Vysoká škola báňská – Technická univerzita Ostrava

Fakulta elektrotechniky a informatiky, katedra telekomunikační techniky

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

Technika pro budoucnost 2.0/Technology for the Future 2.0

CZ.02.2.69/0.0/0.0/18_058/0010212

Fundamentals of photonics

Základy fotoniky

Jazyk: EN

Version 1.5

VŠB TECHNICKÁ UNIVERZITA OSTRAVA | FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Ostrava 2021

Ing. Stanislav Hejduk, Ph.D.

doc. Ing. Jan Nedoma, Ph.D.

Ing. Jakub Kolář

ISBN 978-80-248-4547-0 (on-line)
DOI 10.31490/9788024845470

Content

1. Spectrum of optical radiation - colours

In essence, we perceive light radiation as a visible part of the electromagnetic spectrum in the range of approximately 380nm – 780nm (this range can vary slightly based on the 'quality' of the visual system of each individual). Nevertheless, this is only the part that is visible to human eyes. Regarding technical usage, it is necessary to take into account the infrared part (radiation with too low energy to be captured with the naked eye) and with ultraviolet part of radiation (radiation with too high energy) for optical radiation.

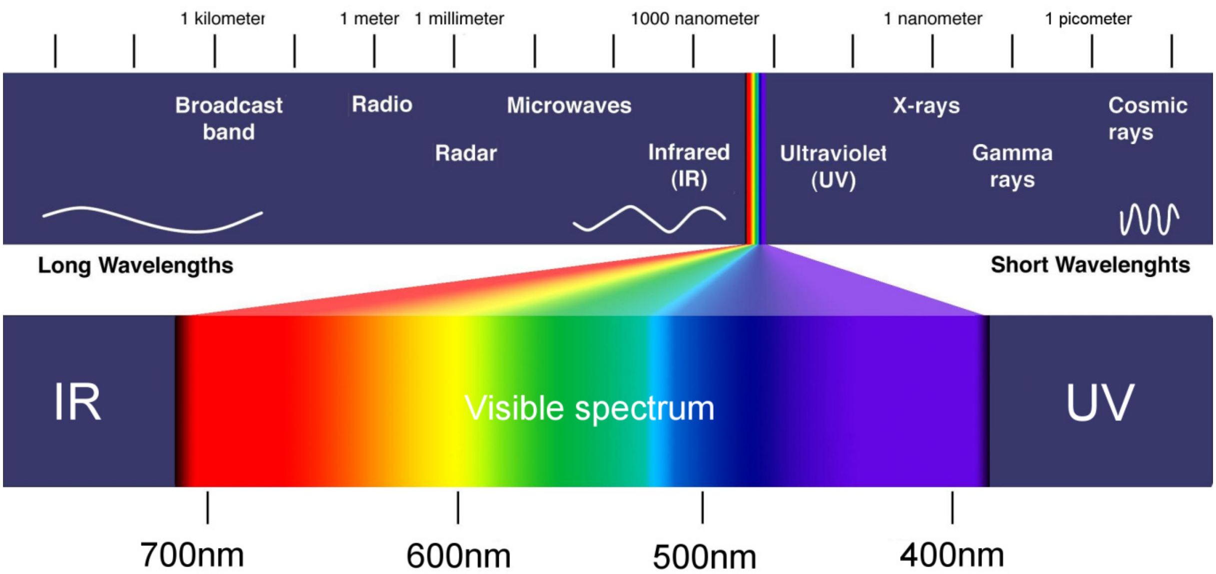


Fig.1.1 Spectrum of the EM radiation.

It is possible to see from the picture that the electromagnetic spectrum also includes radio waves and, therefore, the frequency data can be used instead of the wavelength. We can use the relationship for conversion between the wavelength and frequency.

$$\lambda = \frac{c}{f}$$

where lambda (λ) is the wavelength, *c* is the speed of light ($3 \cdot 10^8$ $m \cdot s^{-1}$) and f is the corresponding radiation frequency.

With increasing radiation energy, the wavelength moves from the infrared spectrum through the visible spectrum further to the UV spectrum. The radiation energy can be described by the relationship for individual wavelengths

$$E = \frac{h \cdot c}{\lambda}$$

where h is the Planck constant (**$6{,}62607015 \times 10^{-34}$ J·s**).

Colorimetry and colour vision

The term colorimetry has basically 2 meanings. In the first case, it is a procedure that compares the intensity of a coloured solution with solutions of known concentrations (however, this is more of a definition for chemically based objects). In the second one, which is more interesting for us, the definition shows the science of colour measurement which allows the objectification of the sensory perception of colours and its capture by numbers.

Colorimetric description of light

$$G=\int p(\lambda)\, R(\lambda) g(\lambda) d\lambda,$$

where the G indicates the trichromatic component, p describes the spectral radiation of the source, R describes the spectral absorption of the reflective material, g describes the physiological intake of the human eye. [1]

The term colour (colour shade) means the sensory perception perceived by the human eye (visual system). As all other sensory perceptions, this perception is subjective. It varies quite considerably from person to person and depends on various factors such as mood, tiredness, but also the age and condition of the sense organs. Thanks to standardized light sources and lighting geometry (viewing angle) and the medium sensitivity of the observer's eye, the colorimetry enables to objectify and express the perception in numbers.

At the beginning, it is necessary to realize that the measurement of the physical properties of the optical radiation itself can be easily performed by a spectrum analyser. Then the received data correspond to specific parts of the spectrum and colour coordinates. The following figure shows an example of the spectrum of a white LED (left) and the corresponding colour coordinates based on the CIE diagram (right).

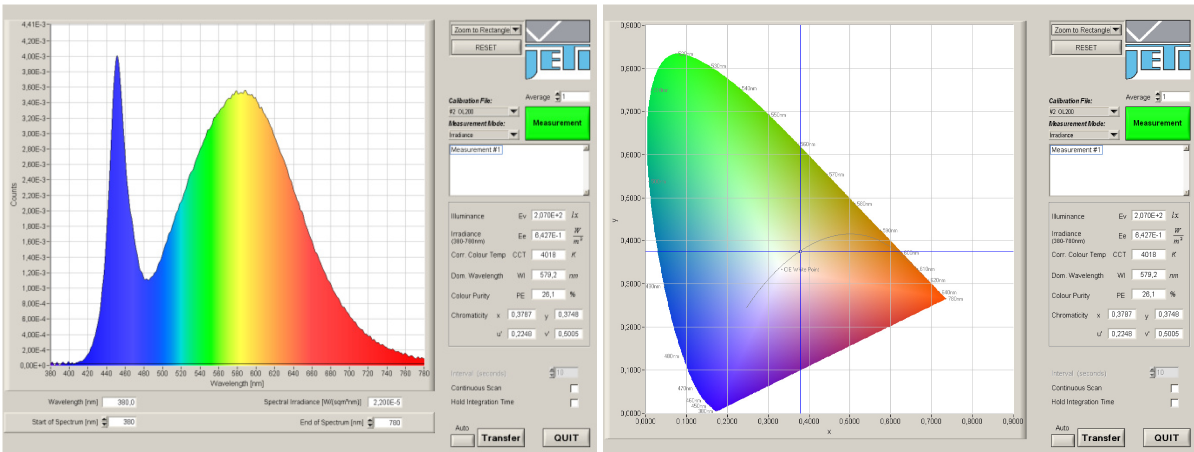


Fig.1.2 White LED spectra (left) and corresponding CIE coordinates (right).

The left figure clearly demonstrates the system of the operation of white LED diodes, which is basically the blue LED diode with a wavelength of about 450nm. A layer of luminophore is applied to it. It covers the missing part of the spectrum in the green, yellow and red parts of the spectrum. The colour seems to be naturally white, despite the fact that the values are different compared to the spectrum of the daylight.

Fig.1.3 Spectrum of daylight.

The use of the spectrometer itself gives information only about the end part of the whole chain. It is possible to use the following example to give an idea:

There is a car with the so-called pearl paint on the road. If we change the position from which we observe the car, the colour of the car will be different. Likewise, the colour will vary based on the angle at which the car is illuminated by the sunlight, or it will also depend on the light source itself. If we had a spectrometer with us in the mentioned cases, it would also show a corresponding 'different' result in any case.

Therefore, if we would like to express the colour of the surface itself numerically, we must take into account 3 factors:

1. Properties of the source of the reference radiation (spectrum, power, ...)
2. Measured colour object (surface properties, viewing angle, …)
3. Sensitivity of the eye (see the chapter on the spectral sensitivity of the human eye)

The substance of the surface colour measurement lies in the fact that the measured object is irradiated with a source of light radiation of known parameters – optimally, it is a broad-spectrum source of radiation. If an important part of the spectrum was missing on the source side, the measured results would be distorted.

It is also necessary to take into account that part of the light is reflected directly from the surface (so-called surface re-emission). The rest of the light falls on the particles of material, where it gradually reflects, absorbs or scatters. The fall on the coloured particle leads to the absorption of a specific part of the spectrum (conversion of radiation to heat). Then only the part of the spectrum that corresponds to the spectrum of coloured particles in the material is reflected back. If the material does not contain coloured particles, the light radiation gets back in the original spectrum. In case of the ideal white, 100% of the original signal would be returned. Similarly, in case of the ideal black, all the incident radiation would be converted to heat. So, we have 2 data available for measurement.

1. Corresponding wavelengths that are reflected (colour)
2. Reflection factor – proportion of light radiation that is reflected back (brightness)

The third component involved in the vision process is the human eye. At the end of 1920s, the comprehensive research of a large number of tested persons was carried out by special devices for mixing spectral colours. They led to the definition of a standard colorimetric observer, CIE standardized, in 1931.

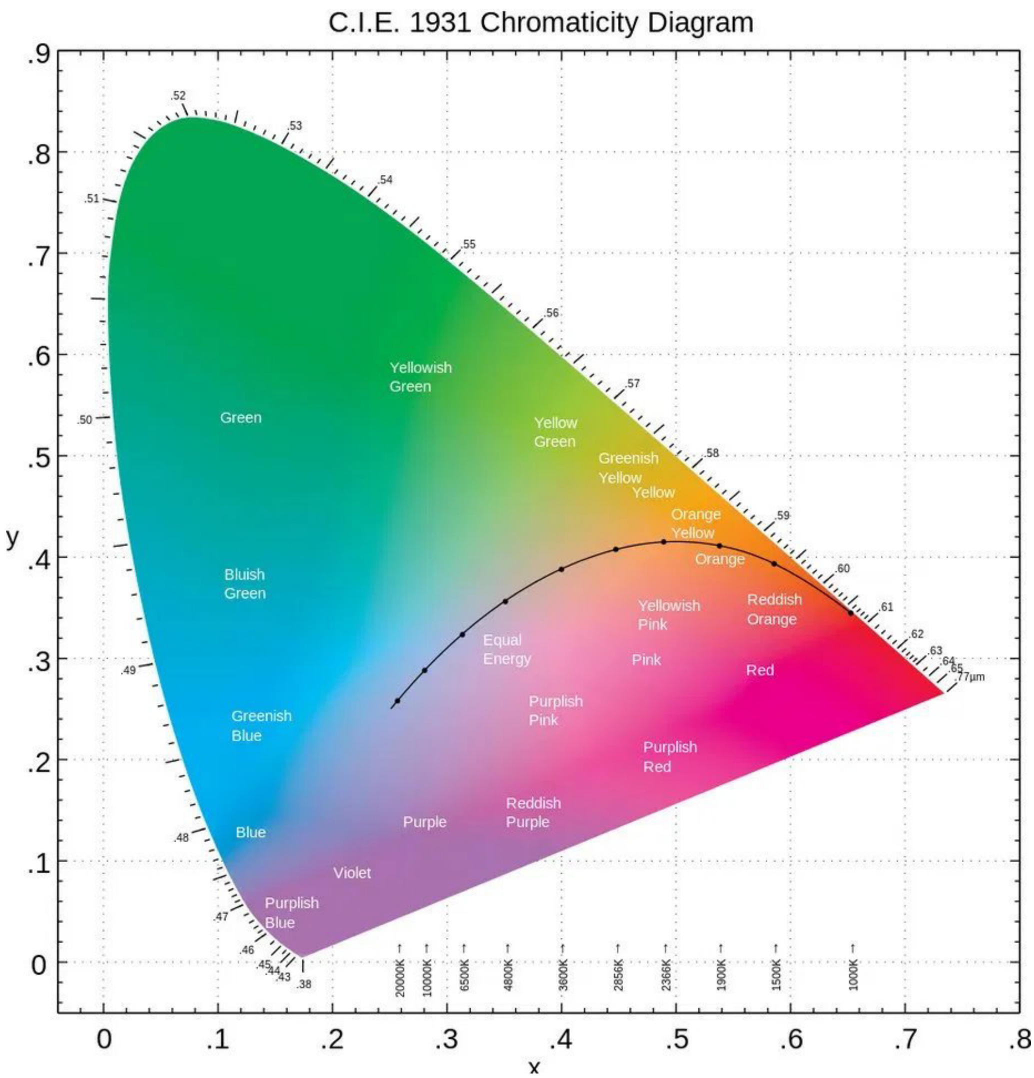


Fig.1.4 CIE diagram [16].

By means of the CIE diagram, we can easily determine the colour for specific colour coordinates (x, y). The wavelength of individual components describes the outer side of the diagram within the visible spectrum in the range of 380nm – 780nm. Therefore, if the resulting colour consists of more than one component of the spectrum, the resulting coordinates are shifted into the colour space (red + green = yellow, etc.) depending on their intensities.

The basic colours come nearer to the following values of the wavelength:

- Purple – 390nm
- Blue – 450nm
- Green – 530nm
- Yellow – 580nm
- Red – 650nm

The black curve inside the diagram describes the so-called radiation of the black body, i.e. the colour coordinates corresponding to the surface temperature of the body in Kelvin. (Thus, it is possible to see how the coordinates shift to the blue part of the spectrum with the increasing temperature).

Another important element is the standardized viewing angle which was set to 2o (standard observer 2o CIE-1931). Based on the new research that was conducted later for a larger viewing angle of 10 o, the standard observer was normalized by the CIE in 1964. It is now used in most colorimetric programs (standard observer 10o CIE-1964). The larger field of vision 10o corresponds better to the conditions of practice for colour patterning.

Different colour swatches are often used for printing and graphic practice.

- **PANTONE Formula Guide** (worldwide)
- **RAL** (Europe)
- **HKS** (Germany)
- **NCS Natural Colour System** (Europe)

Colour systems

Colour perception by the human eye and the term wavelength describe only part of the issue. It is necessary to define several terms for better understanding:

Body colours: colours from illuminated bodies

Light colours: colours from light sources (even reflected from bodies…)

Primary colours: colours that cannot be created by mixing

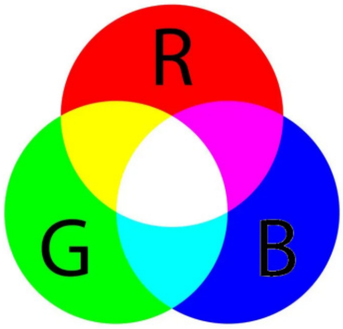
Secondary colours: they are created by mixing two basic colours

Tertiary colours: they are created by mixing three basic colours

Additive colour mix (RGB)

This is the Light colour mixing that describes the composition of colours from light sources (such as monitors and other light sources). In other words, other (different) coloured lights will be added to one coloured light, so that the resulting light has a richer spectral composition than sub-lights. The human visual organ does not have the ability to distinguish individual colours in compound light. The resulting light is always the only final colour for the human eye. The resulting light is always the only final colour for the human eye. However, it is possible to prepare one and the same colour of light in many combinations, i.e. two lights with different spectral compositions can have the same colour.[2]

The basic colours here are: red, green and blue.



Obr.1.5 Additive colour mixing (RGB).

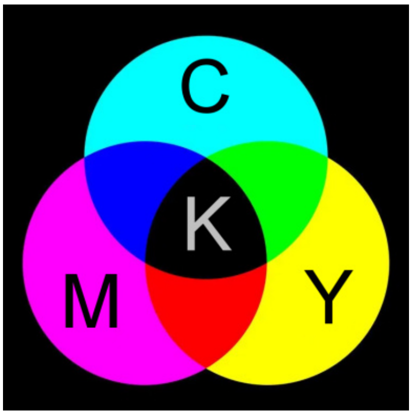
The composition of individual light sources creates other colours when mixing of all colours (in the correct ratio) creates the white colour.

Subtractive colour mix (CMYK)

This is the Body colour mixing when the colours of body surface areas are mixing (for example, the printers, or during painting).

Basic colours: Cyan, Magenta, Yellow a Key (or black…)

Compared to the RGB, we have one additional colour. However, there is a reason for this because mixing all CMY colours does not create exactly black but rather dark brown. For printing purposes, the black is usually more represented, so adding a separate black colour has its advantages.



Obr.1.6 Substractive colour mixing (CMYK).

The interesting fact is that the RGB and CMYK systems are in fact exactly opposite to each other. Therefore, the colours correspond to the so-called negative. (Blue changes to yellow, red to cyan and green to magenta). It is the most evident for the combination of all 3 colours (black – white). Nevertheless, it is possible to see it in case of the combination of 2 colours when the colours rotated with each other.

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. What is the range of wavelengths in the visible spectrum?
2. What is the wavelength of red light (approximately)?
3. What relation is between the wavelength and frequency?
4. How does the white LED work?
5. Which are the base colors for additive color mixing?
6. Which basic colors would you use to create the yellow color of light?
7. Which colors are used for printing?

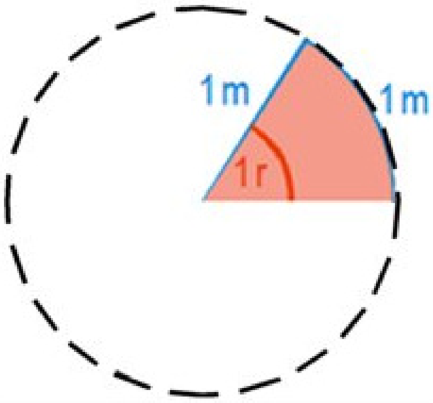
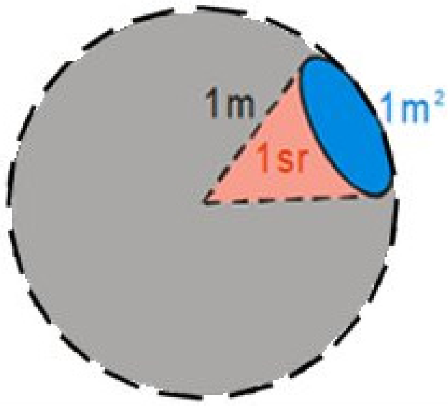
2. Radiometry and Photometry

Light can be seen as an electromagnetic energy covering a specific part of the frequency spectrum. In respect to its nature (possibility of observation by human sight), 2 approaches are used for measurement. Radiometry is the first approach. It describes the light objectively in terms of its energy throughout the whole part of the spectrum. The second approach that describes the light from the perspective of the human sight is called photometry. Each photometric quantity corresponds to a radiometric quantity as shown in Table 2.2. Their relationship depends on the spectral composition of the light and not on the units used. [3]

Related terms:

Plane and spatial angle:

Table 2.1

Plane angle ϕ	Spatial angle Ω
1 radian [1r] – natural angle unit	1 steradian [1sr] – radian on the surface
	
The angle that points an arc in the length of 1 m on a unit circle (a circle with a radius of 1 meter).	The angle that points to a unit sphere area 1m^2.
Circle circumference $O=2\pi r$, the whole circle is 360°=$2\pi r$. 1 The 1 radian corresponds to *57°18'*	Sphere area $S=4\pi r^2$. $\Omega = \frac{S}{r^2}[sr]$
Arc length calculation $l=\phi\, r$	Calculation of the area on the sphere $S=\Omega\, r^2$

Point and planar light source:

The point light source is a source which dimensions are so small in comparison with the distance of the observer (i.e. isotropic emitter). Planar source can be described as an area composed of a large number of individual glowing surfaces (we must take into account the viewing angle.)

Radiometric quantities:

The light can be described by means of electromagnetic waves that are able to transmit some part of the energy. This energy can be described by means of the so-called radiometric, resp. photometric quantities. Radiometric quantities describe the energy transmitted by radiation in the whole spectrum of electromagnetic waves and use objective quantities. In relation to the light, they describe the parts of the spectrum that the human visual system cannot notice (i.e. from the far infrared – FIR to the ultraviolet - UV). [3]

- Radiant energy Q_e– total amount of radiated energy. [J].
- Radiant flux ϕ_e – radiant energy radiated into space per unit of time ('radiant power')

$$\phi_e = \frac{dQ_e}{dt} \left[\frac{J}{s} = W\right]$$

- Radiant intensity I_e – radiant flux emitted by a point light source in a given spatial angle Ω.

$$I_e = \frac{d\phi_e}{d\Omega} \left[\frac{W}{sr}\right]$$

- Irradiance flux density E_e – radiant flux falling on the given area S.

$$E_e = \frac{d\phi_e}{dS} \left[\frac{W}{m^2}\right]$$

- Spectral radiance L_e – radiance from a defined part of the surface ΔS of a planar radiation source at an angle α.

$$L_e = \lim_{\Delta S \to 0} \frac{\Delta I_e(\alpha)}{\Delta S \; cos\alpha} \left[\frac{W}{sr \cdot m^2}\right]$$

- Radiant exposure or radiation dose is the real density that has fallen on a given area in the time interval from t_0=0 to t; it is the product of the mean value of the intensity and the time t when the irradiation interacts ('irradiation dose').

$$H_e = \frac{d\phi_e \cdot t}{dS} \left[\frac{W \cdot s}{m^2}\right]$$

Photometric quantities:

Photometric quantities describe the effects of radiation on the human visual system (the eye is not equally sensitive to the perception of all wavelengths – the eye is the most sensitive to the green-yellow light). Photometry is a measurement of light that is defined as electromagnetic radiation detected by the human eye. That is approximately 380-780nm (depending on the vision mode). The radiation source outside this visible range will therefore show zero values regardless of the actually radiated power.[3]

- Luminous flux Φ – expresses the ability of radiant flux to evoke visual perception. The lumen is the unit.

$$\Phi = K \int_0^\infty \hat{\phi}_e(\lambda) V(\lambda) d\lambda \;\; [lm]$$

- Luminosity I –is defined as the luminosity of a radiation source at a wavelength of 555,2nm and the radiance of 1/683W. The candela is the unit. (The variable K serves for the convert through the sensitivity of the human eye when it represents a value of 680lm/W for photopic vision and 1740lm/W for scotopic vision).

$$I = K\int_0^{\infty} \hat{I}_e\,(\lambda)V(\lambda)d\lambda \quad [cd]$$

- Luminous intensity E – luminous flux falling on a given area. The unit is lux which is defined as a luminous flux of 1lm falling on an area of 1m^2.

$$E = \lim_{\Delta S \to 0} \frac{\Delta\Phi}{\Delta S} \quad \left[\frac{lm}{m^2} = lx\right]$$

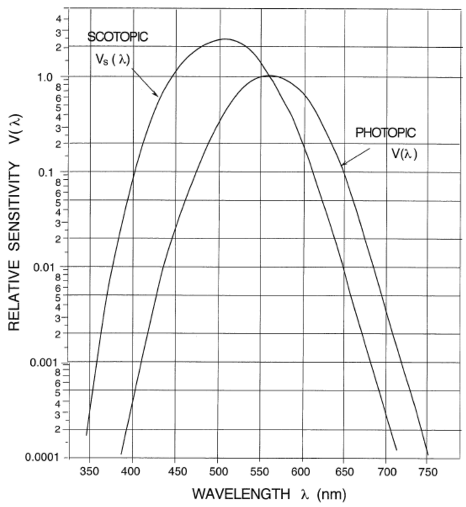
- Luminance L – is determined by the ratio of the luminosity *I* the elementary surface with an area S in the selected direction α. The nit is the unit. It corresponds to a part of the surface of a planar source with a luminosity of 1cd on the area of 1m^2.

$$L = \lim_{\Delta S \to 0} \frac{\Delta I(\alpha)}{\Delta S\ cos\alpha} \quad \left[\frac{cd}{m^2} = nt\right]$$

- Luminous exposure H – the areal density of the amount of light that fell on a given area in a given time interval. The lux-second is the unit.
   - $H = \bar{E} \cdot t \ \ [lx \cdot s]$

Spectral sensitivity of the human eye

The sensitivity of the human eye is not linear, so it perceives different wavelengths with different sensitivity. The peak sensitivity of the human eye under normal conditions is around 550nm (yellow-green colour). Therefore, the spectral sensitivity actually corresponds to the spectrum of daylight.



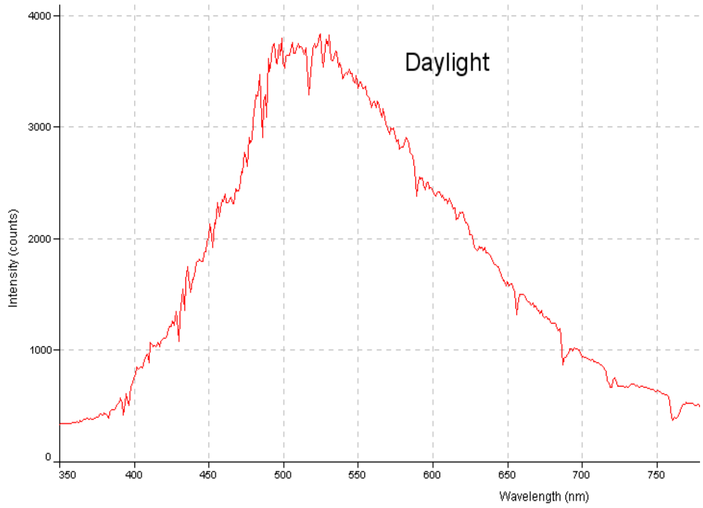


Fig. 2.1 Relative sensitivity of the spectral response of the human eye for photopic and scotopic vision vs. daylight spectrum.

In a normal mode with enough light, the eye works in a photopic mode. If the eye does not have enough light for a long time (10-30minutes), it begins to transmit to scotopic (night) vision. In this mode, the peak of sensitivity shifts closer to the blue part of the spectrum and the sensitivity increases (adaptation to moonlight). In this mode, the peak of sensitivity shifts closer to the blue part of the spectrum and the sensitivity increases (adaptation to moonlight). In scotopic vision, the rod perception starts to prevail in the eye. That is why we stop perceiving colours because the cones are responsible for the perception of colours that are dominant in the photopic mode.

Comparison of radiometric and photometric quantities

Table 2.2: Radiometric and photometric quantities and units. [1]

Radiometric quantity	Symbol	Unit	Photometric quantity	Symbol	Unit
Radiant flux	Φ_e	[W] (watt)	Luminous flux	Φ	lm (lumen)
Radiant Intensity	I_e	[W·sr^{-1}]	Luminosity	I	cd (candela)
Irrradiance flux density	E_e	[W·m^{-2}]	Luminous intensity	E	lx (lux)
Radiant exitance	M_e	[W·m^{-2}]	Luminous exitance	M	lm/m^2
Spectral radiance	L_e	[W.sr^{-1}.m^{-2}]	Luminance	L	nt (nit)
Radiant exposure	H_e	[W.s.m^{-2}]	Luminous exposure	H	lx.s

In the chart we can see the quantity M_e which represents the case where we are interested in the intensity from the point of view of the radiation source. Therefore, the area S in this case represents the area of the radiation source and not the area on which the radiation falls.

The efficacy data will help us with the conversion between radiometric and photometric units regarding the luminous intensity. If we have, for example, the LED diode with the efficiency of 100lm/W we can easily calculate the remaining quantities according to current needs.

Radiation sources

Radiation, such as light, is usually emitted by different sources of different shapes. For simplification, it is easier to imagine replacing such a source by some ideal isotropic emitter. Nevertheless, such a solution cannot always describe the real behaviour of the source. In real cases, the sources show different values depending on the viewing angle. These values are then described by radiation characteristics.

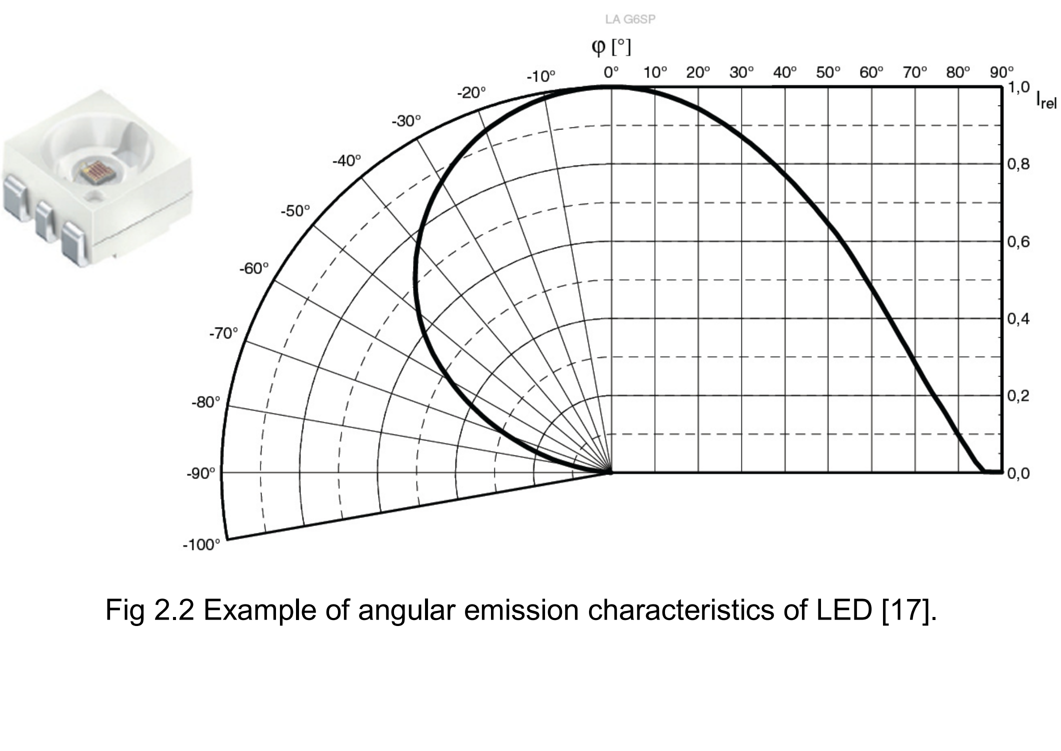
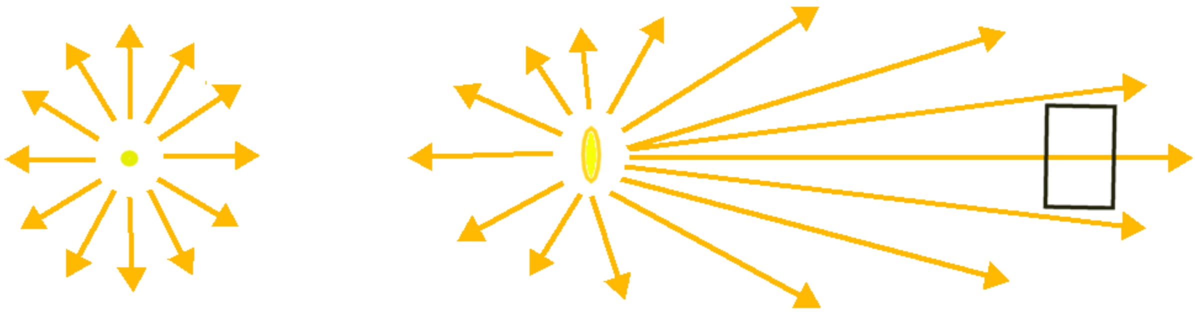


Fig 2.2 Example of angular emission characteristics of LED [17].

The already mentioned radiometric and photometric quantities also take into account this behaviour. They enable to describe the values depending on the power of the source itself, the beam angle and the area from which or on which the radiation falls. Therefore, let´s go through the basic sources to be able to understand this.

Point radiation source

Ideally, it should be a source of very small dimensions that emits radiation evenly in all directions. For simplification, a point source can be also considered as a real source of radiation, which is so far from the place of observation that it will behave as a point source in a defined space. In these cases, we will be able to neglect some characteristics of the radiation source.



An ideal point source.

The real source at a sufficient distance and the small receiving area.

Fig. 2.3 Point radiation source.

If we want to express the radiance L_e, then, only the value of the radiant intensity I_e and the magnitude of the point (radiating from the surface S) will be sufficient. Then we could simplify the calculation by removing the dependence on the beam angle:

$$L_e = \frac{I_e(\alpha)}{S\ cos\alpha} = \frac{I_e}{S}$$

Since, when changing the angle from the perpendicular direction there would be no loss of radiation (*cos* 0°=1). The closest definition to a point source is a light bulb (at a sufficient distance) because the light theoretically emits evenly in all directions. That is, if we neglect some mechanical elements.

Planar radiation source

If we imagine the plane from which the radiation is emitted, it is logical that we record the most energy in the direction perpendicularly to the source. If we change the viewing angle, there will be a gradual loss of energy because the visible area from the point of view of the observer will gradually decrease.

The dependence of the angle on the observed surface is described for an isotropic surface source by the cosine function, where the name cosine emitter or Lambertian emitter come from and where:

$$L_e = \frac{I_e(\alpha)}{S\ cos\alpha}$$

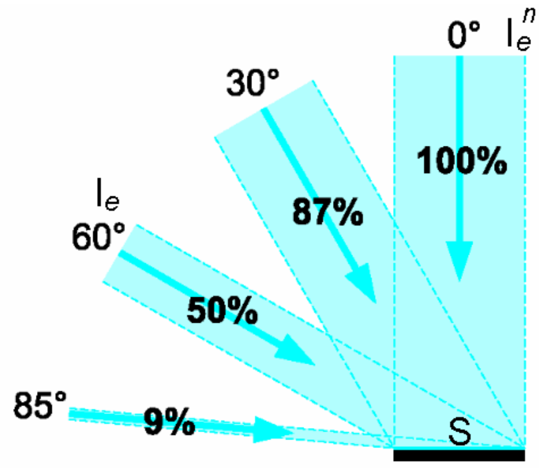


Fig. 2.4 Planar radiation source.

After the adjustment we get the relation for the radiance of an isotropic surface source (Lambert's law):

$$I_e = \ I_e^n \cdot cos\alpha$$

Regarding this emitter, it applies that the radiance decreases with the cosine of the deviation from the perpendicular to the source surface, and thus at an angle of +/-60° from the perpendicular we always have more than 50 % of the power available.

In case of the plane spectral radiance L_ϑ (radiance) the following also applies:

$$L_\vartheta = \ \frac{\Delta I_\vartheta}{\Delta S_\vartheta} = \frac{\Delta I_n}{\Delta S} = L_n$$

This means that the value of the surface radiant intensity of the source does not depend on the slope ϑ from the perpendicular in which we observe the surface because it describes the source or radiation and not its effects on the environment. It depends neither on the target area or its distance.

