Chapter 3

Physics of electromagnetic waves

3.1 Introduction: Basic optical concepts

This chapter reminds some basic principles of electrodynamics which are relevant for the geometric optics and wave optics of astronomical telescopes and instruments. The following treatment is mainly based on the classical textbook of Born and Wolf, "Principles of Optics" where a comprehensive description of the theory can be found.

3.1.1 Geometric optics

Geometric optics is a simplification of the wave optics which assumes that the wavelength is much smaller than the geometric dimensions of the optical system ($\lambda \to 0$). Geometric optics is a valuable tool for the treatment of light propagation (light rays), reflection and refraction. But geometric optics cannot describe polarization, diffraction and interference effects.

Reflection from a flat surface. The incident (i) and reflected (r) rays and the surface normal $\vec{z}$ lie in the same plane. The incident and reflected rays have equal angles (but opposite sign) with respect to $\vec{z}$:

$$\theta_r = -\theta_i \,. \tag{3.1}$$

In general not all light is reflected from a surface. For example, dielectric surfaces like glass reflect only a few per cent because most of the light is transmitted, or an aluminum mirror reflects in the optical only about 90 %. The reflectivity depend also on wavelength.

Refraction and refractive index. Refraction occurs at the boundary between two media with different refractory indices $n_1 \neq n_2$. The index n_i can be defined by

$$n_i = \frac{c}{v_i} \,, \tag{3.2}$$

where $v_i = c_n$ is the light travelling speed in the medium and c the speed of light in vacuum.

Some important properties of the refractive index:

- $n \geq 1$; for example $n = 1.000294$ for air, $n = 1.33$ for water, and $n = 1.4 - 1.7$ for many glasses ($\lambda = 550$ nm),
- n depends on color; $n = n(\lambda)$,
- n is for amorphous materials, like air, water, or amorphous glass, independent of the direction (or isotropic)

– many crystals are **birefringent**, this means that the refraction index depends on the propagation direction (e.g. $n_x \neq n_y \neq n_z$).

Snell's law. Light falling onto an interface between two media with n_1, n_2 at an angle of incidence θ_i to the normal $\vec{z}$ is refracted in the second medium at an angle θ_t (t for transmitted) according to:

$$n_2 \sin\theta_t = n_1 \sin\theta_i \,, \tag{3.3}$$

or the propagation direction is closer to the surface normal in the denser medium.

Total internal reflection. If the light is propagating from the denser medium to the less dense medium $n_1 > n_2$, then there exists a critical maximum angle $\theta_i^{\max}$ where $\theta_t = 90°$ or $\sin\theta_t = 1$. This angle is given according to Eq. (3.3)

$$\sin\theta_i^{\max} = n_2/n_1 \quad \text{for} \quad n_1 > n_2 \,. \tag{3.4}$$

For incidence angles $\theta_i > \theta_i^{\max}$ all light will be reflected back into the denser medium. This is called total internal reflection. Optical light fibers are based on this principle.

3.1.2 Wave optics

Wave optics provides a general, exact and comprehensive description of optical phenomena based on the theory of electromagnetic waves (hereafter also simply called "light"). With wave optics one can treat the propagation of light, interaction of light with a medium, light polarization, diffraction, and interference effects.

Huygens' wavefronts. Christian Huygens proposed in the 17$^{\text{th}}$ century basic principles for the propagation of waves which are also valid for electromagnetic waves:

– As a wave travels, each point along its path makes the same periodic disturbance, but later in time for points further away from the source.

– Each point can be considered as a source of a new spherical wavelet. The wavelets are in phase on a sphere which envelopes them. The **wavefront** is the locus where the wavelets have exactly the same phase of oscillation.

– The wavefronts propagate along straight (radial) lines which are called **rays**.

– For an opaque obstacle secondary wavelets spread round the edge (diffraction).

Interference. Wave interference is another cornerstone of wave optics. Electromagnetic waves are harmonic waves with periodic oscillation along a ray. The oscillations of two waves at a given point are superposed or added. An interference will happen if there are at a given point two coherent waves (= same wavelength) with the same or comparable amplitude a and a constant phase shift over a long time

- if the waves differ in phase by $2m\pi$ then they are in phase and the resulting oscillation has an amplitude of $2a$,
- if the phase difference is $(2m + 1)\pi$ then they are in anti-phase and the resulting amplitude is zero,
- in general the amplitude lies between these two extreme cases.

3.2 Maxwell's Equation

The properties of electromagnetic waves can be deduced from Maxwell's equations. For this we use Maxwell's macroscopic equations which are also called Maxwell's equations in matter. They "factor out" the bound charges and currents so that the resulting equations depend only on the free charges and currents. This requires that auxiliary fields must be introduced besides the electric field $\vec{E}$ and the magnetic field and $\vec{B}$ which depend on the medium (material).

$$\mathrm{curl}\,\vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}\,, \tag{3.5}$$

$$\mathrm{curl}\,\vec{H} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial \vec{D}}{\partial t} \tag{3.6}$$

$$\mathrm{div}\vec{D} = 4\pi\rho \tag{3.7}$$

$$\mathrm{div}\vec{B} = 0\,. \tag{3.8}$$

The used quantities are:

$\vec{E}$	the electric field
$\vec{B}$	the magnetic field
$\vec{D} = \vec{E} + 4\pi\vec{P}$	the electric displacement field in a medium, which is given by the electric field and the material dependent polarization density $\vec{P}$,
$\vec{H} = \vec{B} - 4\pi\vec{M}$	the magnetizing field in a medium, which can be expressed by the magnetic field and the total magnetization $\vec{M}$ in a material,
$\vec{j}$	the free current density,
ρ	the free charge density.

We use here the so-called Gaussian system of units:

- electrical quantities $\vec{E}$, $\vec{D}$, and ρ are given in electrostatic units,
- magnetic quantities $\vec{B}$ and $\vec{H}$ in electromagnetic units.

3.2.1 Material equations

A lot of complicated physics describing the interaction of an electromagnetic wave with a medium is hidden in the equations describing the polarization density $\vec{P}$ and the magnetization of the material $\vec{M}$.

For weak fields a linear approximation can often be made for $\vec{P}$ and $\vec{M}$. This condition is usually fulfilled for electromagnetic radiation from astronomical objects interacting with observing instruments:

$$\vec{P} \approx \chi_e \vec{E} \quad \text{and} \quad \vec{M} \approx \chi_m \vec{H}\,,$$

where χ_e and χ_m are the electric and magnetic susceptibilities. Introducing the electric permittivity $\epsilon = 1 + 4\pi\chi_e$, and the magnetic permeability $\mu = 1 + 4\pi\chi_m$ yields then simple linear relations between the electric field and the displacement field as well as the magnetic field and the magnetizing field:

$$\vec{D} \approx \epsilon \vec{E}\,, \tag{3.9}$$

$$\vec{H} \approx \frac{\vec{B}}{\mu}\,. \tag{3.10}$$

In additions there may be free charges which produce currents in a material. This property is defined by the conductivity σ and Ohm's law:

$$\vec{j} = \sigma \vec{E}\,. \tag{3.11}$$

Conductivity σ. We distinguish between conductors and insulators:

- $\sigma > 0$ are conductors, e.g. metals or "warm" semiconductors, but also ionic solutions or a plasma. Electromagnetic waves in conductors $\sigma > 0$ induce moving charges which cause dissipation due to the production of Joule heat. Therefore, conducting materials are not transparent. However metal surfaces are good reflectors and are widely used as mirrors in optics.
- $\sigma \approx 0$ are insulators or dielectrics. Non-absorbing dielectrics, like glasses, are very important materials for transmittive optics such as lenses.

Permittivity ϵ. The permittivity describes the ability of materials to transmit (or "permit") an electric field. On a microscopic scale the $\vec{D}$ field induces in an insulator charge migration and electric dipole reorientations. These effects act against the $\vec{D}$ field, so that the resulting electric field $\vec{E}$ is weakened $\vec{E} = \vec{D}/\epsilon$ with $\epsilon > 1$, except for vacuum where $\epsilon = 1$.

