About the reduction of reflections for camera lenses

How T*-coating made glass invisible

by Dr. Vladan Blahnik and Dr. Benjamin Voelker

March 2016

Content

Introduction

80 years ago the story of T*-optics, which was to provide ways to 'make glass invisible', began at the ZEISS plant in Jena. The invention of anti-reflection (AR) coatings to reduce reflections from glass surfaces that are in direct contact with air represents one of the most outstanding innovations of technical optics in the past century. AR-coatings have considerably improved the performance of optical systems. It is only due to their invention that the construction of complex optical systems with many lens elements, such as high-quality retrofocus or zoom lenses, has become possible.

We will use photographic material to demonstrate how reflection-reducing coatings enhance the quality of photographic images. Moreover, we will discuss the functioning of antireflection coatings as well as the optimization of the reflection behavior of lenses in the process of optical systems development.

Antireflection coatings and image quality

In an article published earlier in this forum (Blahnik 2014), we presented a side-by-side comparison of pictures taken with two variants of the ZEISS Planar 2,8/80mm lens: one was a normal T*-coated lens, the other specifically built without any coating for test purposes. The present work exhibits the results of an analogous comparison for the ZEISS Distagon 2,8/21 ZE lens, which has 26 air-glass surfaces.

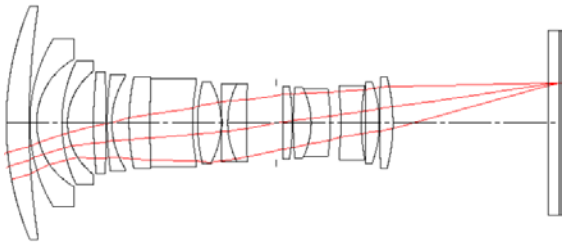


Fig. 1.1: Lens cross section of Distagon 2,8/21 ZE: This lens has 29 refractive surfaces, 26 of them in contact with air.

When looking at such a pair of lenses exposed to lighting, one observes strong reflections (having the same color as the light source, e.g., matt white for a white source) emanating from the uncoated lens. In contrast, the T*-coated lens shows only weak reflections of a slightly purple or green color (see Fig. 1.2). We are going to discuss the colors of reflections later on in the section "The road to T*-optics."

Fig. 1.2: Distagon 2,8/21 ZE, uncoated (left) and T*-coated (right). (Image by courtesy of Andreas Bogenschütz).

Fig. 1.3: Photo taken with Distagon 2,8/21 ZE (left) and with an uncoated demonstration lens (right).

The point spread function (PSF) of a lens

Fig. 1.3 shows a pair of photos taken, respectively, with a Distagon 2,8/21 ZE lens (left) and an uncoated demonstration lens of the same type (right). In case of the uncoated lens, there are lens elements that are strongly reflecting and those elements cause bright ghost images of the candle; the reflected light scatters across a large part of the photo and, as a result, the picture is unusable. By contrast, the reflections from the T*-coated lens are so much weaker that the image information remains nearly completely preserved, even in the darker regions of the photo.

The flame of the candle is very small, nearly "point-like", as compared to the entire field of view. The image of such a light source can therefore be identified, to a reasonable degree of approximation, with the point-spread function (PSF) of the lens for this particular object point.

The number of elements in such a chain-like PSF of a light source correlates with the number of optical elements in the camera lens, it is $(n^2+n)/2$, wherein n denotes the number of refractive surfaces. The exact position, shape, color, and intensity of these elements, i.e. the individual ghost images produced at the various optical lens surfaces, are individually dependent on the optical design of the lens. Figures 1.4 and 1.5 show the optical paths of single reflections as created by reflection from a certain pair of reflecting surfaces.

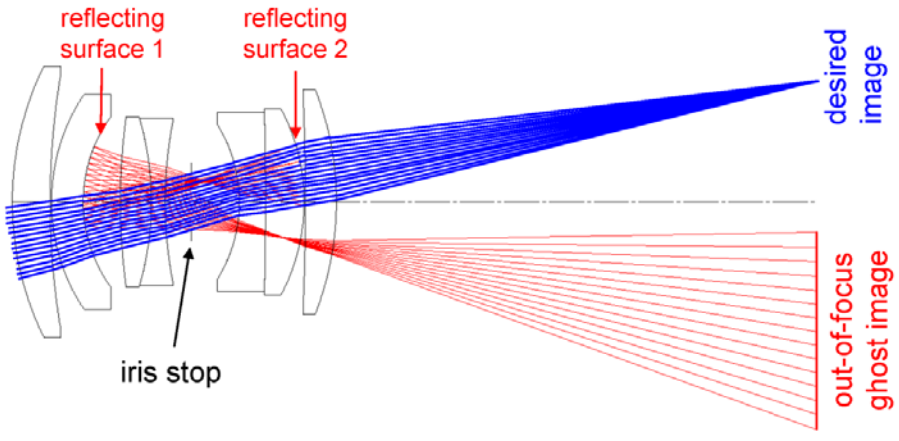


Fig. 1.4: Optical path for an out-of-focus ghost image, shaped like an iris stop (see Fig. 1.6).

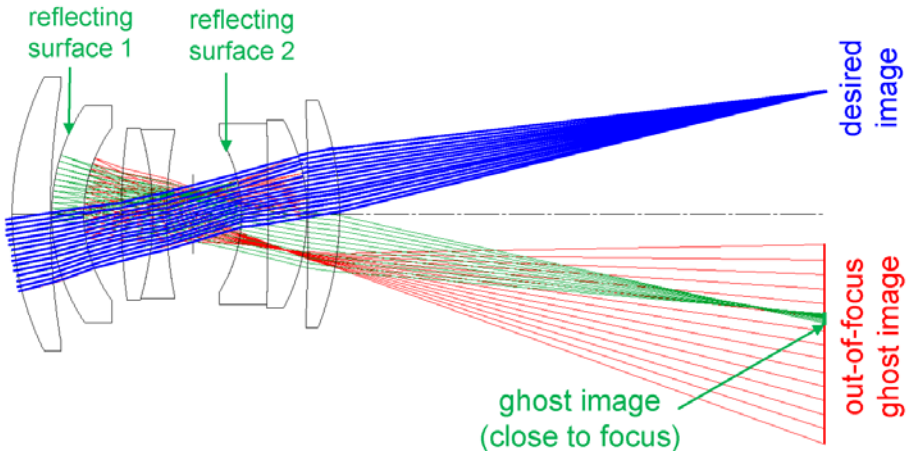


Fig. 1.5: Optical path for a close-to-focus ghost image.

Fig. 1.6 shows these reflections in a real photo taken with a Zeiss 2,8/80mm lens. The close-to-focus ghost image sketched in Fig. 1.5 can be seen as bright, small spot, whereas the out-of-focus ghost image, which is due to the light ray path depicted in Fig. 1.4, stretches out over a relatively large area. The latter has the shape of an iris stop (the iris diaphragm consisting here of five blades), and exhibits much lower intensity than the in-focus reflection.

For a rotationally symmetric camera lens, i.e. one the elements of which are centered at the optical axis, the flare chain extends along a line from the light source through the center of field (Fig. 1.7). The shape of each individual ghost image is symmetrical with respect to this axis. This is because light rays starting from a given source point proceed symmetrically with respect to the so-called tangential plane. The tangential plane is the plane spanned by the optical axis and the chief ray (i.e. the ray going through the center of the stop). This symmetry holds for both refraction and reflection of light at any lens surface.

Fig. 1.6: Ghost images of a real photo taken with an uncoated demonstration lens 2,8/80mm with aperture f/8. In the bottom photo, the ghost images are marked green and red, respectively, corresponding to their ghost ray paths depicted in Fig. 1.4 and 1.5.

Fig. 1.7: Symmetry axis of the point spread function for a rotationally symmetric lens.

A further consequence of the rotational symmetry of a photographic lens is that the point spread function remains invariant when the light source is rotated about the symmetry axis. If the distance of the light source from the center changes, then the lens flare appearance will also change: every single reflection from the different optical surfaces will generally hit at a different point in the image plane and then change its shape (see Fig. 1.8).

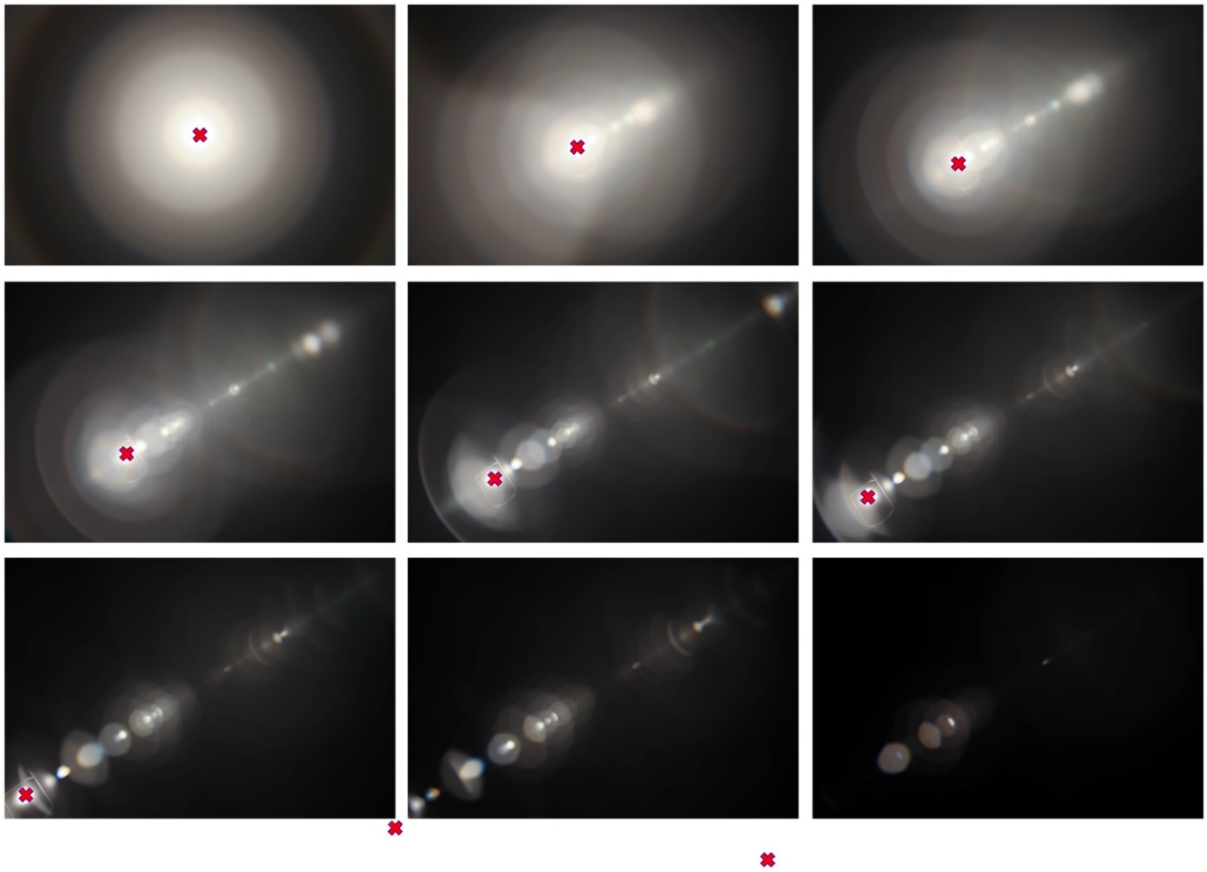


Fig. 1.8: Ghost images for various positions and distances of the light source (marked by a red cross) relative to the center of field.

The ghost image PSF (point spread function) is dependent on aperture size: The number and position of the individual reflections remain constant when the aperture size is decreased (stopping down). However, their size, shape and distribution change individually. With a large aperture, the intensity distribution is often rather inhomogeneous, due to strong aberrations along the ghost ray paths. With large apertures, local light concentrations, so-called caustics can occur. Such light concentrations may have, for instance, the form of a crescent (as in Fig. 1.7 close to the right-hand image border).

When one stops down a lens, the aforementioned intensity distributions become more homogenous and reflections that are strongly out of focus do take on the polygonal shape of the iris stop.

Moreover, it may happen that the rays producing a particular ghost image all pass a small area near the center of the stop before reaching the image plane. The intensity of such a ghost in the image plane will then retain its magnitude even when the lens is stopped down. However, the intensity of the actual image decreases as a result of the smaller aperture (see Fig. 1.10). This means that the ghost intensity increases relative to the useful light intensity. The reflection intensity therefore often increases when stopping down, relative to the intensity of the imaged light source.

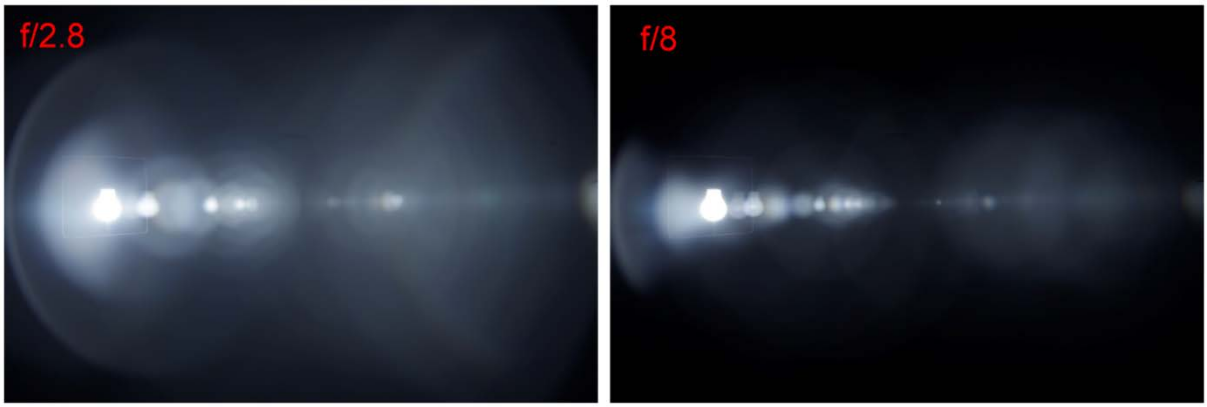


Fig. 1.9: Ghost images with different f-numbers: f/2.8 and f/8 (uncoated demonstration lens Distagon 2,8/21 ZE).