If we assume that the total radiation angle of the surface source is 180° (π) and the value of the surface radiant intensity is constant we can express the intensity of radiation of the cosine emitter as

$$M_e = \ \frac{d\Phi_e}{dS} = \frac{\pi \cdot \Delta I_n}{\Delta S} = \pi \cdot L$$

The intensity of radiation of the cosine emitter is thus π - times greater than the value of its surface radiant intensity.

The term Lambertian emitter can be used especially for LED diodes because they usually emit light radiation from the described area and most of the power is concentrated in the range of 120° (+/-60°). Nevertheless, it is necessary to pay attention to LED diodes with an integrated lens. Basically, it is the Lambertian emitter at the chip level. However, the output beam is adjusted to a different beam angle by the lens.

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. What is the difference between a radian and a steradian?
2. What is the basic difference between radiometric and photometric quantities?
3. What are the two modes of human eye vision and when do they apply?
4. What is the difference between a point and a planar radiation source?
5. Under what conditions can be a surface radiation source considered as a point source?
6. Which basic colors would you use to create the yellow color of light?
7. At what angle does the Lambert radiator radiate 50% of the power?

3. Blackbody radiation

Each body with surface temperature higher than absolute zero (i.e. 0 K, resp. -237,15 °C) emits (radiates) certain amount of radiant energy. The basis of the definition of this behaviour is a so-called ideal blackbody, i.e. a body that absorbs all radiation at all available wavelengths. However, at the same time, this body is also an ideal emitter because out of all the possible bodies of the same temperature it emits the highest possible amount of radiant energy. The amount of energy and the spectrum (range) of the emitted radiation then depends only on the surface temperature of the given body. Both of these parameters can be also described mathematically by means of the Planck blackbody emission law which describes the spectral radiation using the relation [4]

$$r(\lambda) = \frac{hv^2}{\lambda^3 e^{\frac{hv}{kT-1}}},$$

where $\nu = c/\lambda$. Examples of the process for different temperatures are shown in Figure 3.1.

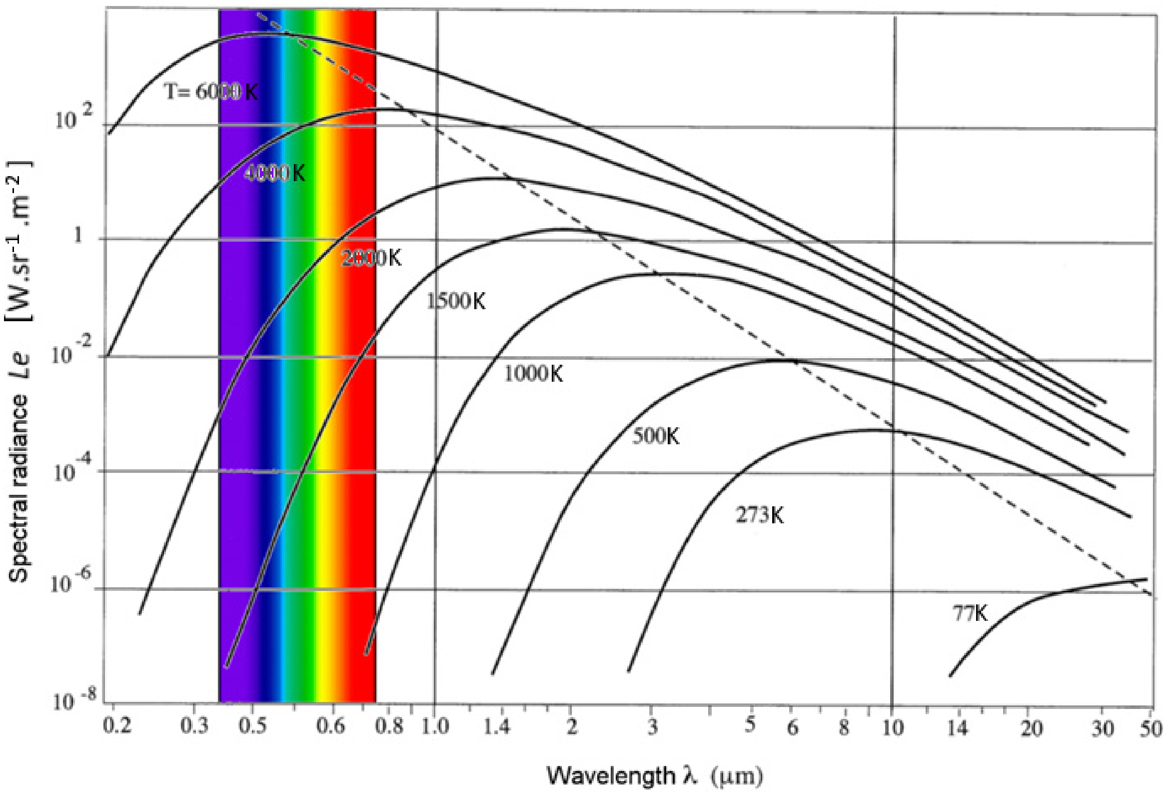


Fig. 3.1. Spectral radiance of the blackbody versus λ.

It is evident from the picture that we are able to perceive temperatures from the so-called Draper's point at the temperature of 798K (525°C) with the human visual system that is able to perceive in the range of 0,38-0,78um (380-780nm). Temperatures below this limit can be then detected, for example, by photodiodes in the IR spectrum (up to approx. 2um) or by pyroelectric sensors (7-14um). The radiation emitted in this way can be easily used for the optical measurement of the body temperature when the total emissivity according to Stefan—Boltzmann's law increases depending on the temperature with the fourth power [4].

$$I = \sigma T^4 \left[\frac{W}{m^2}\right],$$

Where the σ is the Stefan-Boltzmann's constant.

$$\sigma = \frac{2\pi^5 \cdot k^4}{15c^2 \cdot h^4} = 5{,}670400 \cdot 10^{-8}\ [W \cdot m^{-2} K^{-4}]$$

In practice we encounter the concept of the 'grey body radiation' because the measured body will not have ideal properties in real cases. The real body will not only emit and absorb radiation but also reflect or transmit it. In addition, these phenomena can be manifested in different parts of the electromagnetic spectrum so that, for example, the leaf of a tree absorbs the red and blue parts of the spectrum while reflecting the green part back so we can see it. Depending on the surface properties of the material, the radiant energy will always be slightly lower. Lighter or even reflective material will radiate less energy into the surrounding than matte black surface. We will add a variable (ε) into the equation that will describe the emissivity of the given body in relation to an ideal black body. Therefore, after adjustment:

$$I = \varepsilon\sigma T^4 \left[\frac{W}{m^2}\right]$$

The emissivity values ε can be expressed as

$$\varepsilon = \frac{H}{H_b},$$

where *H* corresponds to the integral intensity of thermal radiation of the measured body and H_b corresponds to the integral intensity of thermal radiation of the blackbody. The values ε are usually in the range of 0,90-0,99 for common matte bodies. Nevertheless, in case of the polished aluminium, this value can drop up to 0,05, i.e. depending on the cavity temperature. Therefore, in order to measure the temperature more accurately it is advisable for any measuring system to be calibrated in this respect. The emissivity of a real blackbody depends on the design of the cavity and the material used. The simplest method is the initial comparison of temperatures, e.g. with a contact thermometer.

Wien's law of displacement:

Along with the increasing temperature, the maximum of spectral emissivity also increases which logically shifts from longer wavelengths to shorter ones with the increasing temperature (part of the spectrum with higher energy).

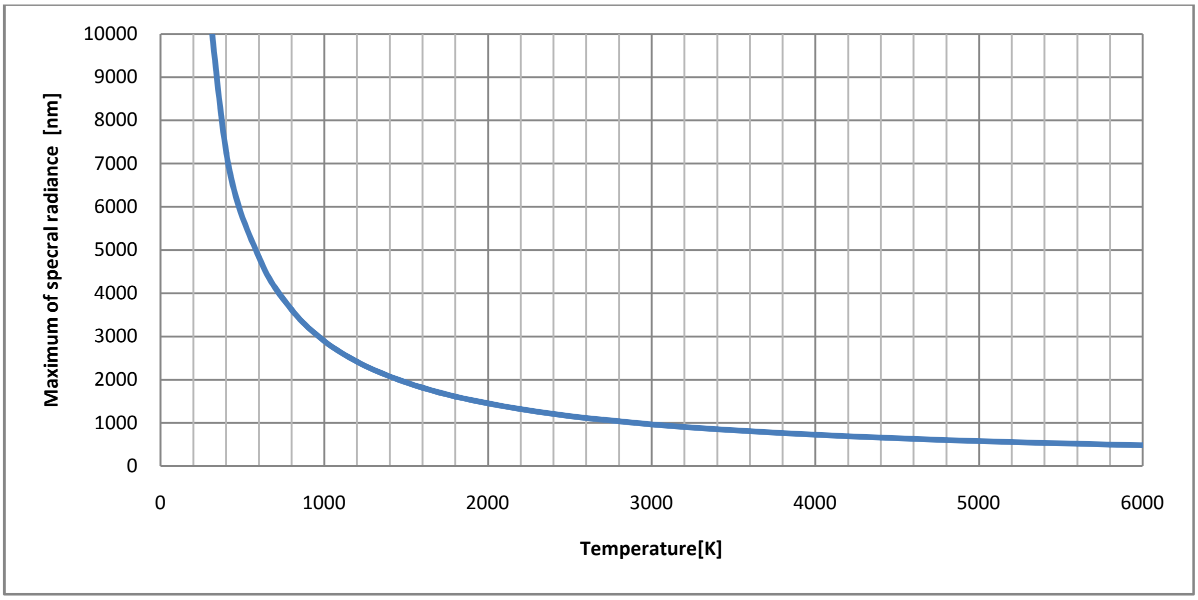


Fig. 3.2. Maximum of spectral radiance versus temperature.

This behaviour can be described by the following equation for the Wien's displacement law:

$$\lambda_{max} = \frac{b}{T},$$

where b is a constant derived from the Planck's law (b=2,89777∗10^{-3} [m*K]) and T is the temperature of the body in Kelvin. Because of this, it is easy to determine the maxima of the spectra emitted by bodies with different temperatures.

Radiating body:	Temperature [K]	Maximum wavelength [nm]
Human body	310	9348
Bulb	2800	1035
Sun	5800	500

For example, this means that a man at ambient temperature of 37 °C that corresponds to the thermodynamic temperature of 310 K emits electromagnetic radiation with a wavelength of $\lambda_{max} = \frac{2{,}89777 \cdot 10-3}{310} = 9{,}35\ \mu m$. And this corresponds (as expected) to thermal radiation.

Based on the above-mentioned findings we can estimate the temperature of the solar surface: The Sun emits yellow-green light with the greatest intensity and this is reflected by the temperature of 5770 °C. Regarding the intensity of radiation of an absolute blackbody Wien derived the following:

$$M_\lambda = \frac{2\pi hc^2}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{\lambda kT}}}$$

where h is the Planck's constant, c is the speed of the light in vacuum and k is the Boltzmann's constant. Wien used an analogy between photons and gas molecules to theoretically explain blackbody radiation. His theory and the results of the experiments of Lummer and Pringsheim fully corresponded only in terms of radiation of small wavelengths. They found certain systematic deviations for longwave radiation. Lord Rayleigh and James

Jeans chose another path to derive the radiation law – the so-called statistical physics. Their formula fully described longwave and shortwave radiation and gave meaningless results which could not be used. The definition is given below:

$$M_{\lambda} = \frac{2\pi kT}{\lambda^4}$$

For example, at any temperature, according to their calculations, the intensity of X-rays should be high, the source should emit an infinite amount of energy. Therefore, it is an obvious nonsense that entered the history of physics under the name of the ultraviolet catastrophe. Therefore, two different laws for spectral radiance were derived, each of which suited only for a certain field or spectrum of wavelengths, the Wien's one for the ultraviolet region, Rayleigh-Jeans' one for the infrared radiation. No law or definition could be applied in the area of medium waves. In 1900 Max Planck combined the two laws. He assumed that the total radiated energy could be divided into individual oscillators. The energy of each oscillator is an integer multiple of some smallest portion of energy. He combined both laws. He assumed that the total radiated energy could be divided into individual oscillators. The energy of each oscillator is an integer multiple of some smallest portion of energy.

$$M_{\lambda} = \frac{2\pi hc^2}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{\lambda kT}} - 1}$$

The Planck's result induced by a desperate step matched the experimental results perfectly. We get the Wien and Rayleigh's laws as approximate relations from the Planck's relation. Regarding the small λT is $e^{hc/\lambda kT}$ >>1 and, therefore, we can neglect 1 in the denominator and we get the Wien's law. For the great λT the exponential in the denominator can be replaced by an approximate relation (the first two terms of the Taylor expansion $e^x = 1+x$) we get the Rayleigh–Jeans Law. Planck further analysed the situation and found that it was no longer possible to think of infinitesimal amounts of energy. The conversion of mechanical energy into radiant or vice versa is not continuous but takes place in whole multiples of a certain quantum of energy, now referred to as the Planck's constant h = $6{,}625 \cdot 10^{-34}$ J s. The energy of each particle is only in certain portions, so-called quantum (from the Latin quantum - amount). The size of this portion is proportional to the frequency of light f and the proportionality constant is one of the fundamental physical constants – the Planck's constant h. [5]

$$E = h \cdot f = h \cdot \frac{c}{\lambda}$$

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. At what wavelengths would you detect the heat of the human body?
2. What is the minimum temperature of the surface in order to be observable by the human visual system?
3. Explain the concept of emissivity and say what values it will reach for the absolutely black body.
4. What is the formula for calculating the photon energy?

4. Photodetectors

In general, each communication system starts with a signal source, continues with the transmission environment and ends with the receiver or detection part of the chain. Basically, the receiving part for the optical domain is called a photodetector, i.e. a device that is capable of converting an optical light signal into an electrical one. It is just the photodetection part of optical systems that represents the greatest challenge, because the detection of a light signal is associated with many limitations and, therefore, fundamental limitations in performance occur. The main effects are especially different types of noise and relationships between sensitivity and communication speed. It will always depend on which signal we want to detect, e.g. measuring intensities at the level of individual photons will be different to receiving energy from the Sun. Similarly, the photodetectors will differ for different parts of the light spectrum. The reason is that covering the range from the UV to IR spectrum is not easy. In this chapter we will examine an overview of principles of individual radiation detection options and their features. [4]

Photocathodes

A photocathode is a photosensitive layer that is able to change the incident flux of photons into the corresponding flux of electrons emitted into the vacuum space with the possibility of their collection or multiplication. Physically, it is a thin semiconductor film with a conductive contact for an external electrical circuit. Electrons must be collected by means of an anode that is maintained at the positive potential.

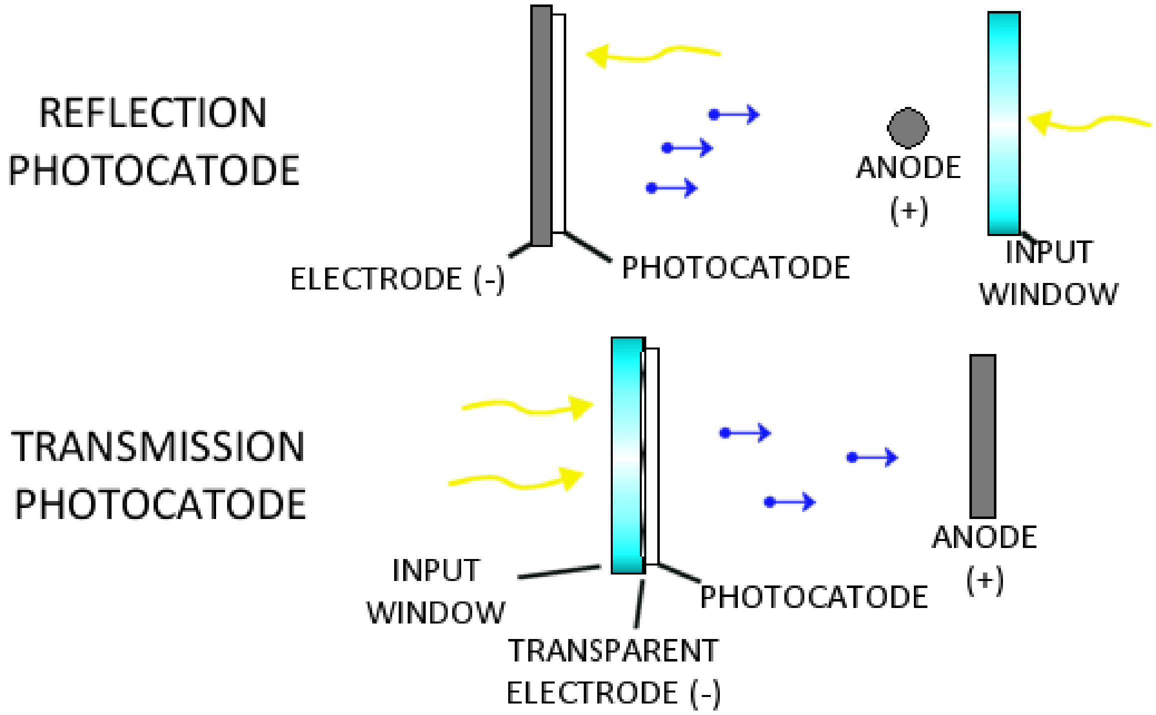


Fig. 4.1 Reflection and transmission photocatodes.

The photodetection process can be described in the following steps:

1. Photon absorption with electron-hole pair formation
2. Electron diffusion on the detector surface
3. Electron emission into the vacuum space

The process of an electron emission does not proceed with 100% success. In addition, if an electron is to leave material it needs enough of energy to do so.

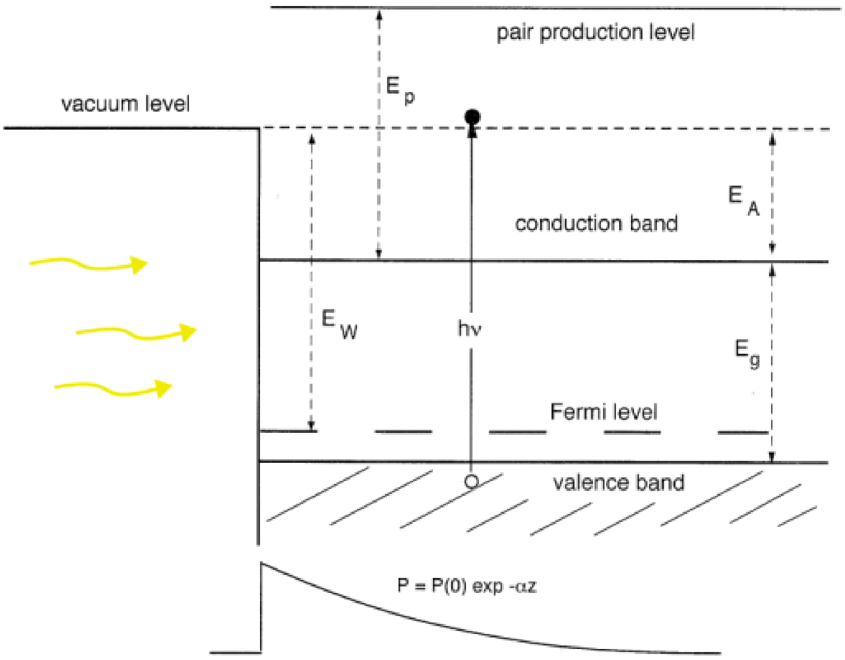


Fig. 4.2 Energy levels of photocatode.

The conduction band and valence band are separated by the bandgap (energy gap) E_g. For example, then the vacuum level represents the minimum energy required for the electron to leave the material.

The working area and activation energy are designated as E_W and E_A.

The energy of a photon corresponds to a relationship

$$E = hv = \frac{h \cdot c}{\lambda}$$

Therefore, the required photon energy must correspond to the relationship

$$hv \geq E_g + E_A$$

It follows that acquisition of the required energy depends on the wavelength of the incident radiation. Thus, we can define the minimum wavelength that is needed to overcome the vacuum band as:

$$\lambda_t = \frac{h \cdot c}{E_{g+}E_A}$$

After overcoming this condition, there is a problem with the environment because the electron is deflected on the way by the remaining atoms in the environment (quasi-elastic scattering). However, this scattering represents only a small influence and the electron reaches its target before it loses the necessary energy. The so-called ionization scattering that causes a great loss (>E_g) is a bigger problem. However, it only applies if the energy is above the E_p level (pair production level).

With these defined rules, we get to the spectral response of photodetection. There will be a zero for $\lambda > \lambda_t$

Starting with $\lambda < \lambda_t$ the response will increase to a maximum and then decrease again to zero when the E_p level is reached.

In the application environment, we use the term **spectral efficiency** to describe this behaviour that is defined as the ratio of the generated photocurrent and the incident power:

$$\sigma(\lambda) = \frac{I}{P} \text{ [A/W]}$$

Alternatively, we may also encounter the term quantum efficiency

$$\eta = \eta(\lambda)$$

that relates to the number of generated photons – to the number of the received photons (*F/F'*).

After connection of those two relations, we can rewrite current as *I=eF'* and power as *P=F'hv*.

Then the quotient $\sigma = I/P$ will be

$$\sigma(\lambda) = \eta \frac{e}{h\nu} = \eta \frac{\lambda}{\frac{hc}{e}} = \eta \frac{\lambda}{1.24} \quad \text{[A/W]}$$

In practice, you will find graphs describing spectral efficiency in datasheets. It is the easiest way to describe graphically the behaviour of any photosensitive element. The features of photodetectors are never perfectly universal which is due to the physical nature of the function. It will always depend on the chemical composition of the components, construction, size, etc.

In the following Figure 4.3 we can see a typical example of such features for different chemical composition. It is obvious that the material here significantly affects the range of usable wavelengths. Moreover, let's note that the spectral efficiency here reaches relatively low values (basically everything is below 0,1A/W).

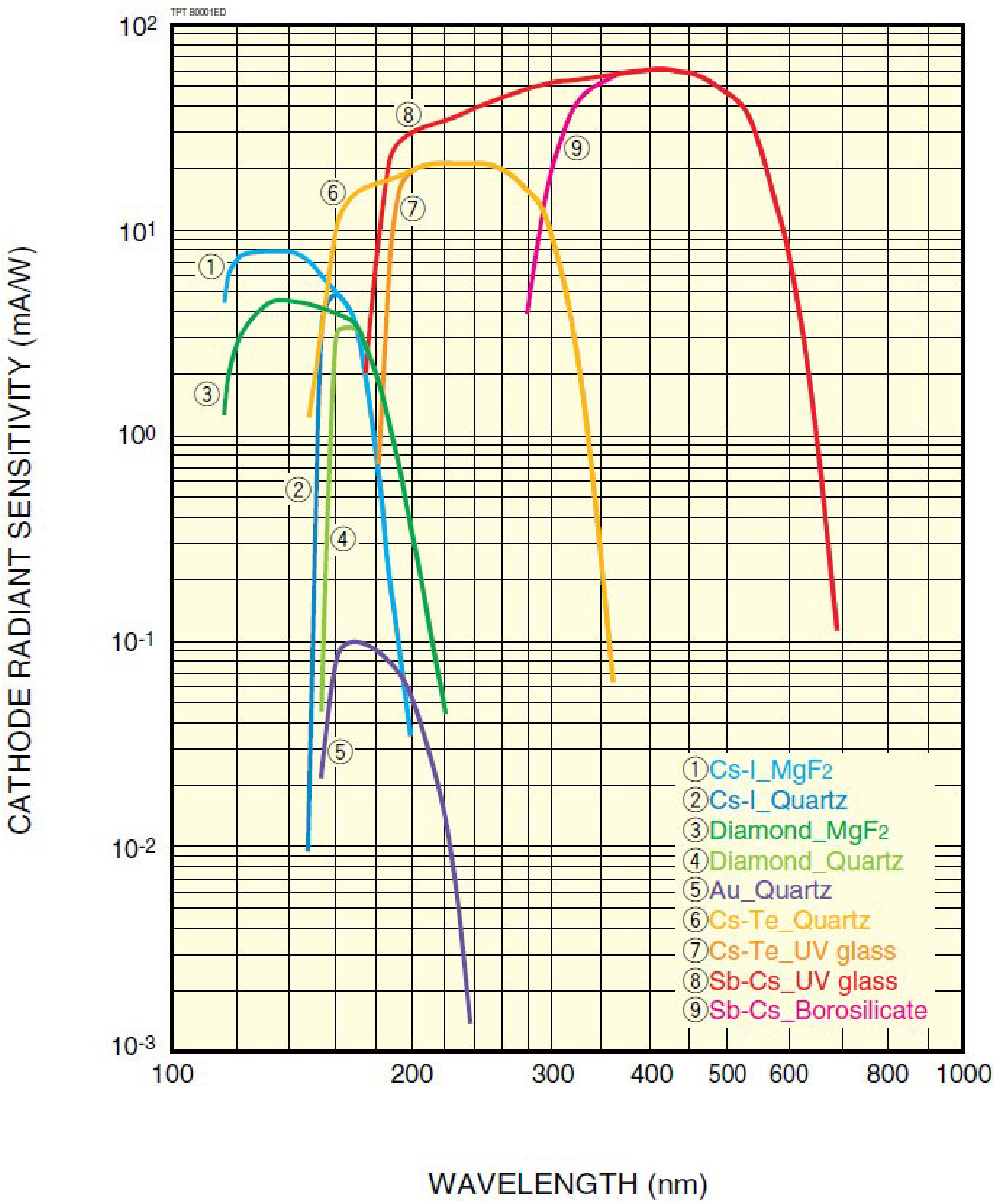
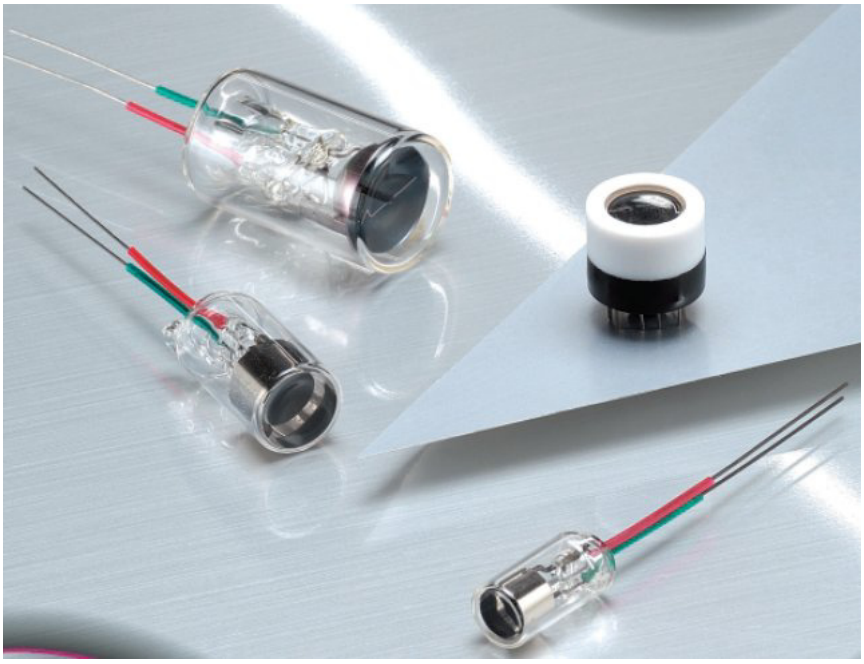


Fig. 4.3 Photocatode efficiency.

If we talk about a photocathode we talk more about the principle. The devices themselves that use this principle are, for example, phototubes or photomultipliers.

Phototubes



Obr. 4.4 Phototubes.

The phototube, sometimes also called the vacuum photodiode, is a simple device in which the photocathode and anode are vacuum sealed inside a transparent case (or a case with a transparent visor).

Advantages:

- Detection also in the UV spectrum (see the previous picture of spectral sensitivity)
- Temperature stability
- Large photosensitive area
- Speed

Disadvantages:

- Not able to view the IR spectrum

Usage:

- Spectrophotometers
- Ozone detection
- Monitoring of sterilization lamps

Wiring example:

The advantage is that operation of some phototubes is sufficient even with a relatively low voltage (15V in this case). However, the values of the closing voltage can reach up to 100V in principle.

The wiring in the Figure is called a transimpedance amplifier. The photocurrent Ip is converted to a voltage here by means of a Rf resistor.

*Note: As the photocurrent is small, a relatively high Rf value Rf is chosen (on the order of MOhm). The filter capacitor Cf is here to achieve better stability and, on the contrary, it is very small (on the order of pF).

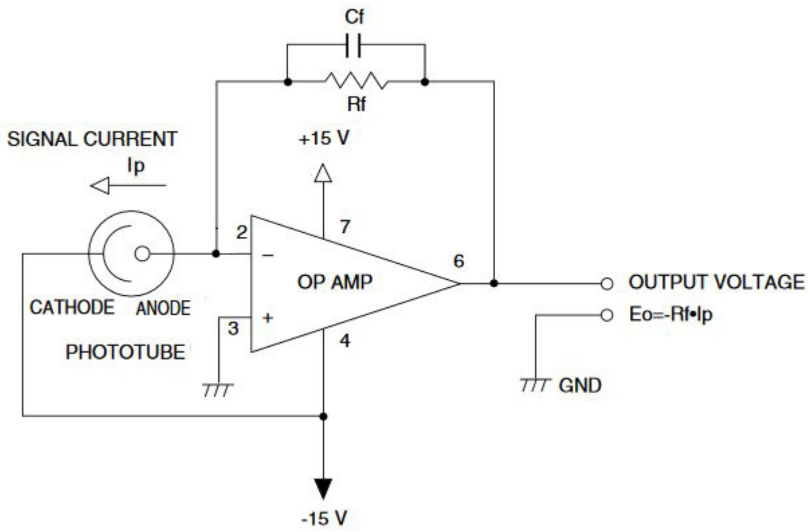


Fig. 4.5 Phototube connection example.

Photomultiplier

Fig. 4.6 Photomultiplier.

This device again uses photocathodes. Nevertheless, there has been a significant modification. Although the photocathode is able to detect individual photons, electrically such a small change is made that the detection at this level is almost impossible. As the name suggests, the photomultiplier has the ability to multiply their number as electrons pass toward the anode. Therefore, a single photon supplies enough electrons to the anode for reliable detection.

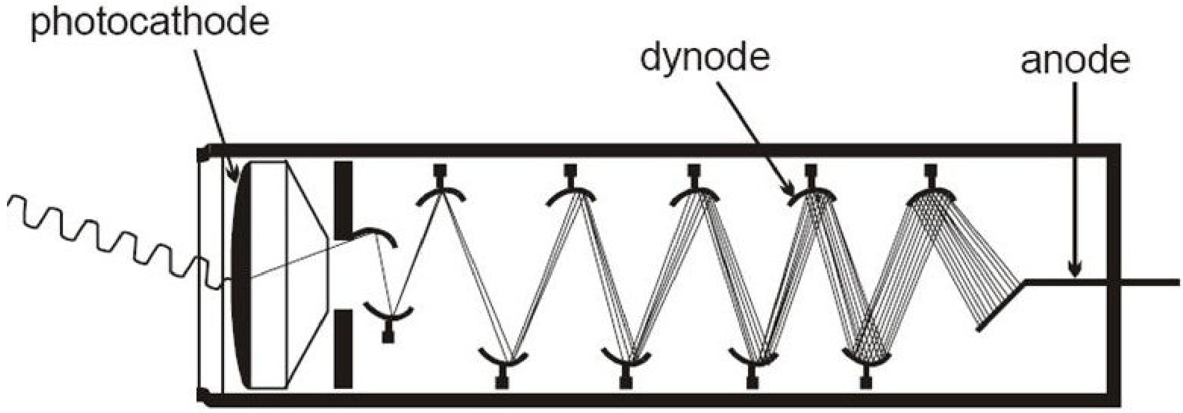


Fig. 4.7 Photomultiplier principle.

We can see the basic scheme of the photomultiplier in the Figure when the photocathode takes care of the conversion of photons to electrons, the electron optics takes care of the focus of electrons on the first dynode of the multiplication chain. The multiplication chain will take care of the necessary high internal profit and, in the end, of course, the anode will take care of the collection of the acquired charge.

High voltage (usually above 100V) is used between the photocathode and the first element of the multiplication chain (1st dynode). Likewise, between other elements of the chain, in order to ensure needed energy for multiplication.

Just to give an idea: Therefore, a single photon can produce photocurrent of the 1mA size which is an easily detectable value.

Newly, we come across the term ‘photon counting’. Thus, the output of the measurement will be the information on the number of received photons per a unit of time, the so-called Counts per second (CPS)

Semiconductor photodetectors

The PN junction uses an internal photoelectric effect when it is able to absorb incident photons and generate charged carriers (electrons). Then these carriers generate corresponding photocurrent on the contacts.

Even here, the energy of the photon (E) must meet the boundary conditions

$$\frac{h \cdot c}{\lambda} \geq E$$

The limit wavelength will be based again on the relationship

$$\lambda_s[\mu m] = \frac{h \cdot c}{E} = \frac{1.24}{E} \qquad \text{[eV]}.$$

From the point of view of the PN junction, the mentioned energy corresponds to the energy of the bandgap ($E=E_g$) which depends on the chemical composition of the junction.

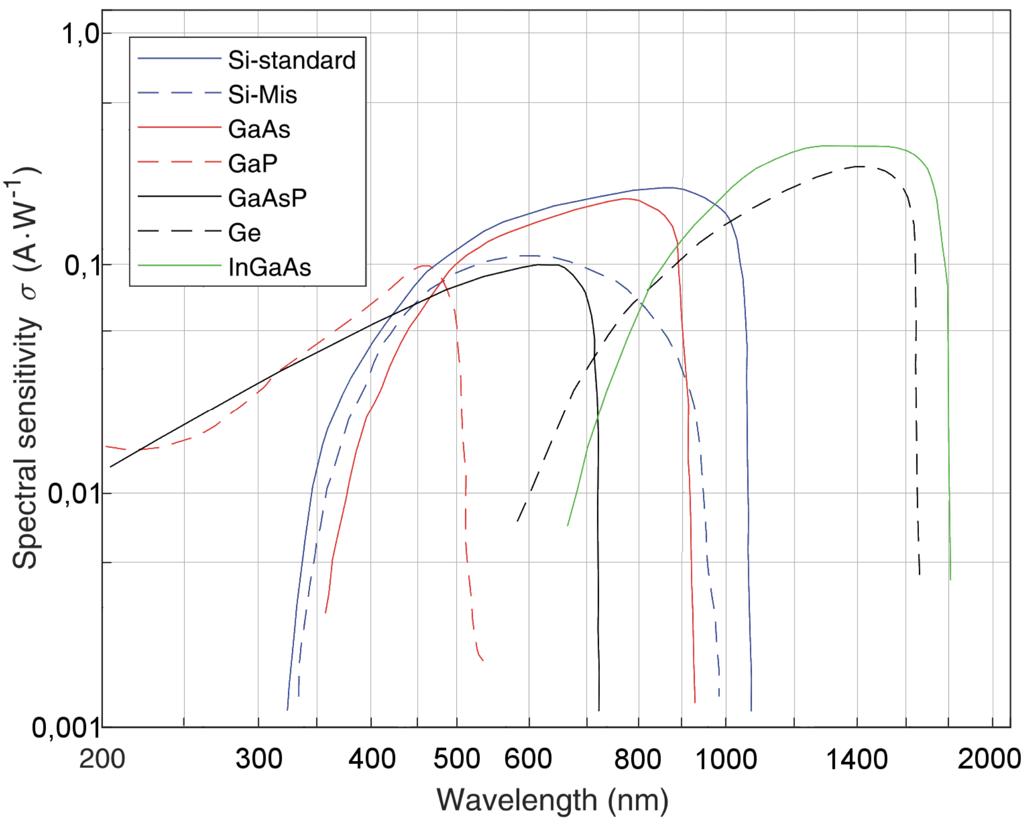


Fig. 4.8 Spectral sensitivity vs. junction material.