For real-world materials the induced electric effects cannot be described exactly with a simple linear law $\vec{D} = \epsilon \vec{E}$ but with more complicated function depending on various parameters: Examples are:

- *Dispersion* in materials depend on frequency. An electromagnetic wave passing through a material induces charge oscillations which are slightly out of phase (delayed) with respect to the driving electromagnetic field. The oscillating charges reradiate a wave but with a phase delay so that the wave travels slower $c_n = c/\sqrt{\epsilon}$ and has a reduced wavelength in the medium but the wave frequency is the same.
- *Absorption* depends on frequency and the wave is weakened when passing through the medium. Absorption means that the permittivity ϵ must be treated as a complex quantity including an absorption component.
- *Anisotropy*, such as birefringence or dichroism can be present in some materials which are often crystals. In this case the permittivity ϵ is described as a second-rank tensor

  $$D_j = \epsilon_{ij} E_i$$

  instead of a scalar as for an isotropic medium.

– *Bi-anisotropic materials* exist for which $\vec{D}$ depends on the $\vec{E}$ and $\vec{H}$ fields:

$$\vec{D} = \epsilon\vec{E} + \xi\vec{H}\,.$$

– *Spatial inhomogeneities* may be present because there are small spatial structures in the medium. Another case is the interaction of waves with a (magneto)-hydrodynamic medium. Also hysteresis effects may produce a heterogeneous structure in space and time.

Permeability μ. The permeability is the ability of a material to support the formation of a magnetic field within itself in response to an applied $\vec{B}$ field. In a microscopic picture such a material rearranges the magnetic dipoles. Materials with $\mu > 1$ are called paramagnetic (e.g. platinum) and with $\mu < 1$ diamagnetic (e.g. bismuth, copper). Magnetic materials show sometimes quite complicated effects which cannot be described with a simple linear μ-law, similar to the case of the electric permittivity. However, for most optical materials the magnetic permeability is practically unity

$$\mu \approx 1\,.$$

and we consider in this lecture only this special but quite common case.

3.3 Electromagnetic waves in a dielectric medium

We consider an isotropic, homogeneous, dielectric medium without free electric currents $\vec{j} = 0$ and charges $\rho = 0$, with a linear permittivity ($\vec{D} = \epsilon\vec{E}$) and a magnetic permeability $\mu = 1$ (or $\vec{H} = \vec{B}$). With these simplifications Maxwell's equations have the form:

$$\begin{aligned}
\mathrm{curl}\,\vec{H} &= \frac{\epsilon}{c}\frac{\partial\vec{E}}{\partial t}\,, && (3.12)\\
\mathrm{curl}\,\vec{E} &= -\frac{1}{c}\frac{\partial\vec{H}}{\partial t}\,, && (3.13)\\
\mathrm{div}\vec{E} &= 0\,, && (3.14)\\
\mathrm{div}\vec{H} &= 0\,. && (3.15)
\end{aligned}$$

For $\epsilon = 1$ we obtain the Maxwell equations for the vacuum.
These equations can be reduced, e.g. for $\vec{E}$ (and equivalent for $\vec{H}$)

$$\mathrm{curl}\,\mathrm{curl}\,\vec{E} = \underbrace{\mathrm{graddiv}\vec{E}}_{0} - \nabla^2\vec{E} = -\frac{1}{c}\,\mathrm{curl}\,\frac{\partial\vec{H}}{\partial t} = -\frac{1}{c}\frac{\partial\,\mathrm{curl}\,\vec{H}}{\partial t} = -\frac{\epsilon}{c^2}\frac{\partial^2\vec{E}}{\partial t^2}\,,$$

into separate differential wave equations for the electric field and the magnetic field:

$$\begin{aligned}
\frac{\partial^2\vec{E}}{\partial t^2} &= \frac{c^2}{\epsilon}\nabla^2\vec{E}\,, && (3.16)\\
\frac{\partial^2\vec{H}}{\partial t^2} &= \frac{c^2}{\epsilon}\nabla^2\vec{H}\,. && (3.17)
\end{aligned}$$

The equation include the speed of light $c = 299'782$ km/s and the electric permittivity ϵ (the dielectric material constant), which is the relevant optical property of "transparent" dielectric materials. The propagation speed of the wave is given by

$$c_n = \frac{c}{\sqrt{\epsilon}} = \frac{c}{n}\,, \qquad (3.18)$$

where n is the refractive index.

Simple solutions for the wave equations

As a reminder we consider simple solutions of wave equations. We start with the 1-dimensional scalar wave equation:

$$\frac{\partial^2 \xi(x,t)}{\partial t^2} = c_n^2 \, \frac{\partial^2 \xi(x,t)}{\partial x^2}$$

where c_n is the propagation speed. The general solution to this equation is:

$$\xi(x,t) = f_1(x - c_n t) + f_2(x + c_n t)\,.$$

Harmonic 1-dim. wave. A special, but extremely useful solution (e.g. for electromatic waves travelling through space and non-absorbing materials) is the harmonic wave:

$$\xi(x,t) = a \cos\left(\frac{2\pi}{\lambda_n}(x - c_n t)\right) = a\cos\left(k_n x - \omega t\right),$$

where λ_n is the wavelength, $k_n = 2\pi/\lambda$ the wave constant, and $\omega = 2\pi\, c_n/\lambda = 2\pi\nu$ the phase velocity with ν as wave frequency. The material properties change the wavelength of the wave but not the frequency (or phase velocity). Thus the material property described by the refractive index n is included in the k_n-parameter, the wavelength λ_n or the propagation speed c_n.

Harmonic plane waves. The form of the 3-dimensional, scalar wave equation is analog to the 1.-dim. case:

$$\frac{\partial^2 \xi(\vec{r},t)}{\partial t^2} = c_n^2 \nabla^2 \xi(\vec{r},t)\,.$$

A special solution of the 3-dimensional scaler wave equation is the harmonic, plane wave:

$$\xi(\vec{r},t) = a\cos\left(\vec{k_n}\cdot\vec{r} - \omega t\right),$$

where $\vec{k_n} = k_n \cdot \vec{s}$ ($|\vec{s}| = 1$) is called the wave vector which points in the direction $\vec{s}$ of the wave propagation.

Harmonic waves as exponential functions. Calculations with harmonic waves can be simplified using exponential functions. A harmonic plane wave may be written as

$$\xi(\vec{r},t) = \mathcal{R}\{u(\vec{r})\mathrm{e}^{-i\omega t}\} \quad \text{with} \quad u(\vec{r}) = a(\vec{r})\,\mathrm{e}^{i(\vec{k}\vec{r}-\delta)}\,,$$

where $\mathcal{R}$ denotes the real part and $u(\vec{r})$ is called the complex amplitude. One can insert this into the 3-dim, scalar wave equation

$$\frac{\partial^2 \xi(\vec{r},t)}{\partial t^2} = (-i\omega)^2 u(\vec{r}) e^{-i\omega t} = c_n^2 \nabla^2 (u(\vec{r}) e^{-i\omega t})\,.$$

and factor out the time-component $\mathrm{e}^{-i\omega t}$. Thus, the complex amplitude satisfies the equation:

$$u(\vec{r}) = -\frac{c_n^2}{\omega^2}\,\nabla^2 u(\vec{r})$$

If the operations on ξ are linear then the symbol $\mathcal{R}$ can be dropped and one can operate with the complex function. The real part of the final expression is then the resulting physical quantity.

For non-linear operation, like squaring for the calculation of the wave intensity, one must take the real part and then operate with these alone.

Spherical harmonic waves. An interesting special case of the 3-dimensional wave equation is a spherical wave equation for which one can use $r = |\vec{r}| = \sqrt{x^2 + y^2 + z^2}$ as space variable. In this case the wave equation can be written as:

$$\frac{\partial^2 (r\cdot\xi(r,t))}{\partial t^2} = c_n^2 \frac{\partial^2 (r\cdot\xi(r,t))}{\partial r^2}\,.$$

The solution for a spherical, harmonic wave with the origin at $r = 0$ (for $t = 0$) is then:

$$\xi(r,t) = \frac{1}{r}\,a\cos\left(k_n r - \omega t\right).$$

One obtains the well-known result that the intensity of a spherical wave decreases proportional the square of the distance from the origin $\xi^2(r,t) \propto a^2/r^2$.

