias conditions.

The total current through the photodiode is the sum of the dark current (current that is generated in the absence of light) and the photo current. The current-voltage relation of the photodiode is thus given by

$$i = i_s \left[\exp\left(\frac{qV}{k_B T} \right) - 1 \right] - i_p. \tag{10.2.22}$$

The photo current i_p is proportional to the number of generated electron-pairs in the depletion region and hence to the incident photon flux Φ.

Photodiodes can be used in different modes of operation:

- In the **p**hotoconductive mode a reverse bias is applied to the photodiode (third quadrant of the current-voltage relation). The current through the diode is a linear

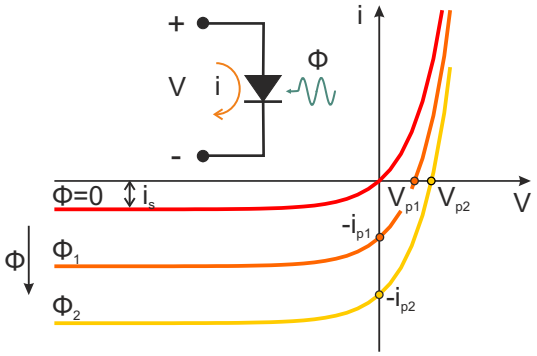


Figure 10.8: Current-voltage relation of an illuminated p-n junction..

function of the photon flux Φ. The reverse bias increases the width of the depletion layer which leads to a larger photosensitive volume. At the same time, the capacitance of the p-n junction is decreased such that the response time is reduced.

- In the **s**hort-circuit mode ($V = 0$), the current through the photodiode is the photo current i_p.
- In the **p**hotovoltaic mode no bias is applied to the p-n junction. The photogenerated carries give rise to a terminal voltage V_p (fourth quadrant of the current-voltage relation), which is a nonlinear function of the photon flux Φ. Solar cells are used in this mode.

The spectral responsivity of a photodiode depends on the semiconductor material it is made of. The following table summarizes the useable wavelength range for some typical semiconductors:

Semiconductor	Wavelength range (nm)
Gallium phosphide	$150 - 550$
Silicon	$190 - 1100$
Germanium	$400 - 1700$
Indium gallium arsenide	$800 - 2600$

Avalanche photodiodes

When a photo diode is operated with a large reverse bias, the electric field in the depletion layer can be so large that some of the photogenerated electrons and holes gain a kinetic energy larger than the band gap energy. These carriers can create new electron-hole pairs by impact ionization. These newly generated free carriers are themselves also accelerated

by the electric field. Consequently, they and the original carrier can potentially give rise to further electron-hole pairs, and so. This so-called avalanche effect acts as a gain mechanism for the photo current, i.e., one incident photon creates several charge carriers. Photodiodes which use this effect are called avalanche photodiodes (APDs).

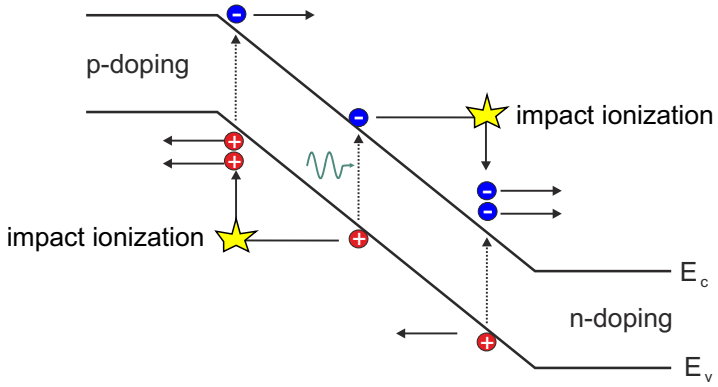


Figure 10.9: Avalanche effect in a strongly reverse-biased p-n junction.

APDs are highly sensitive photodetectors. For example, Si-APDs operated with a reverse bias of -100 V typically exhibit a current gain factor $G \approx 100$.