Photodiodes PN

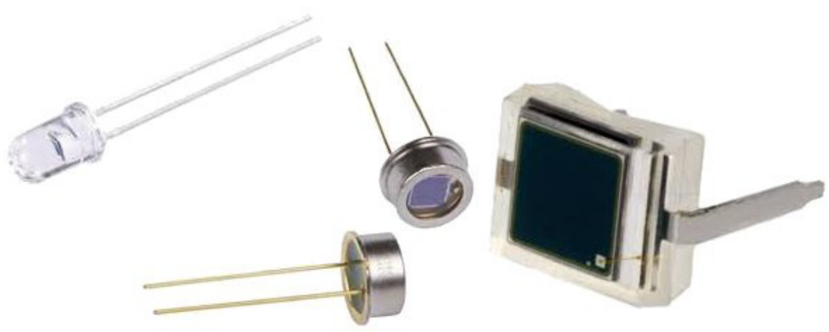


Fig. 4.9. Photodiodes.

Photodiodes can be currently considered to be the most used elements in the area of photodetectors. Especially thanks to prices, durability and high efficiency. The advantages are the spectral characteristics when we cover the complete visible area and IR spectrum up to the boundary of 1100nm with a simple silicon junction (Si).

Structurally, it is the PN junction with a transparent window that allows the passage of photons.

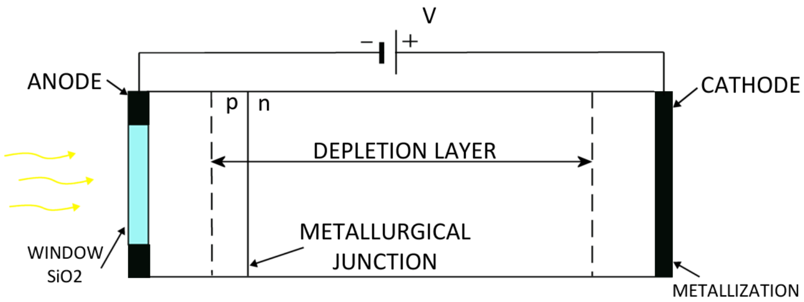


Fig. 4.10 PN photodiode construction.

When the closing voltage is applied, the depletion area then expands, which is done, for example, in order to shorten the reaction time in practice (the capacity of the PN junction decreases and thus the maximum operating frequency increases).

This mode with applied negative voltage is called the photoconductivity mode (III. quadrant of the V/A characteristics). Then, the closing voltage is chosen to be as large as possible based on the photodiode. Therefore, a low value of the transition capacity Cb and a low series resistance Rs are achieved.

The second option is to leave the photodiode without supply voltage in the so-called photovoltaic mode (IV. quadrant). In this case, we will not achieve high operating frequencies, but we will get, for example, a better response for measurement. In this case, the photodiode behaves as the power supply (similar to a photovoltaic panel).

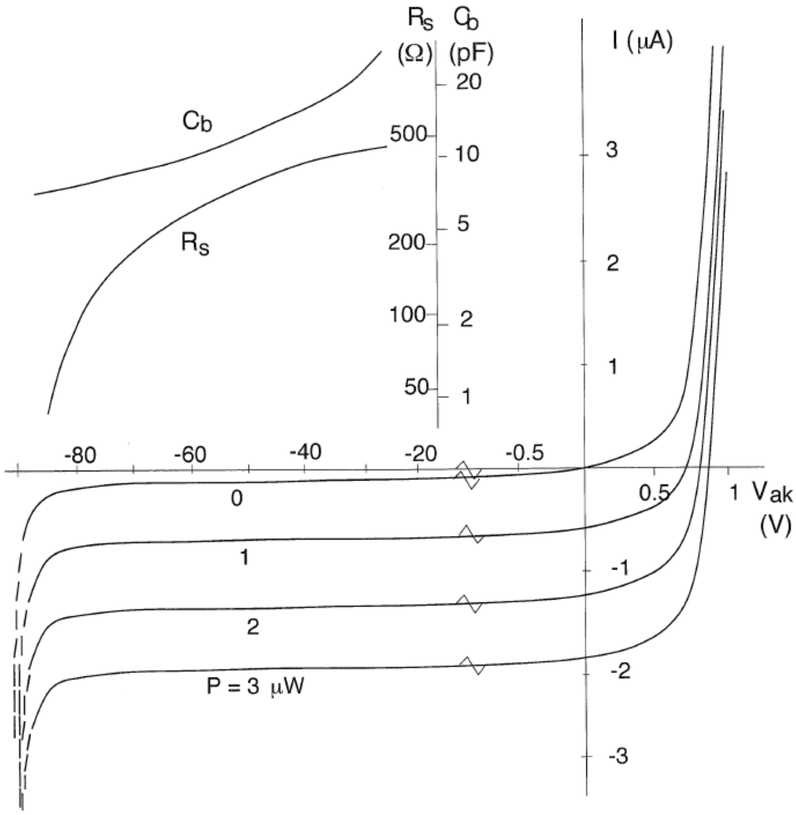


Fig. 4.11 VA photodiode characteristics.

The basic and replacement scheme is shown in the figure 4.12.

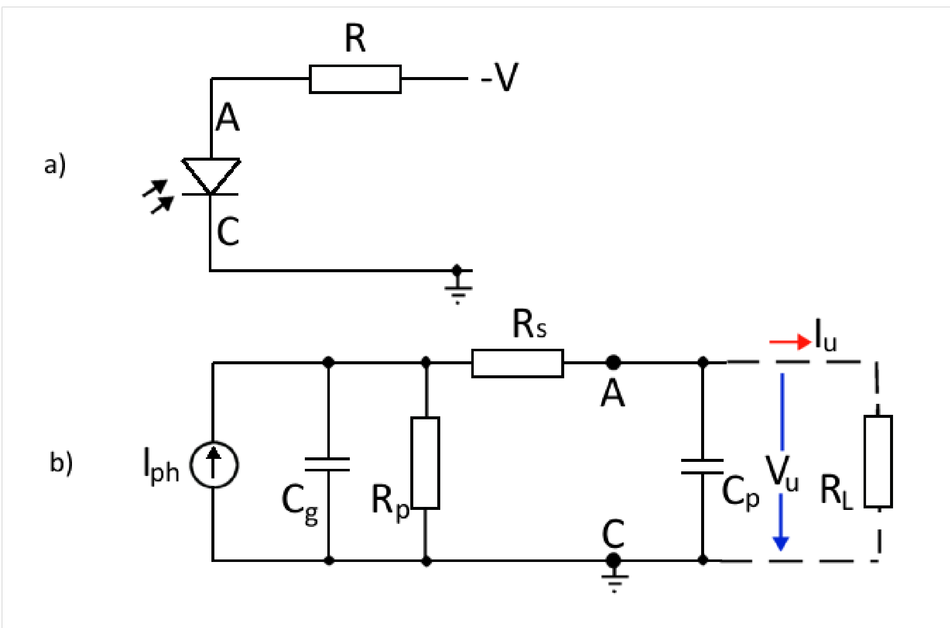


Fig. 4.12 The basic and replacement scheme of photodiode.

The basis shows the part (a) when a photodiode is connected to the closing voltage via a load resistor R. The signal that we are able to detect with the photodiode will be reflected in changes in the Vu voltage. If we analyse the behaviour of the photodiode, we will get to the replacement scheme in the part (b). The basis is the source of the photocurrent I_{ph} that represents the process of converting photons into electrons. Especially for the AC signals, the transition capacity of the junction that is represented by C_g is critical. The faster the signal changes in the received signal, the greater the effect this capacity will have on the signal. The Rp value is usually very high (on the order MOhms). This is the reason why it is advisable to choose amplifiers with FET input circuits in the next stages (they also have a large internal resistance on the order of MOhm, which enables better impedance matching). The value of the series resistance Rs is relatively small according to the V/A characteristics. However, even the load resistor R will have a similar value for higher frequencies that will create unpleasant voltage drops.

The last mentioned is the parasitic capacity C_p. This is formed by the capacity between the contacts of the photodiode and the rest of the conductive paths to the amplifier. If we try to calculate Xc for the frequency 100MHz and the value Cp = 1pF, we will get

$$X_C = \frac{1}{2\pi f C_p} = 1591\Omega.$$

So, even at relatively minimum values, the signal would be affected. Therefore, an important finding regarding CP is that the connection to the photodiode must have the lowest possible capacity (in other words, it must be as short as possible).

PIN Photodiodes

This is a much more desirable type of a photodiode with a modified structure where an intrinsic area is inserted between the P and N material. It brings several advantages:

- The width of the depletion layer W is no longer dependent on the magnitude of the closing voltage (it is now defined by the geometry of the transition). Therefore, we don't need high closing voltage.
- Frequency response doesn't depend on the wavelength
- Final current is less dependent on the final voltage thanks to the high level of Rp.

Thus, compared to the classical PN junction, the PIN photodiode has the advantage of low dependence on the closing voltage.

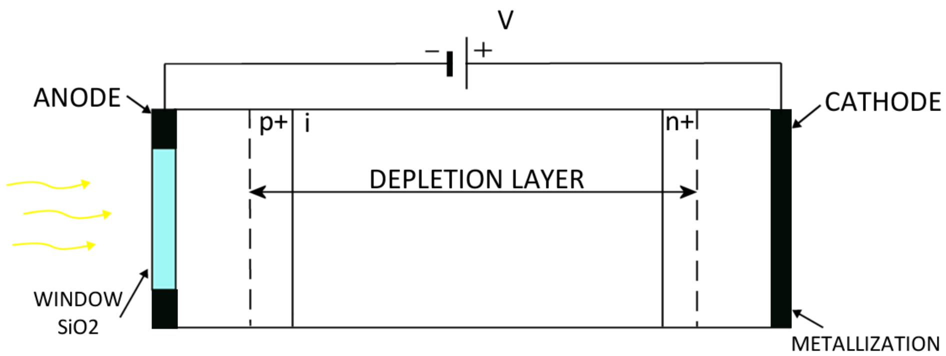


Fig. 4.13 PIN photodiode.

Schottky photodiodes

This type of photodiode uses a metal/semiconductor junction. Here, the photons pass through a partially permeable metal layer and the electrons generated in the depletion region are collected by an electric field in order to increase the photocurrent. Nevertheless, due to the reflections of incident radiation from the metal layer, it has a significantly lower efficiency compared to the PN/PIN photodiodes (especially at longer wavelengths).

This construction is characterized by:

- High speed of response (orders of GHz)
- Easy production

The use is rather around the UV area, or for fast communication with enough power.

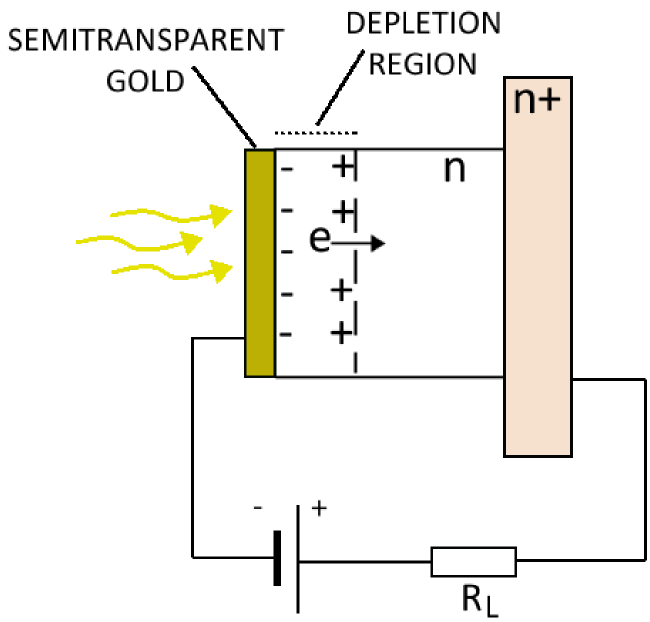


Fig. 4.14 Schottky photodiode.

Avalanche photodiodes (APD)

This photodiode is able to amplify the received signal within its internal structure. The avalanche photodiode consists of four layers: $N+$, P, pure semiconductor and $P+$. Around the $N+$ and P layers, between which the avalanche phenomenon occurs, there is a protective ring made of an N-type semiconductor, which increases the diode's resistance to the surface voltage cross-sections. The avalanche multiplication process works in such a way that the generated electron is gradually accelerated by an electric field in the depletion layer. As soon as it reaches a sufficient speed, it can generate another electron-hole pair in case of collision with the atoms in the semiconductor (impact ionization process). So, now we have 2 electrons which are accelerated again and the next collision will generate 2 more electrons.

The result is the avalanche amplification process when the number of electrons increases exponentially.

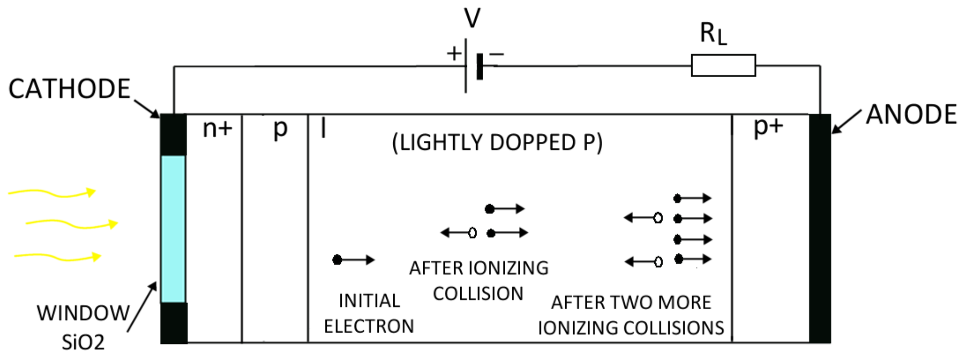


Fig. 4.15 Avalanche photodiode.

The photodiode amplification APD is described by the multiplication factor M that is the ratio of the photocurrent with the amplification to the photocurrent without the amplification. (M is usually around 100 for the Si).

The multiplication principle is basically a solid state similar to a photomultiplier, so even here a relatively high voltage is needed that is able to sufficiently accelerate the acquired electrons. (M is usually around 100-200V for the Si).

Thanks to high voltage, the photodiode was even able to maintain high speed. Nevertheless, the biggest problem is the temperature dependence when the amplification decreases significantly with increasing temperature. When working with variable temperature of the environment, it is convenient to take into account temperature stabilization.

Phototransistor

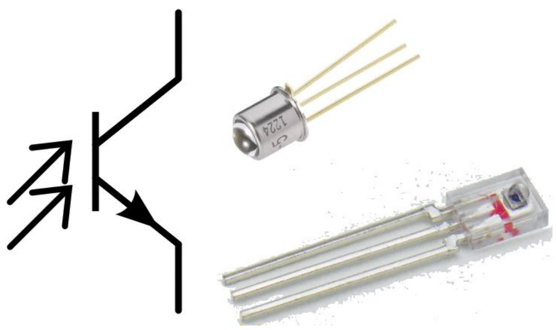


Fig. 4.16 Phototransistor.

Further use of the PN junction and amplifying structure is the phototransistor. Revealing the access basis for light radiation, the photons now generate a photocurrent based on it that is actively amplified by the transistor. Therefore, the current of the collector is h_{21}xlarger than the photocurrent generated at the base-emitter junction.

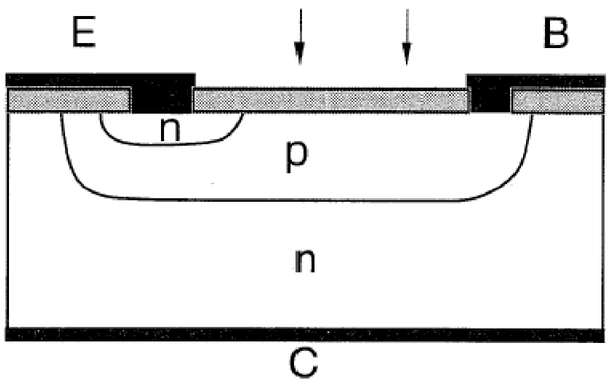


Fig. 4.17 Phototransistor internals.

Moreover, the transistor has a floating base in this mode. Therefore, in the principle, it does not matter whether the load is connected to the side of the collector or the emitter. The following connection shows the moment when the load is connected behind the emitter. If the photocurrent is zero, only a small dark current will flow and the Vu will be almost 0V. As soon as the photons start to fall on the basis, the current flowing through the transistor starts to cause a corresponding voltage drop. (Based on the known formula U=R*I). In the saturation, there will be V_u=+Vbb-U_{CEsat}. This behaviour is, from a point of view of the circuit, an advantage because we can easily adapt the operating logic.

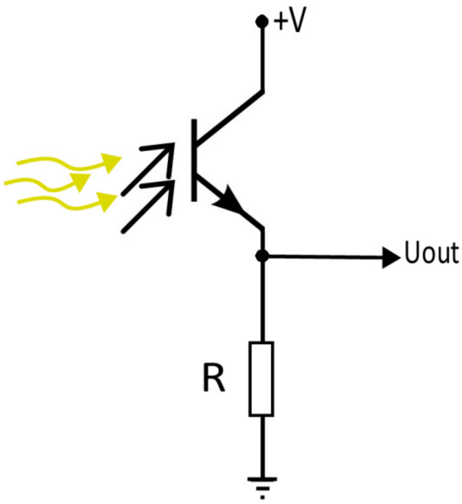


Fig. 4.18 Phototransistor basic scheme.

Photoconductors (Photoresistors)

It is a simple block of semiconductor material (like cadmium selenide CdSe or cadmium sulfide CdS…) that changes its conductivity when it is exposed to the light radiation (also called photoresistance or photoconductance).

Photoresistor has advantages mainly in terms of simplicity of the production and price. Within the electrical connection, we can treat it as a variable resistor, so that its use in various applications is easier.

Nevertheless, in comparison with the photodiodes, it is much slower (leading edge time (t_r) and trailing edge time (t_f) move here in the tens of ms). Therefore, these elements are not suitable for data communication. However, they can be simply used in the area of measurement of lighting and various detectors in the visible and IR areas.

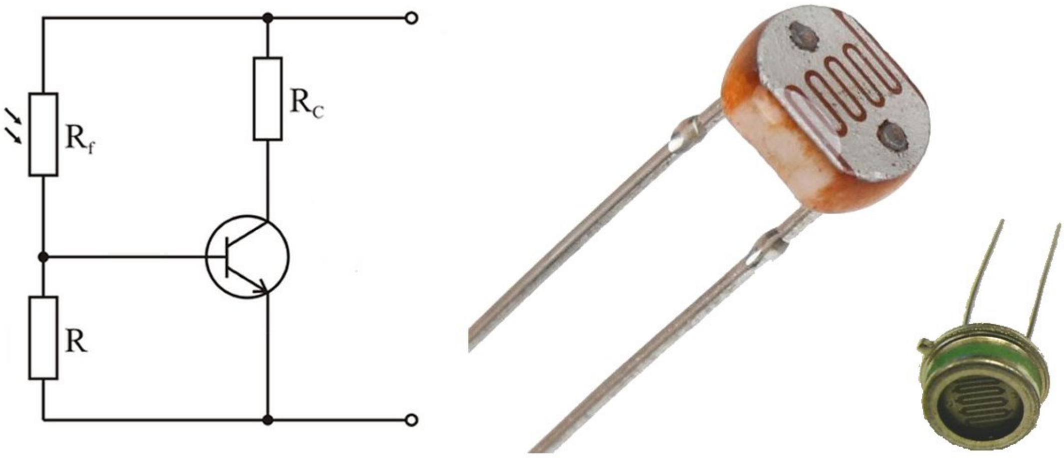


Fig 4.19 Example of photoconductor circuit.

Properties of photodetectors: Summary

In the previous paragraphs we described some basic principles and properties of photodetectors. Now, let´s make a small summary of what we might be interested in regarding these elements in practice.

- **Spectral sensitivity** – the wavelength range in which the selected detector is able to operate with a certain efficiency.
- **Spectral peak** – the wavelength to which the element has the highest sensitivity
- **Dark current** – a value that indicates the amount of current that flows through the element even without incident photons.
- **Temperature dependence –** as the temperature increases, the electrons in the material can move better, so the dark current increases with temperature.
- **Response time** – is described by the leading edge time and trailing edge time that describe the response to a rectangular signal.
- **Bandwidth** – the maximum operating frequency in certain circumstances (usually at specific values of supply voltages and load resistance).
- **Maximum closing voltage** – determines the maximum value of the closing voltage after which the detection element will be damaged.
- **Angle of half sensitivity** – the receiving angle of the element at which it has at least 50 % of the maximum sensitivity compared to the straight direction.

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. What basic types of photodetectors do you know?
2. What type of photodetector would you choose to detect extremely low power?
3. What does the spectral characteristic of a silicon photodiode look like?
4. Is it possible to detect the visible spectrum of light using the InGaAs photodiode?
5. What is the point of supplying a reverse voltage to the PN junction of a semiconductor photodiode?
6. What is the principle of APD operation?
7. What is the basic connection of the phototransistor?
8. Why are photoresistors not suitable for data communications?

5. Optical properties of materials

Various materials have electrical, mechanical or magnetic, as well as optical properties. It applies for each material that the part of the incident light (or electromagnetic) radiation is in different proportions:

- Reflected
- Absorbed
- Released

Reflectivity of the material

It describes how much energy was reflected from the material compared to the incident energy. Based on the optical properties of the material, the spectral characteristics of the reflected radiation may especially differ. The resulting reflected electromagnetic spectrum can then be perceived in the visible spectrum as the colour of the material. There is an example of a red, optically opaque material in the following figure.

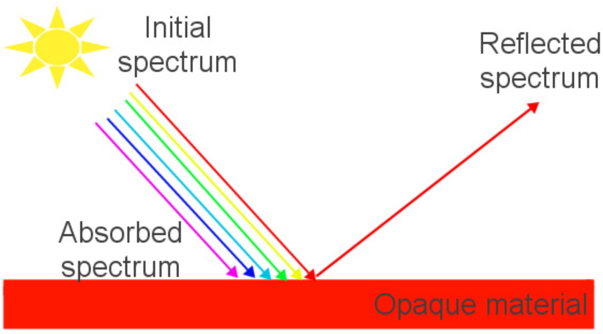


Fig. 5.1 Retro-reflecting material.

The spectrum of solar radiation falls on the material when the red part of the spectrum is reflected. The rest of the radiation is absorbed by the material to which it transfers energy, mostly in the form of heat. Therefore, the material is red for the observer.

Absorption of the material

If the radiation falls on the obstacle from which it does not bounce or pass through, we speak about the radiation absorption. This phenomenon can also occur gradually – for example, it depends on the thickness of the material. During the absorption of radiation, photons transfer energy most often in the form of heat. This means that the material does not retain energy and radiates it to the surroundings in the infrared region. (More in the chapter Blackbody Radiation). In a nutshell, it can be said that the absorbed part of the electromagnetic spectrum is converted to another part of the electromagnetic spectrum with lower energy. That is based on the formula for the photon energy:

$$E = \frac{h \cdot c}{\lambda}$$

Therefore, it can be seen the photon energy decreases with an increasing wavelength. That is why the absorbed energy is radiated to the surroundings in the infrared region.

However, in some cases the composition of the material allows them to radiate energy even in the visible area. This property is called the **photoluminescence** (the light emission by the substance caused by the light). The incident radiation with higher energy (for example in the blue or UV region) will cause the material to start producing secondary radiation of longer wavelength. Likewise, it takes a substance to cumulate the heat, even giving up the energy this way can last for longer period of time. In this case we can recognize 2 basic mechanisms:

- Fluorescence – the effect takes place almost immediately and the energy is reflected to the surrounding quickly, 'afterglow' of the irradiated area lasts less than a millisecond.
- Phosphorescence – the energy stays in the material longer and is released gradually (even for several days). Nevertheless, the intensity of radiation is lower here (it is spread over a longer period of time).

Transparency of the material

The transparency of the optical (electromagnetic) radiation determines the part of the spectrum that is able to pass through the material. Basically, this is the last part of the radiation that was not reflected upon entering the material and was not absorbed by it during the passage.

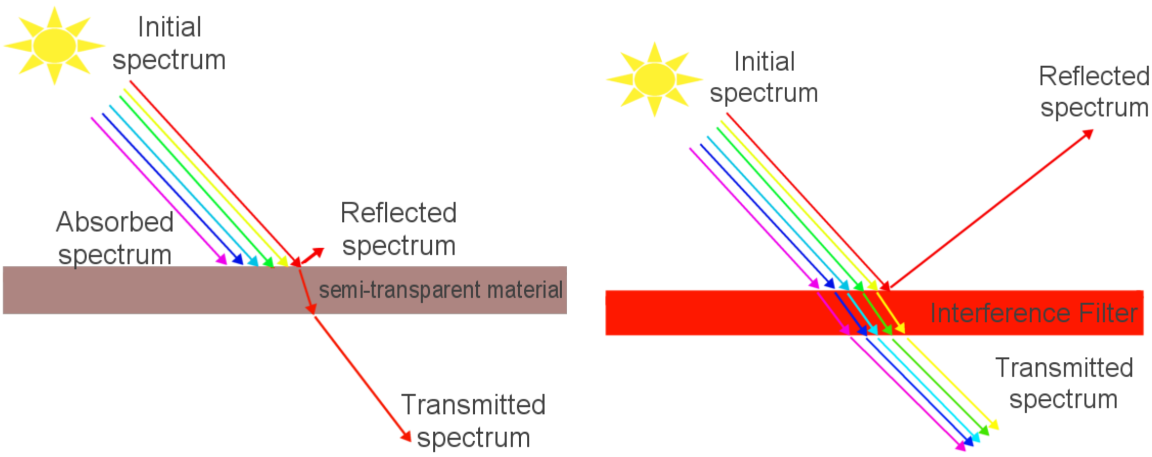


Fig. 5.2 Material transparency.

The picture on the left shows a case when a certain material transmits only a specific part of the spectrum. The picture on the left shows an example of an interference filter that reflects a specific wavelength and the rest of the light can go further. Therefore, it is an optical filtration because we can choose which part of the spectrum of the light radiation the given type of filter releases.

The following example shows how such filters can behave. It is a simple coloured acrylic glass in case of Fig. 5.3. The width of the transmitted spectrum is relatively wide here since this results in a larger part of the energy passing through the filter.

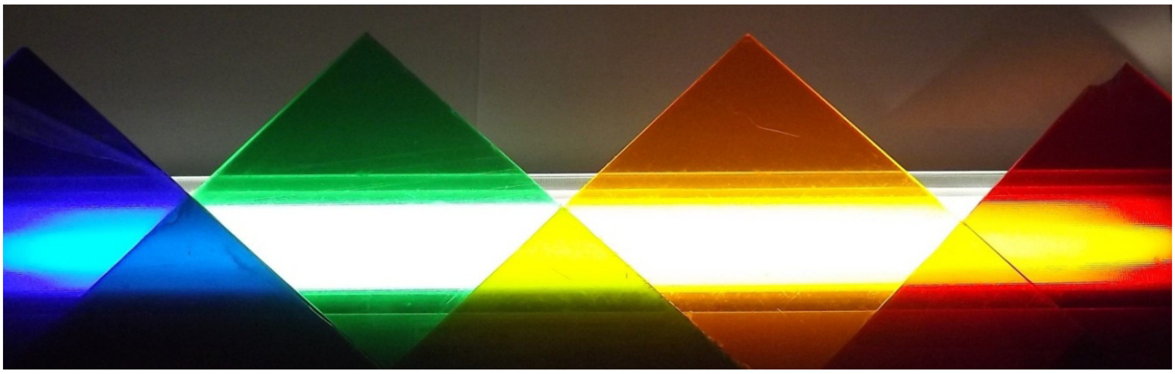


Fig. 5.3 Transparency of acrylic glass.

On the other hand, it can be seen that the windows of transparency clearly overlap. The reason is that in places where 2 filters stand in the way of light, the part of the spectrum of light radiation can still go through/pass. In order to get an idea of how much of the spectrum passes, we can look at the spectral characteristics of the mentioned acrylic glass.

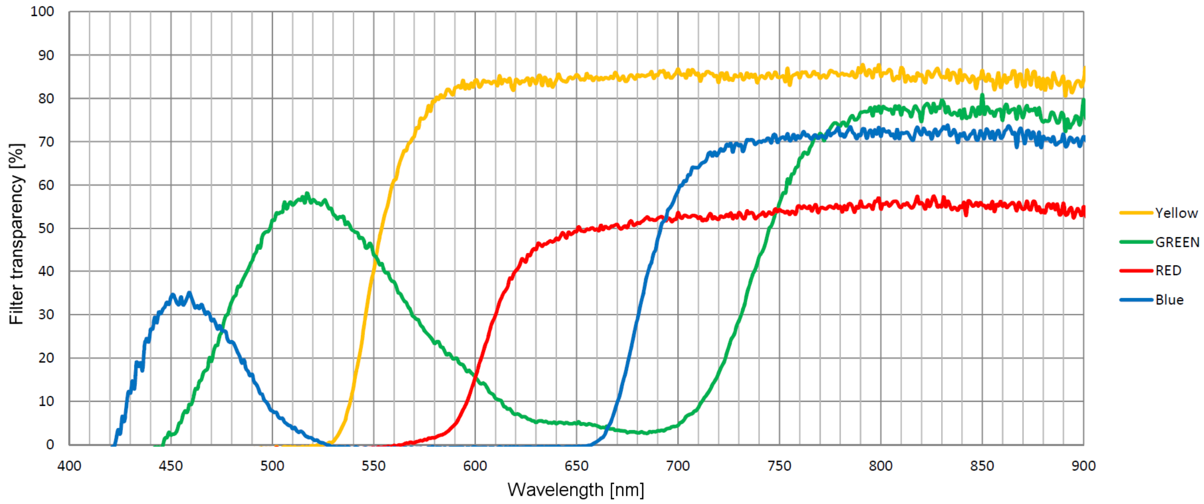


Fig. 5.4 Spectral characteristics of acrylic glass.

Figure 5.4 demonstrates how the individual wavelengths behave as they pass through the material. It is worth mentioning the part in the invisible region (IR part of the spectrum). Regardless of the colours that we can observe with the human eye, all the mentioned filters are transparent for the part of the IR spectrum. Therefore, they hardly affect the transmitted light above 800nm. This is used, for example, to discreetly cover cameras with night vision.

In case of usage of higher quality filters, the results will slightly change.

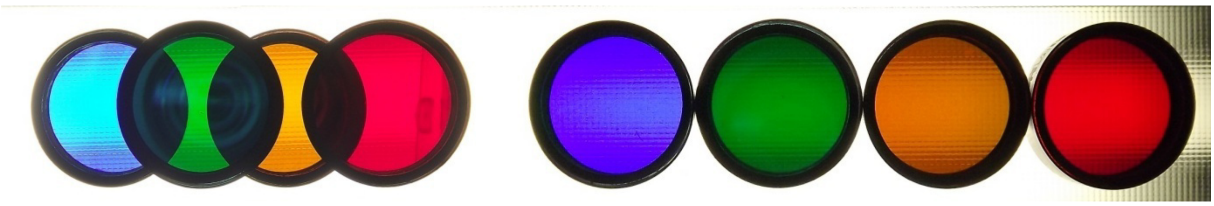


Fig. 5.5 Transparency of the interference filters.

The visible spectrum of light radiation no longer passes in the place of overlap. We really separate the individual wavelengths in this case. When looking at the spectral characteristics, the difference is obvious at first glance. The 2 types of filters with widths of spectral line (FWHM Full Width Half Maximum) 10 and 40 nm are displayed. In addition, these filters work in the near IR spectrum. This information can be very important in practice, some filters do not have to cover this area. In principle, the producers state, for example, the working wavelength, FWHM and the area of wavelengths where the filter operates. It is not always necessary to solve the IR spectrum.

400 - 490 nm Bandpass Filters

Item #	CWL[a]	FWHM[b]	T (Min)[c]	Blocking[d]	Transmission/ OD Data[e]	Laser Line
FB440-10	440 ± 2 nm	10 ± 2 nm	45%	200 - 3000 nm		N/A
FL441.6-10	441.6 ± 2 nm	10 ± 2 nm	60%	200 - 1150 nm		HeCd
FB450-10	450 ± 2 nm	10 ± 2 nm	45%	200 - 3000 nm		N/A
FB450-40	450 ± 8 nm	40 ± 8 nm	45%	200 - 1150 nm		N/A
FL457.9-10	457.9 ± 2 nm	10 ± 2 nm	65%	200 - 1150 nm		Argon
FL460-10	460 ± 2 nm	10 ± 2 nm	65%	200 - 1150 nm		Argon

Fig. 5.6 Example of parameters from the producer.