3.4 Description of electromagnetic waves

Important solutions for the electromatic wave equations (Eqs. 3.16, 3.17) are the vectorial, harmonic, plane waves for the electric and magnetic fields:

$$\vec{E}(\vec{r},t) = \vec{a_E} \cos\left(\vec{k_n} \cdot \vec{r} - \omega t + \delta_E\right), \tag{3.19}$$
$$\vec{H}(\vec{r},t) = \vec{a_H} \cos\left(\vec{k_n} \cdot \vec{r} - \omega t + \delta_H\right). \tag{3.20}$$

From Maxwell's equations follow also the relations

$$\sqrt{\epsilon}\,\vec{E} = -\vec{s} \times \vec{H} \quad \text{and} \quad \vec{H} = \vec{s} \times \sqrt{\epsilon}\,\vec{E} \quad \text{and} \quad \vec{E}\,\vec{s} = \vec{H}\,\vec{s} = 0$$

expressing that the three vectors $\vec{E}$, $\vec{H}$, and $\vec{s}$ form a right-handed orthogonal triad of vectors. Thus we can choose the z-axis in the propagation direction $\vec{s}$, so that there are only electric and magnetic field components in the x- and y-direction The end point of the electric and magnetic vectors is then described by:

$$E_x(z,t) = a_x \cos\left(k_n z - \omega t + \delta_x\right), \tag{3.21}$$
$$E_y(z,t) = a_y \cos\left(k_n z - \omega t + \delta_y\right), \tag{3.22}$$
$$H_x(z,t) = -\sqrt{\epsilon}\,E_y(z,t), \tag{3.23}$$
$$H_y(z,t) = \sqrt{\epsilon}\,E_x(z,t). \tag{3.24}$$

3.4.1 Polarization

An electromagnetic wave or a photon is polarized. In general a wave has an elliptical polarization. Special cases are waves with a linear polarization or a circular polarization. The polarization state of a wave is defined by the wave amplitudes a_x and a_y and the relative phase shift $\delta = \delta_y - \delta_x$) of the wave components $E_x(z,t)$ and $E_y(z,t)$:

linear polarization — the electric wave vector $\vec{E}$ oscillates in one plane
- phase shift: $\delta = m \cdot \pi$ with $m = 0, \pm 1, \pm 2, ...$
- orientation: $(-1)^m \arctan(a_y/a_x)$

circular polarization — $\vec{E}$ rotates around the z-axis with $|\vec{E}|$ = constant
- phase shift: $\delta = (m + 1/2)\,\pi$ with $m = 0, \pm 1, \pm 2, ...$
- e.g. right-handed: $\delta = +\pi/2$; left-handed: $\delta = -\pi/2$

elliptical polarization — $\vec{E}$ oscillates and rotates
- phase shift: $\delta \neq m \cdot \pi/2$

Unpolarized light. A thermal source emits so-called natural or unpolarized light. Unpolarized light stands for many electromagnetic waves with a randomly distributed polarization (elliptical, linear and circular polarization).

3.4.2 Intensity

Optical measurements usually measure the intensity of the radiation, which is the average value of the electromagnetic energy transmitted per unit time through a unit area perpendicular to the propagation direction. This energy is given by the pointing vector:

$$\vec{S} = \frac{c}{4\pi} \vec{E} \times \vec{H} \tag{3.25}$$

In the x, y, z-coordinate system the Pointing vector has only the S_z component given by:

$$S_z = E_x H_y - E_y H_x = (a_x^2 + a_y^2)\sqrt{\epsilon}\cos^2(\vec{k_n}\vec{s} - \omega t + \delta)\,. \tag{3.26}$$

where $E^2 = |\vec{E}|^2 = (a_x^2 + a_y^2) = H^2/\epsilon$. The temporal average relates the wave amplitudes with the pointing vector:

$$\vec{S} = \frac{c}{4\pi}\sqrt{\epsilon}\, E^2\, \vec{s}\,. \tag{3.27}$$

3.5 Reflection and refraction on dielectric interfaces

3.5.1 Boundary conditions

Electromagnetic waves are transmitted and reflected at a boundary between two dielectric media characterized by n_1 and n_2. At the interface the incident (i), reflected (r), and transmitted (t) electric and magnetic waves must be continuous. This defines the boundary conditions:

$$\vec{E}^{(i)}(\vec{r}, t) + \vec{E}^{(r)}(\vec{r}, t) = \vec{E}^{(t)}(\vec{r}, t)\,, \tag{3.28}$$
$$\vec{H}^{(i)}(\vec{r}, t) + \vec{H}^{(r)}(\vec{r}, t) = \vec{H}^{(t)}(\vec{r}, t)\,. \tag{3.29}$$

The reflection and transmission depend on the polarization of the incident ray, the incidence angle θ_i, and the refractive indices n_1 and n_2 (which define also the angle of the transmitted beam θ_t.

Each vector $\vec{E}$ and $\vec{H}$ is resolved into x-, y-, and z-components. The x-axis is defined by the dielectric interface and the reflection / transmission plane, the y-axis lies in the interface perpendicular to the reflection plane, and the z-axis is perpendicular to the interface. The wave vectors can be expressed by components in the x-z-plane or a perpendicular component in y-direction. For the electric vector this is equivalent to the parallel ($\|$ or p-component for "parallel") and perpendicular ($\perp$ or s-component for "senkrecht") polarization components of the electromagnetic wave.

Solving the boundary conditions yields the Fresnel formulae for the parallel amplitudes of the reflected and transmitted wave $r_\|$ and $t_\|$ as function of the initicence angle and equivalent for the perpendicular amplitudes $r_\perp$, $t_\perp$ as function of $a_\perp$.

Detailed derivation: The individual components for the incident, reflected and transmitted waves are:
Incident wave:

$$\begin{aligned} E_x^{(i)} &= -a_\parallel \cos\theta_i \cdot e^{-i\tau_i} & H_x^{(i)} &= -a_\perp \sqrt{\epsilon_1} \cos\theta_i \cdot e^{-i\tau_i} \\ E_y^{(i)} &= a_\perp \cdot e^{-i\tau_i} & H_y^{(i)} &= -a_\parallel \sqrt{\epsilon_1} \cdot e^{-i\tau_i} \\ E_z^{(i)} &= a_\parallel \sin\theta_i \cdot e^{-i\tau_i} & H_z^{(i)} &= a_\perp \sqrt{\epsilon_1} \sin\theta_i \cdot e^{-i\tau_i} . \end{aligned}$$

with

$$\tau_i = \omega\left(t - \frac{\vec{r}\,\vec{s}^{(i)}}{v_1}\right) = \omega\left(t - \frac{x\sin\theta_i + z\cos\theta_i}{v_1}\right)$$

Reflected wave: Note that $\theta_r = -\theta_i$.

$$\begin{aligned} E_x^{(r)} &= -r_\parallel \cos\theta_r \cdot e^{-i\tau_r} & H_x^{(r)} &= -r_\perp \sqrt{\epsilon_1} \cos\theta_r \cdot e^{-i\tau_r} \\ E_y^{(r)} &= r_\perp \cdot e^{-i\tau_r} & H_y^{(r)} &= -r_\parallel \sqrt{\epsilon_1} \cdot e^{-i\tau_r} \\ E_z^{(r)} &= r_\parallel \sin\theta_r \cdot e^{-i\tau_r} & H_z^{(r)} &= r_\perp \sqrt{\epsilon_1} \sin\theta_r \cdot e^{-i\tau_r} . \end{aligned}$$

with

$$\tau_r = \omega\left(t - \frac{\vec{r}\,\vec{s}^{(r)}}{v_1}\right) = \omega\left(t - \frac{x\sin\theta_r + z\cos\theta_r}{v_1}\right)$$