10.3 Noise in photodetection

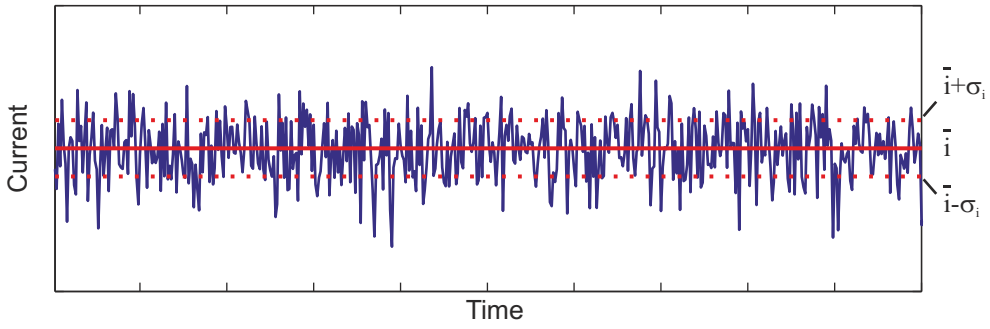


Figure 10.10: Photo current with Gaussian noise.

The photo current generated in a photo detector $i(t)$ always shows random fluctuations (noise) $i_n(t)$ around its mean value $\bar{i}$:

$$i(t) = \bar{i} + i_n(t), \tag{10.3.1}$$

with

$$\bar{i} = \frac{1}{T}\int_{-T/2}^{T/2} i(t)\,dt. \tag{10.3.2}$$

Here, T is the temporal period for averaging.

The temporal average of the noise is zero:

$$\overline{i_n} = \frac{1}{T}\int_{-T/2}^{T/2} i_n(t)\,dt = 0. \tag{10.3.3}$$

In the following, we use the standard deviation of the current σ_i as a measure of noise:

$$\sigma_i = \left[\,\overline{\left(i-\bar{i}\right)^2}\,\right]^{1/2} = \left[\,\overline{i_n^2}\,\right]^{1/2}. \tag{10.3.4}$$

The generalized power associated with the current is given by

$$P(t) = i^2(t). \tag{10.3.5}$$

The corresponding average power $\overline{P}$ can be calculated as

$$\begin{aligned}\overline{P} &= \frac{1}{T}\int_{-T/2}^{T/2} P(t)\,dt = \frac{1}{T}\int_{-T/2}^{T/2} i^2(t)\,dt \\ &= \frac{1}{T}\int_{-T/2}^{T/2} \left[\bar{i}+i_n(t)\right]^2\,dt \\ &= \overline{P_s}+\overline{P_n}, \end{aligned} \tag{10.3.6}$$

where $\overline{P_s} = \bar{i}^{\,2}$ is the average signal power and $\overline{P_n} = \overline{i_n^2} = \sigma_i^2$ is the average noise power.

For several independent noise sources, the average noise power can be calculated as:

$$\overline{P_n} = \overline{P_{n,1}} + \overline{P_{n,2}} + \cdots \tag{10.3.7}$$

The photo current usually fluctuates on different time scales. Therefore, it is useful to consider the noise as a function of frequency:

$$i_n(f) = \int_{-T/2}^{T/2} i_n(t) e^{i2\pi f t} dt. \tag{10.3.8}$$

The noise power can be expresses as

$$\begin{aligned}\overline{P_n} &= \frac{1}{T}\int_{-T/2}^{T/2} i_n^2(t)\,dt = \frac{1}{T}\int_{-T/2}^{T/2}\left[i_n(t)\int_{-\infty}^{\infty} i_n(f)e^{-i2\pi f t}df\right]dt \\ &= \frac{1}{T}\int_{-\infty}^{\infty} i_n(f) i_n(-f) df. \end{aligned} \tag{10.3.9}$$

Since $i_n(t)$ is a real quantity, the relation $i_n(-f) = i_n^*(f)$ holds.

$\overline{P_n}$ can thus be written as

$$\overline{P_n} = \int_0^\infty S_n(f) df, \tag{10.3.10}$$

with the spectral density function

$$S_n(f) = \frac{2|i_n(f)|^2}{T}. \tag{10.3.11}$$

$S_n(f)\Delta f$ is the portion of the noise power that results from a frequency band of width Δf centered around the frequency f. Equation (10.3.10) shows that we can reduce the noise power by limiting the band width of the detector electronics with an appropriate electronic filter.