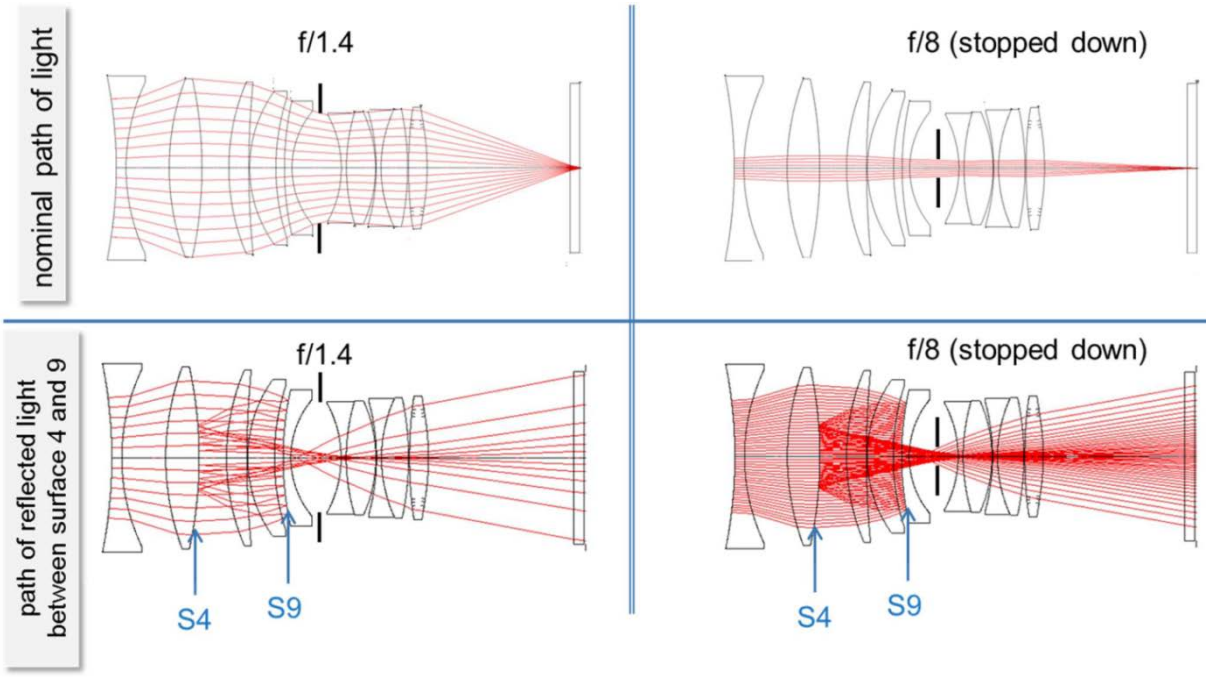


Fig. 1.10: In case ghost ray paths mainly pass through a small portion of the aperture stop, the intensity of the reflected light does not change when the lens is stopped down. The portion of the useful light (cf. top-left and top-right diagrams), however, decreases in proportion to the square of the f-number. As a result, the relative intensity of flare (i.e. of reflected) light increases (c.f. the lower row of diagrams). For the sake of demonstration, we have used the same number of rays in the useful optical path (top) and kept the ray density for the individual ghost ray path.

In addition, the reflection characteristics depend on the object distance. Under variation of object distance the light path through the lens changes. And for zoom lenses the choice of focal length, naturally, has a bearing on the reflection characteristics. Therefore, in our design and development stages, we are testing lenses with different light source positions, f-numbers, object distances and focal lengths (cf. chapter "Simulation and optimization of ghost images in camera lenses" in this article).

In some cases, there might be off-axis ghost images, e.g. when bright light hits dust particles or grease spots and is scattered into the image plane. Fig. 1.11 shows a picture in which several flare light effects with different physical causes may be discerned.

Fig. 1.11: Photo with veiling glare due to different physical causes: 1. Ghost images due to reflections at lens surfaces (lens flare through light source and center of field), 2. Light scattered by dust particles or otherwise (outside the optical symmetry axis), 3. Veiling-glare background. With stopped-down iris stop, the (strongly oversaturated) light source appears "star-shaped" (with a waviness corresponding to the number of blades of the variable iris diaphragm). (Photo by Dr. Hubert Nasse.)

Another type of ghost images that is not rotationally symmetric may be caused by reflections at microlenses located directly above the image sensor (see Fig. 1.12). The purpose of such microlenses is to enhance the efficiency of the image sensors by guiding the light through a substrate area of only a few micrometers thickness, passing conductive structures of integrated circuits, to the photoelectrical layer. These ghost images mostly appear in a regular grid-type arrangement around the light source, in accordance with the regular periodic arrangement of the image pixels. The farther the light source lies from the center of the field of view, the more distorted those ghost images will appear, as a result of the oblique incidence onto the microlenses (cf. Fig. 1.12).

Fig. 1.12: Ghost images of the sun through the microlens grid directly above the pixels of the image sensor.

The appearance of the PSF strongly depends on the intensity of the light source. The relative brightness of the ghost images as compared to the light source and to other (possibly occluded) subjects in the picture is the decisive factor for quality losses of pictures with a high dynamic range or high contrast range.

Fig. 1.13 shows the ghost images of an incandescent lamp (light bulb) used as light source. The light bulb was photographed in the first picture (top left) in such a way that the brightness level range was almost fully exploited, without overexposing any part of the image. When taking this sequence of pictures, the exposure time was successively increased by a factor of 4 (i.e. by 2 EV). "EV" is short for "exposure value", often also called "f-stop". "1 EV" corresponds to a factor of 2. This means, in the pair of images labelled "+ 2 EV", the light source is overexposed by a factor of 4. However, here and in subsequent images, due to saturation of the sensor, it is invariably rendered at the maximum brightness level (e.g. 255 for jpg format with 8 bits), while the ghost images become increasingly visible.

One observes in this image sequence that the T*-coated lens produces discernible reflections only at very high levels of over-exposure of the light source. For example, when the light source is overexposed by a factor of 2^{12}=4096 (i.e. with an exposure time of +12EV), no reflections are visible outside the light source spot. In contrast, the uncoated lens produces bright ghost images that extend over the entire image.

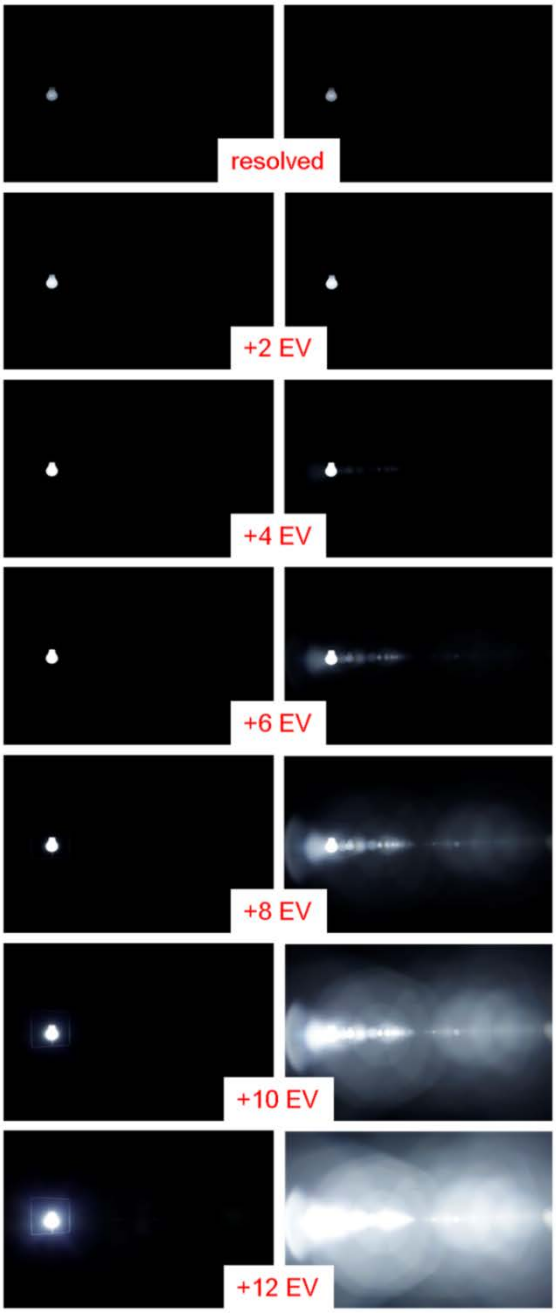


Fig. 1.13: Exposure series of the Distagon 2,8/21 ZE in comparison with an uncoated demonstration lens 2,8/21 both with aperture f/8: Initially, the light bulb is resolved in the brightness range (top left), in subsequent images, exposure times are successively incremented by +2 EV.

The information from this sequence of images can be combined into one single representation of the point-spread function that covers the whole brightness range. Fig. 1.14 shows this PSF on a logarithmic scale: the units on the left axis are f-stops or "exposure values" (EV, i.e. base-2 logarithm), on the right-hand side scale we put relative intensity (base-10 logarithm). For the Distagon 2,8/21 ZE lens, the difference between coated and uncoated lens is roughly eight f-stops or 2.5 orders of magnitude. We will discuss the general behavior of flare light intensity for different lenses in the last chapter of this article.

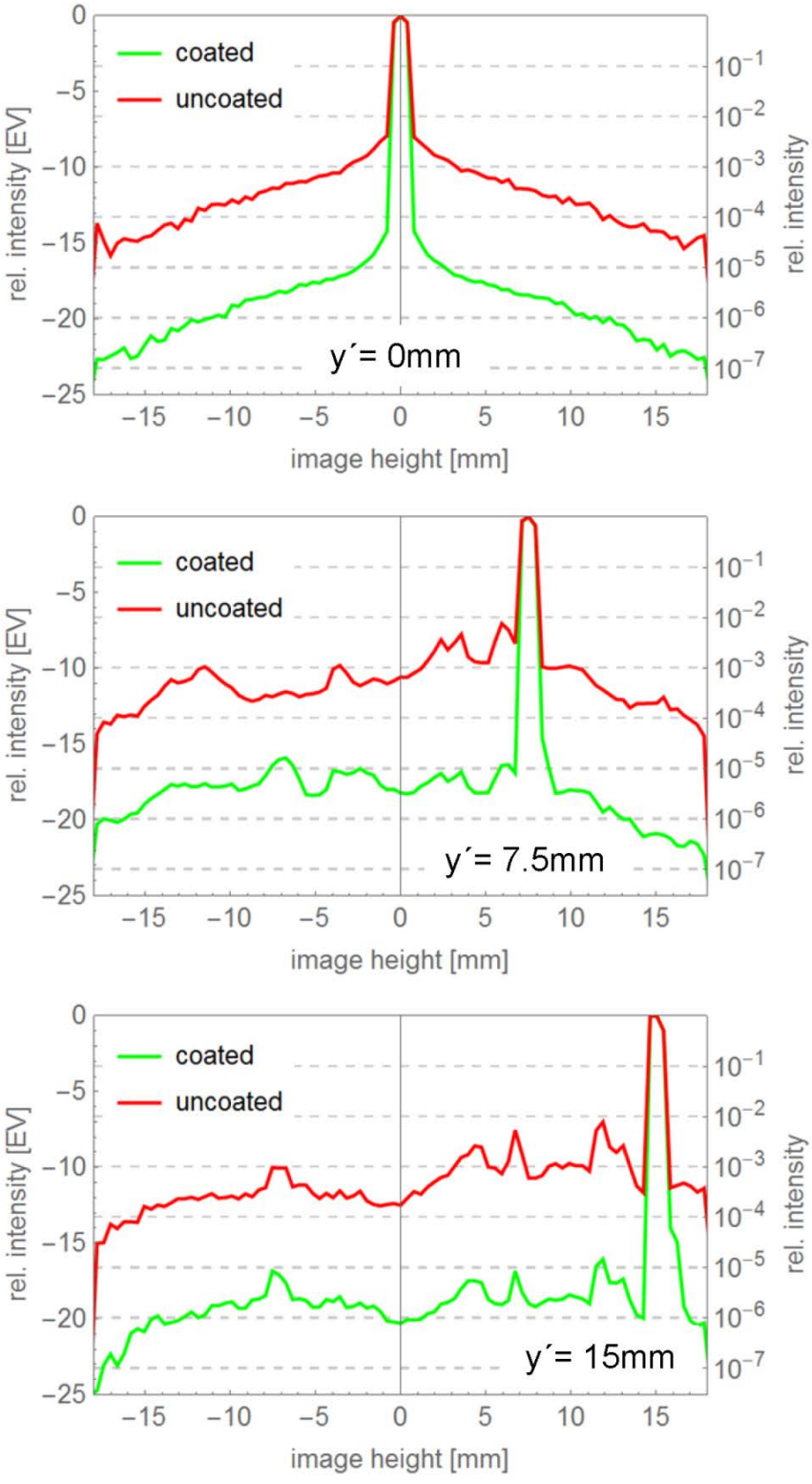


Fig. 1.14: Point spread function (PSF) along the axis through the center of field and the light source of the Distagon 2,8/21 ZE lens in comparison to the uncoated demonstration lens. The left scale shows the exposure values (EV), the right scale directly shows the relative intensity. The light source in these three graphs lies at an image height of 0, 7.5, and 15 mm, respectively.

Comparison of images - coated vs. uncoated lenses

The influence of flare light and ghost images on image picture quality strongly depends on the image motif. This is shown by the series of image pairs we have compiled in Fig. 1.15. Each pair serves as a side-by-side comparison of an image taken with the Distagon 2,8/21 ZE lens and a corresponding image taken with the uncoated variant of the same lens. With diffuse lighting and small brightness differences, lens flare manifests itself predominantly as 'haze', reducing contrast and making for fainter color rendering (Figures 1.15 c, d, g, h, k, m, o). This means a reduction of the "macro-contrast" (see Nasse (2008)) over large parts of the image. However, when bright local light sources are present, such as the sun or street lighting or spotlights at night-time, lens flare will produce ghost images that stretch across large areas of the image in a disturbing way (Figures 1.15 a, b, e, f, i, j, r, s, v).

a)

b)

c)

d)

e)

f)

g)

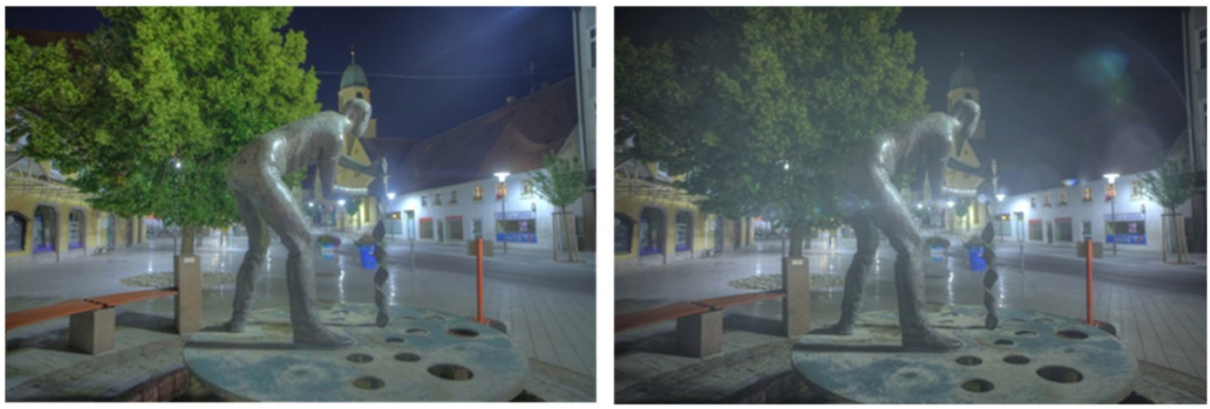
h)

i)

j)

k)

l)



m)

n)

o)

p)

q)

r)

s)

t)

u)

v)

Fig. 1.15: Comparison of images of the Distagon 2,8/21 ZE lens (left) with those taken with the related uncoated demonstration lens.

Contrast range of a scene and dynamic range of image taking and reproduction

In photography, there are often situations in which bright light sources are present in the field of view or nearby. Such light sources are often brighter than other image contents by several orders of magnitude. Typical examples are outdoor photos with the sun, bright street lighting or automobile headlights in images taken at night-time. And in case of indoor photos, the daylight passing through windows may be much brighter than any object inside the room.