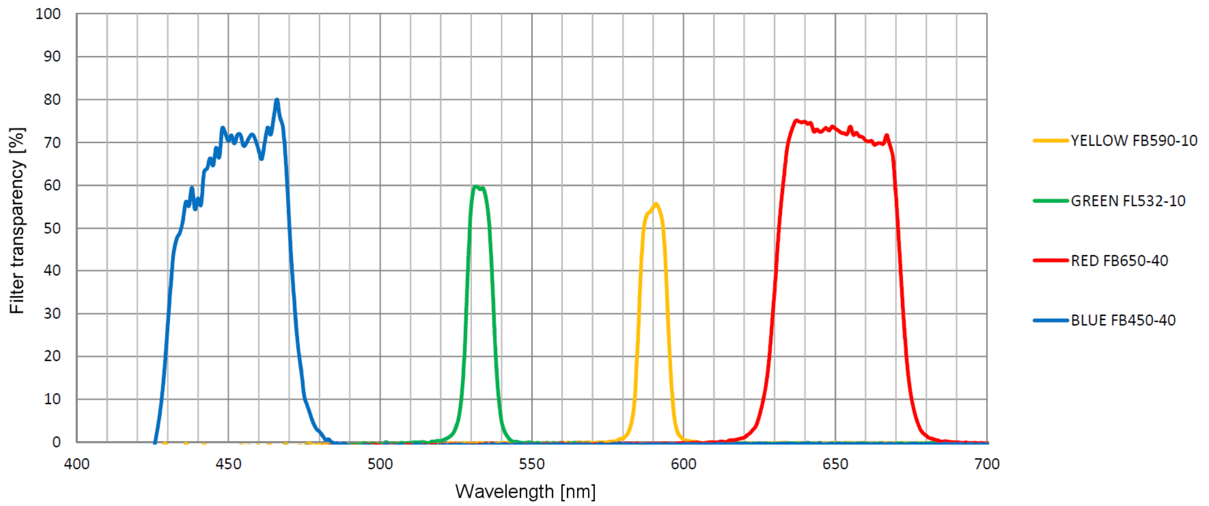


Fig. 5.7 Spectral characteristics of the Thorlabs interference filters.

Types of materials based on the interaction with light

We have already said that materials with incident light radiation react by a combination of different mechanisms. And we can divide materials into 3 basic types based on these combinations.

- **Transparent materials** – light radiation (or at least some of its spectral part) passes through the material and we are able to read through it.
- **Translucent materials** – light radiation is scattered to the surroundings as it passes through the optical environment. This means that in spite of the fact that the light can pass through, the image is distorted. Alternatively, it is not recognizable at all. This

property is used, for example, in illumination, where the covers of the luminaires are able to fulfil the lighting function, but it is not possible to see inside the luminaire itself.

- **Opaque materials** – are materials for which the light cannot pass through. However, the light may still bounce off the surface. Therefore, the material can have parameters such as colour or reflectivity, etc.

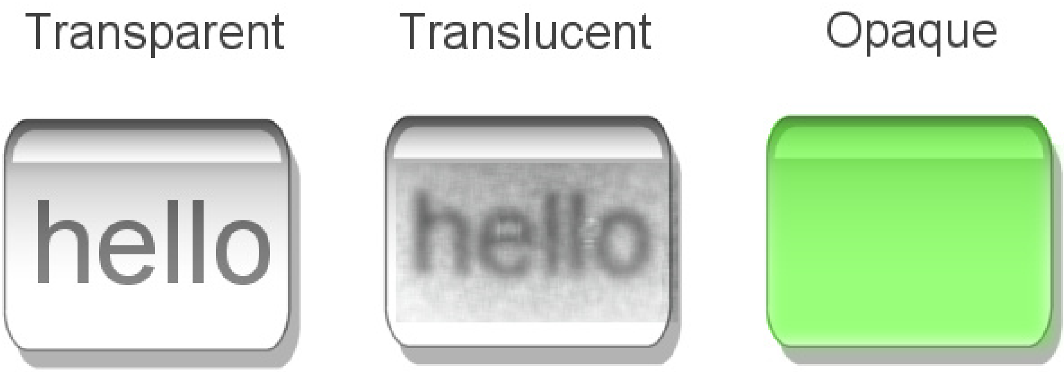


Fig. 5.8 Types of materials based on transparency.

Measurement of properties of optical materials

We have several options for measuring optical materials. First of all, it is necessary to realise what is available for measurement.

Optical power measurement

The basic measurement is the measurement of optical power. If we have a source of optical radiation of known power, we can use the photodetector to measure both the reflected and transmitted power. Everything that is left after adding them up, logically, falls on absorption. The following figure shows an example of a simple principle of this measurement.

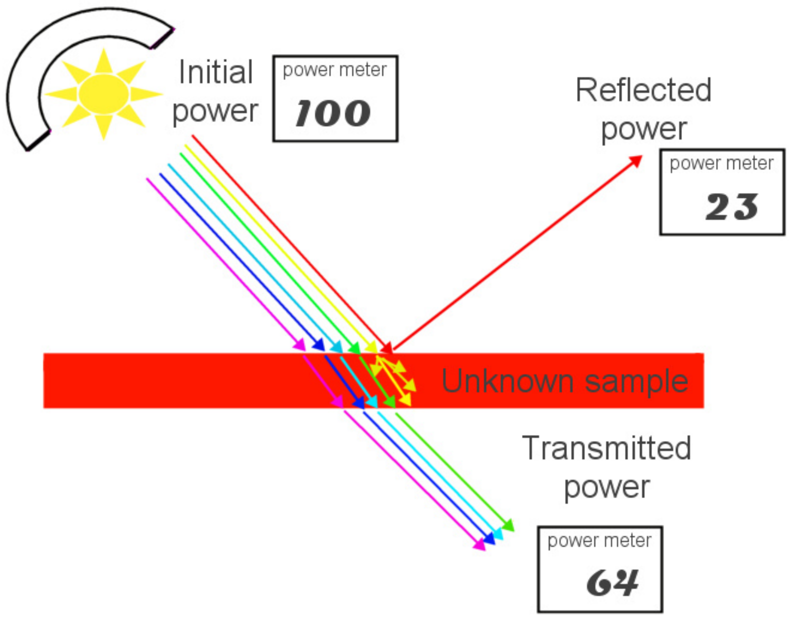


Fig. 5.9 Power measurement of optical parameters.

It is important to note that in Figure 5.9, the radiation measurement is spread over several wavelengths. In this case, we have to measure the individual wavelengths gradually and only finally add up the results. For measurement, we need a source of known power with the ability to set a specific wavelength.

Why is that so?

Let's say that a photodetector as a component does not have a linear response in the whole wavelength range. In principle, it recognizes only the formed electron-hole pairs. However, based on the wavelength of the incident light radiation, the value of the corresponding photocurrent differs. In the following Figure 5.10 we see the spectral characteristics of a silicon photodiode. This material has the greatest response in the area around 900 nm. If we move in the spectrum outside this value, the efficiency of the conversion will decrease and the response will decrease. For example, regarding the blue colour (around 420nm) the measured photocurrent will be only 10% in comparison with the central wavelength of 900nm. Therefore, if we shine a blue light on the detector, it is enough to multiply the measured value by ten.

We can use a photodetector to measure power only if we measure it by using a narrow and known part of the spectrum. Then we can perform an approximation to the corresponding values.

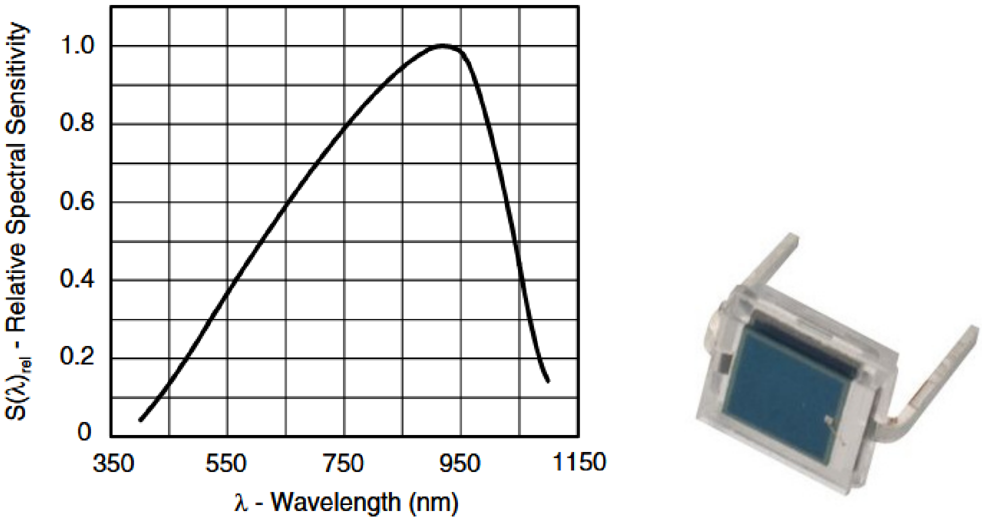


Fig. 5.10 Relative spectral response of the PIN photodiode BPW34.

Spectrum measurement

Spectrum measurement, or also spectral analysis, makes it possible to measure the power for several part of the spectrum simultaneously. The measuring device is called a spectrometer and, thanks to its construction, it first allows the incident light to be decomposed into individual wavelengths and then measured similarly to a photodiode. However, this time the measurement is performed by an array of photodetectors, each of which is set to the exact wavelength.

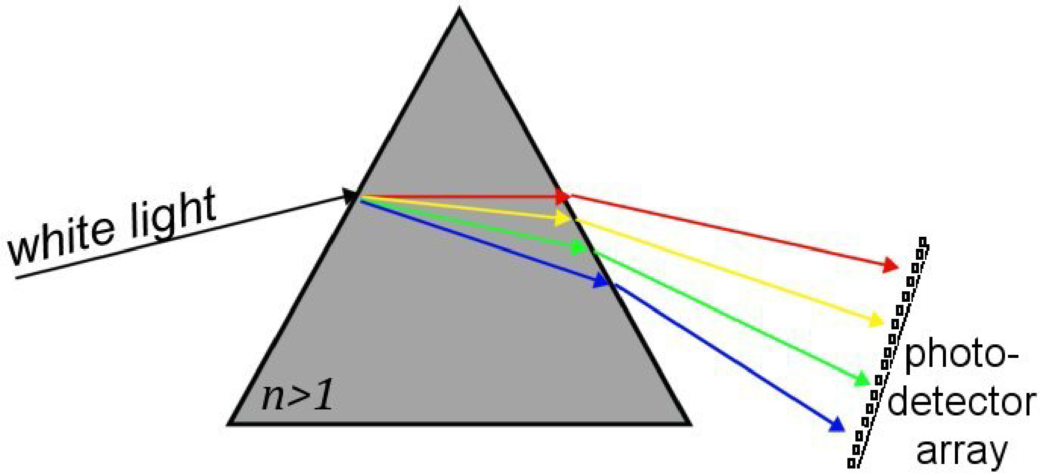


Fig. 5.11 The principle of the spectrometer.

And as we know the wavelength that hits the photodetectors, we can easily calculate the power across the entire measured spectrum. Another advantage is the possibility of graphical display of measured values and therefore an easier analysis, too.

The following Figure 5.12 shows the daylight spectrum measured by a spectrometer for illustration.

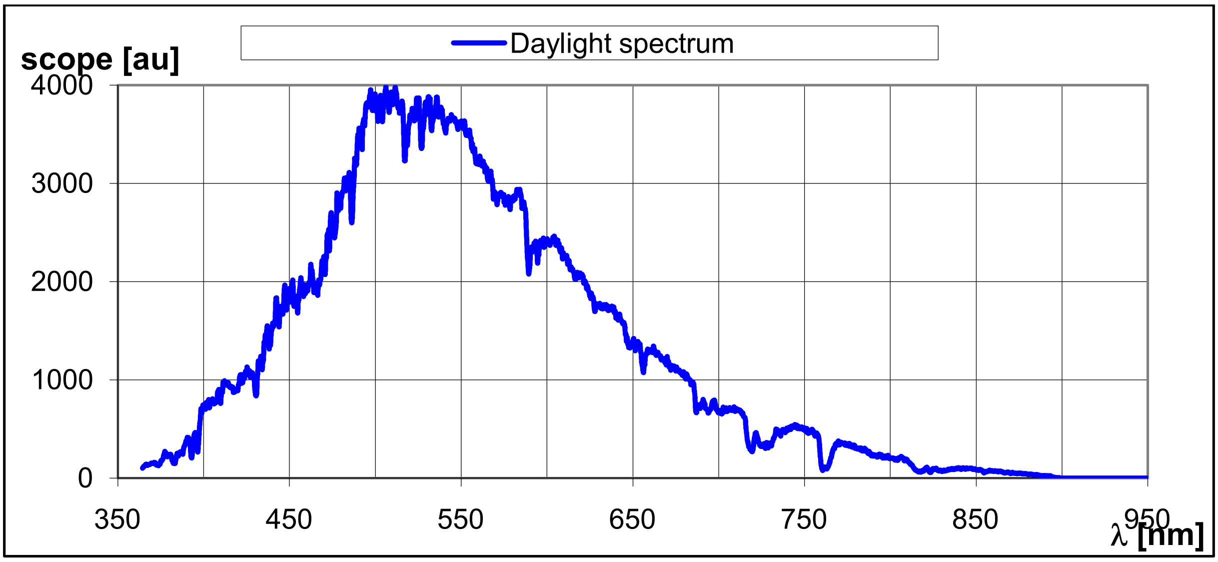


Fig. 5.12 Spectrum of daylight.

Spectrum measurement has many uses. For instance, for measurement within optical communications, the measurement of material properties and, last but not least, it also makes it possible to retrospectively analyse the composition of materials. The elements of the periodic table, or chemical compounds, have their own specific signature that allows them to be identified by analysis of the transmitted or reflected spectrum.

We can use the spectrum of the sunlight from the Figure X 12 above as an example. Individual wavelengths are not visually evenly represented. The Earth's atmosphere and its composition also take the credit. Nevertheless, despite the decrease in the intensity we can see an interesting thing when we take a closer look at the spectrum.

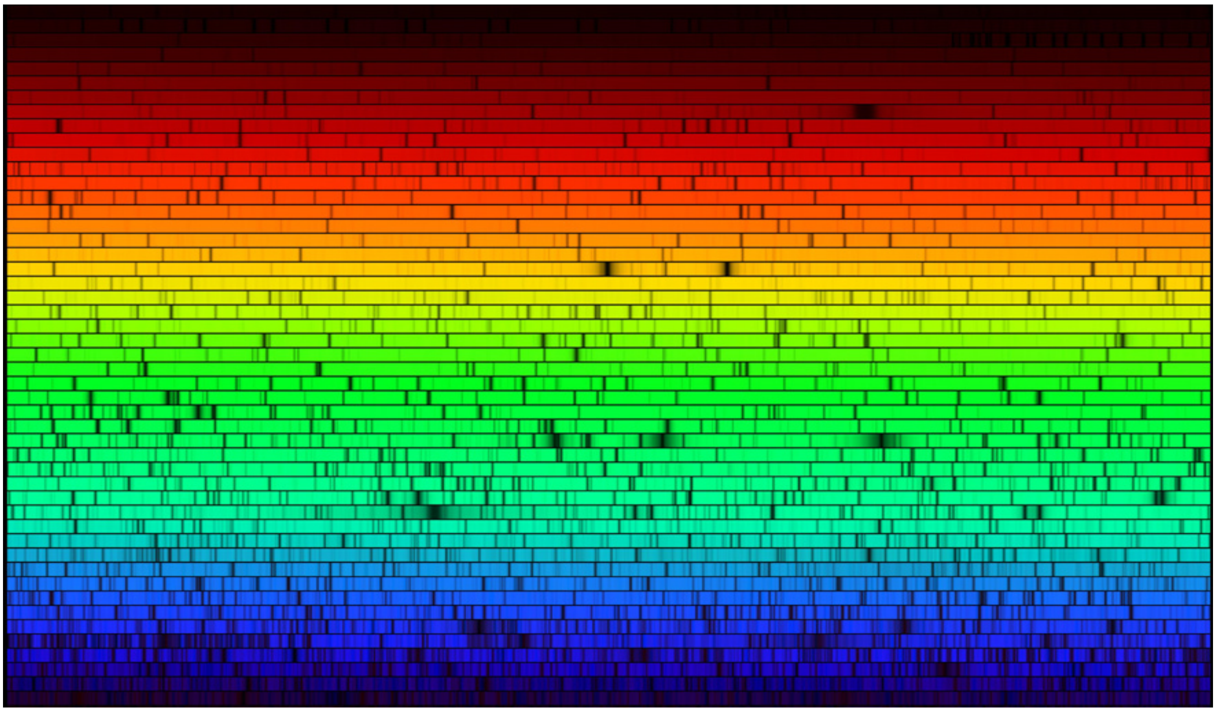


Fig. 5.13 Spectrum of the Sun of the *McMath-Pierce Solar Observatory*.

Some wavelengths are completely missing here. This is due to the content of different elements in its atmosphere that absorb these wavelengths before they are radiated into the environment. The composition for other suns in the Universe will always differ slightly. Therefore, it is a specific signature similar to the analysis of chemical solutions, etc.

Light Dispersion

A phenomenon in which white light decomposes into individual colour components (wavelengths).

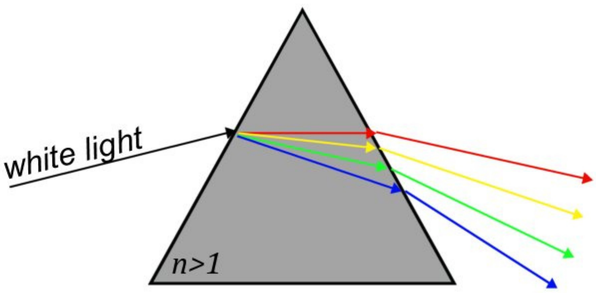


Fig. 5.14 Light dispersion.

Outside the vacuum (in the environments with a non-zero refractive index), the speed of the light in the given environment is dependent on the wavelength. The shorter wavelengths (higher frequency) spread more slowly than longer wavelengths (lower frequency). By assuming the equation for the refractive index:

$$n = \frac{c_0}{v}$$

Therefore, for slower (shorter) wavelengths the refractive index will be higher than for faster (longer) wavelengths.

Fig. 5.15 Dispersion dependence on the wavelength.

Due to the deceleration, the wavelength of the light also changes during the passage through the physical environment. The wavelength will be n-times smaller than in a vacuum.

$$\lambda = \lambda_0 / n$$

However, when you leave the environment, everything returns to normal.

Based on the rules for the reflection and refraction of light: the incident beam is divided into a reflected beam (complying with the law of reflection) and a refracted beam (complying with the law of refraction) at the interface of two optical media with different refractive indices n_1 and n_2.

$$n_1 \cdot \sin\alpha = n_2 \cdot \sin\beta$$

We can see that a change in the refractive index results in a change in the refractive angle of the beam.

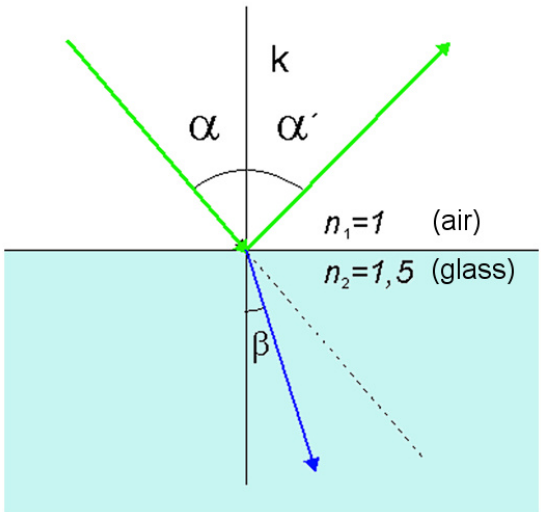


Fig. 5.16 Beam refraction.

Normal dispersion

In the described behaviour when the refractive index decreases with an increasing wavelength, the phenomenon is called normal dispersion.

We welcome this quality in case of spectroscopy. Nevertheless, there are many cases when this phenomenon is undesirable. In optical communications, this phenomenon leads to a reduction in the width of the transmitted frequency band because the light does not only spread to individual wavelengths but also spreads at different speed. Therefore, there is an effort to use monochromatic light on long-distance routes.

The monochromatic light ideally consists of only one wavelength, so it spreads in the environment in a precisely defined way. Nevertheless, we are not able to achieve this quality. According to the used radiation source we always have a finite spectral line width.

In the technical practice, this width is referred to as FWHM (Full Width Half Maximum) and it indicates the width of the spectral line at half of the maximum power.

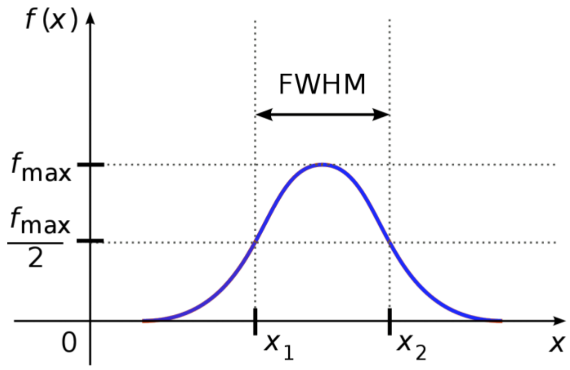


Fig. 5.17 FWHM.

This value is moving in the range of the tens of nm for the LED diodes. Therefore, they represent an advantage over white light and the light dispersion was limited. The following figure shows an example of the spectral characteristics for different LEDs.

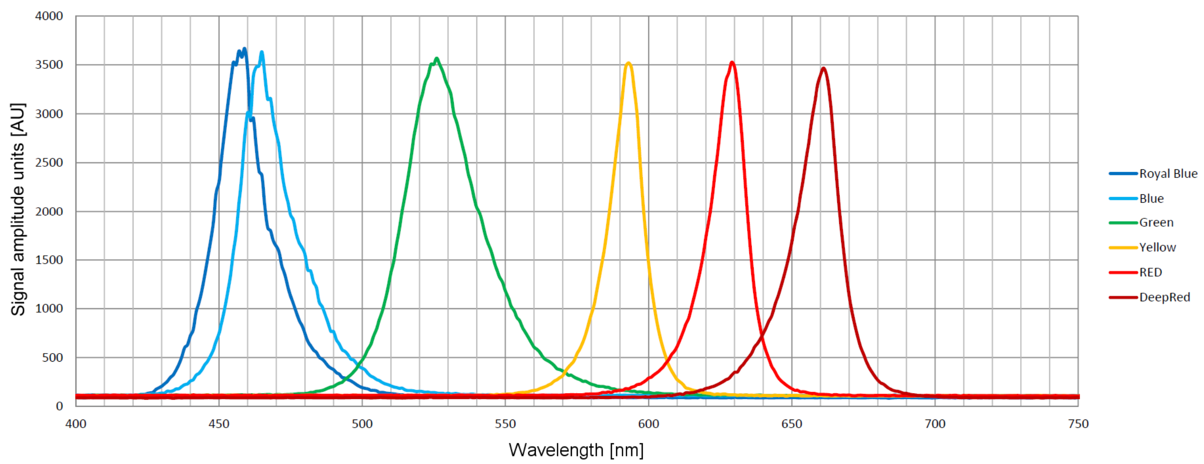


Fig. 5.18 Spectral characteristics of LED diodes.

We can also generate light radiation with the FWHM values around nm unites for more demanding applications. Nevertheless, we need a source of radiation called laser for this purpose. The laser source, by its very nature creates a beam with a very narrow spectrum with high power.

Abnormal dispersion

This dispersion is the opposite of normal dispersion. These are the materials which refractive index increases with the increasing wavelength. Therefore, it is assumed that a suitable combination of materials can compensate for the dispersion behaviour of the optical system.

As a matter of interest: from a mathematical point of view, the anomalous dispersion in the material allows to reach a state where the derivative of the refractive index can be in the absolute value n<1. This would mean that the light spreads in it faster than the speed of light. For example, it is the environment with free electrons such as plasma or the ionosphere. However, even the glass moves toward the value n=0.16 in certain 'remote' parts of the spectrum, which would mean that light spreads here more than 5 times faster than in the case of a vacuum. Nevertheless, it is only the result of the constructive interference of waves and more likely a geometric phenomenon. There is practically no energy transfer and it is only an expression of the phase velocity. The transmission of the energy itself still takes place here with the hitherto known limitation of the speed of light.

In this respect, the issue of light propagation in the material environment is a very complicated issue and is still the subject of intensive research.

The fact remains that, despite the validity of the theory of relativity (and the speed of light), some materials may behave geometrically as if they had a refractive index n<1 (theoretically n<0). And as the following picture shows, in case of glass, we do not encounter this 'anomaly' in everyday life of communications.

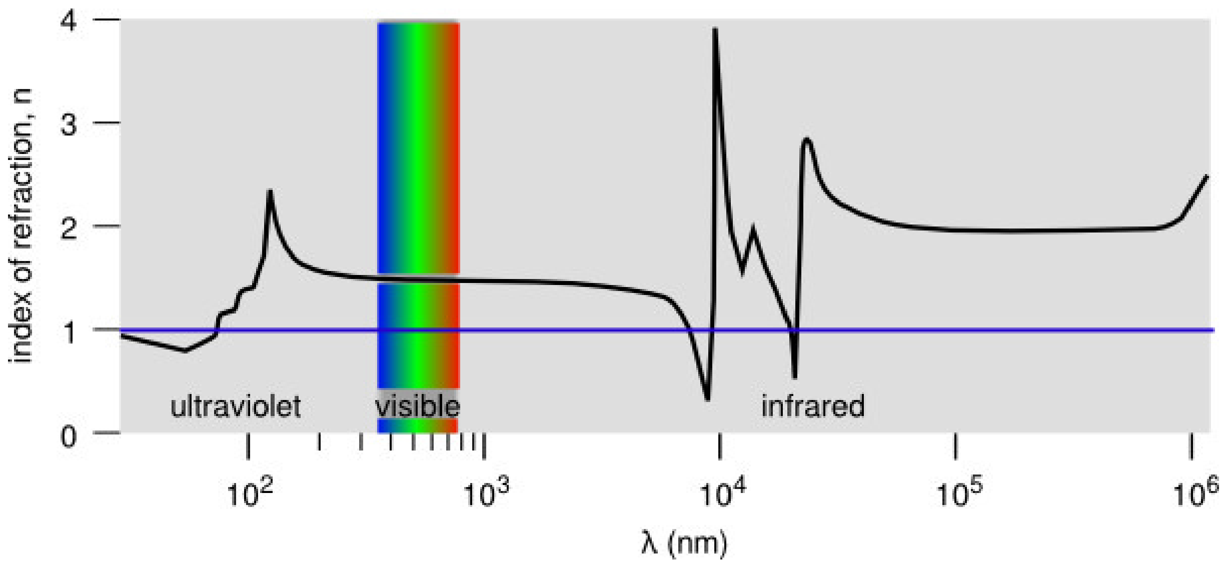


Fig. 5.19 Course of the refractive index of glass. [*6*]

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. What can happen to light rays incident on the unknown material?
2. What is the formula for photon energy?
3. How does acrylic glass behave in the near IR part of the spectrum?
4. What is FWHM?
5. What is the dispersion of light?
6. Describe the refractive index calculation.
7. Which wavelength will propagate faster in a vacuum, 400nm, or 700nm?
8. What is the difference between normal and abnormal dispersion?

6. Basic geometric optics

When the light passes through objects which dimensions are **much larger** than its wavelength of light is, the light can be described by means of linear rays. These rays comply with geometric rules and that is why we use the term **geometric** (ray) **optics**.[7], [8]

The same rules are applied to the reflection of light from obstacles.

Note: Of course, much more is happening from a physical point of view as the behaviour of such simplified rays has an impact on a number of several factors. However, we will focus on this a little later.

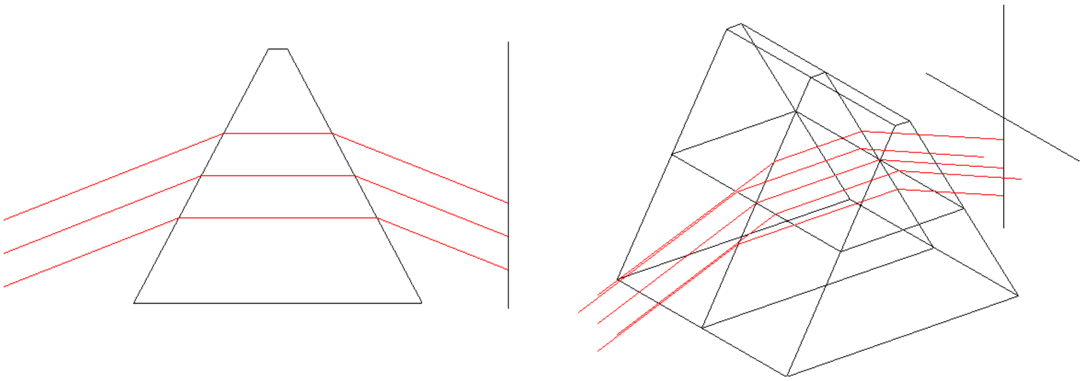


Fig. 6.1 Example of refraction of light rays.

Basic concepts

First of all, let's look at some of the basic concepts that we will need across the chapter and to which we will return repeatedly.

Light is an electromagnetic wave with a specific range of frequencies, resp. wavelengths. More precisely, it is the visible part of the electromagnetic spectrum in the range of approximately 380nm - 780nm.

(Note: It is possible to view the light as a particle - photon - from a physical point of view).

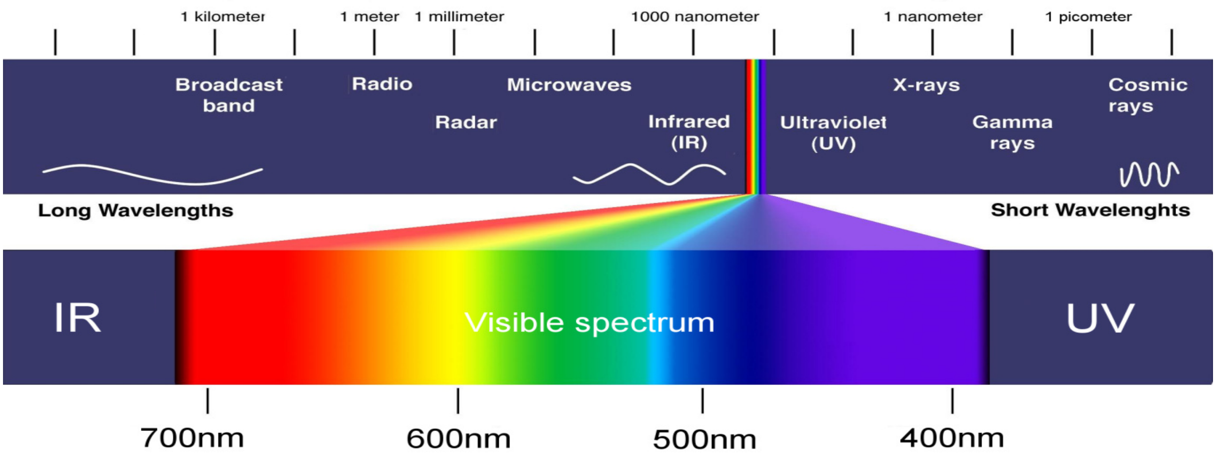


Fig. 6.2 The spectrum of light.

The wavelength is calculated as the ratio of the speed of light and the corresponding frequency of EM radiation.

$$\lambda = \frac{c}{f}$$

The **speed of light (*c*)** is the speed at which a photon (light particle) moves in a vacuum. It is 299 792 458 m/s based on available information. However, for the reasons of simplification we can use the value for indicative calculations:

$$c = 3 \cdot 10^8 \quad m \cdot s^{-1}$$

Regarding the air, this speed can be considered the same because the influence of the atmosphere of the Earth is minimal from the point of view of the speed.

Refractive index – the optical environment can be characterized by a quantity *n* that is called the refractive index. This is the ratio of the speed of light in the vacuum c_0 and the speed of light in a given environment *v*.

$$n = \frac{c_0}{v}$$

Therefore, it always applies that $n \geq 1$

Chart: examples of refractive index for different environments.

Substance	Refractive index n
Vacuum	1
Earth Atmosphere	1,00026
Ice	1,31
Water	1,33
Ethanol	1,36
Glass	1,55 (1,5-1,9)
Salt	1,52
Sapphire	1,77
Diamond	2,42

Attention: Refractive index depends on the wavelength.

Light reflection

A beam of light incident on the reflecting surface under certain angle is further reflected at the same angle.

Therefore, it is true that: $\alpha = \alpha'$

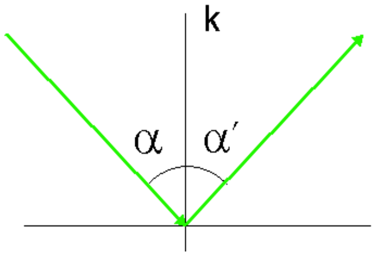


Fig. 6.3 Beam reflection.

A mirror is a typical example of such a reflective surface.

Light refraction

Light refraction – at the interface of two optical environments with different refractive indices n_1 and n_2 the incident beam splits into a reflected beam (complying with the law of reflection) and a refracted beam (complying with the law of refraction).

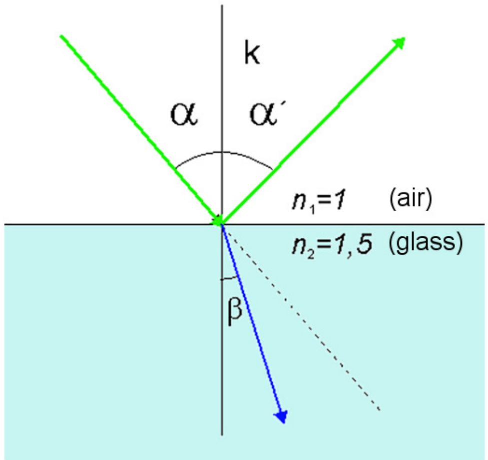


Fig. 6.4 Ray refraction.

Law of refraction – the behaviour of light upon impact on the interface of two optical environments can be expressed by the equation:

$$n_1 \cdot \sin\alpha = n_2 \cdot \sin\beta$$

This implies that the refractive angle will follow the refractive indices of individual environments. If light comes from an optically thinner environment n_1 (e.g., the air) to an optically denser environment n_2 (glass), the value $\sin\beta$ will be smaller to meet the condition in the equation. Therefore, the beam will refract to the perpendicular k.

Absolute reflection – in the opposite case, when the light passes by from an optically denser environment to an optically thinner one, an important phenomenon occurs.

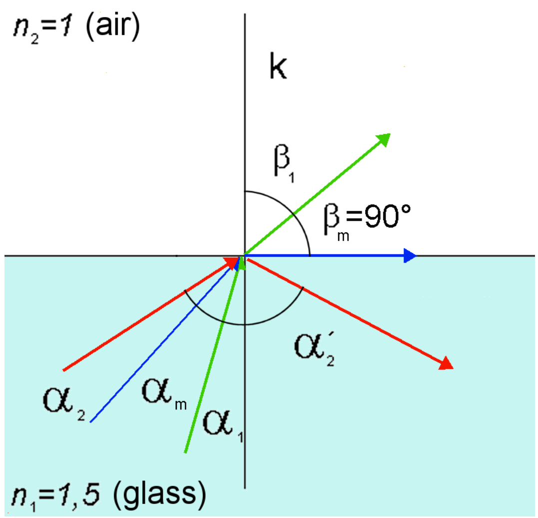


Fig. 6.5 Absolute reflection.