10.3.1 Sources of noise

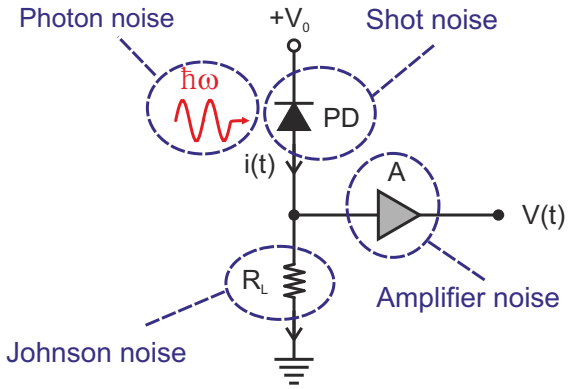


Figure 10.11: Scheme of a typical photo detector circuit.

Photon noise

Photon noise is a result of the quantum nature of light. In the following, we consider an ideal coherent light source[1] that gives rise to a beam with average power $\overline{P}_L$.

[1] A laser emitting a monochromatic light beam in a single mode can - to a good approximation - be considered as such an ideal light source.

The probability $p(m)$ that m photons emitted by a coherent light source impinge on the detector in the time interval T is given by the Poisson distribution:

$$p(m) = \frac{\overline{m}^m \exp(-\overline{m})}{m!}. \tag{10.3.12}$$

Here,

$$\overline{m} = \frac{\overline{P}_L T}{h\nu} \tag{10.3.13}$$

is the mean number of photons detected in the time interval T.

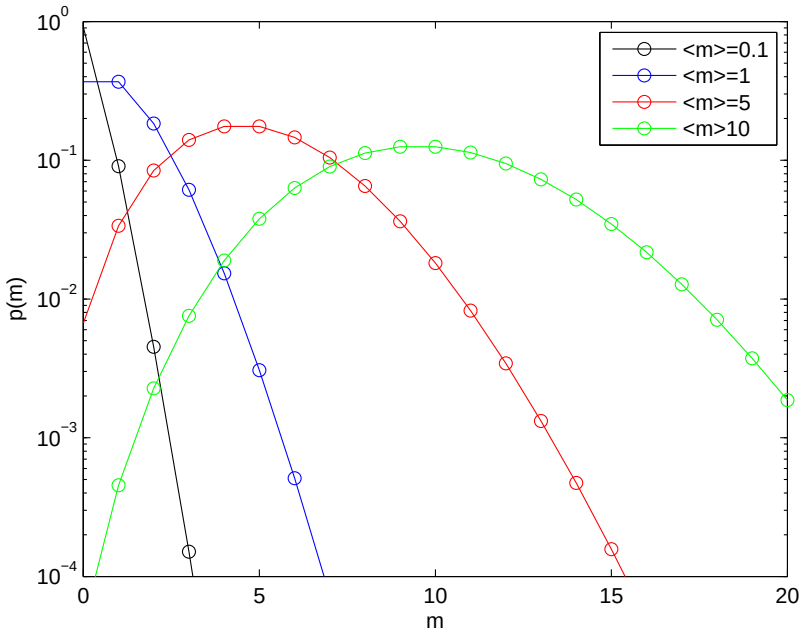


Figure 10.12: Poisson distribution $p(m)$ of the photon number m. The curves are guides to the eye.

The standard deviation of the photon number is given by

$$\sigma_m = \sqrt{\sum_{m=0}^{\infty} (m - \overline{m})^2 \, p(m)} = \sqrt{\overline{m}}. \tag{10.3.14}$$

The relative uncertainty of the photon number of a coherent light source thus scales as:

$$\frac{\sigma_m}{\overline{m}} = \frac{1}{\sqrt{\overline{m}}}. \tag{10.3.15}$$

Shot noise

In the following, we discuss the response of a photo detector to an incident coherent optical signal with optical power P_L. Let us assume that the detector has an an electrical response bandwidth $\Delta f = B$. The corresponding integration time is $T = 1/2B$.

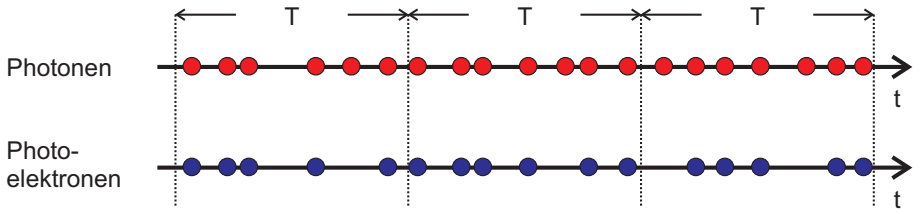


Figure 10.13: Creation of photo electrons by an incident stream of photons.