Whether and to what extent flare light is disturbing a given image, depends on the dynamic range of the photographed scene. Steinbach (2015) has determined typical values for the dynamic range for a number of motifs (see Fig. 1.16). As long as there are no very bright light sources (like the sun) or spotlights in the field of view or its immediate vicinity, the dynamic range of the scene rarely lies above 14 exposure values (EV). However, with strong light sources, the dynamic range may lie significantly above 20 EVs.

In the literature, you can find typical luminance values of different natural and artificial light sources and surroundings (e.g., Grant (2011), Palmer, Grant (2010), Jones, Condit (1941)). Some of them are shown in Fig. 1.17.

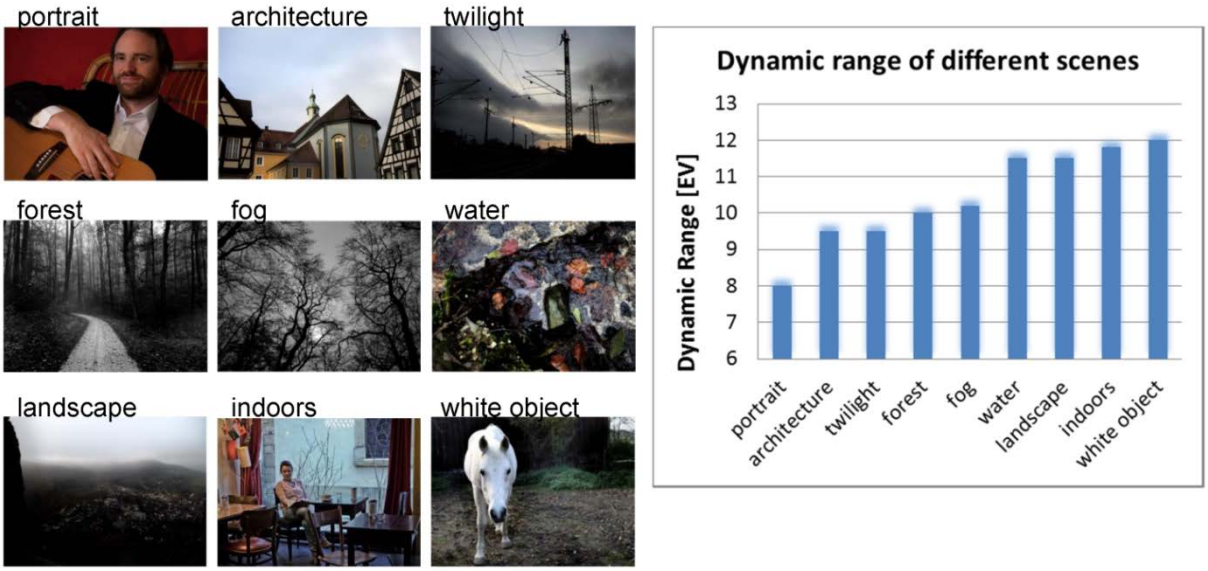


Fig. 1.16: Dynamic range of different image motifs (from Steinbach (2015)).

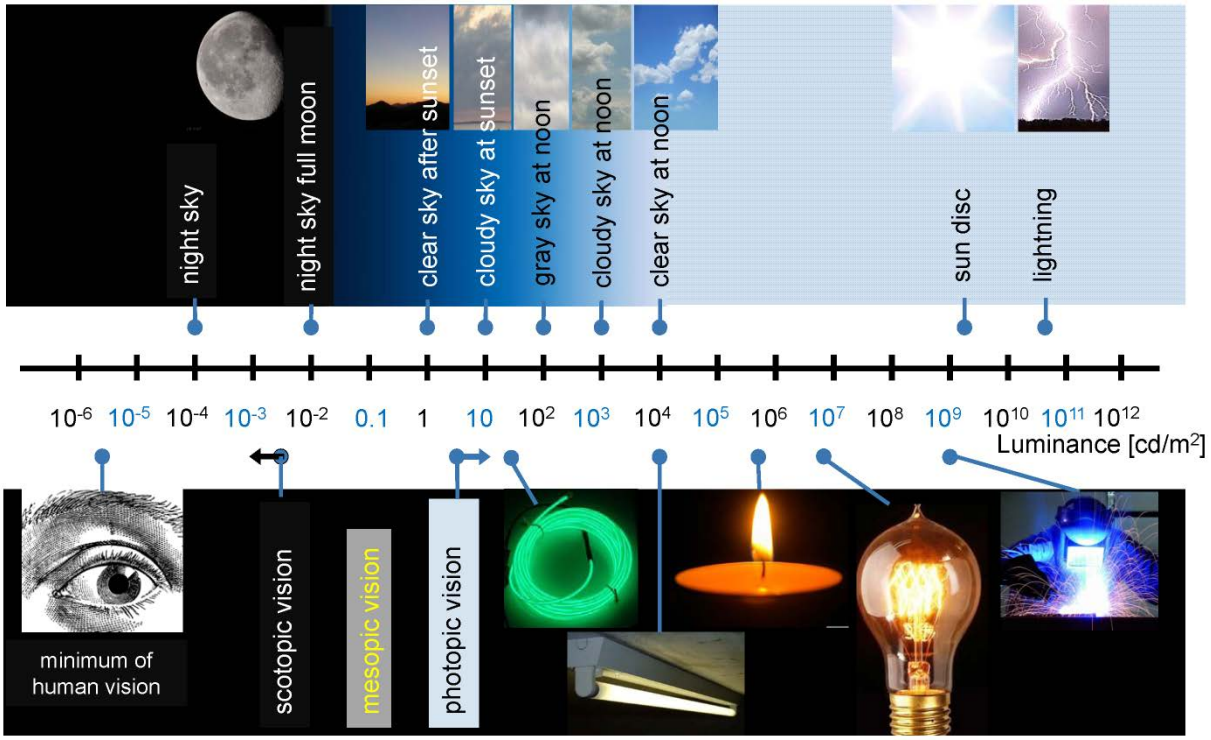


Fig. 1.17: Typical luminance values of different natural scenes and light sources (data from Grant (2011)).

The relation between the physical intensity of the light (horizontal axis) received by the image sensor and the output digital numeric value (vertical axis) is described by the so-called opto-electronic conversion function (OECF). The OECF is often simply designated as 'response curve' or 'characteristic curve.' The exact term for "intensity" in the image plane is "irradiance" as used in radiometry, which is the radiation power per area, measured in Watts per square meter (W/m^2). For visible light, the CIE-standardized luminous efficiency function of the human eye at daylight is used to define the corresponding photometric variable called "luminance", which is measured in Candelas per square meter (cd/m^2).

To represent the sensor signal value D on the vertical axis of the OECF graph, either a normalized scale is used ($0 \leq D \leq 1$), or directly a non-normalized scale ranging to the maximum signal value. Usually, these signal values are expressed as binary numbers. For instance, in the jpg-format used for image compression, 8 bits, i.e. $2^8 = 256$ values per RGB color channel are available. The corresponding figures for raw image formats are significantly higher, for example 14 bits, i.e. $2^{14} = 16384$ values per RGB color channel.

For the medium brightness range, there is often an approximately logarithmic relationship between signal value D and the luminance L, which is $D \sim \log L$, so that the luminance is often represented on a logarithmic scale. Fig. 1.18 shows an example of an OECF of a digital camera.

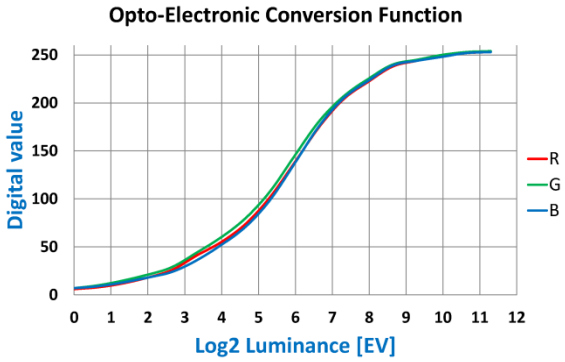


Fig. 1.18: Typical characteristic curve or OECF (opto-electronic conversion function) of an image sensor. The luminance in the image plane, represented in logarithmic units on the horizontal axis (here, logarithm to the base of 2, i.e. exposure values), is transformed (or "converted"), after image taking, to the digital output values on the vertical axis.

For the analysis of dynamic ranges, i.e. of the maximum variation of the luminance in a scene, only relative differences are relevant. Therefore, the physical unit cd/m^2 is mostly omitted, and relative scales or dimensionless quantities are used instead. There is no standard convention in literature and practice as regards the quantification of dynamic range. The following measures are used ("r " denotes the ratio of the maximum luminance and minimum luminance):

- the „contrast ratio" r : 1
- the number of „exposure values", „bits", „f-stops"; dynamic range is expressed in the form „$\log_2$(r) EV"
- the „order of magnitude" $\log_{10}$(r)
- the number of „decibels" (dB); dynamic range is then expressed in the form „$10\cdot\log_{10}$(r) dB"

(Remark: Unfortunately, in the literature there are different conventions for "decibels" involving either $10\cdot\log_{10}$ or $20\cdot\log_{10}$. The reason for these diverging conventions may be an (unsuitable) transfer of definitions from acoustics or electronics. There logarithmic scales are used, on the one hand, for sound-level amplitudes or voltages, and on the other hand for the corresponding energy or power quantities (which are proportional to the square of amplitude or voltage, respectively). By the laws of logarithms, we obtain $10\cdot\log_{10}(x^2) = 20\cdot\log_{10}(x)$. The latter we find the expression "$20\cdot\log_{10}(x)$" in the sensor technology literature (see, e.g., Darmont (2012)), p. 4 ff. and the sources cited therein): For example, there 100 dB corresponds to a luminance difference of 5 orders of magnitude. However, the expression $20\cdot\log_{10}(x)$ is not suitable for photometric quantities that have the dimension of energy or power.)

A practical mnemonic rule for converting orders of magnitude ($\log_{10}$) into f-stops ($\log_2$) is as follows:

Order of Magnitude ≈ 0.3 · (# of f-stops)

or

$3\,\frac{1}{3}\cdot$ (Order of Magnitude) ≈ (# of f-stops).

The exact conversion factor is $\log_2(x)/\log_{10}(x) = 3.3219\ldots$ For example, 3 orders of magnitude correspond to 10 f-stops, approximately.

Tables 1.19 a) and 1.19 b) summarize the conversions with these relative units.

Exposure values	Orders of magnitude	Ratio
$EV = \log_2(r)$	$M = \log_{10}(r)$	R = r : 1
0.0	0	1.00
1.7	0.5	3.16
3.3	1	10
5.0	1.5	31.6
6.6	2	100
8.3	2.5	316
10.0	3	1000
11.6	3.5	3162
13.3	4	10000
14.9	4.5	31623
16.6	5	100000
18.3	5.5	316228
19.9	6	1000000
21.6	6.5	3162278
23.3	7	10000000

Fig. 1.19a: Conversion of measures for the dynamic range (r is short for "range"); starting from order of magnitude M or ratio 10^M:1.

Exposure values	Orders of magnitude	Ratio
$EV = \log_2(r)$	$M = \log_{10}(r)$	R = r : 1
0	0.00	1
1	0.30	2
2	0.60	4
3	0.90	8
4	1.20	16
5	1.51	32
6	1.81	64
7	2.11	128
8	2.41	256
9	2.71	512
10	3.01	1024
11	3.31	2048
12	3.61	4096
13	3.91	8192
14	4.21	16384
15	4.52	32768
16	4.82	65536
17	5.12	131072
18	5.42	262144
19	5.72	524288
20	6.02	1048576

Fig. 1.19b: Like 1.19a, starting from EV, i.e. doubling per step.

Which luminance range can still be imaged with the available brightness values will now be of interest. This ratio of the highest to the lowest image-able luminance value is the dynamic range. For the OECF of a digital camera, the dynamic range is defined in ISO 14524.

The dynamic range, i.e. the range of distinguishable luminance values for image taking ranges from about 10 to 14 EV for today's digital consumer cameras, with low ISO sensitivity (i.e. ISO 100), and decreases towards higher ISO values, typically with 0.6-1.1 EV per ISO step. For example, the dynamic range with a high ISO sensitivity of 6400 mostly ranges between 6 and 9 EV. The website of DxOMark provides a comprehensive database of many cameras with measured data of the dynamic range. Fig. 1.20 shows a development of the dynamic range of different cameras over the past years, according to measurements by DxOMark. So the dynamic range of image sensors is constantly increasing. Moreover, quick sequences of exposure times are realized with an increasingly improving quality; they can be subsequently processed in HDR image processing programs and partly also to HDR images within the camera itself. Meanwhile, many digital cameras, including mobile phone cameras, calculate HDR images from an automatic sequence of exposure times.

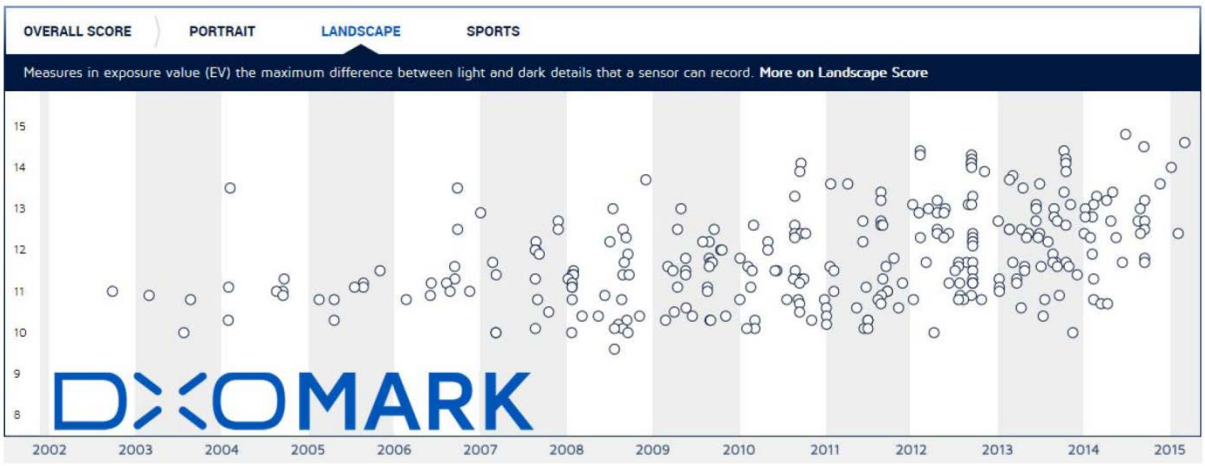


Fig. 1.20: Development of the dynamic range of digital cameras over the past years (measurements by DxOMark (http://www.dxomark.com)); with kind permission of DxOMark.

The extension of the dynamic range by combination of several images taken with different exposure times is not new. Such approaches in photography were made as early as in 1850. Gustave Le Gray created one image out of two negative images with different exposure times, showing the sea and at the same time the much brighter cloudy sky. Such an image (see Fig. 1.21) could not have been realized at that time with one single exposure, because of the limitations of the film material back then.