Regarding the angle α_1, the beam behaves in the opposite way in respect to the previous example. Due to the rotation of the environment, when the light penetrates from a denser environment to the thinner one, the light refracts in the direction from the perpendicular. As a result, at reaching the state of the limit angle α_m, when the angle β reaches the value 90° and the refracted beam begins to spread along the interface. From this point on, as the angle increases, the light cannot break into a thinner environment and all the energy of the beam remains inside the denser environment. When the angle is further increased, only the law of reflection will apply. The advantage of this phenomenon is 100 % of efficiency. All light that meets the condition of absolute reflection continues in an optically denser environment. This phenomenon is mainly used in optical fibres as the light can spread over huge distances with minimal losses.

Light dispersion – a phenomenon in which white light splits into individual colour components.

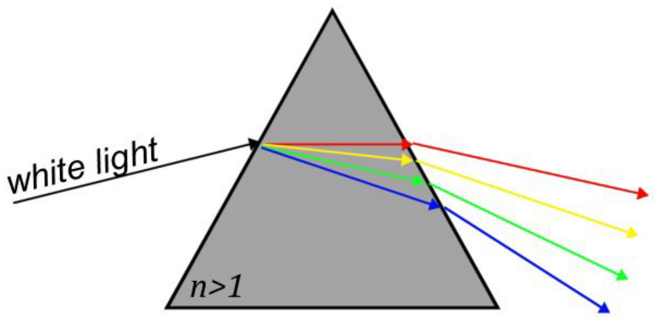


Fig. 6.6 Light dispersion.

Outside the vacuum (in the environments of a non-zero refractive index), the speed of light in a given environment depends on the wavelength. Shorter wavelengths (higher frequencies) spread more slowly than longer wavelengths (lower frequencies). If it is based on the equation for the refractive index:

$$n = \frac{c_0}{v}$$

Therefore, for slower (shorter) wavelengths, the refractive index will be higher than for faster (longer) wavelengths.

Thus, at the refraction of the light, the refraction angles differ for the individual wavelengths.

This phenomenon is used, for instance, in spectrometers, when we can easily decompose the measured light into individual components. After decomposition, the rays will fall on a system of photodetectors, each of which corresponds to a specific wavelength and provides data on the incident power at the same time.

On the other hand, there is the disadvantage of this phenomenon because it also occurs in optical fibres. For this reason, the narrow-spectrum sources are used for optical communication in order to suppress this phenomenon as much as possible. When using the white light, it would be from a longer route.

Mirrors

In principle, it is any reflective surface that can reflect the incident light. It is usually a thin metal layer.

Plane mirror

The image created by a plane mirror is always apparently upright, as large as the object, and symmetrical with the object according to the plane of the mirror (sideways inverted).

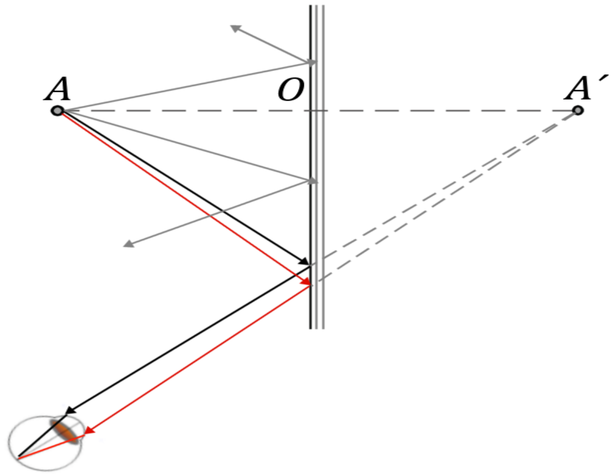


Fig. 6.7 Mirror display.

Spherical mirror

The reflecting surface is applied to a spherical canopy.

From the point of view of the construction, they are divided based on which part of the canopy the reflective layer is applied to. If the reflective layer is located inside the spherical space, it is a hollow (concave) mirror. If the reflecting surface is on the top of a spherical space, it is a convex mirror.

The examples are in the following picture where:

o = optical axis

V = top of the mirror

C = centre of curvature

r = radius of curvature

F = focus

f = focal length

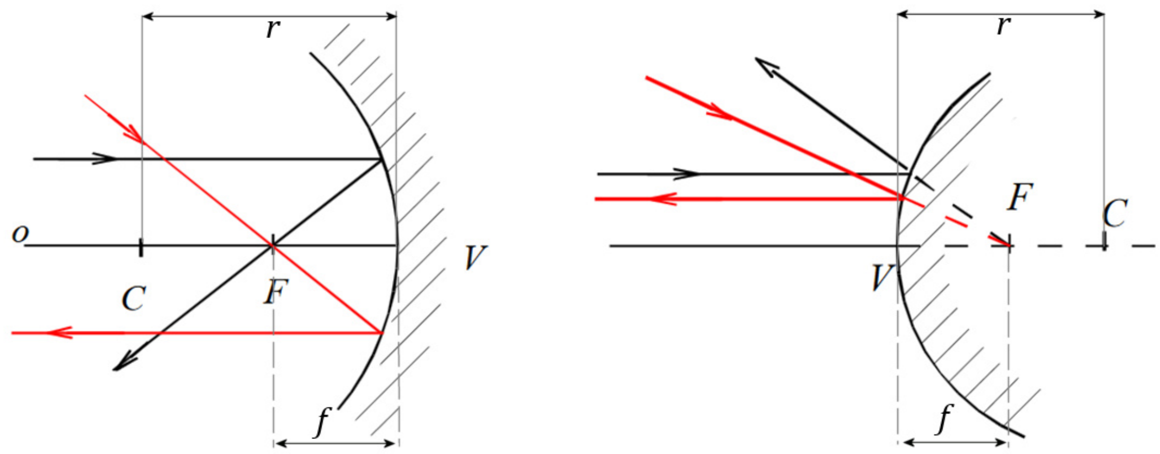


Fig. 6.8 Display by a spherical mirror.

Hollow – concave mirror	Convex mirror
F – real focus	F – apparent focus
Sign convention	
f>0	f<0

We use 3 significant beams for optical imaging.

1. A beam passing through the centre of curvature of the mirror (the beam is reflected back to the source).
2. A beam incident parallel to the optical axis is reflected to the point F (focus of the mirror).
3. A beam passing through the focus F, reflected parallel to the optical axis.

The point F is the so-called focus of the mirror. The distance between the focus and the top of the mirror is called the focal length (f).

The sign convention determines a positive or negative value for the description of individual quantities. Compliance with the convention enables to define optical behaviour and will be particularly important in the design of more complex systems. The values from the top to the left are defined as positive and values from the top to the right are defined as negative.

Mirror materials

A certain part of the energy is always lost at reflection from the mirror surface. During the reflection, the part of the light is, for example, converted into heat or scattered due to microscopic inaccuracies. The reflection efficiency will depend on the material of the reflecting surface and the wavelength.

The limiting factor in optical systems will be the number of reflections and the material. If we reflect the light with an efficiency of 90 %, then after the tenth reflection it will have only 30 % of the original power. Therefore, let us look briefly at the most common materials and their parameters.

Aluminium:

Economical solution for less demanding applications.

Smaller chance of tarnishing than silver in high humidity environment.

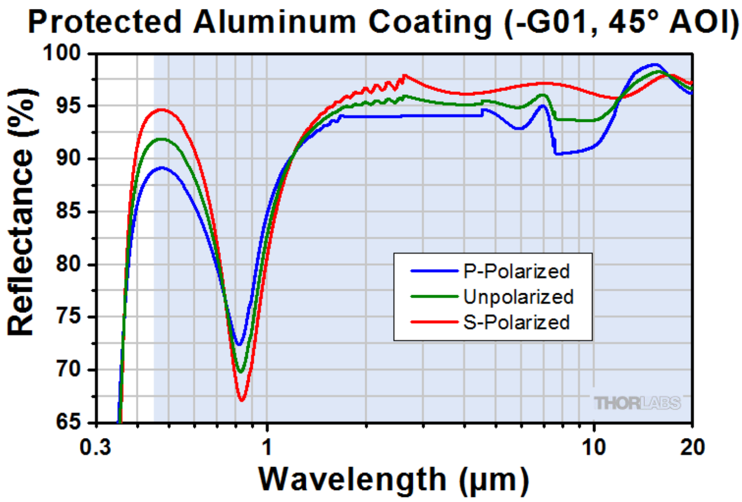


Fig. 6.9 Reflectance parameters of aluminium. (Angle of Incidence 45°)

Silver:

Highest reflectance of all metal coated mirrors.

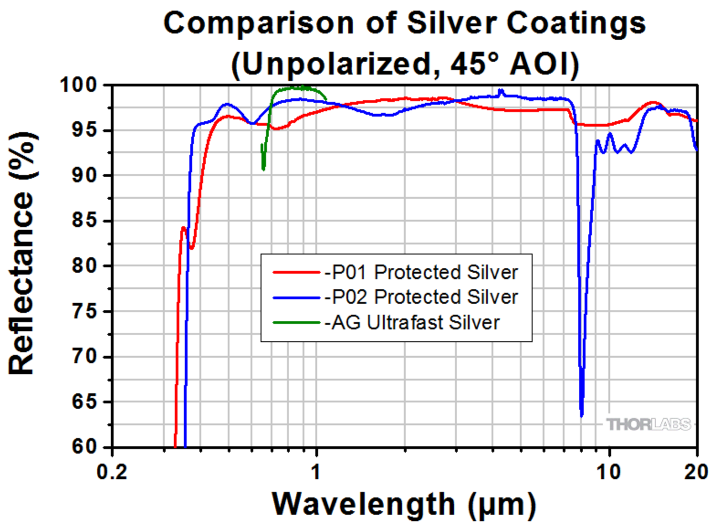


Fig. 6.10 Reflectance parameters of silver.

Gold:

The most efficient reflective coating over the IR (800nm – 20um). However, not good within visible part of the spectrum.

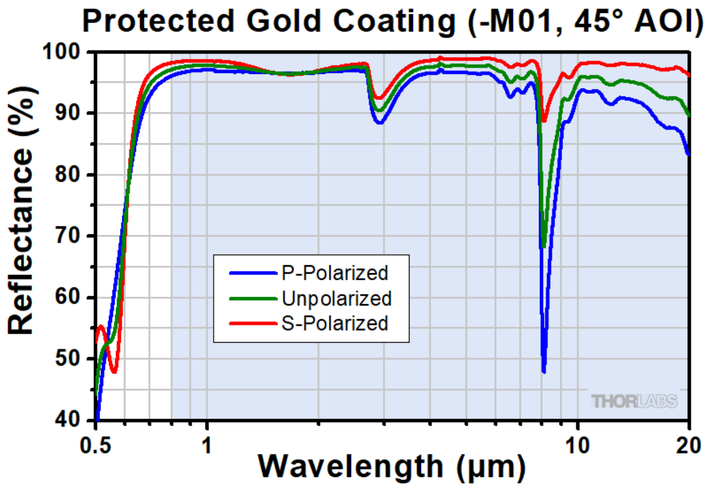


Fig. 6.11 Reflectance parameters of gold.

Dielectric mirrors:

Fused silica dielectric mirrors offer excellent reflectance in specific spectral range (it depends on specific coating).

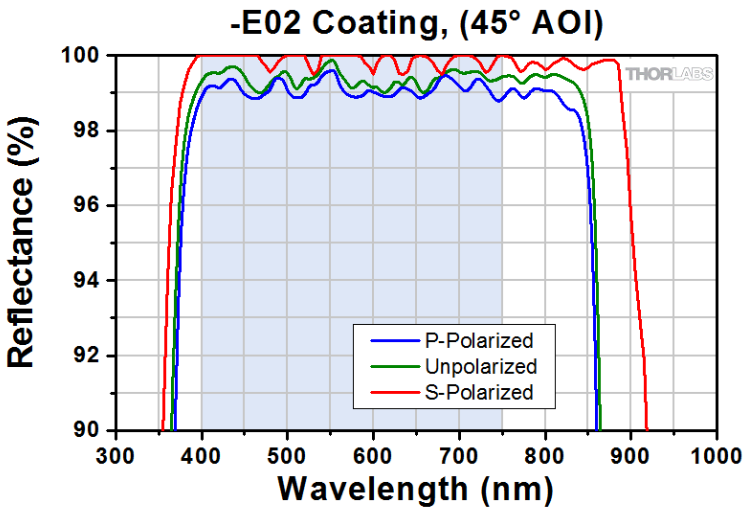


Fig. 6.12 Reflectance parameters of fused silica – coating for 400-750nm.

It is apparent from the values that, in comparison with absolute reflection, the mirrors are less effective. Therefore, we frequently lose valuable performance by repeated reflection of light. Nevertheless, this condition is impossible because we would have to create an exact interface between air and vacuum, or we would have to use a material with s n<1 as a mirror. However, this would mean that light will spread in it faster than the speed of light in a vacuum. And the existence of this material is not currently known.

The required function could be fulfilled by an optical prism. Nevertheless, the light would have to pass repeatedly between the air and the optical environment here, which brings again considerable losses…

Lenses

The lens is a clear optical environment that is limited by two optically smooth areas. Unlike mirrors, there is the display by refraction. The bunch of optical rays passes through the environment of the lens and, depending on its shape and refractive index, the optical rays change their direction. [7]

We can divide lenses according to the shape into two main groups:

- Convergent lenses (they change parallel beams to convergent)

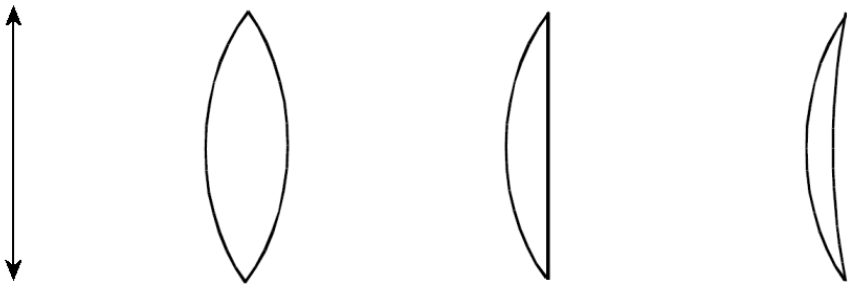


Fig. 6.13 Convergent lenses.

Refractions of significant beams are indicated for passage through the optical axis when the beam passes through the 'centre' of the lens (*O*).

Furthermore, for the passage of parallel beams when the beams meet in the image focus (*F´*).

And, finally, for passing through the object focus (*F*). Then the output beams will spread parallel.

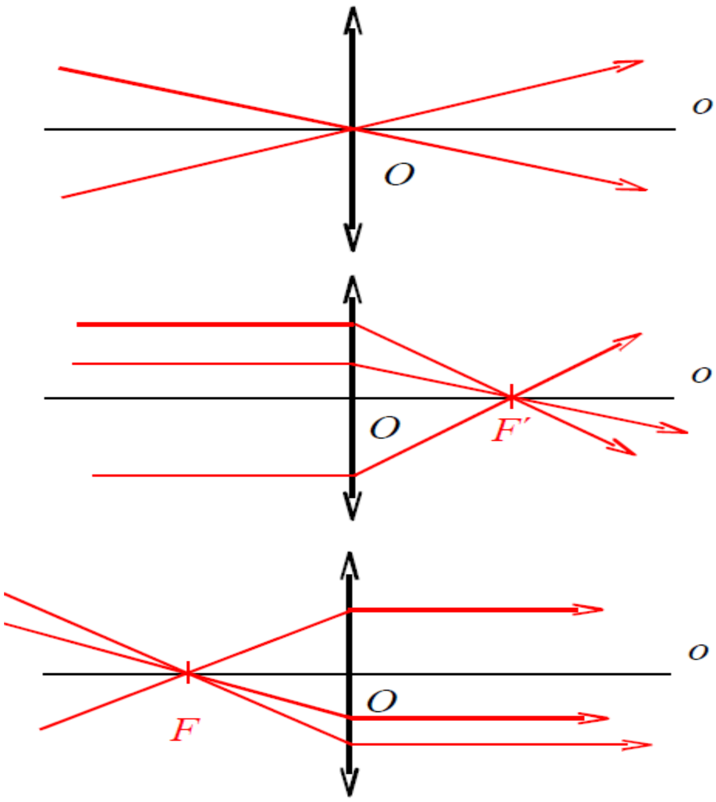


Fig. 6.14 Convergent lenses – refractions of significant beams.

- Concave lenses (they change parallel beams to divergent)

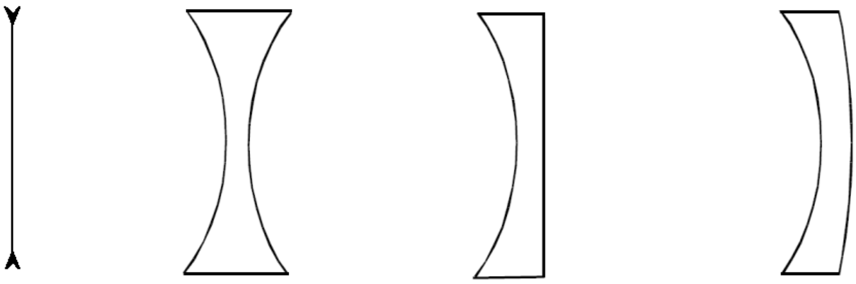


Fig. 6.15 Concave lenses.

The refractions of significant beams for concave lenses show the routes for individual beams as they pass through the centre of the lens. The beams pass through without affecting the route.

In case the beams come horizontally into the concave lens the resulting beams are divergent. Therefore, in this case, the image focus (*F´*) is in front of the lens. It is an unreal image focus.

Similarly, there is a case when the beams enter the concave lens toward an unreal object focus as the beams are parallel at the exit of the lens and never meet.

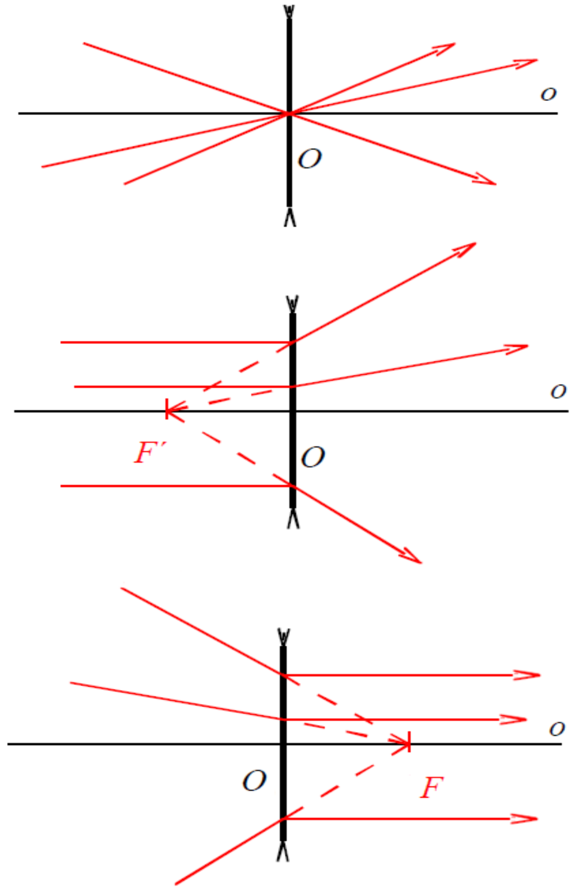


Fig. 6.16 Convergent lenses – refractions of significant beams.

Figures X.14 and X.16 show thin-lens imaging. A thin lens is a limit case when the lens thickness is very small (mathematically close to 0).

The convexity of a thin lens can be described as:

$$\rho = \left(\frac{1}{r_1} + \frac{1}{r_2} \right)$$

The focal length of a thin lens can be simply described as:

$$\frac{1}{f} = \left(\frac{n_2}{n_1} - 1 \right) \left(\frac{1}{r_1} + \frac{1}{r_2} \right)$$

(n_1=refractive index of the environment, n_2=refractive index of lens material, r_1, r_2= radii of curvature of optical surfaces).

Thick lens

There is a case when the optical environment with the refractive index n_1 is bounded by two refractive surfaces with radii r_1 and r_2. Nevertheless, their peaks are at a distance *d* in this case (this is a realistically manufacturable variant of the lens).

It is necessary to know several other variables for the description of the lenses with real parameters because their shape will also play the role in the real environment.

- **Lenses – important concepts**
    - Centres of optical surfaces (C_1 and C_2)
    - Radii of curvature o. of surfaces (r_1 and r_2)
    - Optical axis (connectors C_1 and C_2)
    - Optical centre of the lens (O)
    - Peaks (V_1 and V_2)
    - Subject area (beam entrance area)
    - Image area (beam output area)
    - The refractive index ratio can be simplified in this case of optical systems because the environment is formed by the air and it is almost as a vacuum (n=1).

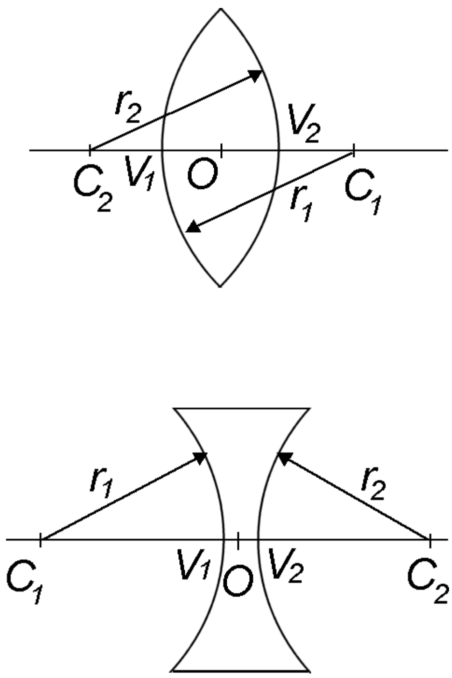


Fig. 6.17 Physical description of lenses.

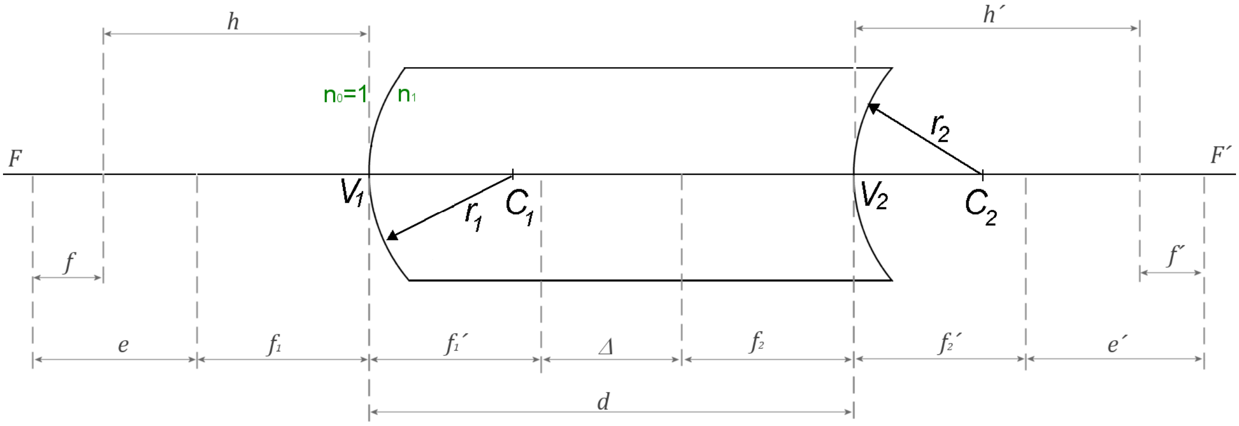


Fig. 6.18 Thick lens.

Determination of focal lengths:

$$f_1 = \frac{n_0 \cdot r_1}{n_1 - n_0} = \frac{r_1}{\frac{n_1}{n_0} - 1} = \frac{r_1}{n-1} \qquad f_1' = \frac{n_1 \cdot r_1}{n_1 - n_0} = \frac{\frac{n_1}{n_0} r_1}{\frac{n_1}{n_0} - 1} = \frac{n \cdot r_1}{n-1}$$

$$f_2 = \frac{n_1 \cdot r_2}{n_0 - n_1} = \frac{\frac{n_1}{n_0} r_2}{1 - \frac{n_1}{n_0}} = \frac{n \cdot r_2}{1-n} \qquad f_2' = \frac{n_0 \cdot r_2}{n_0 - n_1} = \frac{r_2}{1 - \frac{n_1}{n_0}} = \frac{r_2}{1-n}$$

Optical interval Δ:

$$\Delta = d - f_1' - f_2 = d - \frac{n \cdot r_1}{n-1} - \frac{n \cdot r_2}{1-n}$$

Focal lengths of the whole system:

$$f = -\frac{f_1 \cdot f_2}{\Delta} = -\frac{\frac{r_1}{n-1} \cdot \frac{n \cdot r_2}{1-n}}{d - \frac{n \cdot r_1}{n-1} - \frac{n \cdot r_2}{1-n}} = \frac{n}{n-1} \cdot \frac{r_1 \cdot r_2}{d(n-1) + n(r_2 - r_1)}$$

$$f' = -\frac{f_1' \cdot f_2'}{\Delta} = -\frac{\frac{n \cdot r_1}{n-1} \cdot \frac{r_2}{1-n}}{d - \frac{n \cdot r_1}{n-1} - \frac{n \cdot r_2}{1-n}} = \frac{n}{n-1} \cdot \frac{r_1 \cdot r_2}{d(n-1) + n(r_2 - r_1)} = f$$

The focal lengths on both sides of the lens are the same because the optical environment is the same at both ends.

It applies for the distances of the main planes from the peaks of the spherical surfaces of the lens (h):

$$h = f_1 + e - f = f_1 + \frac{f_1 \cdot f_1'}{\Delta} - \left(-\frac{f_1 \cdot f_2}{\Delta}\right) = \frac{f_1}{\Delta}\left(\Delta + f_1' + f_2\right) = \frac{f_1 \cdot d}{\Delta} =$$

$$\frac{f_1 \cdot d}{-\frac{f_1 \cdot f_2}{f}} = -\frac{f \cdot d}{f_2} = -f \cdot d \frac{1-n}{n \cdot r_2} = \frac{n-1}{n} \cdot \frac{f \cdot d}{r_2}$$

$$h' = f_2' + e' - f' = \ldots = \frac{1-n}{n} \cdot \frac{f \cdot d}{r_1}$$

In case of the figure the sign convetion$r_1>0$ and $r_2<0$ is after adjustment:

$$h = \frac{1-n}{n} \cdot \frac{f \cdot d}{r_2} \qquad h' = \frac{1-n}{n} \cdot \frac{f \cdot d}{r_1}$$

$$f = f' = \frac{n}{n-1} \cdot \frac{r_1 \cdot r_2}{n(r_1 + r_2) - d(n-1)}$$

Subject imaging by the lenses

If we have the lens of known dimensions, we can calculate the focus. If we know the focus, we can use significant beams to figure out what the image of the object will look like by means of a lens.

For example, let's have an object with a height 'y' in three different distances 'a' from the centre of the lens:

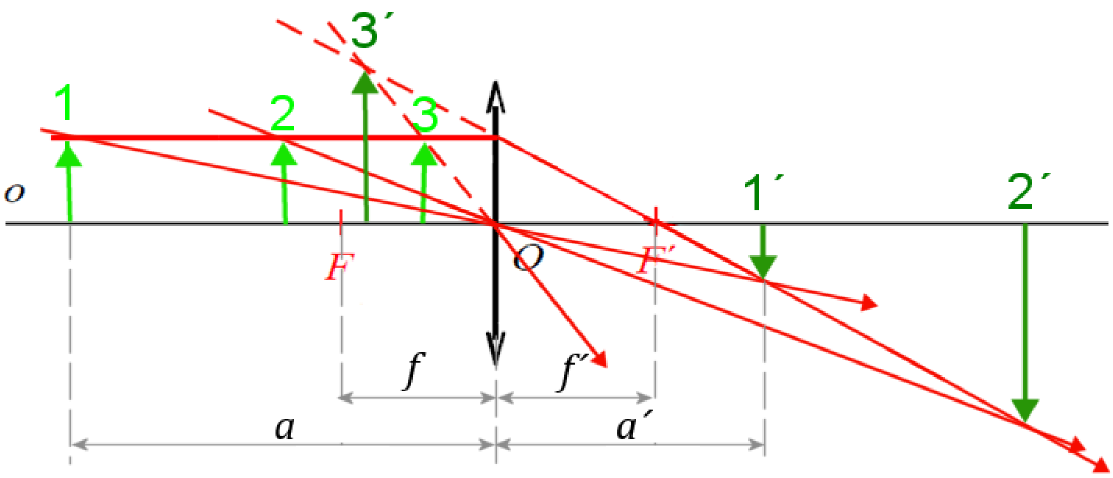


Fig. 6.19 Display by means of a convergent lens.

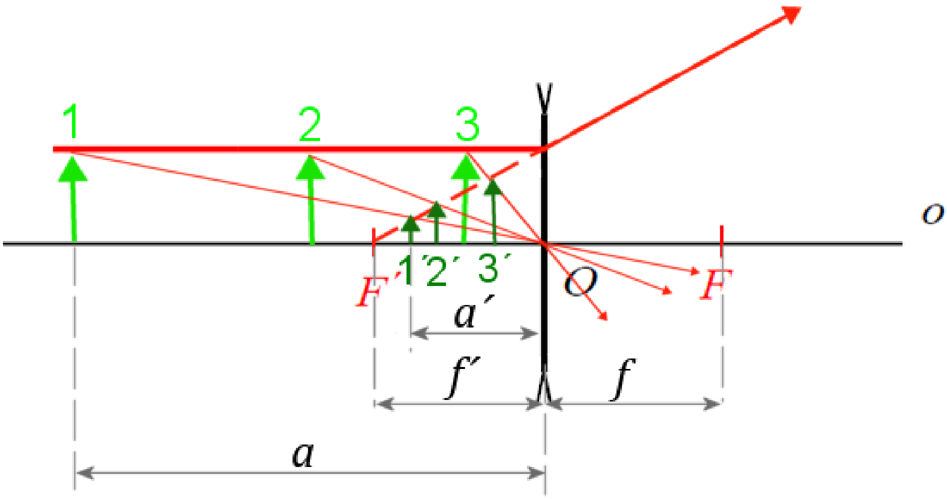


Fig. 6.20 Display by means of a concave lens.

It applies to transverse magnification:

$$Z = \frac{y'}{y} = -\frac{a'}{a} - \frac{a'-f}{f} = -\frac{f}{a-f}$$

The optical power of the lens can be described as:

$$\varphi = \frac{1}{f} \quad \left[m^{-1} = D(dioptric\ \ unit)\right]$$

Imaging equation:

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{a'}$$

where *a* = object distance, *a´*= image distance, *f* = focus distance.

Sign convention:

CONVERGENT LENS	CONCAVE LENS
r_1, r_2 >0	r_1, r_2 <0
φ>0	φ<0
a>0 (in front of the lens), *a*<0 (behind the lens)	
a´>0 (in front of the lens), *a´*<0 (behind the lens)	

Example: Consider an object 10 mm high placed in front of a thin convergent lens with a focal length *f*= 0.1 m. Let's look at images at 3 different distances: a_1=0.3m, a_2=0.15m, a_3=0.05m.

Solution:

$a_1´$=0,15m, Z_1=-0,5. The image is real, inverted and reduced.

$a_2´$=0,3m, Z_2=-2. The image is real, inverted and enlarged.

$a_1´$=-0,1m, Z_1=2. The image is unreal, upright and reduced.

Lens systems

The same system applies here as with a thick lens. Just as we assembled spherical surfaces, we now assemble entire lenses into a centred optical system.

The centred optical system represents two or more optical elements aligned on a common optical axis. As a result, the system equations have the same shape as for a simple spherical surface.

If we create an optical lens system with focal lengths f_1 and f_2 and an optical interval Δ, then the resulting focal length of the system will be:

$$f = f' = -\frac{f_1 \cdot f_2}{\Delta}$$

It is apparent from the equation that only by changing Δ we can create a system with positive ($f = f' > 0$) or negative ($f = f' < 0$) focal length.

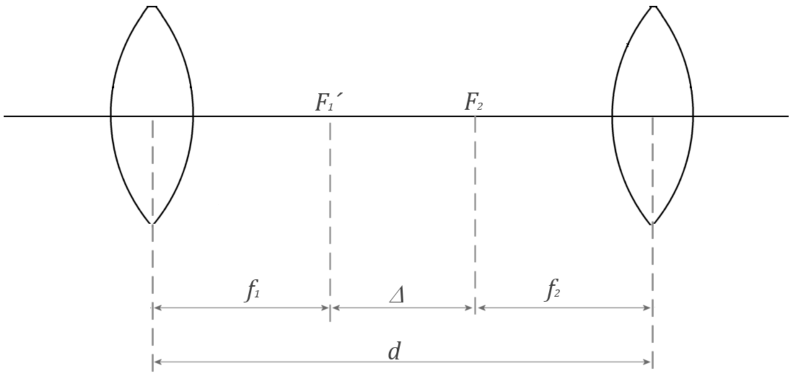


Fig. 6.21 System of two lenses.

For simplicity, let's use the example from the figure 6.21. Let's consider both the lenses to be thin – the main nodal points are in the centre of the lenses.

By substituting we get:

$$f = f' = -\frac{f_1 \cdot f_2}{\Delta} = -\frac{f_1 \cdot f_2}{d - f_1 - f_2}$$

The optical power of a system of two lenses is then calculated as:

$$D = \frac{1}{f} = -\frac{d - f_1 - f_2}{f_1 \cdot f_2} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 \cdot f_2} = D_1 + D_2 - d \cdot D_1 \cdot D_2$$

If the lenses were directly next to each other (d=0) then the relationship is simplified to:

$$D = D_1 + D_2$$

Therefore, the optical powers sum up.

A very special case of two lenses occurs if their interval is Δ=0. This is the case with the telescopic system (telescope).

Telescopes

The telescope is an optical system for observing infinitely distant objects. An object that is at a significantly greater distance from the objective than its focal length can be considered an infinitely distant object.

Refractors

The objective of the refractor is formed by the lens. Historically, these are the first types of telescopes. In principle, these telescopes can be divided into 2 basic types:

The **Galilei telescope** was invented in 1609 and consists of two lenses. The objective – the first lens is formed by the convergent lens that captures light and concentrates it into the focus. The second lens – the eyepiece is the concave lens that again straightens up the rays to be parallel before reaching the focus. This construction produces an upright and undeformed image. Nevertheless, the image suffers greatly from chromatic aberration. This means that individual colour components of the image do not meet in the same focus and the image is distorted by the deceptive colouring of the edges of the observed objects. [9]

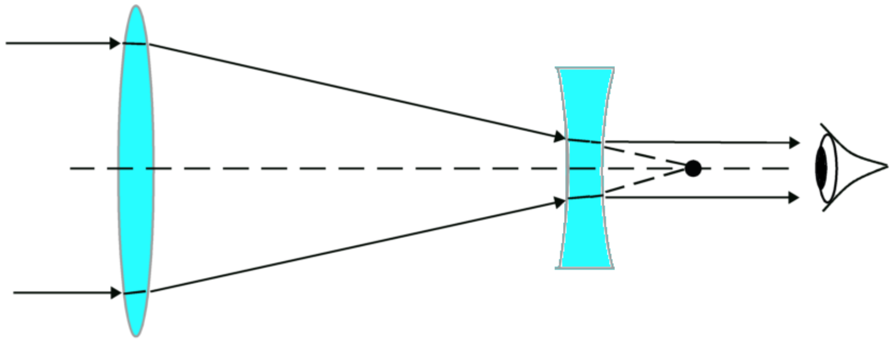


Fig. 6.22 Galileo's telescope.