Transmitted wave:

$$\begin{aligned} E_x^{(t)} &= -t_\parallel \cos\theta_t \cdot e^{-i\tau_t} & H_x^{(t)} &= -t_\perp \sqrt{\epsilon_2} \cos\theta_t \cdot e^{-i\tau_t} \\ E_y^{(t)} &= t_\perp \cdot e^{-i\tau_t} & H_y^{(t)} &= -t_\parallel \sqrt{\epsilon_2} \cdot e^{-i\tau_t} \\ E_z^{(t)} &= t_\parallel \sin\theta_t \cdot e^{-i\tau_t} & H_z^{(t)} &= t_\perp \sqrt{\epsilon_2} \sin\theta_t \cdot e^{-t\tau_t} . \end{aligned}$$

with

$$\tau_t = \omega\left(t - \frac{\vec{r}\,\vec{s}^{(t)}}{v_2}\right) = \omega\left(t - \frac{x\sin\theta_t + z\cos\theta_t}{v_2}\right)$$

The boundary conditions must be fulfilled for the x-, y-, and z- components of $\vec{E}$ and $\vec{H}$:

$$\begin{aligned} E_x^{(i)} + E_x^{(r)} = E_x^{(t)} \quad E_y^{(i)} + E_y^{(r)} = E_y^{(t)} \quad E_z^{(i)} + E_z^{(r)} = E_z^{(t)} \\ H_x^{(i)} + H_x^{(r)} = H_x^{(t)} \quad H_y^{(i)} + H_y^{(r)} = H_y^{(t)} \quad H_z^{(i)} + H_z^{(r)} = H_z^{(t)} \end{aligned}$$

Evaluating these six equation yield two identical pairs and four independent equations:

$$\begin{aligned} \cos\theta_i\,(a_\parallel - r_\parallel) &= \cos\theta_t\, t_\parallel \\ a_\perp + r_\perp &= t_\perp \\ \sqrt{\epsilon_1}\cos\theta_i\,(a_\perp - r_\perp) &= \sqrt{\epsilon_2}\cos\theta_t\, t_\perp \\ \sqrt{\epsilon_1}(a_\parallel + r_\parallel) &= \sqrt{\epsilon_2} t_\parallel \end{aligned}$$

3.5.2 Fresnel formulae: reflected and transmitted amplitudes

The boundary conditions for a dielectric interface yields the Fresnel formulae which describe the reflected r and transmitted t wave amplitudes as function of the initial amplitude a for the two polarization modes $\|$ and $\perp$ independently. There are different ways to express the result:

$$r_{\|} = \frac{n_2 \cos\theta_i - n_1 \cos\theta_t}{n_2 \cos\theta_i + n_1 \cos\theta_t} a_{\|} \quad = \quad \frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)} a_{\|} \tag{3.30}$$

$$r_{\perp} = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t} a_{\perp} \quad = \quad -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)} a_{\perp} \tag{3.31}$$

$$t_{\|} = \frac{2n_1 \cos\theta_i}{n_2 \cos\theta_i + n_1 \cos\theta_t} a_{\|} \quad = \quad \frac{2\sin\theta_t \cos\theta_i}{\sin(\theta_i + \theta_t)\cos(\theta_i - \theta_t)} a_{\|} \tag{3.32}$$

$$t_{\perp} = \frac{2n_1 \cos\theta_i}{n_1 \cos\theta_i + n_2 \cos\theta_t} a_{\perp} \quad = \quad \frac{2\sin\theta_t \cos\theta_i}{\sin(\theta_i + \theta_t)} a_{\perp} \tag{3.33}$$

For the discussion we consider normalized amplitudes, e.g. $\hat{r}_{\|} = r_{\|}/a_{\|}$ and similar for $\hat{r}_{\perp}$, $\hat{t}_{\perp}$, and $\hat{t}_{\perp}$. We distinguish between the incidence from the less dense to the denser medium $n_1 < n_2$ and the other way round $n_1 > n_2$.

$\mathbf{n_1 < n_2}$: Incidence from the less dense medium onto the denser medium. There is $\theta_t < \theta_i$, $\tan(\theta_i - \theta_t) > 0$ and $\sin(\theta_i - \theta_t) > 0$. For the reflection $r_{\|}$ and $r_{\perp}$ there are positive and negative amplitudes. A negative value is equivalent to a phase shift of π with respect to the incident wave. As an illustration the reflected amplitudes for $n_1 = 1$ and $n_2 = 1.6$ are given below in a Table and a Figure. Some results and special cases are:

for the amplitudes of the reflected light

$\hat{r}_{\|} > 0$	for $\theta_i < \theta_B$, where θ_B is the Brewster angle (see below)
$\hat{r}_{\|} = 0$	for $\theta_i = \theta_B$ there is no reflection of the parallel component
$\hat{r}_{\|} < 0$	for $\theta_i > \theta_B$ there is a phase shift of π in this component
$\hat{r}_{\perp} < 0$	for all θ_i a negative vallue = π-phase shift

for extreme incidence angles

$\hat{r}_{\|} = -\hat{r_{\perp}}$	$= (n_1 - n_2)/(n_1 + n_2)$ for $\theta_i = 0°$ or normal incidence
$\hat{r}_{\|} = \hat{r}_{\perp} = -1$	for $\theta_i = 90°$ or grazing incidence there is total reflection

for the amplitudes of the transmitted light

$\hat{t}_{\|} \geq \hat{t}_{\perp} \geq 0$	there is no phase change in the transmitted amplitudes

for extreme incidence angles

$\hat{t}_{\|} = \hat{t}_{\perp}$	$= 2n_1/(n_1 + n_2)$ for $\theta_i = 0°$ or normal incidence
$\hat{t}_{\|} = \hat{t}_{\perp} = 0$	for $\theta_i = 90°$ or grazing incidence there is no transmission

$\mathbf{n_1 > n_2}$: Incidence from the denser medium towards the less dense medium. The amplitudes of the reflected and transmitted light behave very similar to the case $n_1 < n_2$. Two important differences are:

- the valid range for the incidence angle θ_i is reduced to $\theta_i = 0° - \theta_i^{\max}$. For $\theta_i > \theta_i^{\max}$ total "internal" reflection takes place and the Fresnel formulae do not apply.
- the signs for $r_{\|}$ and $r_{\perp}$ are reversed indicating switched phase shifts $(0 \leftrightarrow \pi)$ when compared to the $n_1 < n_2$ case.

Brewster angle. A special incidence angle is the polarizing angle or Brewster angle

$$\theta_B = \arctan\left(\frac{n_2}{n_1}\right) . \tag{3.34}$$

where $r_{\|}(\theta_B) = 0$. For this case the reflected and transmitted rays are perpendicular to each other $\theta_i + \theta_t = \pi/2$ or $\tan(\theta_i + \theta_t) = \infty$. The reflected light is fully polarized linearly in perpendicular orientation.

Total internal reflection. If the light is propagating from the denser medium to the less dense medium $n_1 > n_2$, then there exists a critical maximum angle $i_1^{\max}$ where $i_2 = 90°$ or $\sin i_2 = 1$. This angle is given by (see also Eq. (3.4)

$$\sin\theta_i^{\max} = n_2/n_1 \quad \text{for} \quad n_1 > n_2\,.$$

The angle θ_t is not defined for $\theta_i > \theta_i^{\max}$ and the formulae for the transmission amplitudes $t_{\|}$ and $t_{\|}$ are not valid. Total internal reflection occurs and all light is reflected back into the denser medium.

Example: reflection from a glass surface. As an example we illustrate the case of the transmission and reflection from a glass plate. We use $n_1 = 1.0$ for air and $n_2 = 1.6$ for glass. The Brewster angle for this case is $\theta_B = 58°$.

Reflected amplitudes and intensities as function of the incidence angle for $n_1 = 1$ and a glass plate with $n_2 = 1.6$. $\delta_{\|} - \delta_{\perp}$ is the relative phase shift between $r_{\|}$ and $r_{\perp}$.