The probability that an incident photon creates a photo electron in the detector is given by the quantum efficiency η_e. Since the photon number m_γ follows a Poisson distribution, the same also holds for the number of photo electrons m_e, with

$$\overline{m_e} = \eta_e \overline{m_\gamma}. \tag{10.3.16}$$

Analogous to equation (10.3.14), we find that the standard deviation of the number of photo electrons is given by

$$\sigma_{m_e} = \sqrt{\overline{m_e}}. \tag{10.3.17}$$

Next, we calculate the mean value of the photo current as

$$\overline{i_p} = \frac{q\overline{m_e}}{T} = 2qB\eta_e\overline{m_\gamma}. \tag{10.3.18}$$

The corresponding standard deviation is given by

$$\sigma_{i_p} = \frac{q\sigma_{m_e}}{T} = 2qB\sigma_{m_e}. \tag{10.3.19}$$

With this result, we can finally express the average noise power as

$$\overline{P_{n,shot}} = \sigma_{i_p}^2 = 4q^2B^2\sigma_{m_e}^2 = 2qB\overline{i_p}. \tag{10.3.20}$$

From the inspection of the last equation we can draw two conclusions:

- The average noise power is proportional to the mean value of the photo current.
- The noise is frequency independent (white noise). The corresponding spectral density is given by the Schottky formula:

$$S_{n,shot}(f) = 2q\overline{i_p}. \tag{10.3.21}$$

Dark noise

A photo detector produces a small electric current even without incident optical signal. This so-called dark current i_d results from the excitation of electron-hole pairs by the thermal radiation field. According to the Schottky formula, the average noise power associated with the dark current is given by

$$\overline{P_{n,dark}} = 2qB\overline{i_d}. \tag{10.3.22}$$

The dark current can be often reduced through cooling of the detector.

Excess noise

The inherent amplification process of PMTs and APDs is of statistical nature and thus leads to excess noise. Let us assume that on average every absorbed photon creates $\overline{G}$ charge carriers and that the standard deviation of the amplification process is σ_G. The average value of the photo current then reads

$$\bar{i} = q\overline{G}\eta_e\Phi. \tag{10.3.23}$$

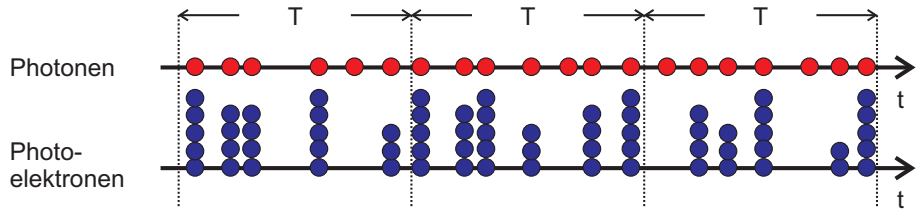


Figure 10.14: In APDs and PMTs, every photon gives rise to a random number of charge carriers.

The corresponding average noise power can be calculated as

$$\overline{P_n} = \overline{i_n^2} = \sigma_i^2 = 2q\overline{G}\bar{i}BF, \tag{10.3.24}$$

where B is the band width of the detection device and F is the so-called excess noise factor

$$F = 1 + \frac{\sigma_G^2}{\overline{G}^2}. \tag{10.3.25}$$

Johnson noise

Johnson noise (also called thermal noise) is a result of the random thermal motion of charge carriers (usually the electrons) inside an electrical conductor. This motion leads

to a fluctuating current i even if no voltage is applied. Since the thermal motion has no preferential direction, one finds

$$\bar{i} = 0. \tag{10.3.26}$$

It can be shown[2] that the mean square value of the thermal noise current for a circuit with resistance R and bandwidth B is given by

$$\overline{i^2_{n,Johnson}} = \frac{4k_BTB}{R}. \tag{10.3.27}$$

In the case of an optical detection system, the resistance R is the total resistance resulting from the internal resistance of the detector and the load resistance at the output of the detector.

[2]See e.g. A. Yariv and P. Yeh, *P*hotonics -Optical electronics in modern communications, sixth edition, Oxford University Press (2007).