Kepler's telescope was invented in 1611. The objective and eyepiece were formed by convergent lenses and foci overlapped inside the telescope body. Nevertheless, the result was an inverted image, making it more appropriate for astronomical observations. Even this construction suffers from chromatic aberration. Nevertheless, it could be reduced by increasing the focal length. (At that time, focal lengths in the order of meters were used and this was reflected in the physical length of these telescopes).

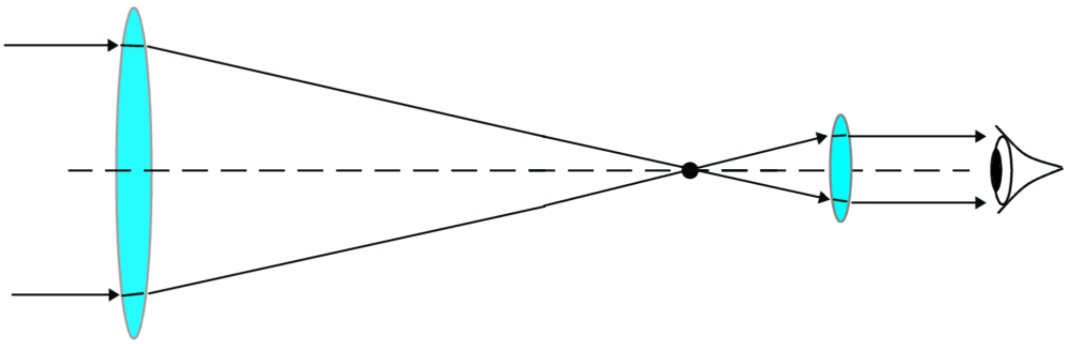


Fig. 6.23 Kepler's telescope.

Chromatic aberration was considered to be irremovable for 200 years. However, an 'achromatic refractor' appeared in the 18th century. This solution used an objective composed of two lenses of different glasses (different refractive indices). As a result, their routes were corrected as the beams passed and beams of different wavelengths met in a common focus. The telescopes could start to shorten.

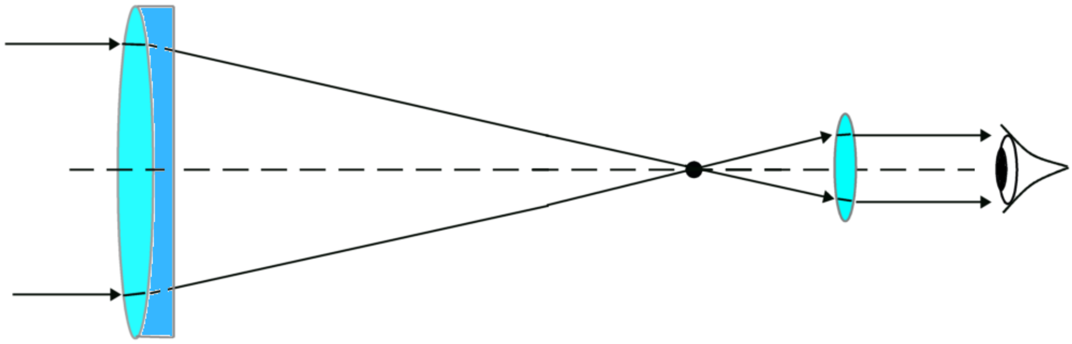


Fig. 6.24 Achromatic refractor.

Reflectors

This is a mirror telescope, in which the objective consists only of mirrors. Its construction concentrates the incident beams into one point (focus), as well as the convergent lens. The main parabolic mirror concentrates its image into the smaller elliptic mirror and from there back through the hole in the centre of the parabolic mirror to the observer. This system is more demanding on accuracy and maintenance, but, on the other hand, it does not suffer from distortion of image in terms of size and colour, as in case of the refractors. The mirror area reflects various wavelengths equally.

Even here, we have a choice of several constructions. The first described is the Gregory's mirror telescope (1663).

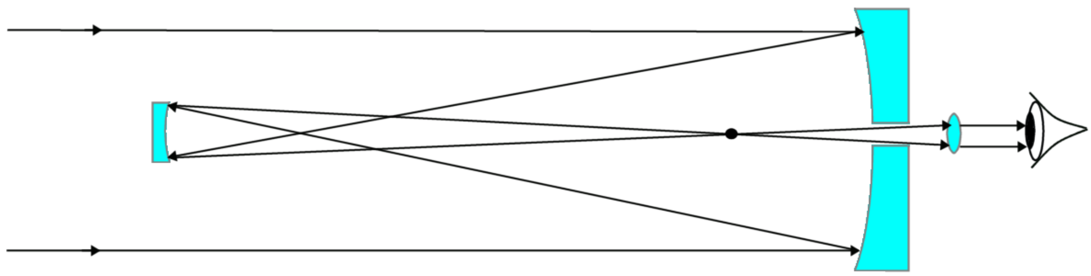


Fig. 6.25 Gregory's mirror telescope.

The second system is the Newton's telescope (1667). This telescope uses a full main mirror of the parabolic or even spherical type (if the ratio of the optically effective area is small). And then a simple flat mirror that directs the beams away from the tube. The advantage is that the production of a spherical mirror is much simpler and, therefore, it was one of the most affordable solutions.

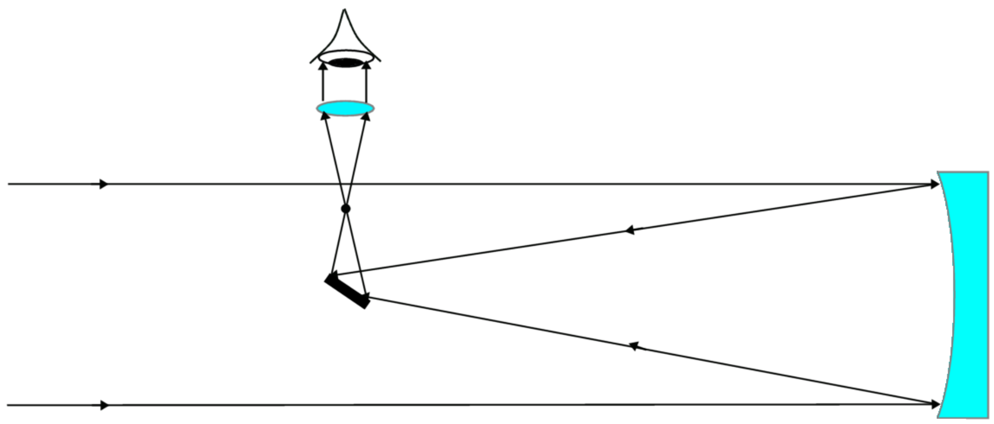


Fig. 6.26 Newton's mirror telescope

Catadioptric telescopes

The objective is formed by the combination of a mirror and a lens. The construction is more complex. However, it combines the advantages of the two previous systems because, thanks to the input lens, it is a closed construction, thanks to the mirror system, it has a high degree of distortion correction. A variant of this system is also used by the Hubble Space Telescope.

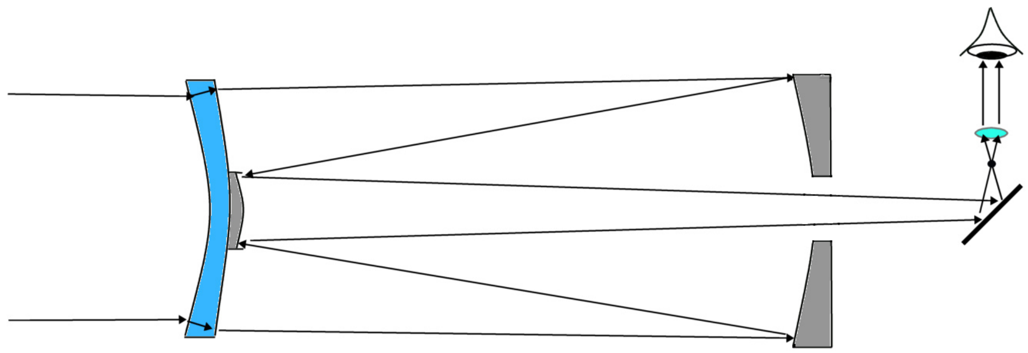


Fig. 6.27 Maksutov-Cassegrain telescope.

Binoculars

Portable binoculars for shorter distances. In principle, it is a Kepler telescope with two convergent lenses and two optical prisms. Thanks to optical prisms the inverted image is turned back to the original state. Moreover, this makes the telescope shorter because the light rays pass through the telescope's body along a longer route.

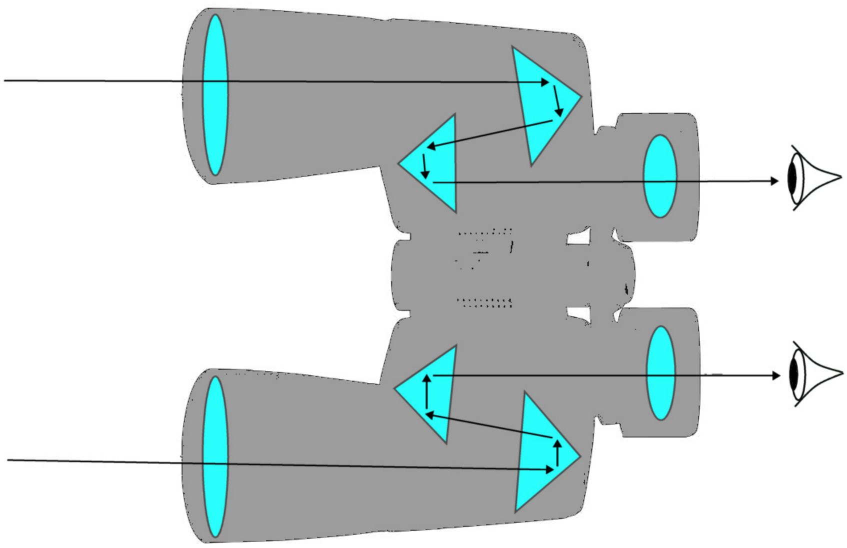


Fig. 6.28 Binocular telescope.

Display defects

Under ideal conditions it applies that the beams passing through the optical system intersect at a common focus. However, in the real world there are many display defects that are caused by anything from geometric inaccuracies, inhomogeneities of materials to the laws of physics themselves (refraction of light).

Deviations from the ideal optical image are referred to as aberrations. They can be divided into two groups:

- Monochromatic – Defects that are caused by beams of single wavelengths. (The result describes geometric defects of the display system).
- Chromatic – Defects caused by the passage of beams of different wavelengths. (It describes defects caused by the refraction of light in different environments).

Monochromatic aberrations

Aperture – the farther from the optical axis the beam will be parallel to the optical axis, the closer from the lens the beam passing through the lens will intersect the optical axis.

The aperture defect is independent of the field of vision and increases with the third power of diameter of the relative aperture. Therefore, the aim is that the beams do not enter the lens close to its edge. In order to reduce this effect, all you have to do is to use a simple circular aperture. [2]

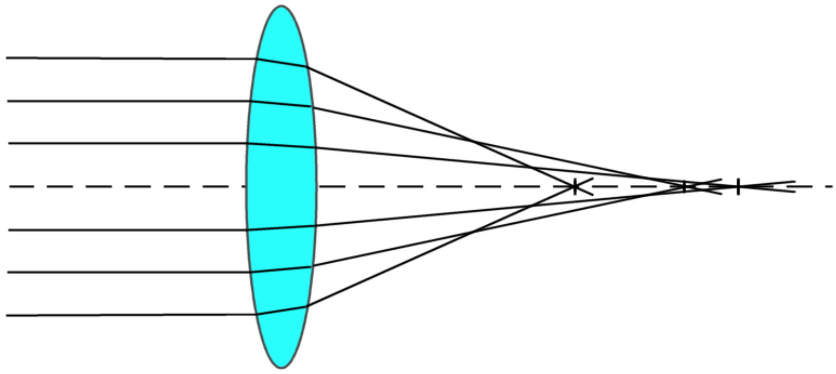


Fig. 6.29 Aperture aberration.

Coma – occurs when displayed with wide off-axis beams. Therefore, if we display a point source of beams, the result will not be the original point, but a larger circular trail. Its area will then grow with the size of the inlet hole. [2]

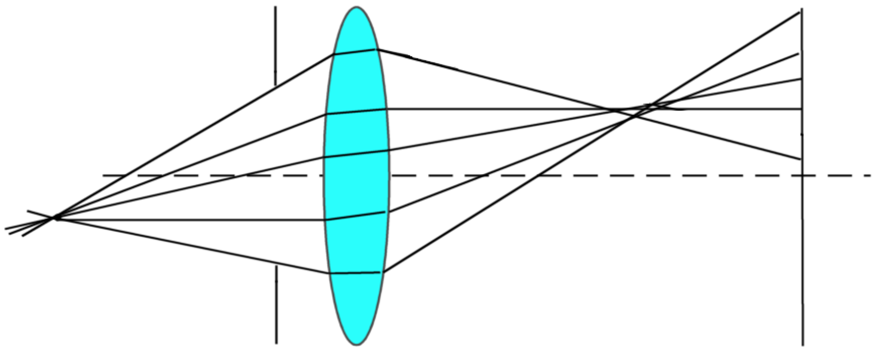


Fig. 6.30 Coma.

Astigmatism – arises from the fact that the extreme beams in two perpendicular planes intersect at different points.

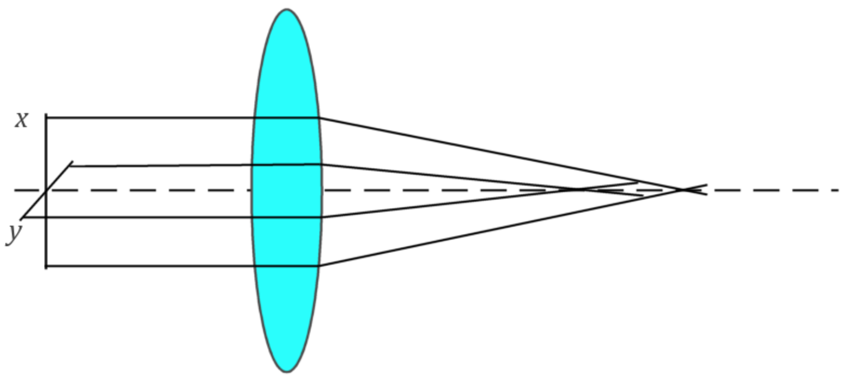


Fig. 6.31 Astigmatisms.

Image distortion – is a defect in which the image is sharp but the distance of the beam from the axis varies depending on the position of the displayed object. In general, this defect is demonstrated by using a square grid that, depending on the value of the distortion coefficient, has the shape of a ‘cushion’ (cushion-shaped distortion) or a ‘barrel’ (barrel-shaped distortion).

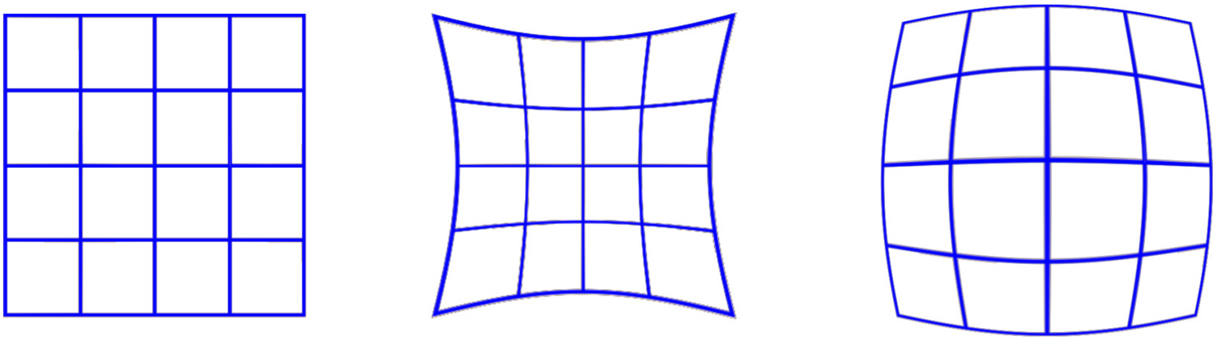


Fig. 6.32 Cushion-shaped and barrel-shaped image distortion.

This defect manifests itself, for example, in compact cameras and mobile phones where relatively small lenses are used. This defect can be corrected by a more suitable lens configuration in order to physically eliminate it. However, in practice this means longer focal lengths and larger lens diameters.

Nevertheless, the time nowadays offers a much simpler solution that is also probably the most used today. As the image is still sharp with this defect, the camera simply captures the image even with distortion. Then the processor unit performs a mathematical operation to return the image to its original shape.

Chromatic aberrations

These are colour imaging defects that are caused by the refraction of light rays of different wavelengths. The index of refraction is dependent on the wavelength and this changes the focal length for the individual beams.

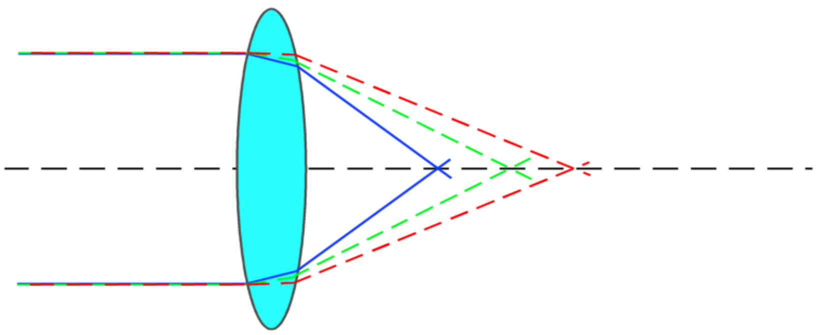


Fig. 6.33 Chromatic aberration.

The chromatic aberration can be minimized in the surrounding of a certain wavelength. For example, by combining two lenses of different materials close to each other. In order to remove this defect, it is possible to use mirrors instead of lenses that do not suffer from chromatic aberration.

Elimination of aberrations

In order to remove already mentioned defects, the systems of multiple lenses from different materials are used. In case of 2 lenses, they are doublets, in case of 3 lenses, they are triplets. The sophisticated SW tools are used to design these systems that are able to optimize the shapes and material of the lenses themselves in order to remove unpleasant aberrations and maintain, for instance, the minimum size of the optical system. Alternatively, they eliminated the need for absolute accuracy during assembly. Thanks to this, we can have cameras in mobile phones with autofocus with minimal dimensions nowadays.

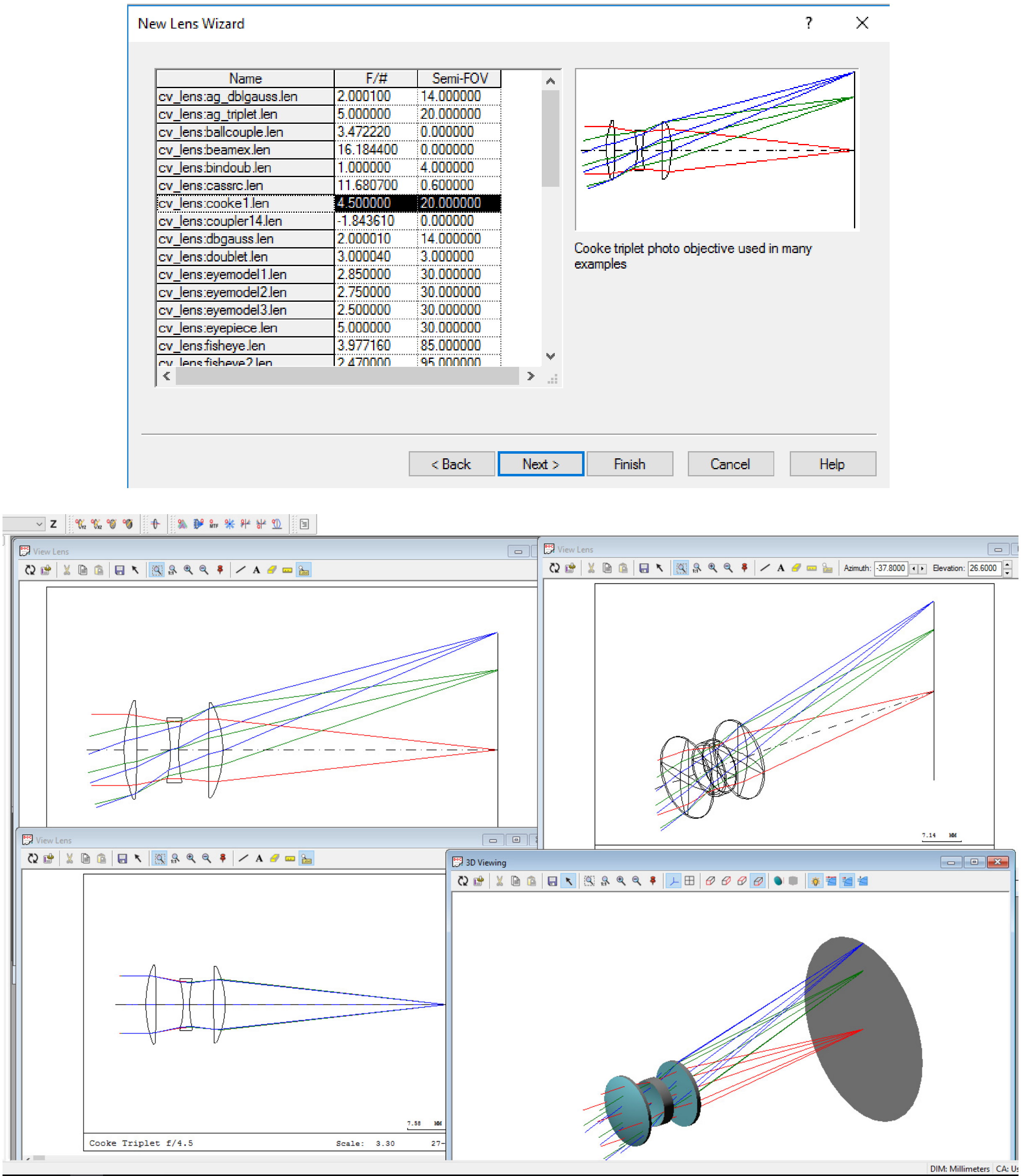


Fig. 6.34 Example of simulation SW Code V.

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. What is light?
2. What is the speed of light in a vacuum?
3. What is the speed of light in an environment with a refractive index n = 1.5?
4. Describe the law of refraction.
5. Under what conditions does absolute reflection occur?
6. What materials for the mirror production do you know?
7. What is the difference between a thin and a thick lens?
8. How does a telescope work?
9. What display defects do you know?

7. Eye

The human visual system (eye) is the sense organ of vision. Its structure forms an optical system that directs the rays of light into the light-sensitive layer in its rear part (retina).

Construction

Concerning other optical systems, the same rules apply even here regarding the terms such as focal length, refractive index, etc. Basically, it is an optical system that we know very well today, for instance, from digital cameras. The image simply enters the eye where the brightness level is adjusted by means of iris. Then, after 'autofocus', the image is displayed on the retina and from here it is converted into electrical signals and sent on for processing in the brain. The construction of the visual system differs among animal species. However, the construction of the human eye is sufficient enough to explain it.

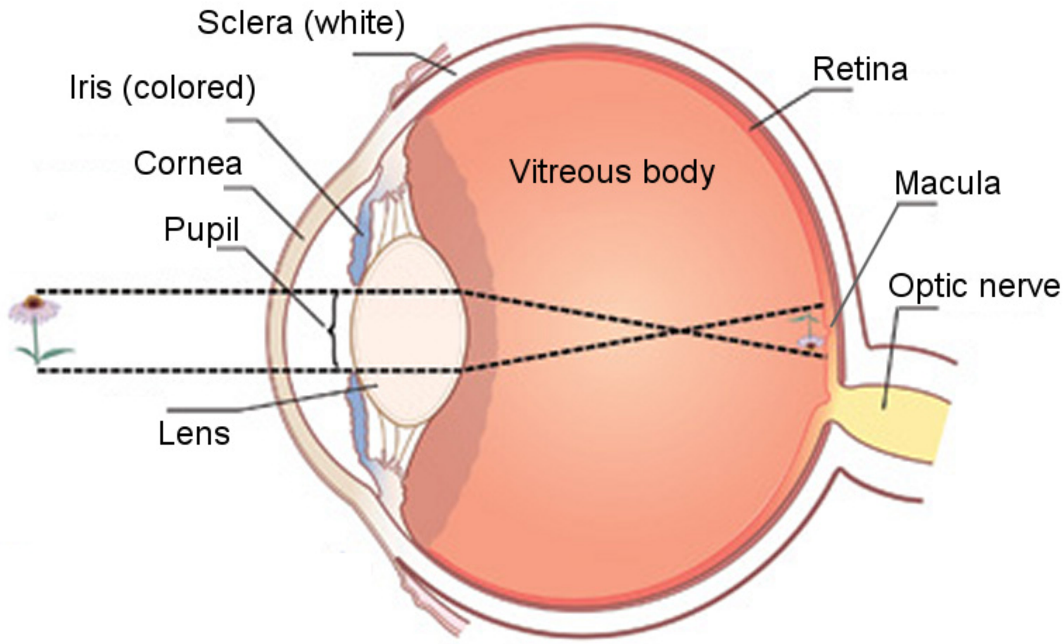


Fig. 7.1 Human eye

The optical system of the eye acts as a convergence lens, the image will be therefore inverted, reduced and real.

The largest part of the optical power of the eye is the **cornea** with a width of less than 1 mm and a refractive index of n=1.377 that forms the interface between the air (n=1) and the rest of the eye.

The **aqueous humour** that fills the space between the cornea and lens serves as a transmission medium and does not significantly contribute to the optical power (it is more concerned about nutrition and maintenance of intraocular pressure).

The **iris** that, based on the incident light flux, changes the diameter between 2 and 6 mm, is located in front of the lens. Then the resulting hole is called the **pupil.**

The second largest part of the optical power is in charge of the lens that is variably adjusted as needed. It is up to 4 mm thick and its refractive index is approximately 1.42.

The vitreous humour represents a dense intraocular fluid which certainly forms part of the optical system. However, it serves more as a medium for ray transmission, nutrition and for maintaining intraocular pressure. Its refractive index is approximately 1.336.

The resulting image is displayed on the **retina** where there are 120 millions of rods and 6 millions of cones. The **rods** are sensitive to light intensity and can also perceive the impact of individual photons. However, the light intensity information does not provide colour information. Therefore, the rods themselves allow only black and white vision. However, their extreme sensitivity allows them to see even in very dim light. The so-called **cones** are used for completion of information. They are less sensitive than rods, so they do not work in low light level of lighting and we stop perceiving colours. The cones occur in the form of 3 types in the eye and differ in colour pigment (red, green and blue). By means of these 3 colours, we can perceive a large number of colour shades in the range of approximately 390-790 nm. Using a photochemical reaction, the information is then converted into electrical impulses that are transmitted to the brain via the optic nerve.

The place of the sharpest vision is located in the axis of the field of view, where the so-called yellow spot (**Macula**) is located. The cones dominate in it and give us the most, while we perceive the environment only peripherally (out of focus). In order to create a broader picture of the environment we need to move the eye so that the brain has enough information to create this short-term illusion. It is the reason why we have a constant tendency to move the eye and observe the surrounding. Additionally, there are no rods or cones in the area where the optic nerve leaves the eye. This place is also called the **blind spot**. The brain only guesses information from this area based on the available information. We can use the following test for illustration:

The test is very simple. At first, we will cover the left eye and focus on the **+** shape with the right eye. Then we will start zooming in/out until the dot on the right side disappears. The brain has no information about the dot at this moment, but it has information about what the surrounding looks like. For this reason, if fills in a blind spot based on the surrounding (we can see a fictitious square background) and at first glance we have no idea that we are 'blind' at this point.

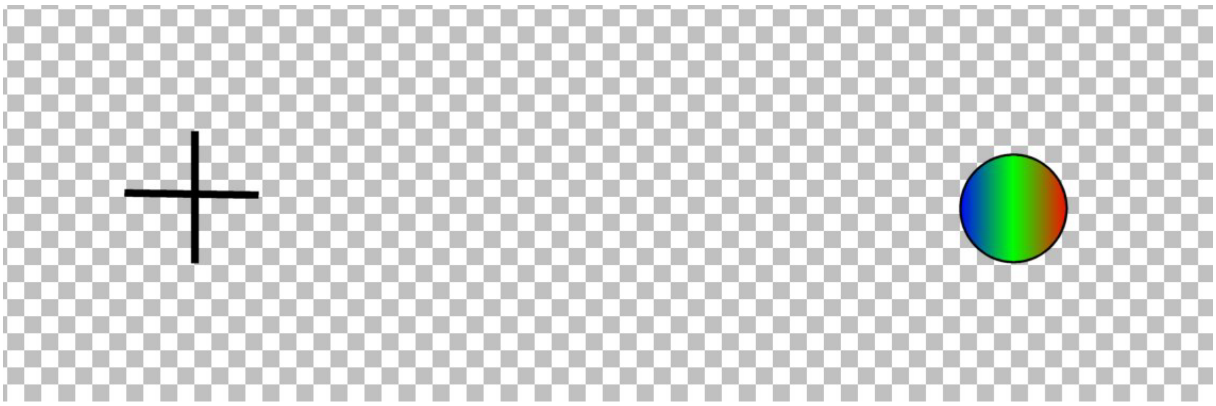


Fig. 7.2 Illusion of a blind spot

The advantage is that the second eye has a blind spot on the other side, so if the brain has information from the other eye, it can complete the image correctly ant the field of view is not disturbed.

Vision modes

As the difference between light conditions at night and day is too big, the eye uses the day (photopic) mode and night (scotopic) mode. The eye has enough light for rods and cones during the day and can easily perceive colours. The most sensitive region of the spectrum in this case is around 555 nm (yellow-green colour). During the night, we perceive mainly with the help of sensitive rods and the perception of colours is limited. Sensitivity to the colour spectrum shifts closer to the blue colour. The transition to night mode occurs only in really dark conditions and is very slow – it can take over ten minutes.

The blue part of the spectrum has another purpose in the brain as the brain uses it to recognize day and night. The blue light in the brain activates the production of serotonin which positively affects brain activity and forces us to be awake and alert. The opposite is melatonin that reduces activity and forces us to sleep. At night, when there is not enough blue light, the melatonin prevails and we start to feel tired. Nevertheless, the recent problem is the arrival of the LED lighting. The basis of white LED diodes is the blue LED diode and the rest of colour spectrum is completed by luminophores on the surface. Based on the spectrum we define the colour temperature that corresponds to this composition. As a reminder, the following figure shows the spectrum of white LEDs for different temperatures.

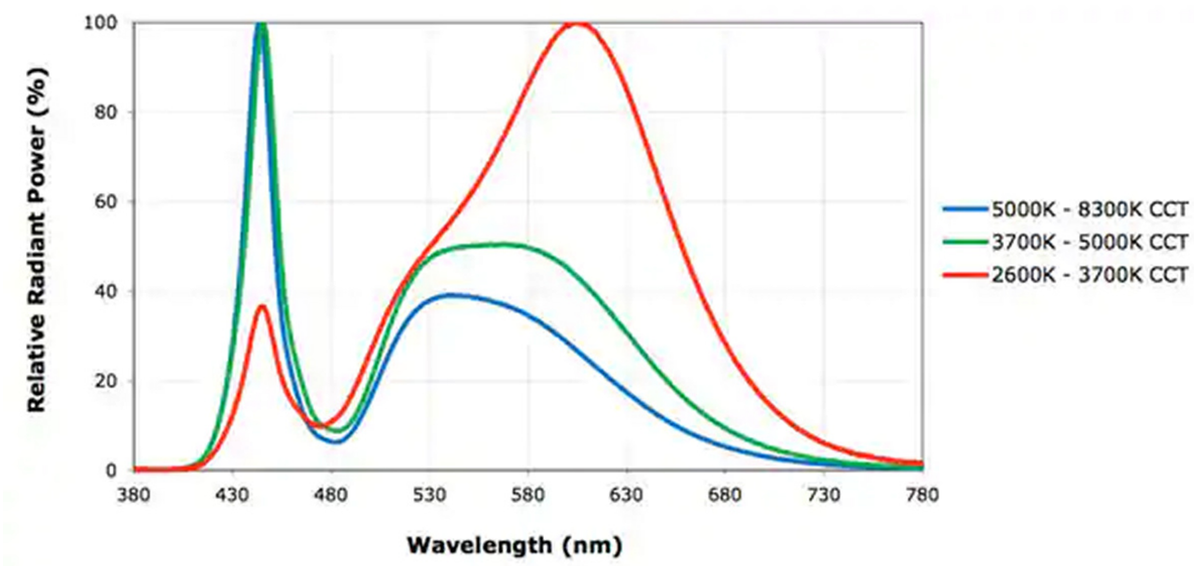


Fig. 7.3 Spectrum of white LEDs.

The higher temperatures around 7000K are described as cold white and have a large share of the blue component. As a result, this light basically tries to keep a person awake, so he or she is ready to work. Lower temperatures around 3000K are called as warm white and consist much less of the blue component. So, they fit into the home environment when a person is getting ready to sleep. Then, degradation of large amounts of serotonin can take several hours for more sensitive individuals.

Eye defects

When the eye is functioning correctly, the image is accurately displayed on the retina. Therefore, from an optical point of view the parallel rays that are entering the eye meet at one point of the retina (focus). However, if everything is not accurately set in the eye for some reason, or it is not within the required tolerance, the image is not focused on the retina. In principle, there are problems with bad characteristics of fracture surfaces (refractive errors) or bad eye geometry when the eye has different lengths (axial defects). Of course, combinations of both can also occur. Eye defects can be divided into spherical and aspherical. [10]

Spherical defects

They are corrected by spherical lenses (lenses are symmetrical in both axes) and include:

Myopia – Parallel rays that are entering the eye are projected in front of the retina.

- Spherical cause: the eye is physically too long.
- Refraction cause: the optical environment of the eye has greater refraction.