θ_i	θ_t	$r_{\|}/a_{\|}$	$r_{\perp}/a_{\perp}$	$\mathcal{R}_{\|}$	$\mathcal{R}_{\perp}$	$\delta_{\|} - \delta_{\perp}$
0°	0°	+0.231	−0.231	0.053	0.053	π
15°	9.3°	+0.221	−0.241	0.049	0.058	π
30°	18.2°	+0.187	−0.274	0.035	0.075	π
45°	26.2°	+0.116	−0.340	0.013	0.115	π
60°	32.8°	−0.025	−0.458	0.001	0.210	0
75°	37.1°	−0.316	−0.662	0.100	0.439	0
90°	38.7°	−1.0	−1.0	1.0	1.0	0

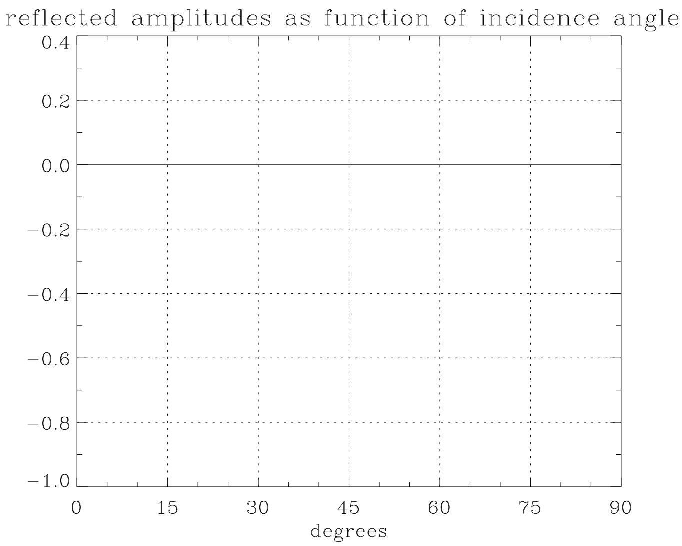

Figure 3.1: Reflected amplitudes for the parallel and perpendicular polarization $r_{\|}$ and $r_{\perp}$ directions as function of incidence angle for an air / glass ($n = 1.6$) interface.

3.5.3 Reflected and transmitted intensities.

The reflected $\mathcal{R}_\parallel$, $\mathcal{R}_\perp$ and transmitted $\mathcal{T}_\parallel$, $\mathcal{T}_\perp$ intensity fractions of the initial intensities $\mathcal{I}_\parallel$ and $\mathcal{I}_\perp$ are then obtained from the squares of the amplitudes. The parallel and perpendicular wave components can be treated independently. For the reflectivity we obtain:

$$\mathcal{R}_\perp = \left(\frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t}\right)^2 = \left(\frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i - \theta_t)}\right)^2 \tag{3.35}$$

$$\mathcal{R}_\parallel = \left(\frac{n_1 \cos\theta_t - n_2 \cos\theta_i}{n_1 \cos\theta_t + n_2 \cos\theta_i}\right)^2 = \left(\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i - \theta_t)}\right)^2 \tag{3.36}$$

Since there is no energy lost at a dielectric boundary the transmittivity follows from the energy conservation:

$$\mathcal{T} = 1 - \mathcal{R}, \quad \text{or} \quad \mathcal{T}_\perp = 1 - \mathcal{R}_\perp \quad \text{and} \quad \mathcal{T}_\parallel = 1 - \mathcal{R}_\parallel\,. \tag{3.37}$$

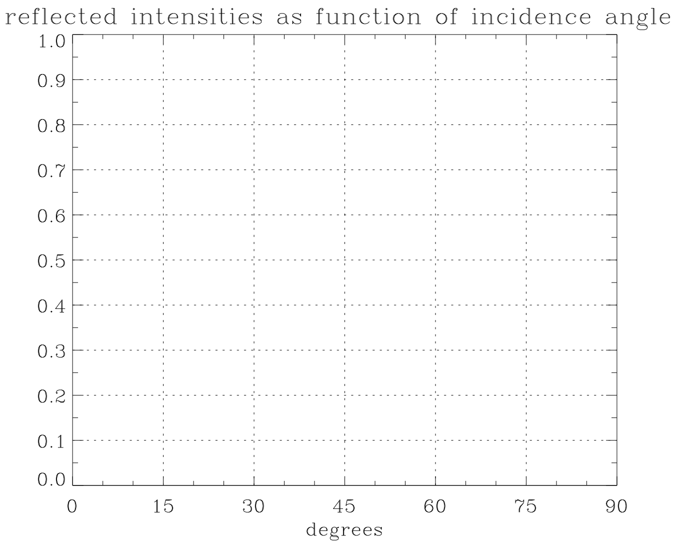

Figure 3.2: Reflected and transmitted intensities for the parallel and perpendicular polarization directions $\mathcal{R}_\parallel$ and $\mathcal{R}_\perp$ as function of incidence angle for an air / glass ($n = 1.6$) interface.

3.5.4 Special case: normal incidence.

For normal incidence $i_i = 0°$ the reflectivity $\mathcal{R}$ and transmittivity $\mathcal{T}$ are

$$\mathcal{R} = \frac{(n_2 - n_1)^2}{(n_2 + n_1)^2} \quad \text{and} \quad \mathcal{T} = \frac{4 n_1 n_2}{(n_2 + n_1)^2} \tag{3.38}$$

This equation shows that the reflection is larger for strong jumps in the refractive indices. For $n_1 \to n_2$ the reflection goes towards zero $\mathcal{R} \to 0$.

Example: The reflected intensity increases rapidly for an interface between air ($n_1 = 1$) and glass n_2 with increasing refraction index n_2, e.g. $\mathcal{R}(n_2 = 1.5) = 0.040$, $\mathcal{R}(n_2 = 2.0) = 0.111$, $\mathcal{R}(n_2 = 2.5) = 0.184$.

3.5.5 Transmission and ghosts from dielectric components

Transmitting, dielectric components are frequently used in astronomical instruments. Typical components are:

- lenses,
- broad and narrow band filters,
- beam splitter plates or beam splitter cubes,
- prisms,
- masks on substrates (e.g. coronagraphic masks).

Transmitting components produce reflections at the entrance and the exit surfaces. Often the components themselves are made up of two or more sub-components with eventually optical glue layers between them. Each interface can cause reflections which reduces the overall transmission of the instrument. More harmful are often back-and-forth reflections (or retro-reflections) between two interfaces which may produce ghost images or enhance the diffuse light in the instrument. One can distinguish between:

- internal reflections: two interfaces in a transmitting component reflect light back into the beam. This produces often ghost images, because the optical path difference between the direct image and the retro-reflected are small.
- external reflections: surfaces from one component reflect light backwards which might be retro-reflected by the surface of another component in the system. Sometimes external reflections produce also ghost images. However, often the ghost image is strongly defocussed or deflected because the path difference of the retro-reflected light is significant. In any case external reflections can add harmful diffuse light.

The reduction of diffuse light and ghosts is an important part of an instrument design. There are many tricks to reduce diffuse light and to avoid ghost images. Ghosts from internal reflections can be deflected by using wedge shaped components while external reflections and diffuse light can be reduced by tilting plane surfaces. A most important way to enhance the transmission and to reduce reflections are anti-reflection coatings, which are discussed in the next section.

Transmission and ghosts from one component. We consider a transmitting component or sub-component which has an entrance and exit interface characterized by the refractive indices n_1-n_2 and n_2-n_3, respectively. For each interface the formulas for the reflection $\mathcal{R}$ and transmission $\mathcal{T}$ as given in Eqs. (3.37), (3.35), and (3.36) apply. For the discussion here we assume that the reflections are quite small $\mathcal{R}_{12}\,, \mathcal{R}_{23} < 0.1$, $\mathcal{T}_{12}$, $\mathcal{T}_{23} > 0.9$ and we do not consider explicitly the dependence on incidence angle and polarization.

If also interference effects are neglected then the reflected and transmitted intensities can be expressed as series which consider the direct reflection or transmission through the interfaces and the higher order terms which describe one or more back-and-forth reflection in layer n_2.

The reflected intensity is

$$\mathcal{R} = \mathcal{R}_{12} + \mathcal{T}_{12}\mathcal{R}_{23}\mathcal{T}_{21}(1 + \mathcal{R}_{21}\mathcal{R}_{23} + (\mathcal{R}_{21}\mathcal{R}_{23})^2 + \ldots) \approx \mathcal{R}_{12} + \mathcal{R}_{23}\,(\mathcal{T}_{12}\mathcal{T}_{21}) \tag{3.39}$$

Because of $\mathcal{R}_{21}\mathcal{R}_{23} < 0.01$ higher order terms in the series can be neglected for a first approximation.