Fig. 1.21: Extension of the dynamic range by combining two images for the simultaneous representation of the cloudy sky and the sea (taken with a much longer exposure time) (Gustave Le Gray, 1850).

The simple image structure with a subdivision of the image along the horizon simplifies image composition. For more complex image compositions, the stitching of different image details from different brightness ranges is extremely cumbersome. For example, it took Ansel Adams several painstaking days in the darkroom to expose negative images locally with different exposure times (dodging and burning) in order to obtain as much brightness information as possible on the positive image. Adams' impressive landscape images produced in this way can be found, among others, in his photo book "The Print" published in 1950, where he exactly describes this method of local tone mapping.

In digital photography, methods were developed that replaced this manual local tone mapping by global processes (Mann, Picard (1995), Debevec, Malik (1997)). In HDR (high dynamic range) image formats, to each pixel a number (or vector in case of RGB) is assigned, which is taken from a very large set of digital values (e.g. 2^{16} = 65536 tonal values for 16 bits). Tone mapping reduces this large digital value range to a much smaller set of values (i.e. 2^8 = 256 for 8 bits). In this way, the information of existing compressed standard formats like "jpg" on image

playback media like computer screens or paper printouts, is typically displayed with a similar number of tonal values (apart from still expensive HDR displays). The dynamic range of image display media is more and more increasing, similar to the increasing dynamic range of image taking with digital cameras (see Fig. 1.22).

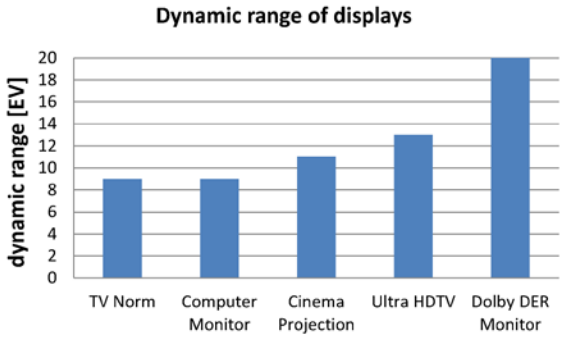


Fig. 1.22: Dynamic range of different image display media.

For viewing the image on a monitor, the signal value is translated into a radiometric light intensity. The characteristic curve of the monitor, i.e. the dependence of the radiant exitance from the signal value, is approximately a logarithmic correlation again. If the monitor is not capable to distinguish the signal value range directly, the signal values of the camera have to be rescaled to a smaller value range, which means that information and thus image quality will get lost. Details on transfer functions of digital image formation and reproduction and their adaptation to human image perception can be found in the book of Reinhard, Ward, Pattanaik, Debevec, Heidrich, Myszkowski (2010). The quality of the algorithms of this tone mapping determines the quality of the HDR image converted to the brightness range shown.

Influence of flare light on the dynamic range of a scene

Using the following example, we are going to examine in more detail how the brightness distribution within an image changes in the presence of flare light: If you take a photo on a sunny day out of a room through a window, the room will appear nearly completely dark, if the outdoor area is correctly exposed within the dynamic range of the image sensor. When the exposure time is increased, so that the brightness values of the interior are well resolved, the outside area is overexposed, i.e. it appears as white area. Fig. 1.23 shows comparison images, taken with the Distagon 2,8/21 ZE lens and the uncoated demonstration lens. It can be seen that there are only slight quality differences with regard to the representation of the outdoor area, despite the considerable flare light difference of both lenses. However, with the longer exposure time for the interior, large part of the image information, above all directly below the window, is nearly totally covered by flare light caused by the outdoor area. Whereas the interior photographed with the T*-coated lens remains well visible.

In this example, outdoor area and interior cannot be imaged simultaneously, which is due to the limited dynamic range of the image sensor. Flare light from the lens reduces the image quality in the dark areas of the image.

When an HDR image is calculated from the images of a sequence of exposure times (Fig. 1.24), the image information otherwise covered by flare light remains visible in the darker image areas (Fig. 1.25).

Fig. 1.23: Images out of a room taken on a sunny day to the outside, using the Distagon 2,8/21 ZE lens (left) and the uncoated demonstration lens (right) with a stop of f/8. Owing to the limited dynamic range of the image sensor, interior and outdoor area cannot be imaged simultaneously. In case of the uncoated lens with long exposure, much flare light intrudes from the overexposed area (window) and overexposes large interior areas.

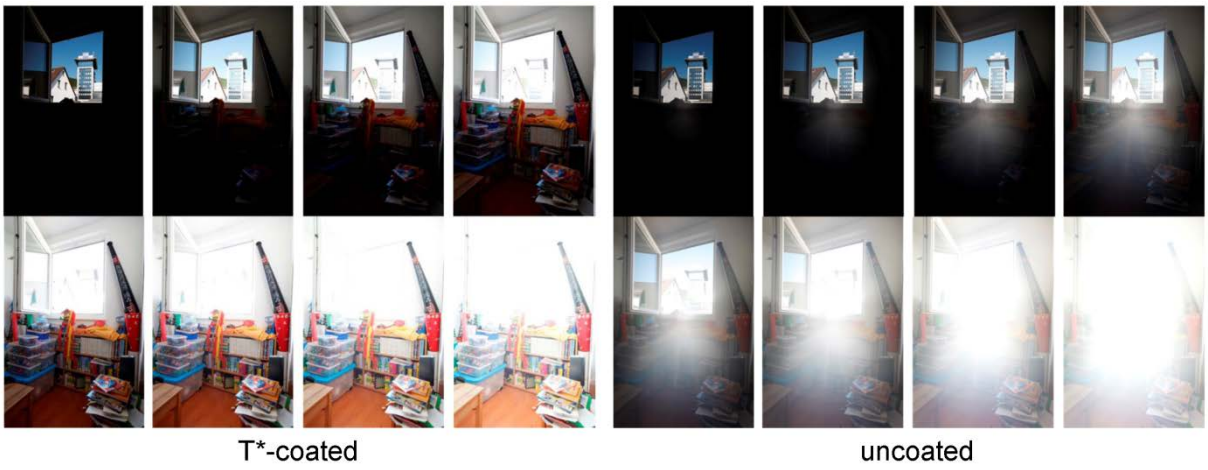


Fig. 1.24: Sequence of exposure times of the same scene as in Fig. 1.23.

Fig. 1.25: HDR images calculated from the sequences of exposure times. In case of the uncoated lens, the image quality decreases dramatically.

The image quality losses strongly depend on the distribution of brightness in the image and also on the point spread function of the lens over the image field. This is shown using the same example image, by considering the brightness differences in different areas of the image. These local regions are shown in Fig. 1.26.

For each of these regions in the image, we write down the RGB values (red-green-blue values) as captured by the sensor; this is done for each image of the exposure sequence shown in Fig. 1.24. Since we are only interested in gray levels, we convert this RGB value into a gray-scale value. For each of these examined points we obtain the sensor characteristic (OECF) approximately. There are small deviations from the exact curves, because straylight from other image regions increases effective exposure in a non-linear manner. Characteristic curves taken at different image positions are shifted horizontally with respect to each other (see Fig. 1.27). For any two image positions the difference in (physical) irradiance that would result in the same gray-level can be determined from the respective shift value.

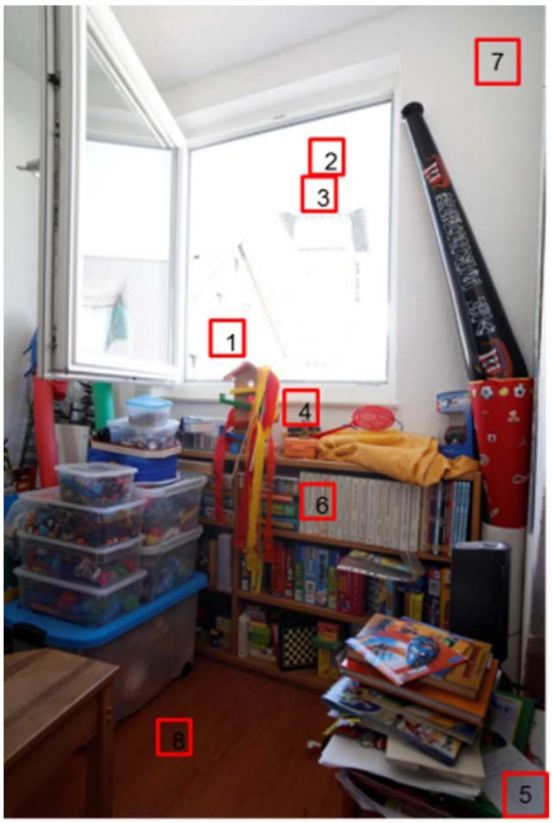


Fig. 1.26: Local areas within the image for analyzing relative brightness values.

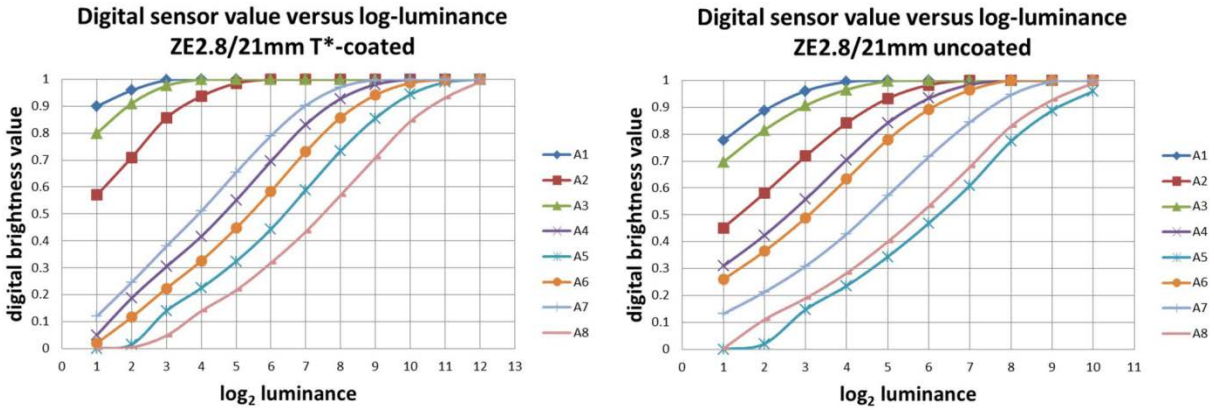


Fig. 1.27: Characteristic curves obtained from the images of the sequence of exposure times for 8 areas, i.e. A1, ..., A8 in the image field (see Fig. 1.26): using the Distagon 2,8/21 ZE lens (left) and the uncoated demonstration lens (right).

From Fig. 1.27, the differences in irradiance can be determined for the different image regions, by reading the difference of the characteristic (in x direction) for a certain brightness value. These differences are shown in Fig. 1.28.

It can be seen that in the image taken with the uncoated lens, the characteristics for the interior points (A4-A8) lie much closer to the characteristics for the outdoor area (A1-A3) than is case in the image taken with the T*-coated lens. This means that the brightness differences shown in the image, i.e. the dynamic range, are smaller when using the uncoated lens. Furthermore we note an unequal dependence of brightness differences on position within the room. For example, the difference at A7 (top right edge of the image) compared to the outdoor light area A1-A3 is similarly high in both images, whereas the difference at A6 (below the window) is clearly different in both images. This is due to the form of the PSF (see Fig. 1.8), according to which the flare light preferably distributes to the opposite side of the image, which is why more flare light gets to A6 than to A7.

We have established that the intensity of the flare light strongly depends on the size and position of bright areas, particularly in dark image areas. An interesting analysis based on test charts with very different brightness distributions can be found at McCann, Rizzi (2012).

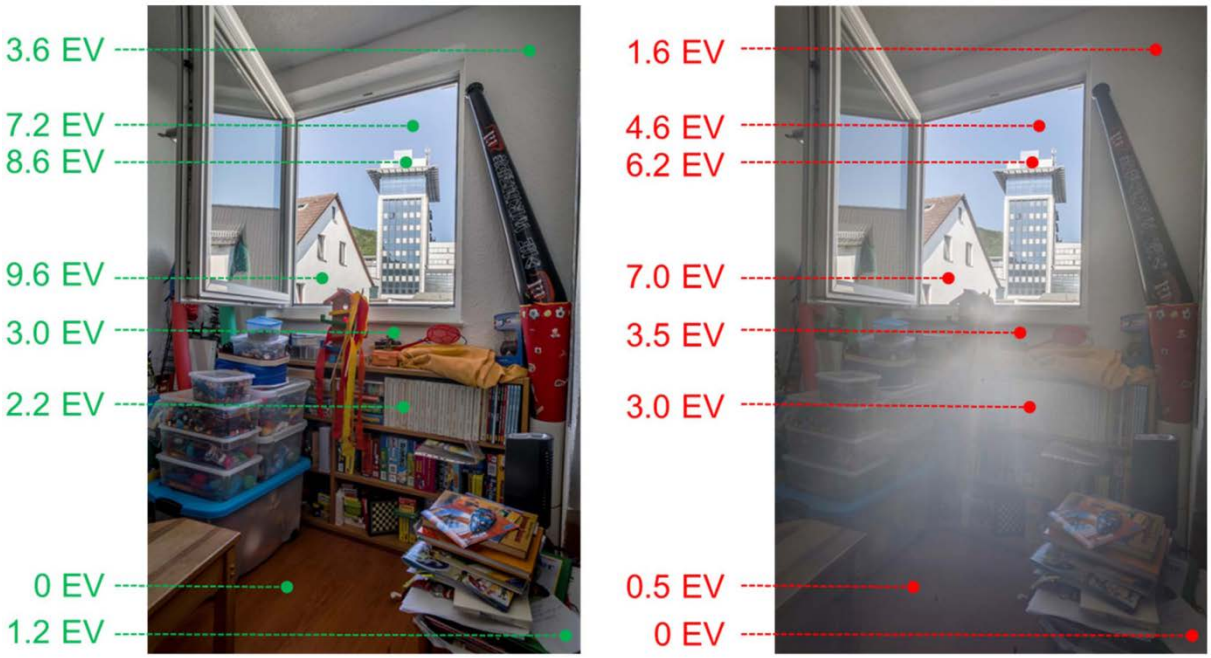


Fig. 1.28: Differences in relative irradiance between different field points. The values are referenced each to the darkest point within the image.

Aptitude of digital image correction methods

The only means to remove strong ghost images that are due to bright, local light sources, is cumbersome manual retouching. When important image contents are covered by deep ghost images, the entire image is corrupted.

By contrast, when large areas of the image are similarly bright, there is less local overexposure. Instead there is a haze effect covering the image. In addition to the change in brightness due to the flare light, the colors stemming from different areas of the field mix with each other and the image colors appear muted (see Fig. 1.29): As dark areas are covered by bright flare light from other image regions, the portion of dark areas decreases; black turns into gray. This can also be seen as compression of the value range in the image histogram.