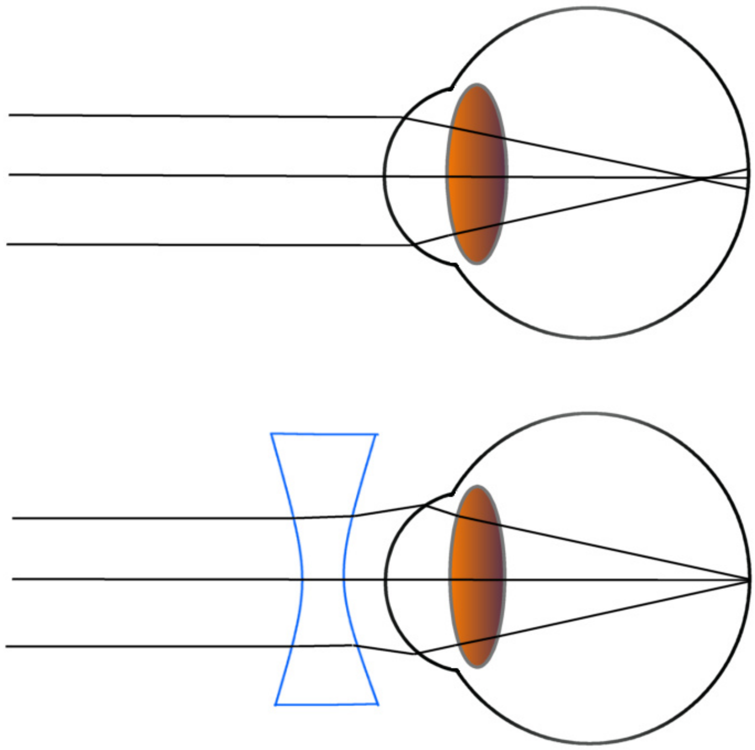


Fig. 7.4 Correction of myopia. (concave lens)

As a result, the eye can see poorly at long distances (however, at the same time it is able to see at short distances).

Concave lens is being used for correction. They move the focus further with retina.

Farsightedness (hypermetropia) – Parallel rays that are entering the eye are projected behind the retina.

- Spherical cause: the eye is physically too short.
- Refraction cause: the optical environment of the eye has less refraction.

Due to this defect, the eye sees poorly at short distances (in principle, this defect appears at longer distances to some extent).

We can use convergence lens that will bring the distant focus back to the retina for the correction.

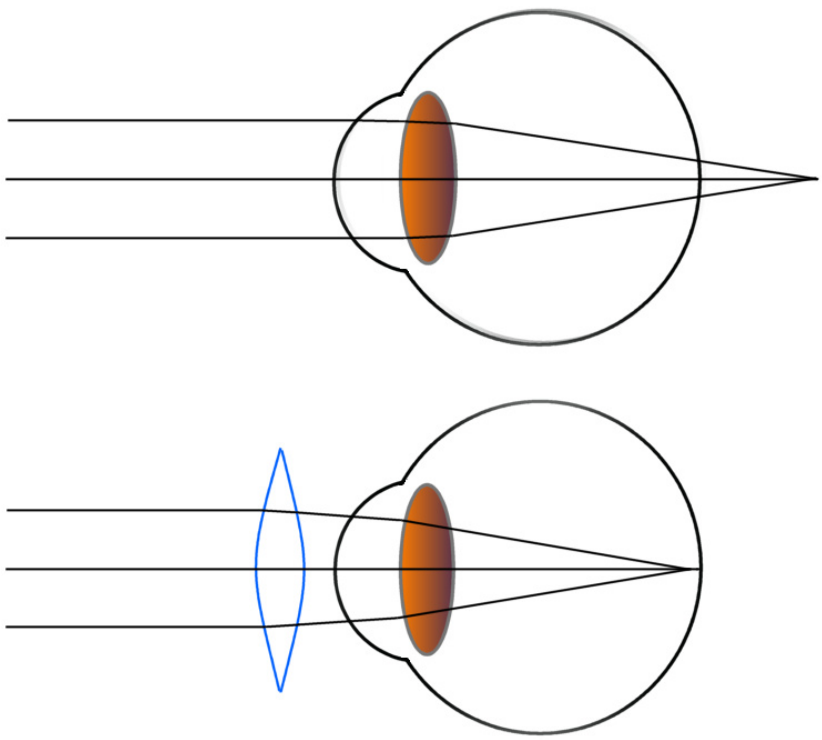


Fig. 7.5 Farsightedness correction (convergence lens).

Aspherical defects

They are corrected by aspherical lenses (lenses are not symmetrical in both axes): and include, for instance:

Astigmatism – Parallel rays that are entering the eye are projected in two different foci.

- Cause: The cornea has a different curvature in two perpendicular axes (less often the poor lens centering).

The eye focuses poorly in different axes and it is necessary to design the lens so it would fit both foci to a single point.

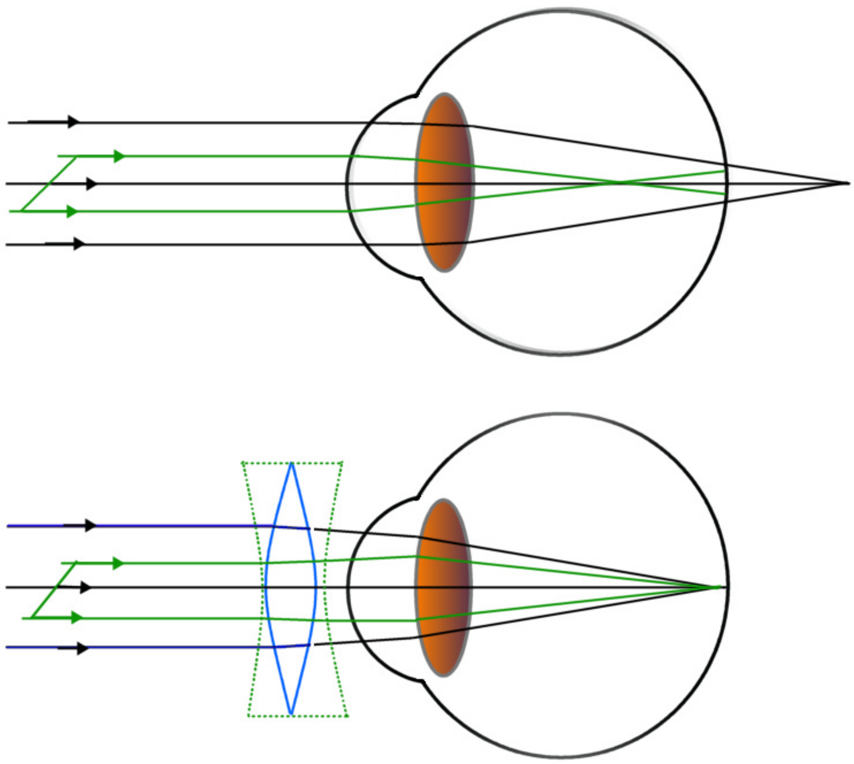


Fig. 7.6 Astigmatism correction (lens with different foci in the *x* and *y* axes).

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. What is an eye and how does it work?
2. Describe the function of the yellow spot?
3. How does the eye perceive colors?
4. What light temperature would you choose for work and what before sleep?
5. What are the basic defects of the eye?

8. Polarization of light

Light can be understood not only as particles (photons) but also as electromagnetic (EM) waves. It means that it has its electrical (E) and magnetic component (B) that are perpendicular to each other. We can exclude the magnetic component for simplification because we are mainly interested in the electrical component. The polarization describes the parameters that are related to the course of the EM light component. The reason is that each particle (photon) is its own source.

Linear polarization

To begin with, let's imagine a simple course of an electromagnetic wave of light radiation from the following figure:

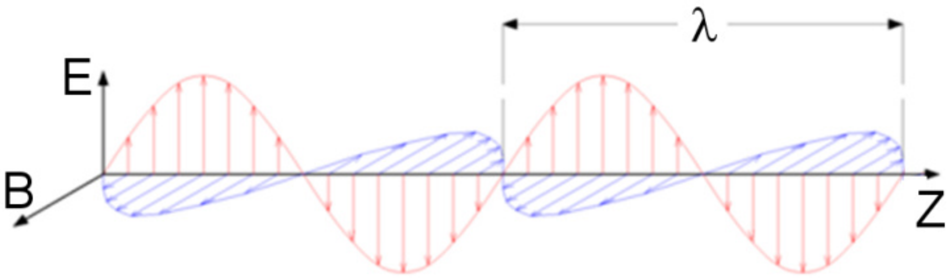


Fig. 8.1 EM component of light radiation

Now let´s imagine that we are observers on the Z axis and we look at the course of the electrical part from the front side where the waves remain constant. The observed course would be as follows:

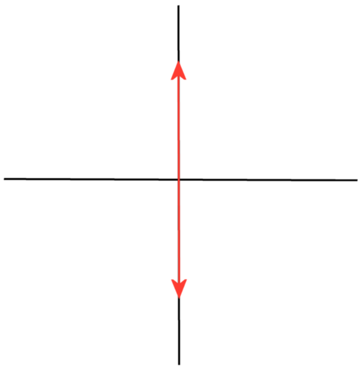


Fig. 8.2 Vertically polarized light radiation (I would illustrate the axes and also the E component)

Since the angle of the electrical component does not change, it would be the so-called polarized light radiation in this case. In addition, the polarized light radiation is exactly vertical in this case. If the electrical component were rotated 90° with respect to the axis Z, it would be the light polarized horizontally (see the following figure).

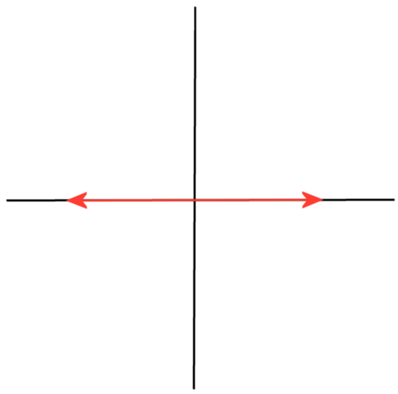


Fig. 8.3 Horizontally polarized light (the same even here)

And, of course, if the electrical component is rotated at any different angle, it will also be the polarized light.

Nevertheless, the polarized light is basically a very specific case. The light that normally surrounds us, e.g. (light rays of the sun, light bulbs and even LED diodes), is unpolarized. Individual photons are created randomly here (either due to heat or spontaneous emission). Therefore, each photon is polarized randomly regardless of the others. Thus, we see all possible polarizations mixed together on the receiving side. The intensities can differ within individual types of polarization.

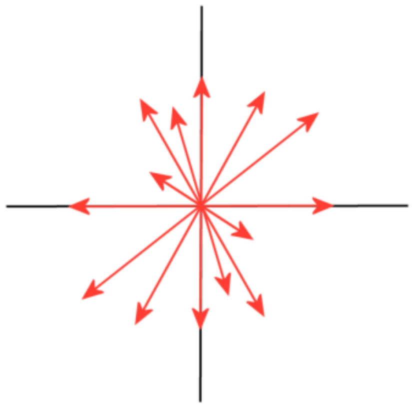


Fig. 8.4 non-polarized light radiation

In this case, it is not possible to say what polarization the next photon will have as the polarization is random at every moment.

The fact is that we are normally surrounded by rather unpolarized light. However, the polarized light has many advantages that we use in the world nowadays.

- Polarized light can be added or subtracted. (more photons in the phase will cause an increase in intensity and vice versa)
- Polarized light cannot pass through the material with opposite polarization (this is used, for instance, in display devices, such as LCD monitors, projectors…)

How do we create the polarized light?

- We will use a radiation source that generates the polarized light automatically. A laser is a typical example of this source.
- We will use a common source of non-polarized radiation and use a simple and cheap device in order to select only the requested part of the EM spectrum for a given polarization. This device is called a polarizer. Basically, it is an optical filter that releases only light of a specific polarization and absorbs or reflects the rest.

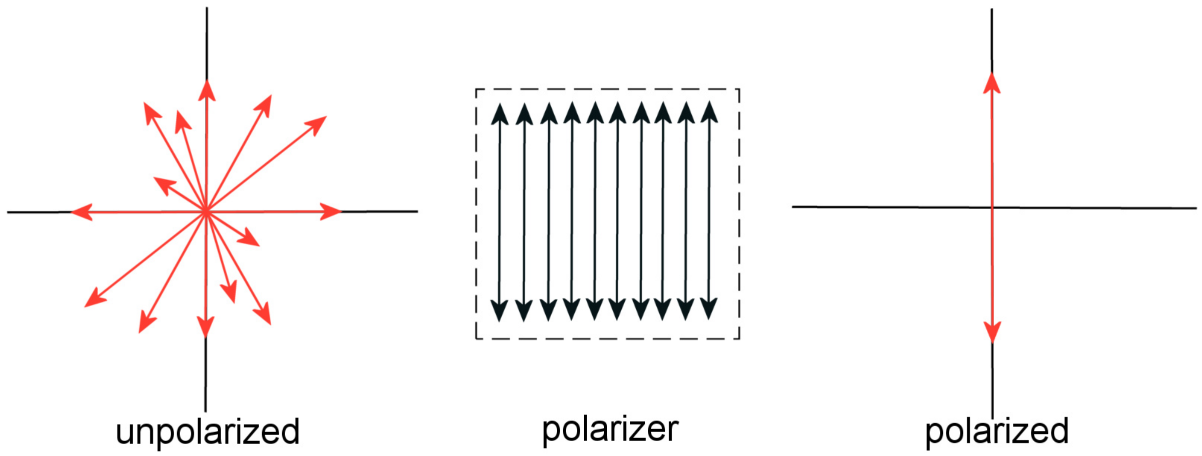


Fig. 8.5 Polarizer

We can rotate the polarizer an any angle. Thus, we can easily produce radiation of any other polarization from an unpolarized radiation source.

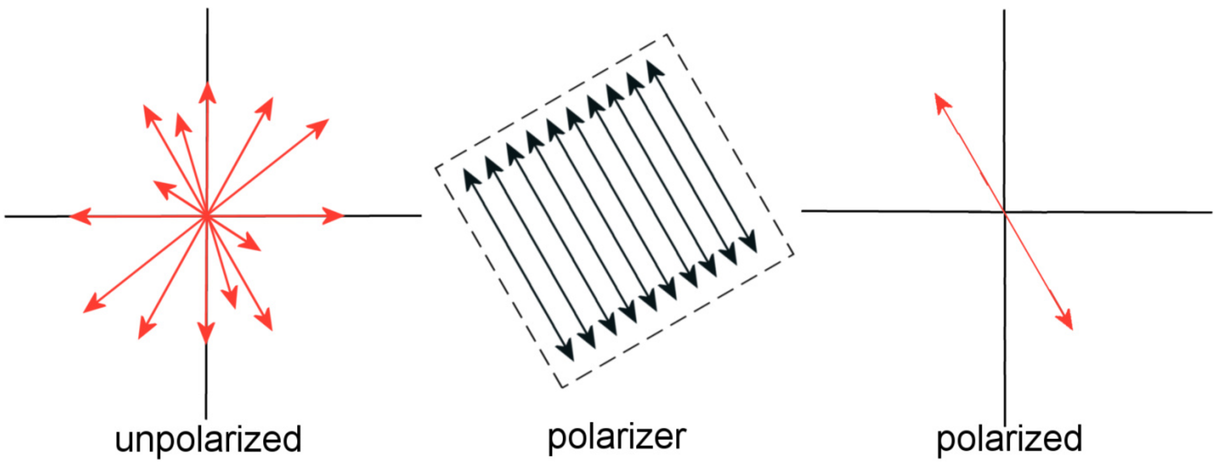


Fig. 8.6 polarizer rotated 30°.

The polarized light radiation is also generated by reflection from reflective surfaces. The light rays of the sun falling on the water surface show a simple example. The light that falls on the reflecting surface is polarized (or at least partially polarized) at an angle equal to the plane of the reflecting surface. Therefore, if we stand by the water surface the light, which is reflected from it against us, is partially horizontally polarized (efficiency is usually not 100%, but most of the light will be polarized). It means that if we use the so-called polarizing glasses (with vertical polarization), the intensity of the sun's rays reflected from the water surface will be significantly reduced.

The most well-known use of polarization in everyday life is probably watching 3D movies. The person uses two simultaneous views from slightly different angles in order to perceive a 3D image, thanks to which the human visual system has information not only about the image, but also about the distances. It is necessary to compose 2 images from different viewing angles and ensure that each eye sees only one in order to create a 3D effect.

Both the images will be projected on the projected screen at the same time, but with different polarizations. Then, the image is separated for each eye separately by means of the polarized glasses.

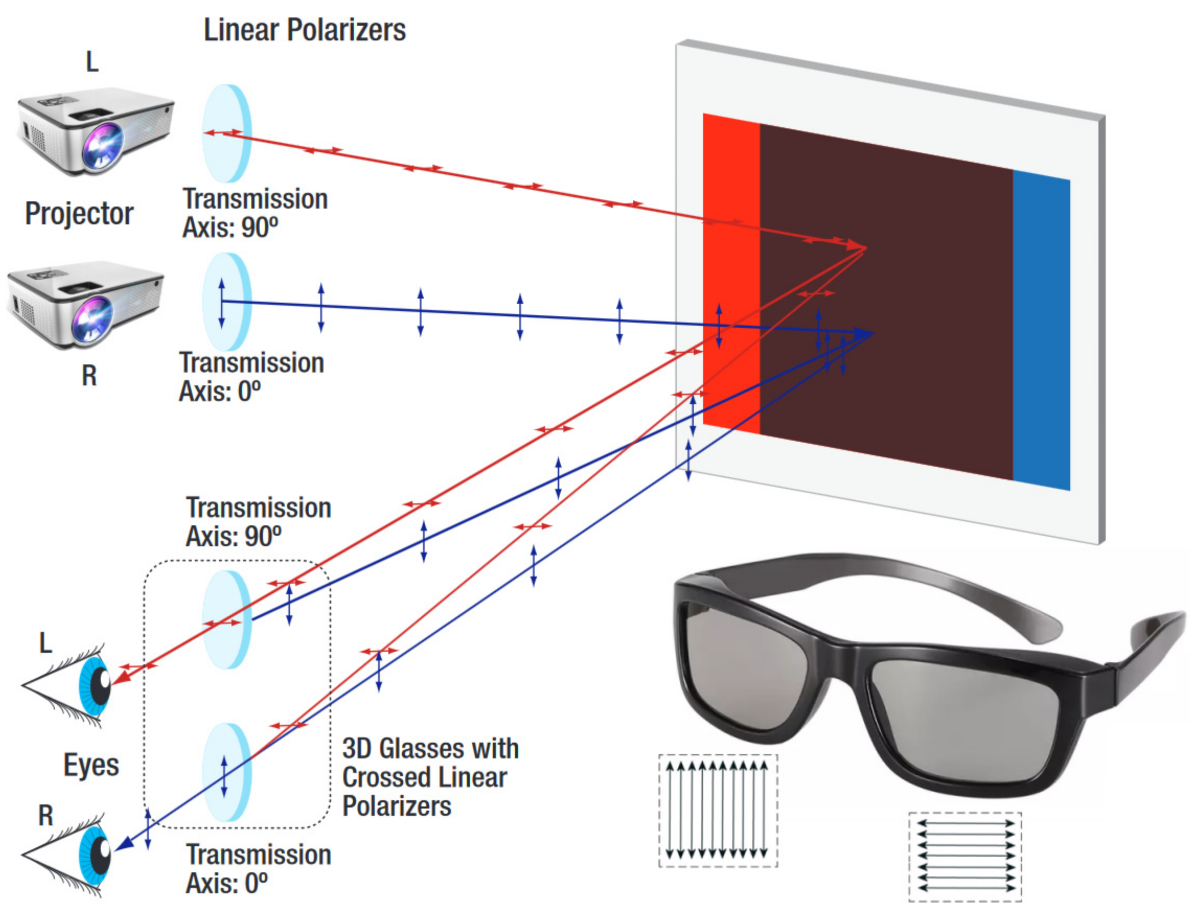


Fig. 8.7 The principle of 3D movies. [11]

All polarization states that have been mentioned so far are linear. They always have a given state that does not change. Nevertheless, there are also cases where the state of polarization changes over time.

Circular polarization (Circular polarization)

In order to understand circular polarization, let's consider a state when we have multiple electromagnetic waves of the same frequency (or the wavelength) at the same time and they are in phase with each other. However, they have simultaneously different polarization. [12]

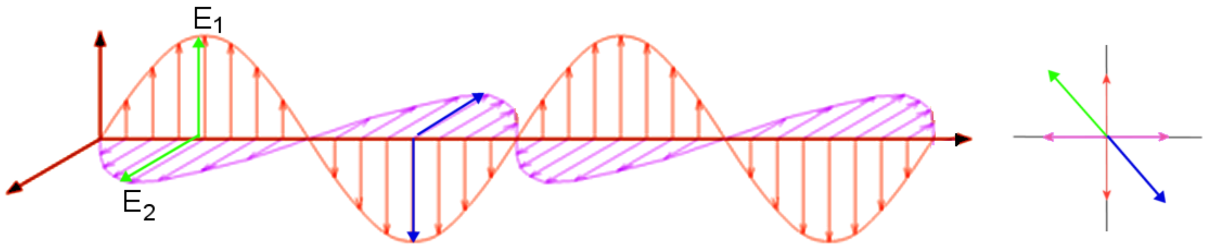


Fig. 8.8 two mutually perpendicular waves.

In case of the Figure 8.8, this is just the case when we have 2 EM waves in phase but their polarization is rotated 90° with respect to each other. The vector of the resulting electrical component will be the sum of these waves. It is still a linear polarization in this case, but the change occurs when the described waves will not be in the phase. If we move the waves by 90° in the axis, we get the following course:

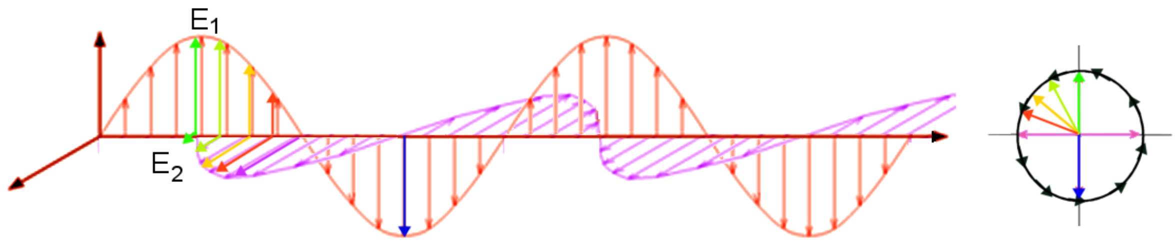


Fig. 8.9 two mutually shifted waves.

The polarization will keep constantly changing. The change will take place so that the resulting polarization vector will constantly rotate in one direction and copy the circle.

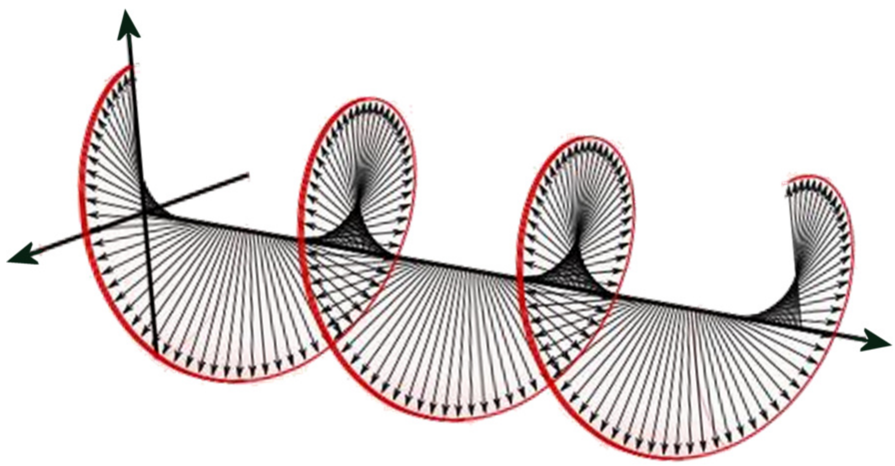


Fig. 8.10 the resulting course of circular polarization.

The circular polarization can be clockwise or anti-clockwise. It depends on the fact whether the second wave shifts 90° or -90° (Pi/2, or – Pi/2).

The creation of this polarization is again possible by means of a ‘filter’ or more precisely by means of a quarter wavelength plate that ensures the delay of one axis by Pi/2.

If the waves are not shifted exactly, or if they have different amplitudes, the resulting vector will copy an ellipse (the case of elliptical polarization).

What is it for?

Let’s imagine a 3D film projection for simplification. Provided that we use linear polarization, we will have to keep our head in the exact position. If there is a deviation from the ideal state, then the second image will begin to partially penetrate the glasses (if we tilt our head by 45°, so that we would see both images with each eye…). If we use circular polarization, we will remove this effect. Therefore, the position of the glasses will no longer matter and watching the film will be much more pleasant.

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. What is the difference between polarized and non-polarized light?
2. How can we produce linearly polarized light?
3. How can we produce circularly polarized light?

9. Interference of light

Light can be seen as a particle (photon) – quantum optics and as electromagnetic waves at the same time. And it is the principle of waves that carries with it the possibility that individual light waves can affect each other. The result is the ability to add and subtract individual waves. Just as in the case of mechanical waves when the instantaneous deviations are added, so in the case of light the electrical and magnetic components of electromagnetic waves are added and subtracted.

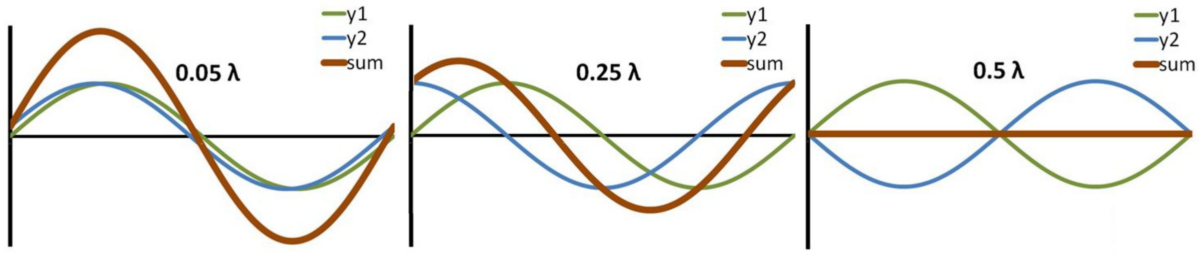


Fig. 9.1 Light interference.

If the light waves are in a common phase or are not too shifted mutually, the effect of addition is noticeable. If they are shifted by 0.25 λ, the effect of the subtraction also becomes apparent. The zero value of the resulting signal occurs at the moment when the positive and negative parts of the interfering waves are identical. The resulting signal for 0.25 λ is as a result larger than in the case of one wave, but the amplitude is already noticeably lower than in case of the waves in the phase. And, along with further shifts of the phases, the amplitude of the resulting course will further decrease. In the case of the shift of 0.5 λ, the waves will cancel each other out and their amplitude will be zero (the light literally disappears). Therefore, if the light comes in different paths, we can observe interference patterns that are similar to the following figure.

Fig. 9.2 The example of the interference pattern.

Based on these rules, we should observe similar patterns all around us. The reason why this is not the case is simple. The light coming out from radiation sources (sun, candle, light bulb, LED diode…) is of random nature. Individual waves differ in the rotation of the electrical and magnetic components (polarization), the direction of propagation and they contain more wavelengths (different wave frequencies). This means that the resulting sum of these waves will usually be non-zero. For simplicity, the nature of ordinary light is more like the noise that stays around ‘stable’ mean. Such a chaotic arrangement is called incoherent waving. As a result, we are able to observe interferences of incoherent radiation only under certain conditions.

Coherent waves

Coherent waves can be understood as the mutual connection of the phase and the wave amplitude. It can come either from different sources or just from one (with a certain time interval). The substance is the ‘arrangement’ of light that has the same phase in a given place which does not change over time.

Then the light behaves in the same way as in Figure 9.1. If the waves pass mutually in the phase, they are added up (constructive interference). If the waves are in the counter-phase, they are attenuated (destructive interference).

The typical source of coherent waves is a laser. The light produced is mutual in the phase in terms of the nature of the function. The second condition is the identical frequency (wavelength) of individual waves, the so-called monochromatic radiation. In case of the laser the spectrum is narrow enough that we can consider this condition to be met under certain conditions (at least on a limited path). At longer distances, of course, even minimal deviations in frequency will cause visible interference of the laser trace (Speckle patterns).

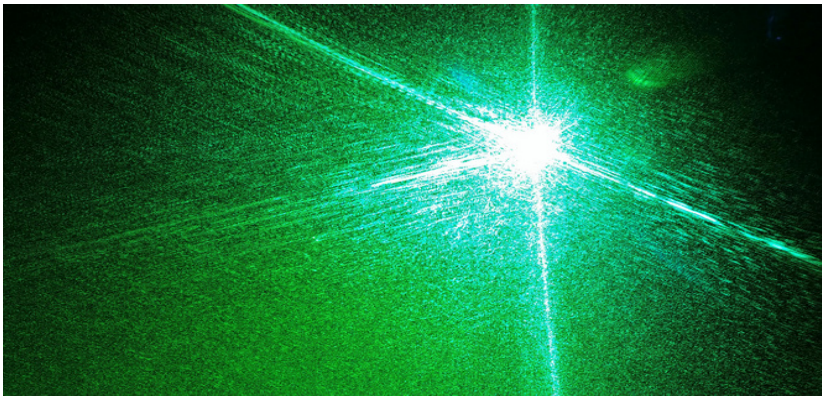


Fig. 9.3 Laser light interference.

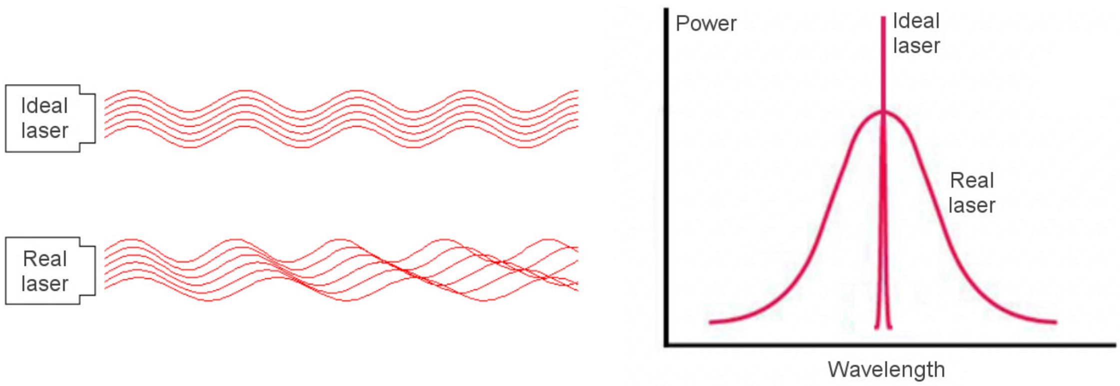


Fig. 9.4 Comparison of real and ideal laser.

When it comes to the laser output it will always depend on how quality laser we use. Cheaper variants with a spectral line width of around 1-2nm may not always be suitable.

For example, let´s imagine a laser having an FWHM of 1nm and a wavelength of, for example, 650nm. The frequency of the light radiation is then calculated simply by a formula.

$$\lambda = \frac{c}{f} \qquad f = \frac{c}{\lambda}$$

After adjustment:

If we substitute wavelengths of 649,5 and 650,5nm, we receive the frequency that differ by 7,1*10^{11} Hz. That is the difference in the extreme emitted frequencies of 710 GHz which is an almost unimaginable value.

On the other hand, nowadays there are so-called single-frequency lasers that reach FWHM well below 1nm and are thus able to reduce the differences in emitted frequencies to the values in the order of MHz (DFB lasers), or even units of kHz (He-Ne lasers). That is the value that is very close to the ideal state. Nevertheless, it also means that there will be a thousand interferences in a single second. (However, in terms of the speed of light, it will be a sufficiently long distance).

Instead of the frequency difference, a coherence time τ_k that corresponds to the value Δt can be used for the quantification which is the value of the time after which the interference occurs.

$$\tau_k = \Delta t = \frac{1}{\Delta f}$$

If we know the speed of wave propagation we can calculate the coherence l_k based on the formula.

$$l_k = c \cdot \tau_k = \frac{c}{\Delta f}$$

After adjustment

$$\Delta f = c \cdot \frac{\Delta \lambda}{\lambda^2}$$

We can make a simplification to:

$$l_k = \frac{\lambda^2}{\Delta \lambda}$$

Then the coherent length will range from millimetres to kilometres – depending on the type of laser.

The interference on a thin layer

We have already said that the ordinary light is incoherent. However, now we see that even in terms of the radiation it is possible to express the value of the coherent length. However, it will be very small. The second condition for interference is the identical polarization of the light waves. The partial polarization in nature occurs, for example, by reflection from the water surface (see the Chapter on polarization…).

The result is the possibility of observing the light interferences if, for example, there is a thin layer of oil on the water surface, possibly on soap bubbles, etc.

Fig. 9.5 Example of interference on a thin layer.

From a physical point of view, the following happens:

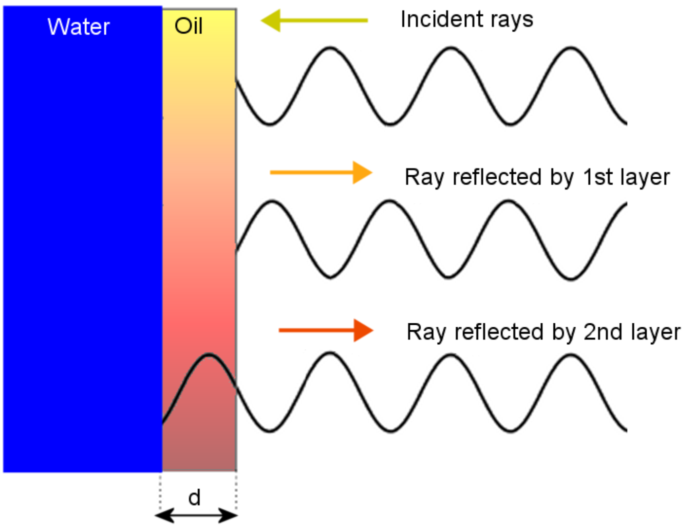


Fig. 9.6 Example of a thin layer of interference.

The result of the interference depends on the difference of the route Δs between the rays reflected from 1st and 2nd layer.

$$\Delta s = 2 \cdot n \cdot d + \frac{\lambda}{2}$$

Where *d* is the thickness of the layer, *n* is the refractive index of the thin film and λ is the wavelength of light.

Condition of Interference maximum: for the moment of a maximum amplification when the waves are in the phase, the following applies:

$$\Delta s = k \cdot \lambda$$

Condition of Interference minimum: for the moment of greatest attenuation when the waves are in counter-phase:

$$\Delta s = (2k + 1) \cdot \frac{\lambda}{2}$$

where *k* is the order of the interference maximum/minimum.