The transmitted intensity, including the back and forth reflected light is given by:

$$\mathcal{T} = \mathcal{T}_{12}\mathcal{T}_{23}(1 + \mathcal{R}_{23}\mathcal{R}_{21} + (\mathcal{R}_{23}\mathcal{R}_{21})^2 + \ldots)\,. \tag{3.40}$$

The directly transmitted intensity is

$$\mathcal{T} = \mathcal{T}_{12}\mathcal{T}_{23}$$

which is the value to be used for the calculation of the component transmission for the instrument sensitivity. For a multi-component system the final transmission is the product of all transmissions at all interfaces.

The ghost intensity from this single layer is then as first approximation ($\mathcal{T}_{12}\mathcal{T}_{23} \approx 1$) at a level of:

$$\mathcal{S}_{\mathrm{ghost}} \approx \mathcal{R}_{12}\mathcal{R}_{23}\,. \tag{3.41}$$

Transmission and ghosts of a multi-layer components. For a multi-layer component the transmission through each interface must be considered. As an example we consider a component made of two different dielectric materials A and B bonded with an optical glue. Examples for such components are doublet lenses or polarimetric retarder plates. This yields three-layers with 4 interfaces:

– interface 1: air – component A: $n_1 = 1.0$ and $n_2 = n_{\mathrm{A}}$,
– interface 2: component A – glue: $n_2 = n_{\mathrm{A}}$ and $n_3 = n_{\mathrm{glue}}$,
– interface 3: glue – component B: $n_3 = n_{\mathrm{glue}}$ and $n_4 = n_{\mathrm{B}}$,
– interface 4: component B – air: $n_4 = n_{\mathrm{B}}$ and $n_5 = 1.0$.

Transmission: The total transmission is just the product from all interfaces:

$$\mathcal{T} = \mathcal{T}_{12}\,\mathcal{T}_{23}\,\mathcal{T}_{34}\,\mathcal{T}_{45}$$

The transmission can be enhanced by minimizing the jumps in the refractive index at the interfaces. For this reason it is useful to bond two plates (lenses) with $n_{\mathrm{A}} = 1.5$, $n_{\mathrm{B}} \approx 1.7$ with a glue with $n_{\mathrm{glue}} \approx 1.6$ instead of having an air gap with $n_3 = 1.0$. The glue enhances the transmission by about 10 % when compared to a component with air gap.

Ghosts: For the ghosts the retro-reflection from each interface pair must be considered. The following combination of back-and-forth scatterings are possible:

– retro-reflections $\mathcal{R}_{xy}\mathcal{R}_{12}$ from the first interface: $\mathcal{R}_{45}\mathcal{R}_{12}$, $\mathcal{R}_{34}\mathcal{R}_{12}$, $\mathcal{R}_{23}\mathcal{R}_{12}$,
– retro-reflections $\mathcal{R}_{xy}\mathcal{R}_{23}$ from the second interface: $\mathcal{R}_{45}\mathcal{R}_{23}$, $\mathcal{R}_{34}\mathcal{R}_{23}$,
– retro-reflections $\mathcal{R}_{xy}\mathcal{R}_{34}$ from the third interface: $\mathcal{R}_{45}\mathcal{R}_{34}$.

The total ghost intensity is the sum of all these contributions:

$$\mathcal{S}_{\mathrm{ghost}} = \mathcal{R}_{45}\mathcal{R}_{12} + \mathcal{R}_{34}\mathcal{R}_{12} + \mathcal{R}_{23}\mathcal{R}_{12} + \mathcal{R}_{45}\mathcal{R}_{23} + \mathcal{R}_{34}\mathcal{R}_{23} + \mathcal{R}_{45}\mathcal{R}_{34}$$

Only some ghosts are really strong, those involving interfaces with large jumps in the refractive indices. For $n_\mathrm{A} = 1.5$, $n_\mathrm{glue} = 1.6$, and $n_B \approx 1.7$ the retro-reflection from the external surfaces dominate

$$\mathcal{S}_\mathrm{ghost} \approx \mathcal{R}_{45}\mathcal{R}_{12} \approx 0.040 \cdot 0.067 = 0.0027$$

If the components A and B are separated by an air gap instead of a glue layer then all six retro-reflections are more or less equally strong and the ghost and diffuse stray-light level is enhanced by a factor of about 6 and the final value is of the order $\mathcal{S}_\mathrm{ghost} \approx 0.015$. This illustrates that strong jumps in refractive indices at dielectric interfaces can be harmful sources for ghosts and straylight.

3.6 Dielectric films

Reflected light from thin films, i.e. surfaces which are close together, can produce interference effects. The light reflected by one surface may be reinforced by the second or if the phase difference is $\pi + n\,2\pi$ then partial cancellation can occur. The effect of thin dielectric films are used in many optics applications:

- anti-reflection coatings,
- high reflectivity coatings,
- beam splitters, filters, polarizers.

Evaporation techniques allow the production of films with very accurate thickness and consisting of many layers.

We consider again a geometry with three layers n_1, n_2, n_3 and two interfaces like in Section 3.5.5. But his time the intermediate layer is a thin film with a thickness d of the same order as the wavelength $d \approx \lambda$. Such a thin layer can be applied onto a flat plate or a curved surface of a lens.

Ray geometry. If the incident angle θ_1 and the refractive indices n_1, n_2, and n_3 are known then the angles θ_2 and θ_3 can be determined from Snell's law:

$$\sin\theta_3 = \frac{n_2}{n_3}\sin\theta_2 = \frac{n_2}{n_3}\frac{n_1}{n_2}\sin\theta_1 = \frac{n_1}{n_3}\sin\theta_1 \tag{3.42}$$

This indicates that the refraction angle and the light path geometry (geometric optics of the system) of an interface $1-3$ remains unchanged if a thin intermediate layer "2" is added.

3.6.1 Transmission and reflection from an interface with coating

The amplitudes and intensities of the reflected and transmitted waves from a thin film has two contributions:

- the reflections and transmissions from the two interfaces like for a plate,
- the interference term which depends on the phase difference between the reflected light from the first interface $1-2$ and the second interface $2-3$.

The phase difference depends for a given wavelength on the light path difference Δx and the differential phase change δ introduced by the reflection on the first and second interface.

The light path difference is given by the film refractive index $n_2 = n_{\text{film}}$, the film thickness d, and the angle θ_2 (a function of n_1, n_{film} and θ_1):

$$\Delta x = 2\, n_{\text{film}} d \,\cos\theta_2\,. \tag{3.43}$$

A phase difference $\Delta\phi$ due to the extra light path Δx includes in addition the wavelength of the light

$$\Delta\phi = \frac{2\pi}{\lambda}\,\Delta x = \frac{2\pi}{\lambda}\,2\, n_{\text{film}} d\,\cos\theta_2 \tag{3.44}$$

Not included is the differential phase difference which may be introduced by the two reflections at interface $1-2$ and $2-3$.