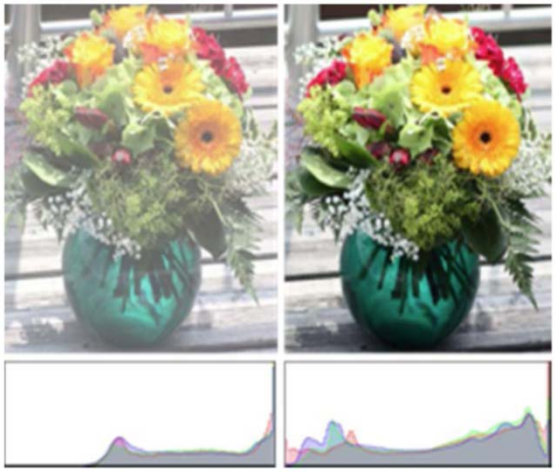


Fig. 1.29: Comparison of images taken with the ZEISS Planar 2,8/80mm lens; left: uncoated, right: T*-coated. The histogram shows that the dynamic range of the left image is much smaller. The dark brightness tones were removed by the flare light haze.

When using image processing software, this reduced brightness range can be re-expanded, so that the brightness range of the image covers all available brightness values (see Fig. 1.30). Dark gray pixels are simply replaced by black ones, and light gray pixels by white ones. In addition to the global adaptation of the tonal value range, there are more complex tonal value corrections methods, which are performed locally over image areas, context-sensitively (depending on the brightness gradient) and over different brightness ranges. Descriptions of such methods can be found in text books on digital image processing under the term "histogram equalization" or "histogram matching" (e.g. Gonzalez, Woods (2008)).

A tonal value correction often helps increase the image quality. However, problems might occur when areas with small brightness differences are so strongly intensified that extraneous abrupt contours arise (Fig. 1.30).

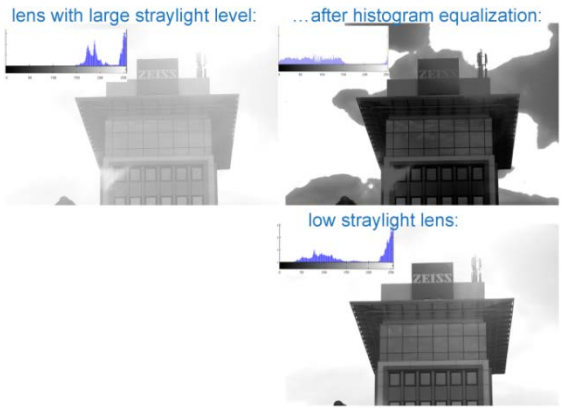


Fig. 1.30: Strong extraneous contours in the areas of the sky after expansion of the tonal value range: top left: image taken with an uncoated lens with much flare light; top right: after histogram equalization the whole tonal value range is utilized. Bottom right: image taken with little flare light (Planar T* 2,8/80mm CFE).

Even with a tonal value correction, the quality of an image taken with a high-quality-coated lens can never be reached. The reason being that information of an object point, i.e. its brightness and color, would be automatically distributed because of the flare light to a whole range of different image points. In this way, the information of different object points is amalgamated in the image plane, which reduces crispness and color fidelity of the image.

Artistic image design with reflections

In the past sections, we have shown examples, where ghost images and flare light reduced the image quality. It should not go unmentioned that reflections are also used as an important artistic device. Lens flare of bright local light sources is intentionally utilized in professional movies as well as in photography to add expressiveness to the scenes and let them appear more dynamic. Reflections emphasize the irradiation emanating from very bright or hot sources, which, because of saturation of the brightness signal, are rendered in flat white in the image. The dullness of deep space appearing in a science fiction movie can be enriched to a large extent thanks to the flare-rich view through an optical system. As of today, such lens flare effects can be emulated computationally with a very good level of detail and applied subsequently in the post-production stage (see Hullin, Eisemann, Seidel, Lee (2011)). Computer-graphical models of fictitious lenses have been in use for some years for animation movies and computer games.

Nearly all camera lenses are rotationally symmetric and their ghost images show the features mentioned. Anamorphic lenses are different in that regard. Nowadays, about one third of all Hollywood blockbusters are shot with professional anamorphic cine lenses. With these lenses, the image information is squeezed in horizontal direction (by a factor of 2 in most cases) and de-squeezed to the original ratio for projection on the cinema screen (or nowadays electronically in the digital image data) (for more details see: ARRI Anamorphic De-Squeeze white paper, 2011). Such lenses contain, apart from spherical lens elements, also cylinder-shaped ones. In connection with bright light sources, these anamorphic lenses sometimes create horizontally or vertically oriented patterns or stripes in the image field. Such "horizontal streaks" are a popular stylistic device used by cinematographers, which can be seen in many science fiction or action movies.

Fig. 1.31: ARRI/ZEISS Master Anamorphic including Flare Set (additional module in the front-end and rear area).

To achieve maximum brilliance of the images, we optimize ZEISS lenses to a minimum flare level. To offer further creative freedom to professional filmmakers for vivid and creative images, Flare Sets are offered for ARRI/ZEISS Master Anamorphic lenses (see Fig. 1.31), which consist of exchangeable glass elements for the front-end and rear-end lens areas, with a clearly higher reflectance of the glass surfaces. Fig. 1.32 shows images from a take by Tom Fährmann made with ARRI/ZEISS Master Anamorphics with Flare Set (Fährmann, 2015).

Fig. 1.32: Images from the movie “Master Anamorphic Flare Sets showreel” by Tom Fährmann, taken with an ARRI/ZEISS Master Anamorphic with Flare Set and an ARRI ALEXA camera.

The road to "T*-optics"

Principle of reflection-reducing coatings

When light travelling through air strikes a glass surface, most of the light is transmitted through the glass, while part of it will be reflected.

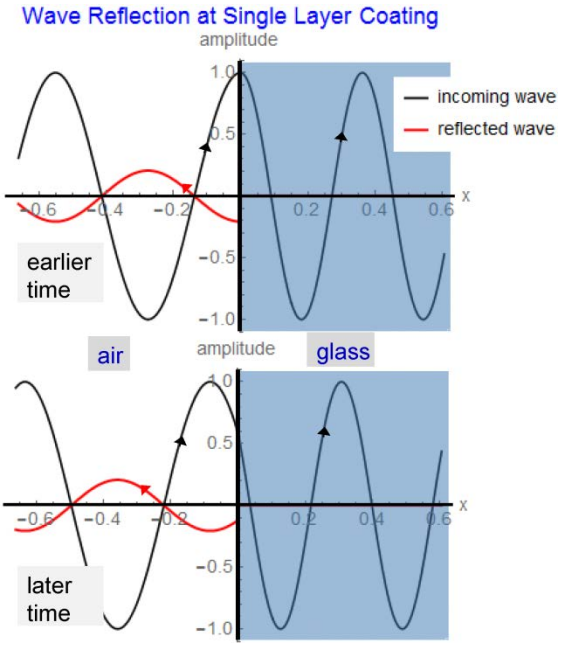


Fig. 2.1: A light wave passes from air into glass; part of the light (i.e. the portion $((n-1)/(n+1))^2$) being reflected at the surface.

In case light hits a surface between air (refractive index 1) and glass (refractive index n) straight on (parallel to the surface normal), the reflectance R is given by

$$R = \left(\frac{n-1}{n+1}\right)^2.$$

Fig. 2.2 shows the reflectance values of some glasses that result from this formula and the respective indices of refraction.

	n	R [%]
CaF_2	1.44	3.3
BK7	1.52	4.3
SF5	1.67	6.3
LAF7	1.75	7.4
LASF46A	1.9	9.6
Silicium	3.54	31.3
Germanium	4	36.0

Fig. 2.2: Refractive indices n and reflectance R of optical glasses for the visible (blue) and infrared spectral range (red).

When reading some text books on optics, one may be induced to believe that the refractive index of glass is always 1.5. However, the mean value of the refractive indices of all available optical glasses in the visible spectrum is more likely in the range of 1.7.

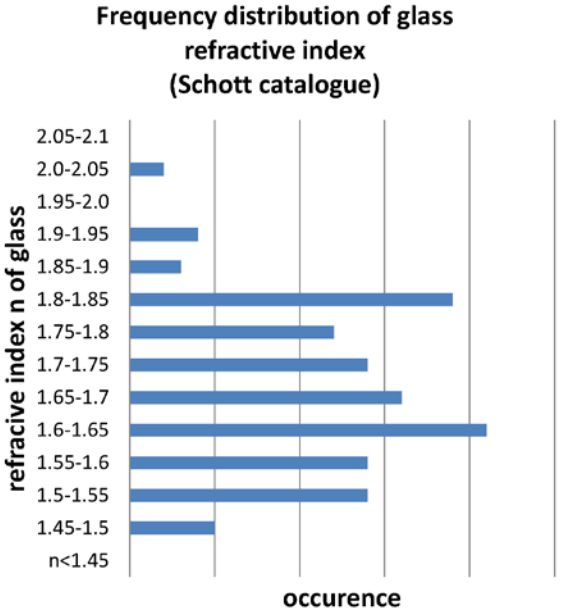


Fig. 2.3: Frequency distribution of refractive indices of currently 124 glasses as listed in the SCHOTT catalogue (SCHOTT glass data 2014).

Fig. 2.3 shows the frequency distribution of the refractive indices of all glasses listed in the current SCHOTT catalogue: The refractive index of most of the glasses ranges between 1.48 and 1.93. As far as glasses for the visible spectrum are concerned, the reflectance values vary between 4% and 10%,

with the average being around 7%. For the silicon and germanium glasses used for infrared optical systems, the reflectance values are even above 30%.

The reflectance at the surface transition in a cemented lens is considerably lower. Here the following applies:

$$R = \left(\frac{n_1 - n_2}{n_1 + n_2}\right)^2.$$

For example, at a surface between two glasses with n_1=1.7 and n_2=1.5, respectively, the reflectance is only about 0.4%. Fig. 2.4 shows a comparison of different combinations of refractive indices of cemented lenses.

	1	1.5	1.6	1.7	1.8	1.9	n_2
1	0.00%	4.00%	5.33%	6.72%	8.16%	9.63%	glass-air
1.5	4.00%	0.00%	0.10%	0.39%	0.83%	1.38%	glass-glass (cemented lens element)
1.6	5.33%	0.10%	0.00%	0.09%	0.35%	0.73%	
1.7	6.72%	0.39%	0.09%	0.00%	0.08%	0.31%	
1.8	8.16%	0.83%	0.35%	0.08%	0.00%	0.07%	
1.9	9.63%	1.38%	0.73%	0.31%	0.07%	0.00%	
n_1	glass-air	glass-glass (cemented lens element)					

Fig. 2.4: Reflectance values of different combinations of refractive indices with perpendicular light incidence.

In order to interpret and describe how reflection-reducing coatings work, one has to consider the wave nature of light: to wit, the interference (i.e. superposition – see Fig. 2.5) of the waves reflected at the surfaces of the coating and the lens glass, respectively.

Reflection at the air-glass interface may be reduced by destructive interference, when the waves reflected at the surfaces of glass and coating, respectively cancel each other (partially). We consider the example of a single-layer coating consisting of only one coating layer. Fig 2.6 shows how the two waves reflected at the air-glass interface and coating-glass interface, respectively, cancel each other out (if thickness and refractive index of the layer have been selected appropriately).

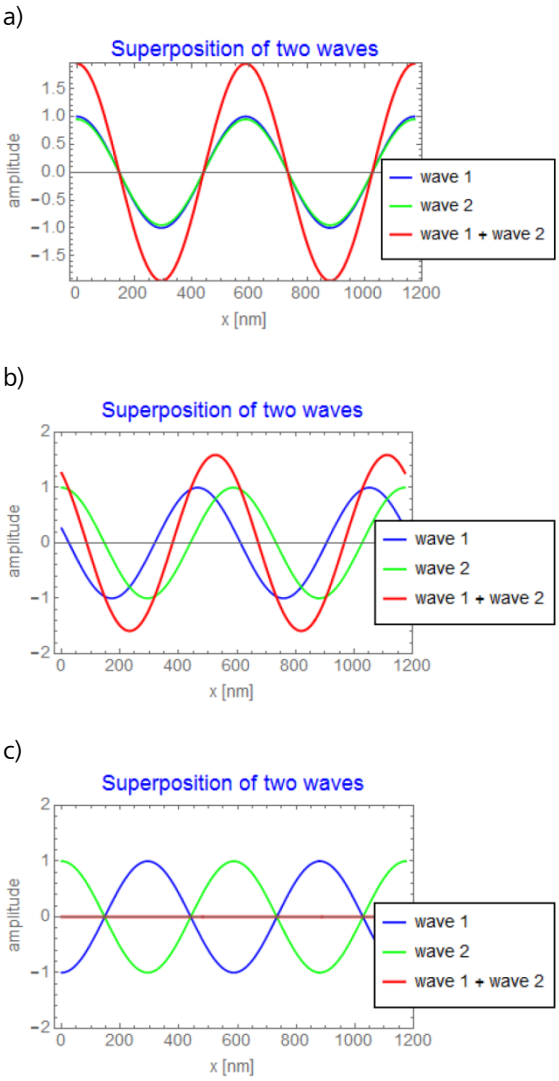


Fig. 2.5: Superposition of two waves (green and blue) of the same amplitude and wavelength: a) constructive interference (phase difference of 0°), b) general case, c) destructive interference (phase difference of 180°).

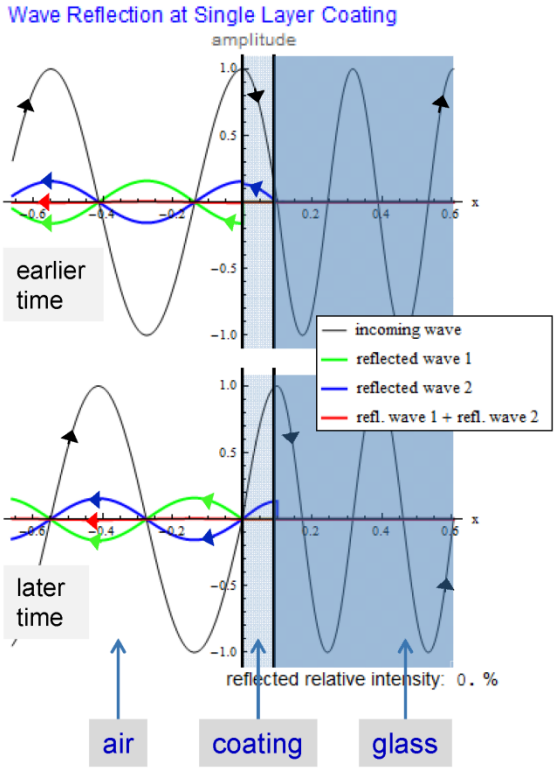


Fig. 2.6: Principle of an antireflection coating using the example of a single-layer coating (green: wave reflected at the air-coating interface, blue: wave reflected at the coating-glass interface). If the thickness and the refractive index of the material have been selected in such a way that the reflected waves oscillate with opposite phases and the same amplitude, they will extinguish each other.

The physical principles of wave optics have their roots in Christiaan Huygens' work on the propagation of light (Huygens, 1690) and Isaac Newton's observation of interference fringes (called "Newton's rings" today) and the description of various diffraction phenomena (1704). Major progress was achieved at the beginning of the 19th century, above all through the work presented by Thomas Young (1802) on the interference of light and by Jean Augustin Fresnel (1823). To Fresnel we owe the mathematical theory of diffraction and propagation as well as the polarization-dependent equations for the transmission and reflection at optical interfaces.