This explains both the decrease in light intensity and the colour change. The formula shows that this phenomenon depends on the layer thickness, the refractive index of the layer and also the wavelength.

This phenomenon, apart from admiring soap bubbles and observing natural disasters, can be used in practice. It allows you to create, for example, anti-reflective layers on the lenses of the objectives or glasses.

A thin layer of the material with a refractive index less than the refractive index of the lens is created on the lens. Then the thickness of the material is chosen so that for the light of a specific wavelength, an interference minimum occurs after reflection from the 1^{st} and 2^{nd} layer. As this principle can be only used for a specific wavelength, a value of 550 nm (i.e. the maximum of solar radiation) is usually chosen to limit the reflection in the visible spectrum. The result is a clearer image with less reflections from ambient light sources.

Note: If the light is reflected from an optically denser environment, its phase changes to the opposite. If the light is reflected from an optically thinner environment, then the phase does not change. [13]

Interferometers

When we know the basic principles of the origin of interferences, we can proceed to other possibilities of practical use. It is, for example, the possibility to measure precisely the changes of distances or the refractive index. Thanks to the ability of resolution at the level of the wavelength we are able to record literally microscopic changes.

These devices are generally called interferometers. And there are several basic types that differ in the organization of the optical system.

Michelson interferometer

This interferometer was invented in 1887 and consists of a combination of mirrors and semi-transparent mirrors.

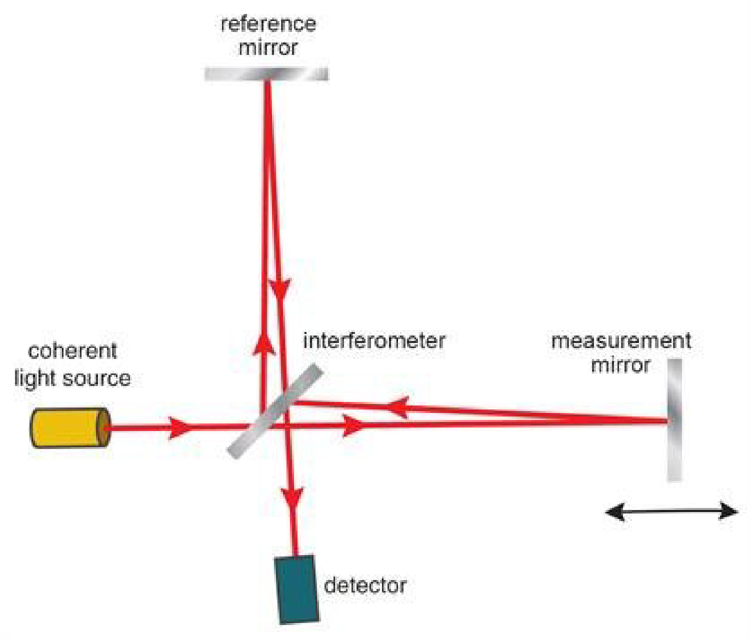


Fig. 9.7 Michelson interferometer. [14]

This interferometer divides the input beam of coherent light (laser) into two identical beams (50:50). Then each beam continues on a different route to the mirrors where it is reflected back to the semi-transparent mirror. As soon as the two optical beams meet, the interference occurs due to the phase shift which is captured by the photodetector. In this way, the system is able to detect anything that has affected its phase along the route of the light beam. Apart from the position changes, there can be, for example, air temperature, relative humidity, pressure and temperature within the atmosphere. Nevertheless, this behaviour is not always desirable and may need to be compensated.

Nowadays, we can easily utilize the system by using optical fibres. And, as the fibre bending changes the route of the beam inside, we get a sensor capable of detecting the slightest change in the optical route. The system can be so sensitive that the optical fibre picks up the sound waves of the human voice in its vicinity. And last but not least, thanks to the optical fibre, the system can work over longer distances.

Mach-Zender interferometer

It was developed in 1891-1892 by Ernst Mach and Ludwig Zehnder. This interferometer is similar to the Michelson´s one because it also splits the light beam into two arms. One serves as a reference and one as a measured one. The changes in the beam route between the reference and measured sections then show themselves as the interference on the side of the photodetector.

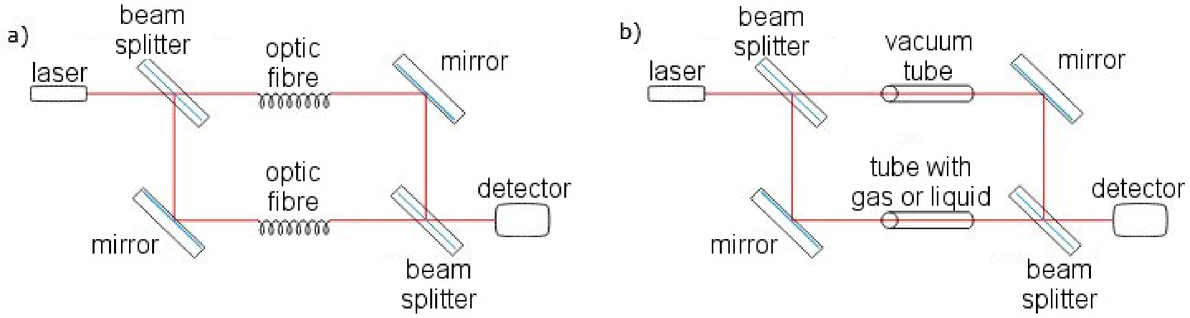


Fig. 9.8 Mach-Zehnder interferometer.

Even here, we can use it to measure the atmosphere or optical fibres. In case of optical fibres, we use a 50:50 optical divider instead of mirrors.

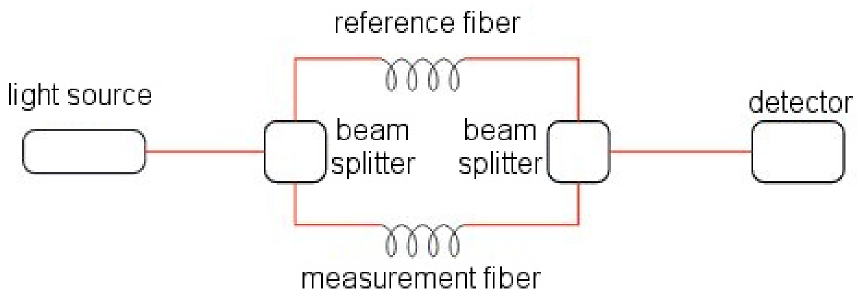


Fig. 9.9 The fibre optic Mach-Zehnder interferometer.

Sagnac interferometer

It was invented by Georges Sagnac in 1913. Even here the light beam is divided into two arms. However, the divided beams spread against each other across the same optical path. The laser beam enters the optical splitter and splits into two opposite directions that interfere with each other.

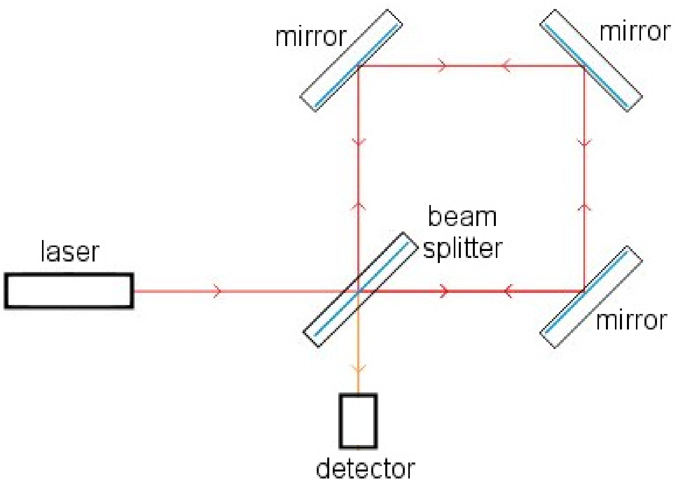


Fig. 9.10 Sagnac interferometer.

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. How does light interference work and in which cases does the interference maximum and minimum occur?
2. When does interference occur?
3. Under what conditions can we easily observe light interferences?
4. What are the basic types of interferometers?
5. What is the general principle of interferometers?

10. Light bending (diffraction)

The beam of the light spreads in straightforward based on the laws of geometric optics. However, this definition assumes that light spreads only as a particle. The obstacle between the point light source and the shade should therefore create a precisely defined shadow. The light should not fall into the resulting area of the shadow.

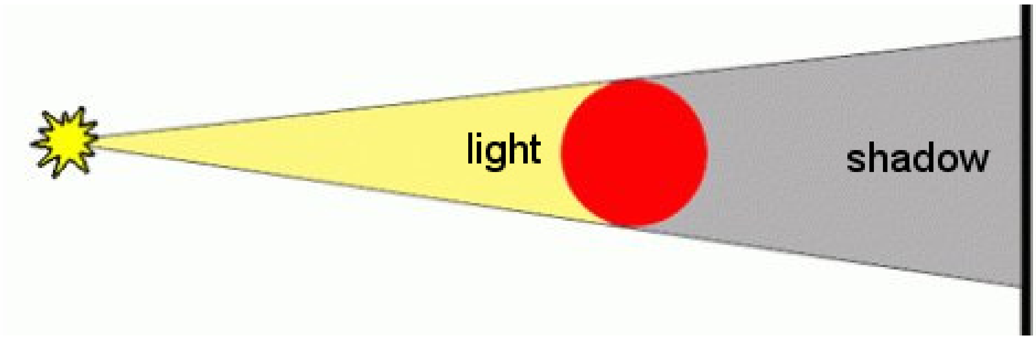


Fig. 10.1 Geometric propagation of light.

However, the light can be seen as waving at the same time. As a result, the light also propagates according to the Huygens principle. The light, when it is spreading, creates wavefronts, and each point of this wavefront into which the wave has reached can be considered as another source of elementary wave. Out from this source, the light propagates further in the form of elementary wavefronts. Therefore, the light can also propagate into the area of the geometric shadow behind the obstacles.

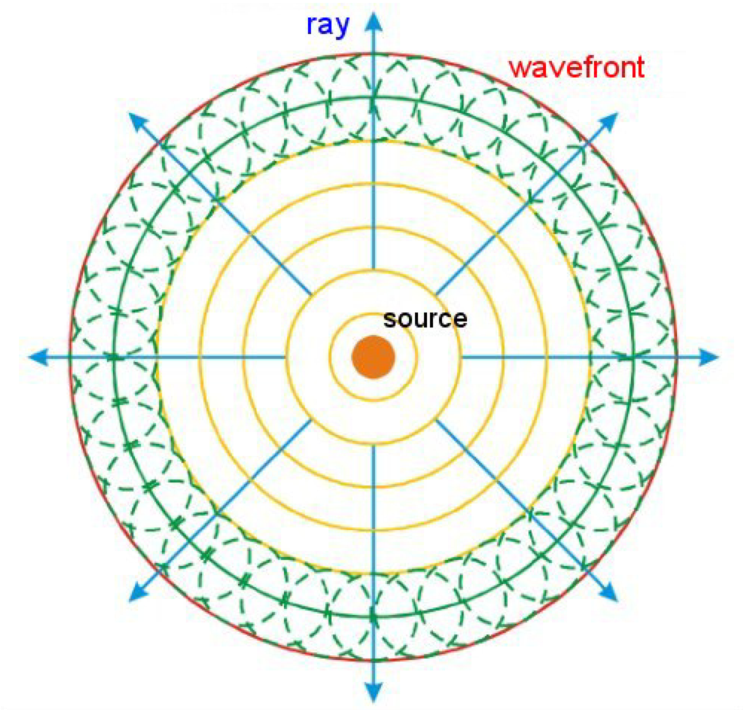


Fig. 10.2 Propagation of beams and wavefronts.

Based on the Huygens principle, the light can propagate not only on spherical wavefronts, but also in the form of planar wavefronts. In case the light propagates from a remote source and falls to a small area, we can approximate the spherical propagation for simplicity.

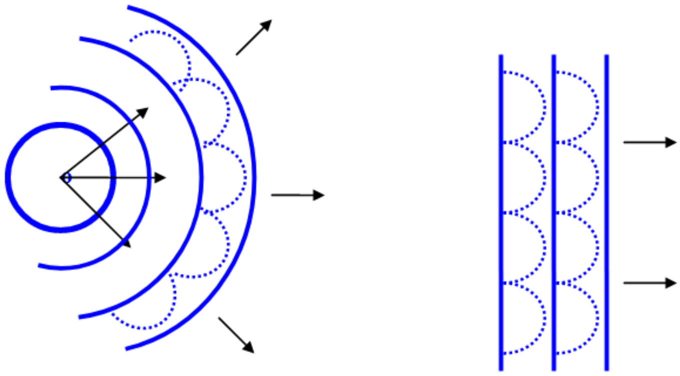


Fig. 10.3 Propagation of spherical and planar wavefronts.

Diffraction grating

Let´s assume we have a wave that falls on an obstacle with a hole which dimensions are comparable to the wavelength. The part of the wave that passes through the hole will spread (bend) into the area behind the obstacle.

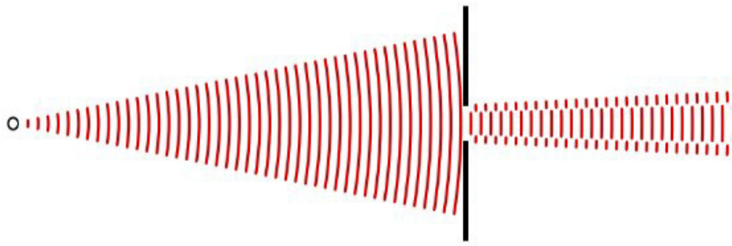


Fig. 10.4 Diffraction of waving.

The diffraction grating consists of a system of thin bands with very small dimensions that are comparable to the wavelength. This divides part of the incident radiation into several parts, in steps at an angle ϑ_m. Whereas, the following relationship applies:

$$\sin\vartheta_m = \frac{\lambda}{d}$$

Therefore, the value of this angle depends on the wavelength of the radiation λ and distance d of the individual strips on the grating (so-called Grating Parameter). We can easily calculate the corresponding distances of individual maxima from the geometric properties that are displayed on the screen at distances x_m and at the distances from the grating Δ.

$$\tan\vartheta_m = \frac{x_m}{\Delta} \quad \Rightarrow \quad \sin\vartheta_m = \frac{x_m}{\sqrt{x_m^2 + \Delta^2}}$$

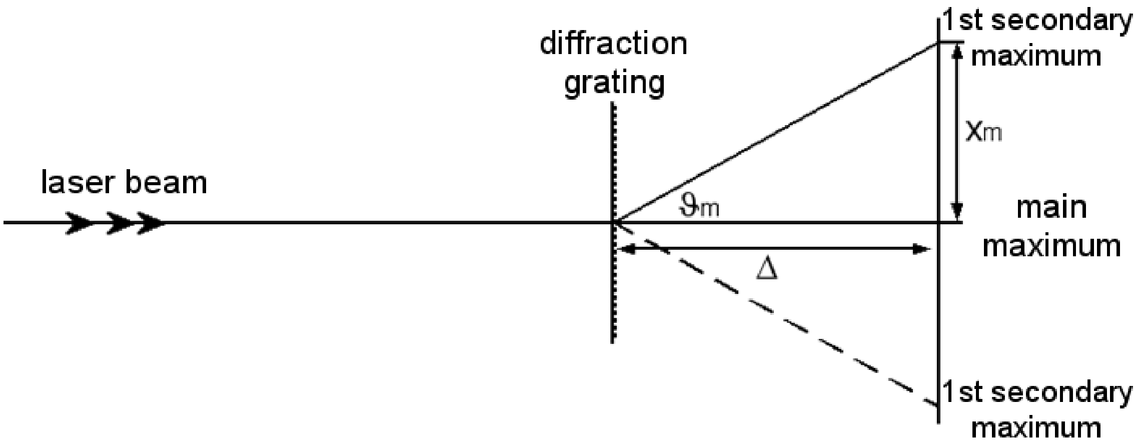


Fig. 10.5 The Fraunhofer diffraction in a grating.

The following figure shows an experiment with a diffraction grating in which the light beam of the laser is divided into major and minor maxima.

In this case, it is the Fraunhofer diffraction that describes the case for small holes and large distances. In practice, this means that we consider almost parallel beams of light.

If we know the grid parameter and the distance from the screen, we can also easily determine the wavelength of the laser radiation.

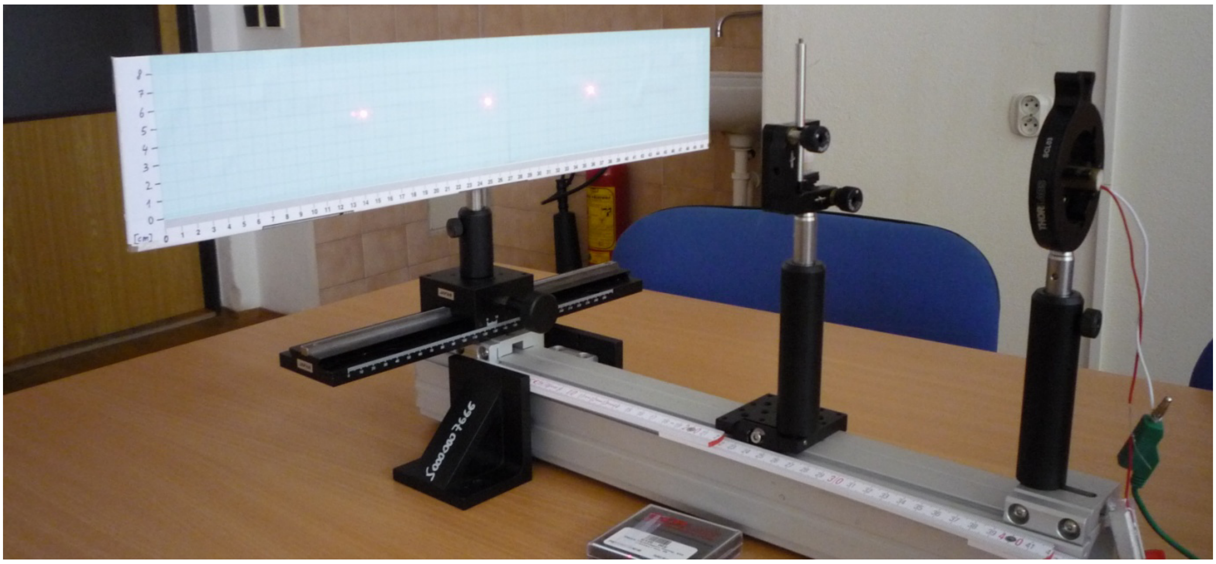


Fig. 10.6 Diffraction grating experiment.

In the picture of the experiment with the diffraction grating, it is not so evident that, apart from the main and secondary maximum, other points for higher order diffractions are formed on the screen. Their intensity is declining rapidly compared to the main maxima.

In order to give you an idea, let´s show another example of the Fraunhofer diffraction. Let´s assume that we have a rectangular hole of small dimensions with a sufficient distance from the light source and the screen. Then the diffraction will take place in the x and y axes and will create a regular pattern on the screen.

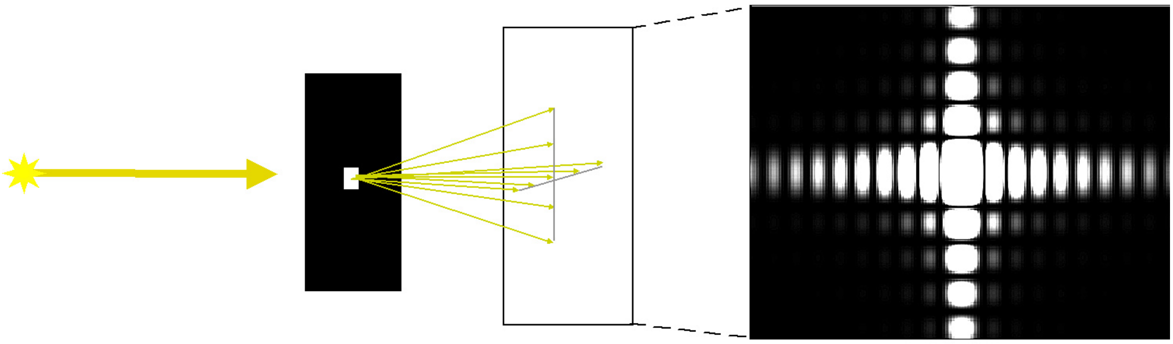


Fig. 10.7 The Fraunhofer diffraction on a rectangular hole.

Note: For illustrative purposes, the intensity of the points on the screen was increased to make the diffraction more observable.

The Fraunhofer diffraction only describes a specific case when we have sufficient distances and a small hole. The pattern on the screen will be sharp enough. The mathematical description of the Fraunhofer diffraction will take the form of the Fourier transformation:

$$A_0(n_x, n_y) = \left(\frac{k}{2\pi}\right)^2 \iint\limits_{-\infty}^{\infty} \psi_0(x_M, y_M) \exp\left[-\mathrm{i}k(n_x x_M + n_y y_M)\right] \,\mathrm{d}x_M \,\mathrm{d}y_M$$

In terms of larger holes and smaller distances, we use the description using the Fresnel diffraction. The Fresnel diffraction also takes into account members of a higher order and thus describes the propagation at higher angles and diverging beams of light (in contrast to the specific case of the Fraunhofer diffraction). In this case, it is a convolution of the wave function $\psi_0(x_M, y_m)$ in the plane of the diffraction screen and propagator that is characterizing the transmission of waves from the screen to the plane of observation.

$$\psi(x, y, z) = -\mathrm{i}\frac{k}{2\pi}\frac{\exp(\mathrm{i}kz)}{z} \iint\limits_{-\infty}^{\infty} \psi_0(x_M, y_M) \exp\left\{\frac{\mathrm{i}k}{2z}\left[(x - x_M)^2 + (y - y_M)^2\right]\right\} \mathrm{d}x_M \,\mathrm{d}y_M$$

The result of the Fresnel diffraction on a rectangular hole will be the so-called Fresnel diffraction pattern.

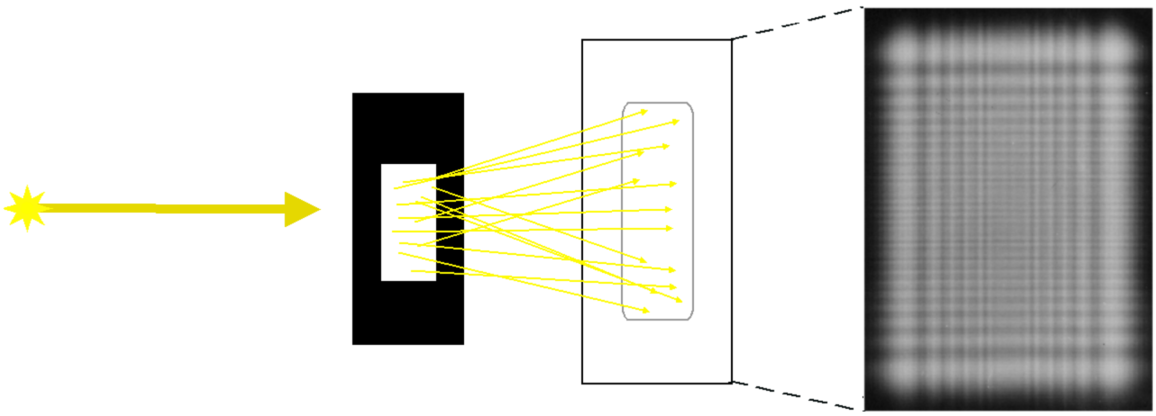


Fig. 10.8 The Fresnel diffraction on a rectangular hole.

Interference of two plane light waves

Imagine two-point sources of monochromatic light of the same frequency. If we place these sources next to each other, the light will propagate in the wavefronts towards the screen and gradually interfere. This creates observable areas with minimum (L) and maximum (H) intensity.

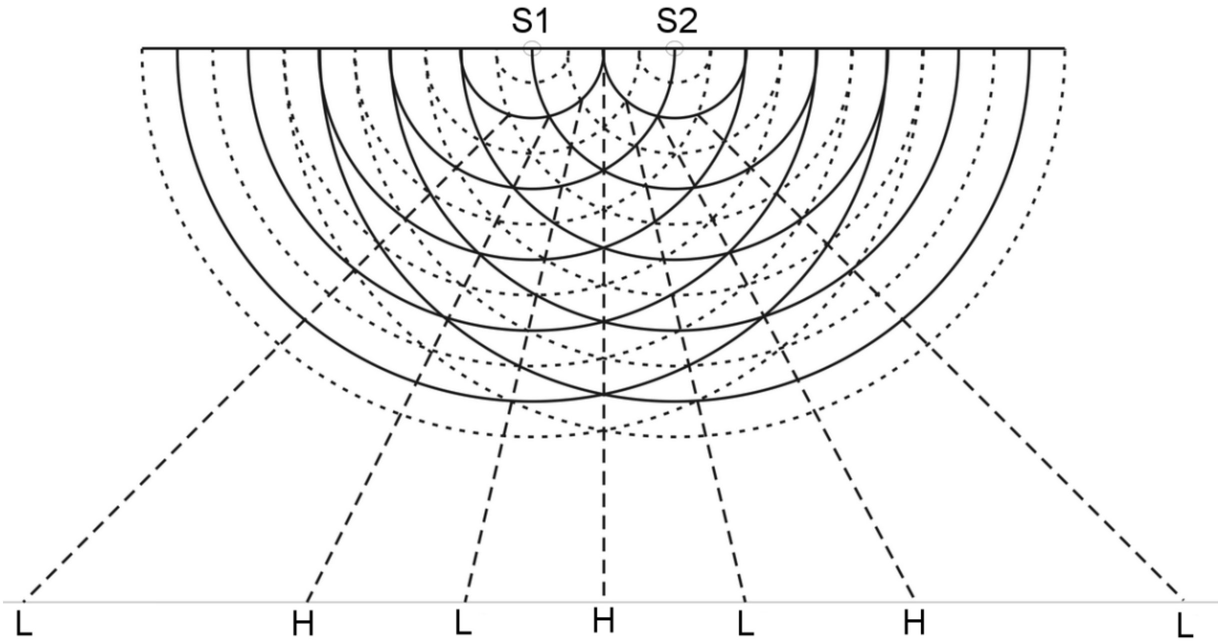


Fig. 10.9 Propagation of 2 plane waves.

The solid lines represent positive maxima of light waves and, on the contrary, the dashed lines represent their minima. If both waves are mutually in the phase (they meet at maxima) it corresponds to the paths along the lines H. On the L paths, where the waves are shifted by 180° (minima and maxima), there is the interference and a decrease in the intensity. Therefore, the light 'bends' to the specific places on the screen.

In order to have a better idea of the diffraction of light, let´s look at the so-called Young´s experiment. It starts at the radiation source which proceeds to an obstacle with one slit and then to an obstacle with two slits.

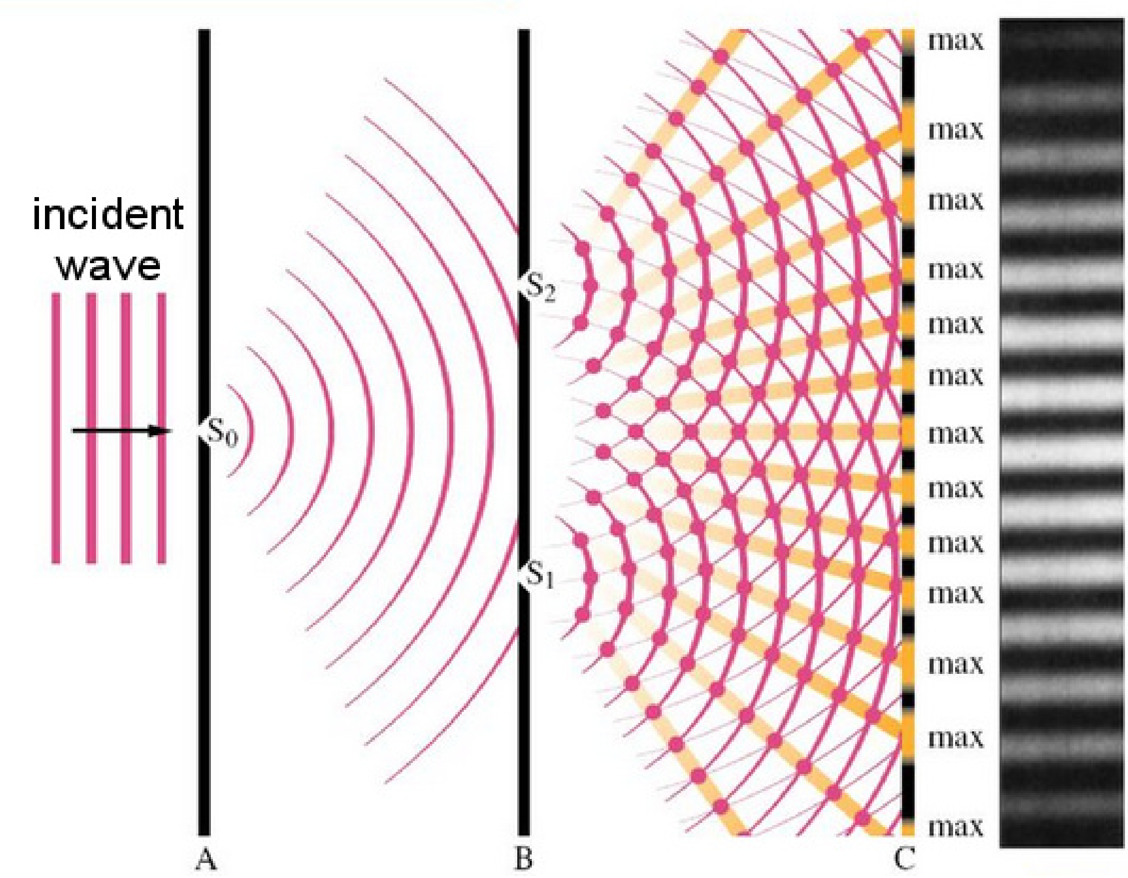


Fig. 10.10 Young´s experiment. [15]

In the Young´s experiment, we observe similar behaviour as in case of the light diffraction. However, it is the consequence of the interference of the light waves that are propagating from the points S_1 and S_2. Geometrically, the light should not fall on the screen at all. Nevertheless, due to the wave nature, after reaching the first slit, the slit itself becomes an elementary source of waving, and after the fall on two second slits, these slits also become a source of waving.

Nevertheless, similar rules apply for individual maxima and minima as in case of diffraction in a grating:

Maxima:

$$d \cdot \sin \vartheta = m \cdot \lambda$$

where m=0, 1, 2, 3, …

Minima:

$$d \cdot \sin \vartheta = (m + 1/2) \cdot \lambda$$

where m=0, 1, 2, 3, …´

Definition of the concepts of interference, diffraction and dispersion

The manifestations of individual mechanisms may look similar, but the way they arise and behave is slightly more specific. So, let´s summarize when we will use the individual terms.

Interference

The concept of interference is used for the specification of the composition of any number of waves with respect to their phase. Therefore, it is the result of the mutual interaction of waves from a finite number of sources.

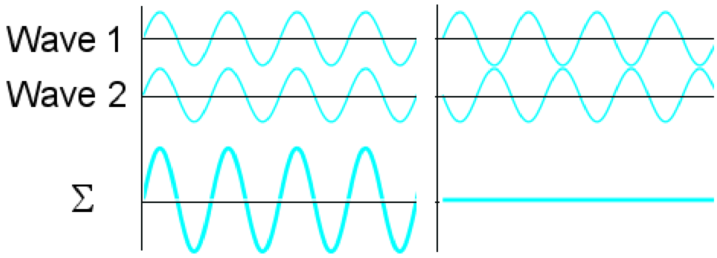


Fig. 10.11 Light interference.

Diffraction

Diffraction (bending of light) describes the interaction of light with a two-dimensional object, alternatively, with a three-dimensional object with a regular structure. The individual parts are larger than the wavelength of the light.

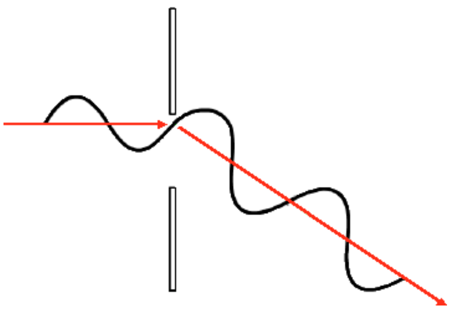


Fig. 10.12 Light diffraction.

Dispersion

This is the most general description that can be used anytime the light waves interact with two/three-dimensional obstacles whose details are smaller than the wavelength.

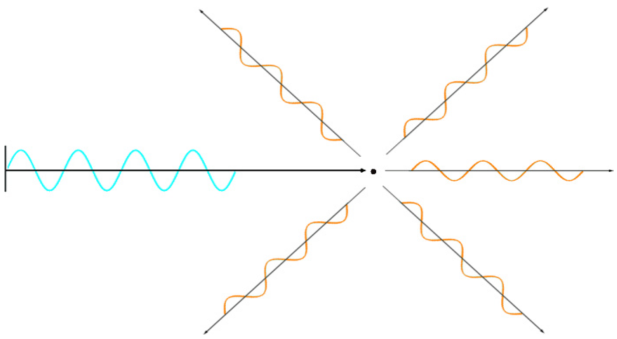


Fig. 10.13 Light dispersion.

Review questions:

After studying this chapter, you should know the answers to the following questions:

1. How does the Huygens principle work?
2. How is it possible that light propagates into the area of geometric shadow?
3. Describe the difference between diffraction and dispersion.

References:

[1] https://cs.wikipedia.org

[2] http://fyzika.jreichl.com/

[3] http://fu.mff.cuni.cz/biomolecules/files/courses/Fotometrie_a_radiometrie.pdf

[4] Silvano Donati, Photodetectors, Devices, Circuits and Applications. ISBN 0-13-020337-8.

[5] https://edu.techmania.cz/cs/encyklopedie/fyzika/kvanta/zareni-cerneho-telesa

[6] [Kitamura R. et al.: Optical constants of silica glass from extreme ultraviolet to far infrared at near room temperature. Appl. Opt. 46, 8118-8133 (2007).]

[7] http://www.fyzika007.cz/

[8] https://www.studopory.vsb.cz/

[9] https://www.levenhuk.cz/clanky/jak-si-vybrat-teleskop/

[10] https://www.wikiskripta.eu

[11] https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=8204

[12] https://khanovaskola.cz/video/7/257/2567-linearni-a-kruhova-polarizace-svetla

[13] http://www.fyzika007.cz/optika

[14] https://www.renishaw.pl/pl/interferometry-explained--7854

[15] http://homel.vsb.cz/~uhl72/VLNOVA_OPTIKA_FII_upravene_2.pdf

[16] https://www.hisour.com/cie-1931-color-space-24840/

[17] datasheet Osram - LA G6SP.01