Reflected and transmitted intensity. The intensities reflected from the interfaces of the thin layer follows from the normalized, reflected amplitudes $\hat{r}_{12}$ and $\hat{r}_{23}$ (Fresnel forumlae). The result which is not derived here explicitly is:

$$\mathcal{R} = \frac{\hat{r}_{12}^2 + \hat{r}_{23}^2 + 2\hat{r}_{12}\hat{r}_{23}\cos\Delta\phi}{1 + \hat{r}_{12}^2\hat{r}_{23}^2 + 2\hat{r}_{12}\hat{r}_{23}\cos\Delta\phi} \tag{3.45}$$

This expression includes in the nominator the reflection from the two surfaces $\mathcal{R}_{12} = \hat{r}_{12}^2$, $\mathcal{R}_{23} = \hat{r}_{23}^2$ and the interference term $2\hat{r}_{12}\hat{r}_{23}\cos\Delta\phi$. The denominator accounts for the corrections due to the non-perfect transmission $\mathcal{T}_{12} < 1$ and $\mathcal{T}_{23} < 1$. This equation holds for both polarization directions $\mathcal{R}_{\parallel}$ and $\mathcal{R}_{\perp}$ independently.
Energy conservation requires that the transmitted intensity is

$$\mathcal{T} = 1 - \mathcal{R}\,.$$

For $\mathcal{R}_{12}$, $\mathcal{R}_{23} < 0.1$ we can use the approximation (error less than 1 %):

$$\mathcal{R} \approx \hat{r}_{12}^2 + \hat{r}_{23}^2 + 2\hat{r}_{12}\hat{r}_{23}\cos\Delta\phi\,.$$

Maximum and minimum reflectivity. For a given thin film there are maxima and minima for the reflected intensities which are equivalent to the maxima and minima of the interference term. Maxima and minima occur for the path length phase difference $\Delta\phi = n\,2\pi$ or $\Delta\phi = \pi + n\,2\pi$. Which of these two cases applies for a maximum or minimum reflectivity depends on the sign of the product $\hat{r}_{12}\hat{r}_{23}$ which depends on the refraction index jumps at the interfaces. Two cases must be distinguished. In case 1 the thin film has an intermediate refraction index and the differential phase shift is $\delta = 0$ and $\hat{r}_{12}\hat{r}_{23} > 0$. In case 2 n_{film} is larger or smaller than index of the two media and $\delta = \pi$ or $\hat{r}_{12}\hat{r}_{23} < 0$.
Case 1: a thin film with an intermediate index: $n_1 < n_{\text{film}} < n_3$ or $n_1 > n_{\text{film}} > n_3$.
The same phase shift is introduce by the reflections at the first and second interface because the refractive indices change in the same way at both interfaces.
The maximum reflectivity is obtained for $\cos\Delta\phi = 1$

$$\mathcal{R}_{\max} = \frac{\hat{r}_{12}^2 + \hat{r}_{23}^2 + 2\hat{r}_{12}\hat{r}_{23}}{1 + \hat{r}_{12}^2\hat{r}_{23}^2 + 2\hat{r}_{12}\hat{R}_{23}} = \left(\frac{\hat{r}_{12} + \hat{r}_{23}}{1 + \hat{r}_{12}\hat{r}_{23}}\right)^2\,. \tag{3.46}$$

The minimum reflectivity is obtained for $\cos\Delta\phi = -1$

$$\mathcal{R}_{\min} = \frac{\hat{r}_{12}^2 + \hat{r}_{23}^2 - 2\hat{r}_{12}\hat{r}_{23}}{1 + \hat{r}_{12}^2\hat{r}_{23}^2 - 2\hat{r}_{12}\hat{r}_{23}} = \left(\frac{\hat{r}_{12} - \hat{r}_{23}}{1 - \hat{r}_{12}\hat{r}_{23}}\right)^2\,. \tag{3.47}$$

Case 2: a thin film with an refraction index which is larger or smaller than the two dielectric media $n_1 < n_{\text{film}} > n_3$ or $n_1 > n_{\text{film}} < n_3$. A phase shift difference of π is introduced between the first and second interface (or $\hat{r}_{12}\hat{r}_{23} < 0$) because there is one transition from a higher n to a lower n and one opposite transition. In case 2 the formulas for $\mathcal{R}_{\max}$ and $\mathcal{R}_{\min}$ are exchanged when compared to case 1.
The maximum reflectivity is obtained for $\cos\Delta\phi = -1$:

$$\mathcal{R}_{\max} = \frac{\hat{r}_{12}^2 + \hat{r}_{23}^2 - 2\hat{r}_{12}\hat{r}_{23}}{1 + \hat{r}_{12}^2\hat{r}_{23}^2 - 2\hat{r}_{12}\hat{r}_{23}} = \left(\frac{\hat{r}_{12} - \hat{r}_{23}}{1 - \hat{r}_{12}\hat{r}_{23}}\right)^2 . \tag{3.48}$$

The minimum reflectivity is obtained for $\cos\Delta\phi = +1$:

$$\mathcal{R}_{\min} = \frac{\hat{r}_{12}^2 + \hat{r}_{23}^2 + 2\hat{r}_{12}\hat{r}_{23}}{1 + \hat{r}_{12}^2\hat{r}_{23}^2 + 2\hat{r}_{12}\hat{R}_{23}} = \left(\frac{\hat{r}_{12} + \hat{r}_{23}}{1 + \hat{r}_{12}\hat{r}_{23}}\right)^2 . \tag{3.49}$$

3.6.2 Special case: normal incidence

For normal incidence $\theta_i = 0$ the expressions for the maximum and minimum reflectivity $\mathcal{R}_{\max}$ and $\mathcal{R}_{\min}$ are particularly simple and the results are polarization independent. The reflectivity can be expressed with the refractive indices n_1, $n_2 = n_{\text{film}}$, and n_3:
Case 1: a thin film with an intermediate index: $n_1 < n_{\text{film}} < n_3$ or $n_1 > n_{\text{film}} > n_3$.
The maximum reflectivity is obtained for path differences $\Delta x = \lambda + n\lambda$ and the reflectivity maximum is:

$$\mathcal{R}_{\max} = \left(\frac{n_1 - n_3}{n_1 + n_3}\right)^2 . \tag{3.50}$$

This is equivalent to the case of a single interface where the intermediate layer "2" would just be absent.
The minimim reflectivity is obtained for path differences $\Delta x = \lambda/2 + n\lambda$ and the minimum reflectivity is

$$\mathcal{R}_{\min} = \left(\frac{n_1 n_3 - n_2^2}{n_1 n_3 + n_2^2}\right)^2 . \tag{3.51}$$

Complete cancelation of the reflected light can be achieved if the thin film has a refractive index of $n_2 = \sqrt{n_1 n_3}$ which would be in the range $n_2 \approx 1.25 - 1.30$ for typical glasses. In the "real world" there exists the problem that there is no "ideal" anti-reflection coating with $n_{\text{film}} \approx 1.25$ available. The most frequently used low refraction coating is MgF_2 with $n = 1.37$

Case 2: a thin film with refractive index larger or smaller than the two dielectric media $n_1 < n_{\text{film}} > n_3$ or $n_1 > n_{\text{film}} < n_3$.
The maximum reflectivity is obtained for an integer phase shift $\Delta\phi = n \cdot 2\pi$ composed of a light path difference $\Delta x = \lambda/2 + n\lambda$ and a differential phase shift of $\lambda/2$ from the two reflections. The maximum reflectivity is:

$$\mathcal{R}_{\max} = \left(\frac{n_1 n_3 - n_2^2}{n_1 n_3 + n_2^2}\right)^2 . \tag{3.52}$$

The minimum reflectivity is obtained for a path difference $\Delta x = n\lambda$ plus a differential phase shift $\lambda/2$ and the minimum reflectivity is:

$$\mathcal{R}_{\min} = \left(\frac{n_1 - n_3}{n_1 + n_3}\right)^2 . \tag{3.53}$$

Illustration off different examples for the normal incidence reflectivity $\mathcal{R}(\theta_i = 0°)$ as function of the optical path difference of an interface between air $n_1 = 1$ and glass $n_3 = 1.6$ with a thin film with refractive index $n_2 = n_{\text{film}}$.

- for $n_{\text{film}} = 1.0$ or $n_{\text{film}} = 1.6$ the refractive indices of the film layer is equal to n_1 or n_3 respectively and this is equivalent to the single interface case with $\mathcal{R} = 0.053$.
- for $n_1 <_{\text{film}}< n_3$ the thin film decreases the reflectivity when compared to a bare glass surface with maxima for $\Delta x = n\,\lambda$ (same reflectivity as bare glass) and minima for $\Delta x = (n{+}1/2)\,\lambda$. The minimum reflectivity is obtained for $n_{\text{film}} = \sqrt{n_1 n_3} = 1.26$.
- for $n_1 <_{\text{film}}> n_3$ the thin, high refraction film, enhances the reflectivity of the glass plate. In this case a differential phase shift $\delta = \pi$ is introduced between the reflections from interface $1-2$ and $2-3$. Therefore the minimum reflectivity with $\mathcal{R}$ like for a bare glass plate occurs for a path difference of $\Delta x = n\,\lambda$ while the maximum reflectivity is at $\Delta x = (n + 1/2)\,\lambda$. The reflectivity increases rapidely for high density coatings

3.6.3 Colour dependence of a thin film

We consider the reflection for normal incidence of a glass plate $n_3 = 1.6$ with a film MgF_2-coating $n_{\text{film}} = 1.37$ with a thickness $n_{\text{film}}d = 125$ nm. This film produces a phase shift of $\Delta\phi = \pi/2$ for a wavelengths of $\lambda = 500$ nm, an almost perfect case for a broad-band anti-reflex coating. For an 8 times thicker film with $n_{\text{film}}d = 1$ μm the reflectivity of only certain wavelengths are suppressed while other wavelengths are reflected like for an uncoated surfaces. Low reflections occurs e.g. for wavelengths 444 nm $(9\lambda/2)$, 571 nm $(7\lambda/2)$, or 800 nm $(5\lambda/2)$, with high reflectivity wavelengths in between.