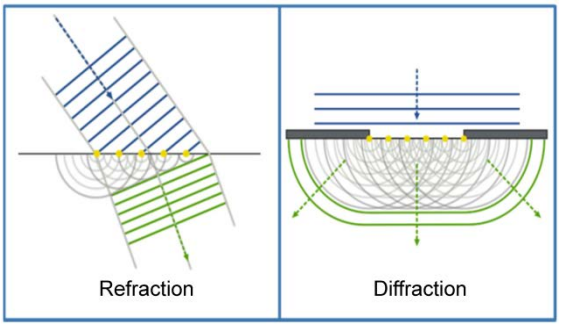


Fig. 2.7: Wave-optical model of refraction and diffraction (Huygens, Fresnel).

In 1864, James Clerk Maxwell demonstrated – as part of his theory of the dynamic electromagnetic field – that light is an electromagnetic wave. Heinrich Hertz confirmed this experimentally in 1886.

Ernst Abbe at Carl Zeiss in Jena, took this body of fundamental work in physical optics further by extending and applying it to an imaging theory for optical devices (see Abbe (1873)). The mathematical formulation of the theory was refined by his students (c.f. Lummer, Reiche (1910) or Czapski (1904)). According to Abbe, any optical image is created by superposition, i.e. interference, of the waves that proceed from the lens pupil in different directions to the image plane. As a consequence of this theory, Abbe derived a formula for the resolution limit of optical systems: to wit, the resolution power of any optical system is limited by the ratio of the light wavelength and the numerical aperture of the lens. Following preparatory work performed by Max von Laue (1907) and Erwin Schrödinger (1920), this image-formation theory was then generalized by Max Berek (1926), Pieter van Cittert (1934), Frits Zernike (1938) and, later on, Harold Hopkins (1953). They developed the optical coherence theory, which includes propagation and imaging equations for partially coherent light, i.e. light capable to partially interfere in the object plane.

Interference is of great importance at many other places for the production of optical systems: shape, thickness and shape devia-

tions (i.e. the deviations of the surfaces of optical components from the target shape), as well as the wave front of entire optical systems over their image field are measured interferometrically with high precision.

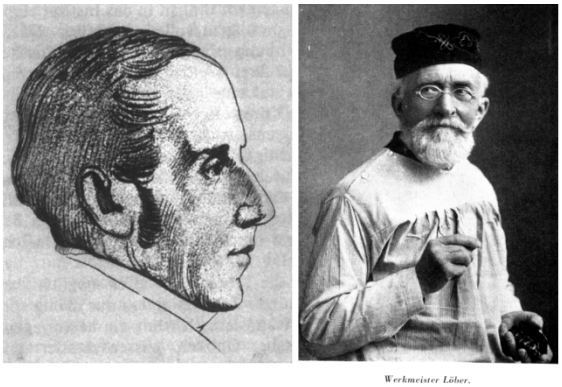


Fig. 2.8: Introduction of interferometric testing of optical surfaces. Left: Josef Fraunhofer (1787-1826); right: August Löber (1830-1912).

Around 1810, Josef Fraunhofer was presumably the first to measure optical surfaces interferometrically using "test plates". This method fell into oblivion for some decades after Fraunhofer's early death; it was revived around 1860 by August Löber, master workman at ZEISS in Jena: The light refracted at the optical surface to be tested interferes with the light reflected at the test glass, and the interference fringes deliver information about the surface deviation (see Fig. 2.9).

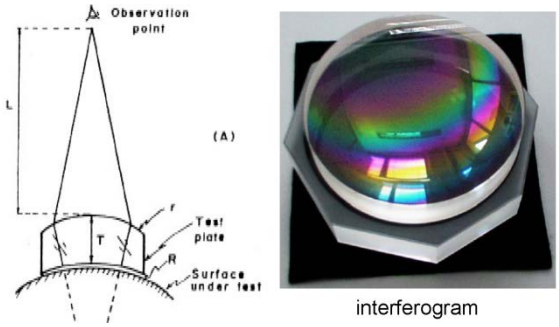


Fig. 2.9: Simple visual interferometric test of an optical surface using a test glass. The interference fringes appear colorful in white light, since the periodicity of the fringe patterns depend upon wavelength.

This simple visual test allows for measuring accuracies in a range from about λ/4 up to λ/2 (where λ denotes the light wavelength, amounting to about 0.5 micrometers). This testing method has been used up to the present day in every optical workshop as a standard for the determination of radius and shape deviations. With today's interferometers it is possible to detect and measure surface deviations that are in the range of fractions of wavelengths (cf. Dörband, Müller, Gross (2012)).

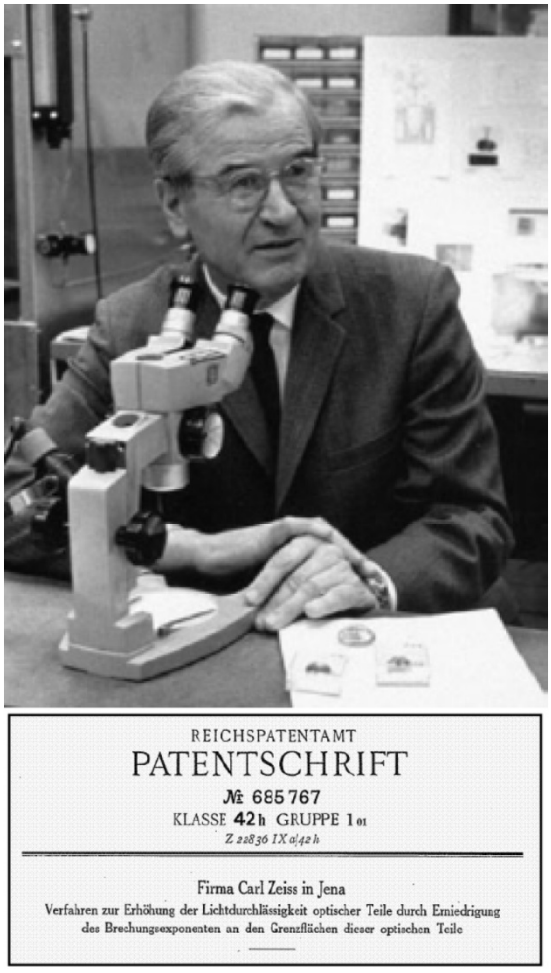
About the history of T- and T*-optics

The application of optical coatings or "optical thin films" is a very broad field: Today, antireflection coatings are applied to virtually all optical devices – from eyeglass lenses to binoculars to microscopes. By now, even sheet glass for windows in buildings is very often provided with an antireflection coating. Reflection-reducing coatings serve to increase the efficiency of solar cells, to suppress stationary waves in photoresist for the lithographic production of microchips, or to enhance the efficiency of fiberglass cables or light-emitting diodes. In addition to performing an antireflection function, coatings often also act as protection against dirt, water or ultraviolet radiation, and they can regulate the penetration of infrared heat radiation. We note that coatings are used not only as an antireflection means, but also to achieve quite the opposite effect, e.g. to increase the reflectance of mirrors, to adjust reflectance factors in partially transparent beam-splitter mirrors or wavelength-selective filters, to increase light absorption and improve heat removal or act as polarization components. Depending on the purpose of use, the spectral range or radiation load, there is a great variety of coating designs, which may consist, for some applications, of far more than 100 single layers.

The wish to remove disturbing reflections from optical glasses existed hundreds of years before the introduction of AR coatings

into the optical industry at ZEISS. There were numerous ideas and attempts to suppress reflections from glasses: In 1817, Josef von Fraunhofer noticed, as did Lord Rayleigh in 1886, that polished glass surfaces become more translucent over time. In 1896, Denis H. Taylor observed the same phenomenon and assumed that this was due to the formation of a porous surface with an effectively lower refractive index owing to atmospheric influences. Consequently, Taylor tried to develop a similar but accelerated process in the laboratory. In 1904, he obtained an English patent for a process to form a reflection-reducing surface from the existing glass in an acid bath. However, this method did not work well under production conditions: In the chemical etching process, ultrafine and otherwise invisible abrasion shadows caused by lens polishing were coarsened and became optically disturbing scratches. The surface of some types of glass was practically destroyed. As a consequence, research into this approach was discontinued and it fell into oblivion.

The term used today for reflection-reducing coatings is "antireflection coating" or "AR coating". Instead of speaking of the "avoidance of reflections", we could use the term "improved transmission", taking inspiration from the Russian designation for AR coatings: "просветление оптики" (pronounced: "prosvetlenie optiki" and meaning a "better shining-through of optics." The positive neologism of "transparency optics" or "T-optics" for short or "T-coating" for reflection-reducing coatings by ZEISS may have to be attributed to the roots of the inventor Alexander Smakula (1900-1983), who was born and raised in the Ukraine. In the 1930s, Smakula was head of the Optical Surface Laboratory II of the R&D department at ZEISS in Jena, and he developed, with suggestions from A. König and H. Bauer, a process for the application of abrasion-resistant, durable coatings on lenses to reduce reflections.



REICHSPATENTAMT

PATENTSCHRIFT

№ 685767

KLASSE 42h GRUPPE 1 01

Z 22836 IX a/42 h

Firma Carl Zeiss in Jena

Verfahren zur Erhöhung der Lichtdurchlässigkeit optischer Teile durch Erniedrigung des Brechungsexponenten an den Grenzflächen dieser optischen Teile

Fig. 2.10: Dr. Alexander Smakula (1900-1983) and his patent granted in 1935 for the "process to increase the transparency of optical components by reducing the refraction at the interfaces of these optical components".

ZEISS was the first company to produce such coatings and place them on the market. After the patent issuance to Smakula on November 1st of 1935, several years had to pass, mainly needed to introduce and improve the production process, until coated lenses were eventually brought to market. One major difficulty for the coating process is that any impurities, grease and moist residues on the glass, which could disturb the coating adhesion, have to be thoroughly removed prior to evaporation. Glass surfaces are always covered with a fine adsorbed water film from the atmospheric humidity. This very thin water film has considerable influence on the coating quality and must

therefore be removed by glowing in the vacuum plant directly before coating is started. Such controlled production of thin coatings was possible thanks to the progress made in high-vacuum technology in the 1930s, which in turn was stimulated by the rapid development of thin films. In 1938, 17 coating plants were already in operation - however, because of a government directive, for military-related optical devices only. In 1940, the first coated camera lenses of the types Biotar and Sonnar were presented at the Leipzig Spring Fair. In 1941, nearly 300 lenses of the Sonnar 1,5/50mm type for the CONTAX camera were coated per month. Generally, at the beginning of the 1940s, more than 100 vacuum deposition plants were developed and optimized in conjunction with W. C. Heraeus (Onstmettingen). By 1950, all ZEISS lenses were coated and marked with a letter T in red print. At the same time, starting around 1943, tests were made for double and triple layer coatings. Such multi-layer coatings were first used in the 1950s for wide-angle lenses that often consisted of 18 or more optical surfaces. These multi-layer coatings were labelled "MC" in Jena and "T*" in Oberkochen. As of 1972, each ZEISS camera lens has been provided with a multi-layer coating (Dr. Hans Sauer, ZEISS Press Information of February 20, 1973). Further historical accounts dealing with the T-coating can be found in publications by R. Richter (1940), K. Leistner (1941), A. Smakula (1942), J. Flügge (1960) and B. Gänswein (1997).

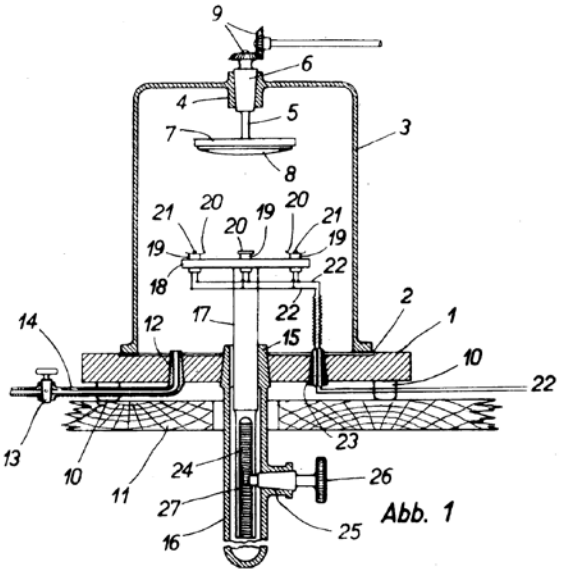


Fig. 2.11: Vacuum deposition plant from Smakula's patent: The glasses to be coated (8) are mounted on a holder (7) under vacuum in a glass dome (3). The material (21) filled into the small tungsten bowls (20) is evaporated by electrical heating. The holder is constantly rotated, so that the material evenly deposits onto the lens.

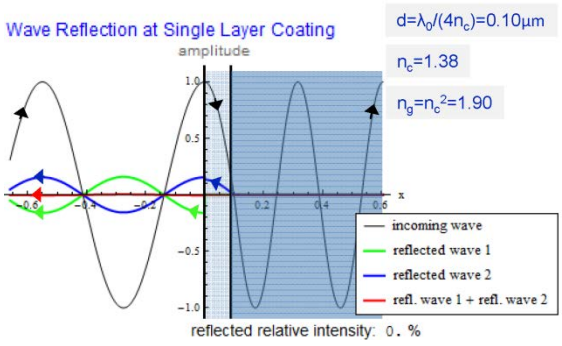
Fig. 2.12: Evaporation plant for lenses (photo of 1961).

In the following sections, we are going to examine single-layer, double-layer, triple-layer and general multi-layer antireflection coatings, in terms of their respective principles, capacities and limits. Readers who wish to see in detail how we obtain the various reflectance curves that are presented in the following are kindly asked to refer to the Appendix.

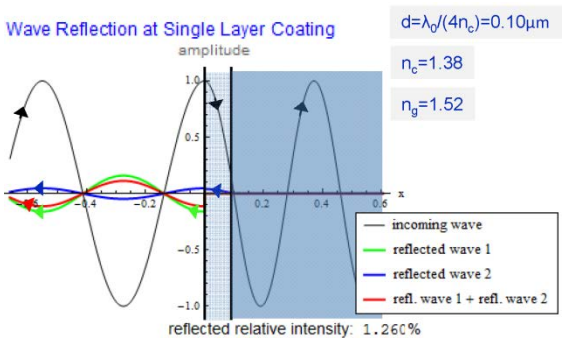
Single-layer coatings

When a single-layer coating is placed onto a glass surface, there are two reflected waves: one at the upper side of the coating, the other at the coating-glass interface. To achieve perfectly destructive interference, i.e. complete mutual cancellation of these two waves, for any given wavelength, the phases of the two waves must be offset by exactly 180° (or π), and their amplitudes must be identical (see Fig. 2.13).

a)



b)



c)

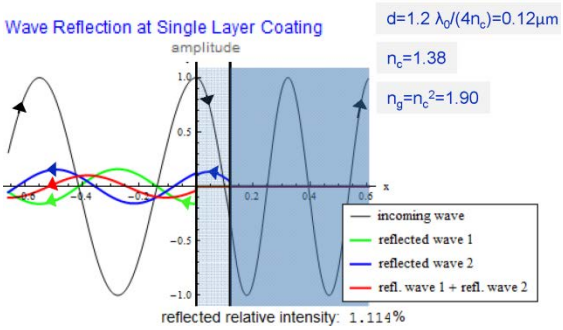


Fig. 2.13: Incoming wave (black), wave reflected at the front side (green), and wave reflected at the interface to glass (blue). The sum of the two reflected waves is shown in red, with its intensity indicated below. The wavelength is λ_0=0.55µm. a) Phase and amplitude conditions both fulfilled; b) Phase condition fulfilled, amplitude condition violated (i.e. the refractive indices are not in the relation $n_g=n_c^2$); c) Amplitude condition fulfilled, phase condition violated (i.e. the layer thickness is not equal to $\lambda_0/(4n_c)$).

The optical path difference (OPD) of both reflected waves for a single-layer coating is depicted in Fig. 2.14:

$$OPD = 2\, n_c\, d\, cos i'.$$

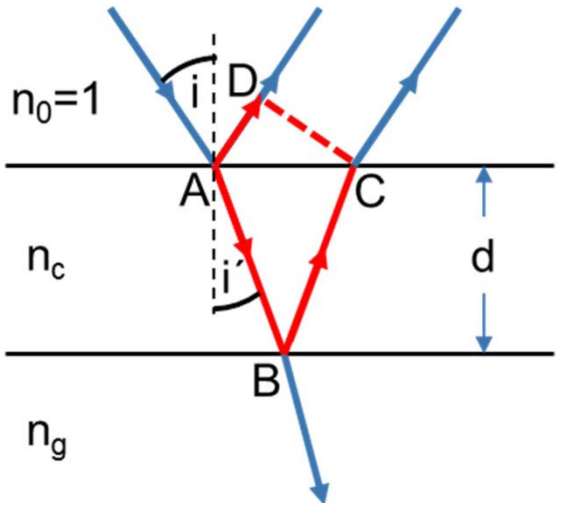


Fig. 2.14: The optical path length ABC is $OPL_{\overline{ABC}} = 2\, n_c\, d/\cos i'$, or AD is $OPL_{\overline{AD}} = \sin i \cdot \overline{AC} = \sin i \cdot 2d \tan i' = 2d\, n_c\, \sin i' \tan i'$. Thus, the following results for the OPD of the two reflected waves: $OPD = OPL_{\overline{ABC}} - OPL_{\overline{AD}} = 2d\, n_c\, (1 - \sin^2 i')/\cos i' = 2d\, n_c \cos i'$.

For perpendicular incidence (i=0°), this equation yields:

$$OPD(i = 0°) = 2n_c d.$$

The phase difference of the two waves in units of the wavelength λ is

$$\varphi = \frac{2\pi}{\lambda} OPD.$$

So, the phase condition that both waves cancel each other and thus interfere destructively is

$$\frac{2\pi}{\lambda_0} 2n_c d = \pi.$$

The reference wavelength λ_0 is chosen to lie approximately in the middle of the visible spectrum, the latter ranging from approx. 420 nm to 670 nm. More precisely, since the reflectance with single-layer coatings increases more rapidly towards the blue end of the spectrum, the value of λ_0 is more appropriately chosen to be 510 nm rather than the algebraic mean of approx. 550 nm.

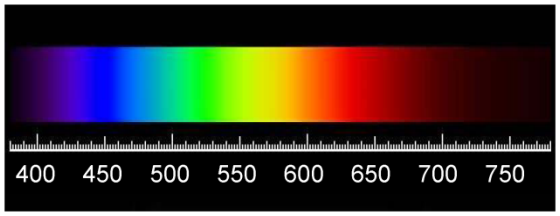


Fig. 2.15: Spectrum versus wavelength.

Solving the above equation for the variable d yields an expression for the layer thickness at which destructive interference occurs (in the case of perpendicular incidence i=0°):

$$d = \frac{\lambda_0}{4\, n_c}\,.$$

With perpendicular incidence 0°, Fresnel's amplitude coefficients r for reflection and t for transmission are given by

$$r(n, n', 0°) = \frac{n - n'}{n + n'}\,,$$

$$t(n, n', 0°) = \frac{2n}{n + n'}\,.$$

The amplitudes of the reflected waves are

$$\rho_1 = r(1, n_c),$$

$$\rho_2 = t(1, n_c)\, r(n_c, n_g)\, t(n_c, 1).$$

We neglect the transmission loss, i.e.

$$t(1, n_c)\, t(n_c, 1) = \frac{4n_c}{(1+n_c)^2} = 1 - \left(\frac{1-n_c}{1+n_c}\right)^2 \approx 1.$$

(For instance, this expression takes on the value of 0.975 instead of 1.0 for n_c=1.38.) Using the amplitude condition,

$$\rho_1 = \rho_2 \quad \text{or} \quad \frac{1-n_c}{1+n_c} = \frac{n_c - n_g}{n_c + n_g},$$

we obtain an expression for the sought-after refractive index n_c of the coating:

$$n_c = \sqrt{n_g}.$$

This means that for this value of the coating's index of refraction, the reflected intensity disappears completely at the wavelength λ_0, if the phase condition is fulfilled at the same time.

The examples in the following sections will demonstrate that the refractive index of the glass greatly influences the extent to which reflection reduction through coatings is achieved. For the SCHOTT glass BK7 with a refractive index of n_g=1.52, the condition $n_c = \sqrt{n_g}$ yields an optimal refractive index of n_c=1.23 for the coating layer. For the high-index glass type LASF9 (n_g=1.85), however, the corresponding value is n_c=1.36. There is only a limited choice of suitable, i.e. above all robust materials with such a low refractive index. The most important material because of its high robustness is magnesium fluoride MgF_2, with a refractive index of 1.38.

Generally, for single-layer coatings as well as for multi-layer coatings, the reflectance values depend on the refractive index of the coated glass. This means that the thicknesses

of the coating layers are adapted to the refractive index of the glass to be coated.

Since the necessary optical path length is $n_c \cdot d = \lambda/4$, single-layer coatings are also simply called "λ/4-coatings." With the wavelength λ_0=550nm ("green light") and n_c=1.38, a layer thickness of about 100nm or 0.1µm results from the phase condition $d = \lambda_0/4n_c$.

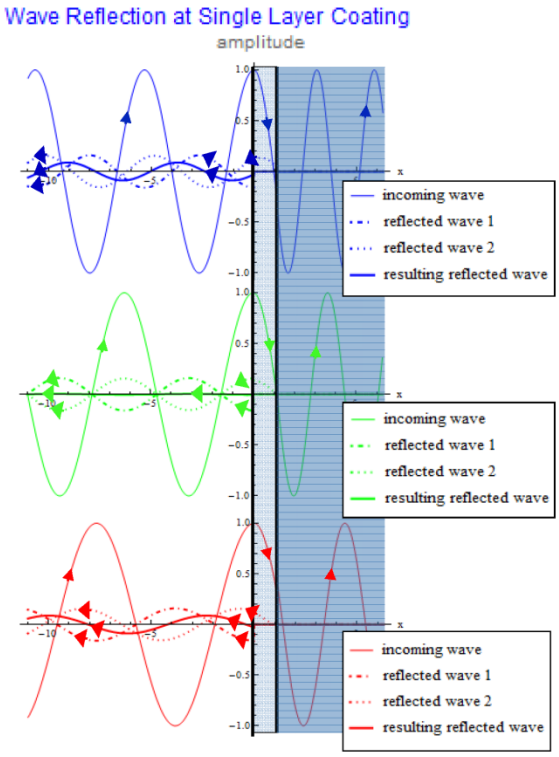


Fig. 2.16: The thickness of the single-layer coating is chosen in such a way that the two reflected waves (dashed) for green light are inversely phased, so that the intensity of the green reflected light is minimal. For this layer thickness, the optical path in the coating is too long for blue light and too short for red light. This means that there remains nonzero reflectance for blue as well as red light.

Since one single value has to be chosen for the layer thickness the optical path length in the coating for light of any other wavelength differs from the value λ/2. For instance, the optical path length for blue light (λ = approx. 450nm) is too long by the factor 550/450≈1.22, and that of red light (λ = approx. 650nm) is too short by a factor of approx. 550/650≈0.84. This means that in either case the two reflected waves are not inversely phased to 180°, hence the reflectance for red and blue light is greater than zero (cf. Fig. 2.16).

The wavelength-reflectance functions that result when a single-layer MgF_2 coating is placed on different glass types are depicted in Fig. 2.17, each in comparison to the respective reflectance of an uncoated lens.

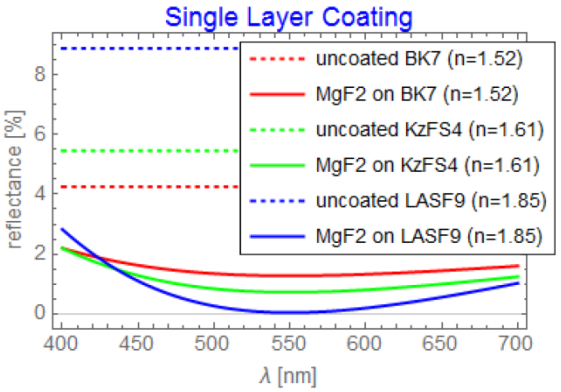


Fig. 2.17: Spectral reflectance profiles of single-layer MgF_2 coatings on glasses with different refractive indices, in comparison to the uncoated interface, respectively (dashed).

For our schematic diagrams, we ignore dispersion, i.e. the wavelength dependence of the refractive indices. For all glasses, the refractive indices are slightly increasing from the red to the blue wavelength range. For instance, the index of refraction of the BK7 glass of SCHOTT exhibits an increase from n=1.514 for a wavelength of 656 nm to 1.527 for 435 nm. Taking this into account, one observes a further slight increase of the reflectance value towards the blue end of the spectrum.

As a consequence of increased reflectance values for blue and red light compared to green light, the coating appears purple-colored, i.e. as mixed color of red and blue (see Figs. 2.18 and 2.19).

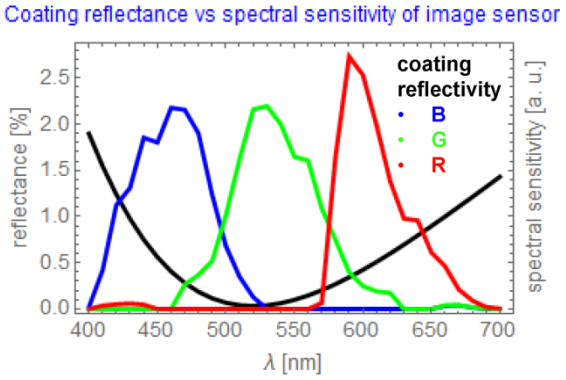


Fig. 2.18: Reflectance curve (black) of a single-layer coating (MgF_2 (n_c=1.38) on LASF9 (n_g=1.85), centered at λ_0=520nm) and spectral sensitivities, respectively, of the three color channels (RGB) of a typical SLR camera image sensor.

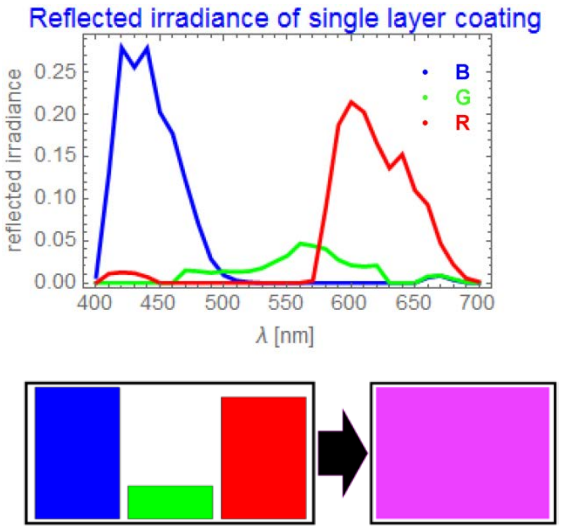


Fig 2.19: The spectral irradiance resulting from the example in Fig. 2.18 (above), the resulting integral irradiance per red/green/blue channel (bottom left), and the corresponding additive color mixture. The single-layer coating appears purple.

If there are deviations of the layer thickness from the target thickness, for example due to defects during the coating process, the minimum values of the spectral reflectance curves are offset corresponding to the changed optical path length. Layers that are too thick appear with a stronger blue, too thin layers with a stronger red hue (cf. Fig. 2.20).

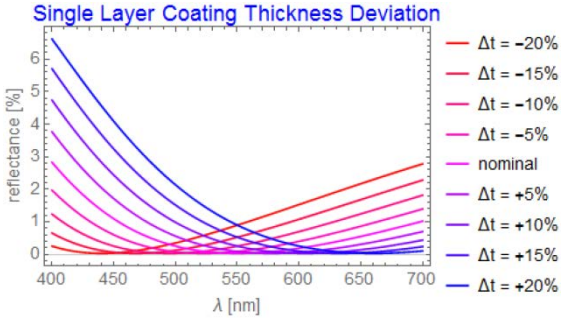


Fig. 2.20: Single-layer coating with n_c=1.38 on glass with n_g = 1.85 with deviating layer thicknesses. Deviations of the layer thickness lead to shifted reflectance profiles. A relative thickness deviation of Δt = 10% corresponds to a thickness defect of only 10 nanometers.

If the layer thickness deviates from its nominal value or if the thickness varies across a lens, image quality may deteriorate considerably. Thickness deviations have to be avoided especially when multi-layer coatings are manufactured.

In real optical systems, the angles of incidence on the refractive surfaces in the nominal optical path may amount to 40° or even more. For back reflections, the angles of incidence may even be larger. If the light hits a surface at an angle i, the optical path difference of the two reflected waves is reduced by the factor $\cos i'$ according to the formula given before:

$$OPD = 2\, n_c\, d\, cos\, i'.$$

Expressing this in terms of the angle of the incoming ray, the factor is $cos\, i' = \sqrt{1-(\sin i/n_c)^2}$. Thus, for oblique incidence, the two reflected waves are no longer inversely phased even for the wavelength λ_0 (green light), and the reflectance rises. Blue light, however, (for which the optical path length is too long with normal incidence, cf. Fig. 2.16) benefits from the reduction of the optical path length for oblique incidence: the phase difference between the two reflected waves is reduced. For red light then, the optical path difference becomes even smaller than in the case of

normal incidence, and the reflectance increases further. As an overall effect, the reflected light assumes increasingly reddish hues for rising angles of incidence (see Fig. 2.21).

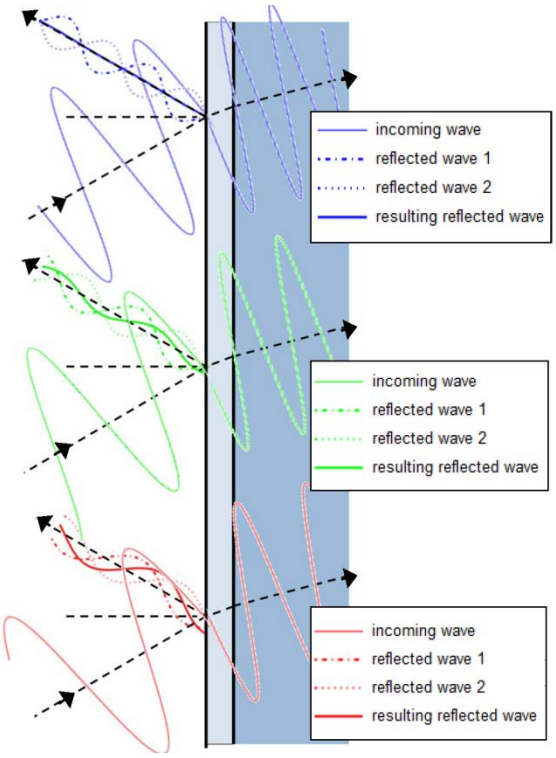


Fig. 2.21: Spectral displacement of the reflectance profile for oblique incidence.