3.7 Electromagnetic waves in metals

An electromagnetic wave interacting with a conducting metal surface will produce moving charges and Joule heat. This destroys electromagnetic energy (attenuates the wave) and therefore metals are opaque. Metals are important in optics because of their high reflectivity so that metallic surfaces are used as mirrors.
Maxwell's equation for conducting, isotropic media with $\mu = 1$ have the form:

$$\mathrm{curl}\,\vec{E} = -\frac{1}{c}\frac{\partial \vec{H}}{\partial t}\,, \tag{3.54}$$

$$\mathrm{curl}\,\vec{H} = \frac{4\pi}{c}\,\sigma\vec{E} + \frac{\epsilon}{c}\frac{\partial \vec{E}}{\partial t}\,, \tag{3.55}$$

$$\mathrm{div}\vec{E} = 0 \tag{3.56}$$

$$\mathrm{div}\vec{H} = 0\,. \tag{3.57}$$

For these formulas the following changes were made with respect to general form of Maxwell's equations (Sect. 3.2): Ohm's law $\sigma\vec{E} = \vec{j}$, $\epsilon\vec{E} = \vec{D}$, and $\mu\vec{H} = \vec{B}$ were used to get only a set of formulae for the electric field $\vec{E}$ and the magnetic field $\vec{H}$. There is $\mathrm{div}\vec{E} = 0$ because there are no charged areas (high conductivity) in a metal which could produce sources for a static electric field.

These Maxwell equations are identical to the case for a dielectric medium except for the diffusion term

$$\frac{4\pi}{c}\,\sigma\vec{E}$$

in Eq. (3.55) which descibes the production of currents and absorption via the production of Joule heat.
A general solution for these equations has the form

$$\vec{E} = \vec{E_0}e^{-i(\vec{k}\vec{s}-\omega t)}\,, \tag{3.58}$$

$$\vec{H} = \vec{H_0}e^{-i(\vec{k}\vec{s}-\omega t)}\,. \tag{3.59}$$

with the temporal derivatives $\partial\vec{E}/\partial t = -i\omega\vec{E}$ and $\partial\vec{H}/\partial t = -i\omega\vec{H}$. In this case Eq. (3.55) can be rewritten as

$$\mathrm{curl}\,\vec{H} = \frac{1}{c}\left(\epsilon + i\frac{4\pi\sigma}{\omega}\right)\frac{\partial\vec{E}}{\partial t} = \frac{\bar{\epsilon}}{c}\frac{\partial\vec{E}}{\partial t} \tag{3.60}$$

This equation is formally identical with the corresponding equation for non-conducting, dielectric media, if the permittivity ϵ is replaced by a complex permittivity

$$\bar{\epsilon} = \epsilon + i\frac{4\pi\sigma}{\omega}\,. \tag{3.61}$$

The differential wave equations have then the same form like for the dielectric case

$$\frac{\partial^2\vec{E}}{\partial t^2} = \frac{c^2}{\bar{\epsilon}}\,\nabla^2\vec{E}\,, \tag{3.62}$$

$$\frac{\partial^2\vec{H}}{\partial t^2} = \frac{c^2}{\bar{\epsilon}}\,\nabla^2\vec{H}\,. \tag{3.63}$$

Because of the complex $\bar{\epsilon}$ we also get a complex refraction index $\bar{n}$ and a complex propagation speed $\bar{c_n} = c/\bar{n}$:

$$\bar{n} = n(1 + i\kappa) = \sqrt{\bar{\epsilon}} \quad \text{and} \quad \bar{n}^2 = n^2(1-\kappa^2) = \bar{\epsilon}\,,$$

where n is the real part of the refraction index and $n\kappa$ the complex part or the attenuation index. These two quantities are often used to characterize metals.

3.7.1 Reflection from a metal surface

We consider the reflection of a wave from an interface between a dielectric medium with n_1 and a metal with $\bar{n_2}$. For this case Snell's law for refraction can be generalized:

$$\sin\bar{\theta}_t = \frac{n_1}{\bar{n_2}}\sin\theta_i\,. \tag{3.64}$$

Because the refractive index $\bar{n_2}$ for metals is a complex quantity also the transmission angle $\bar{\theta}_t$ is a complex quantity.

Fresnel formulas for metals: For the calculation of the wave reflected from a metal surface the same boundary conditions as for a pure dielectric interface given in Eqs. (3.28, 3.29) must be fulfilled. Therefore we get also the same form of the Fresnel formula for the reflected amplitudes. However the refractive index $\bar{n_2}$ and refraction angle $\bar{\theta}_t$ are now complex quantities so that also the reflected amplitudes have a complex component $e^{i\delta_\parallel}$ and $e^{i\delta_\perp}$ describing a phase retardation introduced by the interaction with a conductor:

$$r_\parallel e^{i\delta_\parallel} = \frac{\bar{n_2}\,\cos\theta_i - n_1\,\cos\bar{\theta}_t}{\bar{n_2}\,\cos\theta_i + n_1\,\cos\bar{\theta}_t}\,a_\parallel \quad = \quad \frac{\tan(\theta_i - \bar{\theta}_t)}{\tan(\theta_i + \bar{\theta}_t)}\,a_\parallel \tag{3.65}$$

$$r_\perp e^{i\delta_\perp} = \frac{n_1\,\cos\theta_i - \bar{n_2}\,\cos\bar{\theta}_t}{n_1\,\cos\theta_i + \bar{n_2}\,\cos\bar{\theta}_t}\,a_\perp \quad = \quad -\frac{\sin(\theta_i - \bar{\theta}_t)}{\sin(\theta_i + \bar{\theta}_t)}\,a_\parallel \tag{3.66}$$

These equations illustrate that metal surfaces introduce:

- a differential attenuation of the $r_\parallel$ and $r_\perp$ wave amplitude, introducing a linear polarization like dielectric interfaces,
- a differential phase shift $\delta = \delta_\parallel - \delta_\perp$ or retardation between the two polarization components. This introduces a cross-talk (a partial conversion) between the linearly and circularly polarized wave modes.

The optical constants for metals can be derived by the measurement of the reflection ratios $r_\parallel/r_\perp$ and the differential retardation $\delta = \delta_\parallel - \delta_\perp$ with a so-called ellipsometer.

Reflectivity of different metals. Frequently used metals for mirrors are aluminum, silver, and gold. These materials can be evaporated as a thin layer on polished glass surfaces. Some properties of these three metal coatings are:

Al: Aluminum is the most common coating for telescope mirrors. It provides a reflectivity of about 92 % from about 200nm to 1 μm with a distinct minimum ($\sim$ 86 %) around 800 nm. The reflectivity is $>$ 95 % in the near-IR and mid-IR up and beyond 20 μm.

Ar: Silver has a higher reflectivity $>$ 95 % than aluminum in the optical but a short-wavelength reflectivity cutoff below 400 nm. Reflectivities of up to 99 % are reached with high-reflectivity coatings on silver mirrors. However, the production of such mirrors is more delicate and sometimes caused problems for large telescope mirrors.

Au: Gold is a high reflectivity metal $\mathcal{R} > 98$ % but only for wavelengths $\lambda > 600$ nm (hence its yellow colour). Therefore it is an ideal coating for infrared instruments and telescopes.