For oblique incidence, Fresnel's reflectance changes as well. Its increase is comparatively larger at the air-coating interface than at the coating-glass interface (see Fig. A.3 in Appendix A.1). Thus, the amplitude condition will be violated even more. As a consequence, the minimal reflectance increases for larger incidence angles (see Fig. 2.22). This "displacement towards red" for oblique incidence is also observed in multi-layer coatings, by the same mechanism.

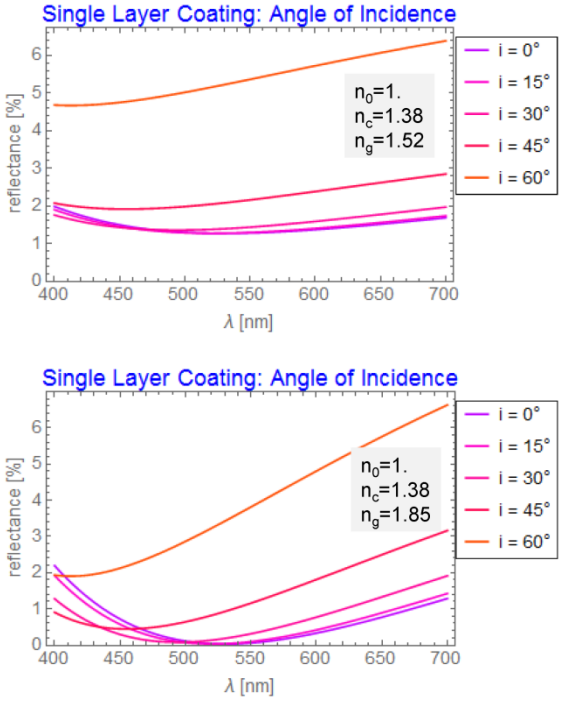


Fig. 2.22: Reflectance profiles for different angles of incidence i. Top: MgF_2 (n_c=1.38) on BK7 (n_g=1.52); Bottom: MgF_2 (n_c=1.38) on LASF9 (n_g=1.85).

So far, we have tacitly ignored the fact that the two reflected waves remain inversely phased if integer multiples of 2π are added to the right-hand side of our phase condition $\frac{2\pi}{\lambda_0} 2n_c d = \pi$: $\frac{2\pi}{\lambda_0} 2n_c d = \pi + 2\pi j$ (j=1,2,…). This generalized condition leads to layer thicknesses that are equal to λ_0/n_c multiplied by $1/4$, $3/4$, $5/4$, … For layer thicknesses greater than the minimum value $d = \frac{\lambda_0}{4\, n_c}$ the light has to cover larger distances in the medium. Hence, for any given wavelength (other than λ_0), the phase shift is larger. As a consequence, the wavelength range within which low reflectance values can be achieved becomes narrower (see Fig. 2.23). Therefore, in order to achieve broadband antireflection properties, coatings having the minimum thickness $d = \frac{\lambda_0}{4\, n_c}$ are normally used.

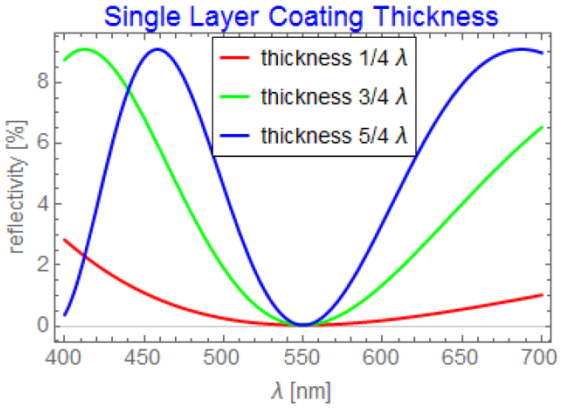


Fig. 2.23: The phase condition at λ_0 (550nm here) is fulfilled also for integer multiples of the phase shift 180°+360°=540° (corresponding to a thickness of 3/4 λ_0), 180°+720°=900° (corresponding to 5/4 λ_0), etc. Reflectance becomes more narrowband then.

A disadvantage of the magnesium fluoride single-layer coating is that the reflectance cannot be completely brought down to zero with low-refracting glasses. For the BK7 glass (n_g=1.52), for example, a residual reflectance of about 1.3% remains. Another disadvantage is that the reflectance remains rather high in the blue and red regions of the visible spectrum (cf. Fig. 2.17). With single-layer coatings it is not possible to avoid this 'V' shape of the spectral reflectance; the reflectance function of a single-layer coating has one minimum only.

Double-layer coatings allow for more flexible material choices or achromatizing

AR coatings with two layers, one lying on top of the other, produce three reflected waves (see Fig. 2.24). (For the AR coatings examined here, any more than two reflections within one coating layer can be neglected to good approximation – see Appendix A.2.)

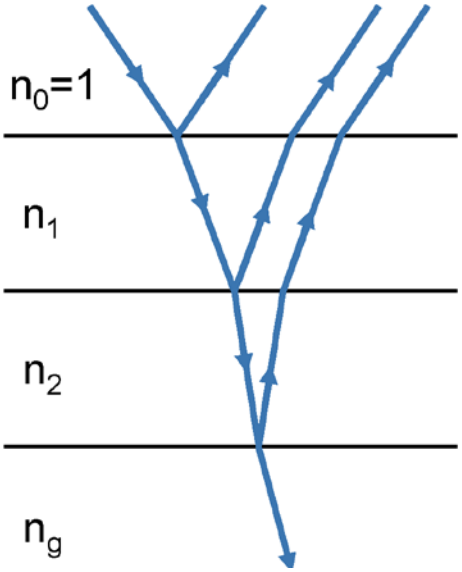


Fig 2.24: With a double-layer coating, three reflected waves are created (while neglecting multiple reflections), which interfere with each other.

For single-layer coatings, there is one unique combination of amplitude and phase resulting in the cancellation of the two reflected waves. In contrast, there is a great number of possibilities with a double-layer coating (i.e. three reflected waves) to achieve a cancellation of the reflected intensity as a whole at a certain wavelength. Two examples are shown in Fig. 2.25.

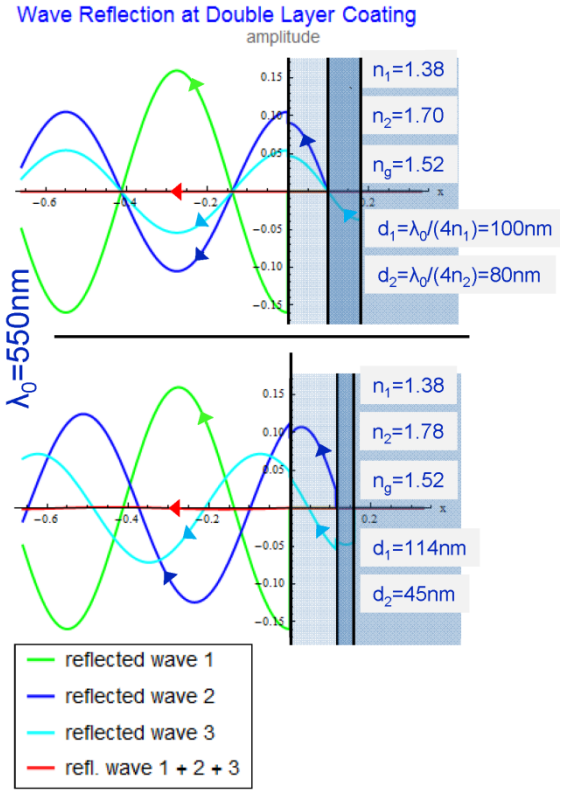


Fig. 2.25: Two examples of a resulting disappearing amplitude.

The combination of two single-layer coatings yielding the same optical path length $n_1 d_1 = n_2 d_2 = \lambda_0/4$ but having, possibly, different refractive indices n_1, n_2 is called a "λ/4-λ/4-coating". With a λ/4-λ/4-coating, the following amplitude condition (see Appendix A.4) results for the reference wavelength λ_0 (again with n_0=1):

$$\frac{n_2}{n_1} = \sqrt{n_g}.$$

It follows that the necessary value $\sqrt{n_g}$ for low-refracting glasses, such as $\sqrt{1.52} = 1.23$ with BK7, can be obtained by adjusting the ratio of the refractive indices n_2/n_1 of the layers. This is in propitious contrast to the case of single-layer coatings, where the refractive index itself has to be adapted; as we have seen, this may be thwarted by posed problem of available robust materials. For BK7 glass, suitable combinations are, for instance, n_1=1.38, n_2=1.70 (MgF_2/MgO), or n_1=1.65, n_2=2.1 (CeF_3/ZrO_2). However, the reflectance curves resulting from such anti-reflection layers are clearly V-shaped (as for single-layer coatings). Thus the AR effect applies to a relatively narrow wavelength range, resulting in the typical purple sheen.

Material	Material name	Refractive index n
Na_3AlF_6	Cryolite	1.35
MgF_2	Magnesium Fluoride	1.38
SiO_2	Silicon Dioxide	1.45
Si_2O_3	Disilicon Trioxide	1.55
CeF_3	Cerium Fluoride	1.63
Al_2O_3	Aluminum Oxide	1.65
MgO	Magnesium Oxide	1.70
Nd_2O_3	Neodymium Oxide	1.95
ZrO_2	Zirconium Oxide	2.05
CeO_2	Cerium Oxide	2.22
ZnS	Zinc Sulfide	2.36
TiO_2	Titanium Dioxide	2.32
ZnSe	Zinc Selenide	2.65

Fig. 2.26: Selection of materials for coatings in the visible spectrum.

The refractive indices of most coating materials fall into one of the following value bands: 1.35 – 1.45, then 1.63 – 1.76, and approx. 2 – 2.5 (see Fig. 2.26): Thus the quotients of values from adjacent bands range from 1.2 to 1.4 (i.e. from $\sqrt{1.5}$ to $\sqrt{1.8}$, approximately), as required by the condition for a low reflectance level.

A viable alternative to the layer structure just described is obtained by replacing the λ/4-layer that is in immediate contact with the glass surface by a λ/2-layer. At first, this may appear surprising, since for this thickness the light wave oscillation amounts to a complete wave cycle of the reference wavelength. Hence, the wave reflected from the glass interface is in phase with the wave reflected at the surface between the first and second layer. Consequently, as far as the reference wavelength is concerned, the reflectance of a "λ/4-λ/2-coating" is identical to that of the λ/4-single-layer coating, e.g. R=1.3% for MgF_2 on BK7. However, the additional λ/2-

coating lowers the reflectance in the blue and red range of the visible spectrum, provided that the index of refraction n_2 is chosen to be larger than n_g: the profile gets W-shaped (see Fig. 2.28). The mechanism of this "achromatizing" effect of the λ/2-coating is shown in Fig. 2.27: The phase shift it causes compensates the residual phase differences for red and blue produced by the λ/4-coating.

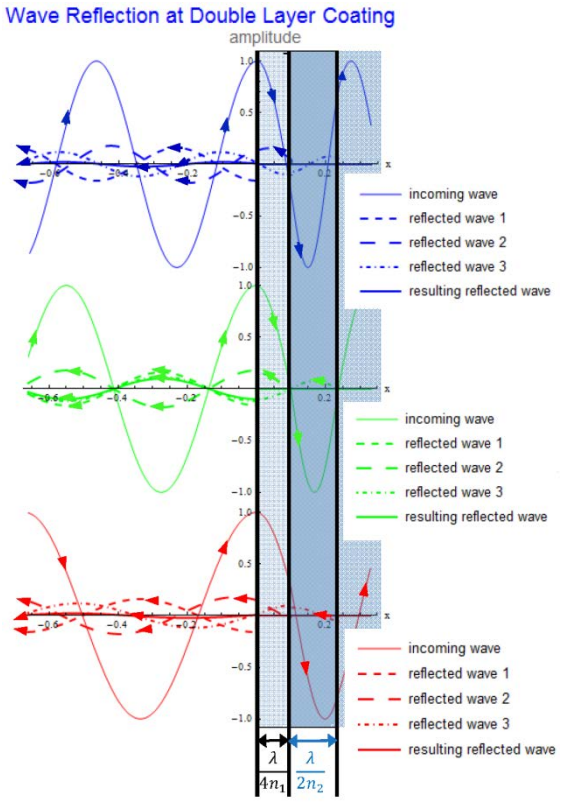


Fig. 2.27: Representation of the phase shift of the λ/4-λ/2-coating: The phase is compensated in the red and blue spectral range by the opposite sign; for the green reference wavelength λ_0, the λ/2-coating does not change anything: The reflectance at λ_0 is thus only defined by the λ/4-coating layer and the refractive index of the glass.

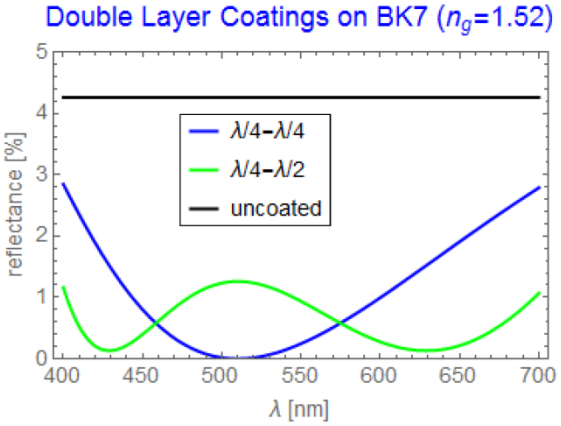


Fig. 2.28: Reflectance profile of a λ/4-λ/4 double-layer coating with n_1=1.38, n_2=1.7 and a λ/4-λ/2 double-layer coating with n_1=1.38, n_2=2.0, both on glass with n_g=1.52. Reference wavelength: λ_0=510nm.

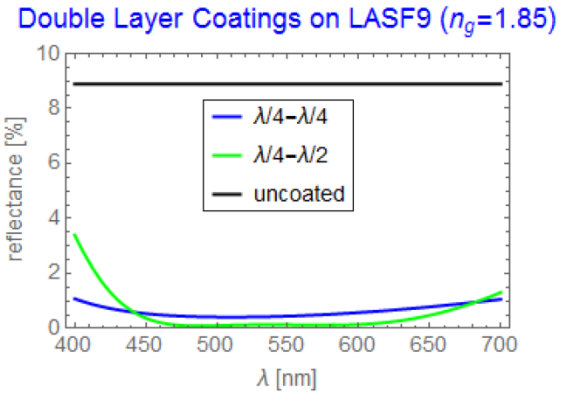


Fig. 2.29: Double-layer coatings on higher-refracting glass (n_g=1.85). Blue curve: λ/4-λ/4 double-layer coating with n_1=1.38, n_2=1.78. Green curve: λ/4-λ/2 double-layer coating with n_1=1.38, n_2=2.0. Reference wavelength: λ_0=510nm.

The triple-layer coating represents another major improvement for broadband anti-reflection coatings, especially for low-refracting glasses.

Triple-layers make broadband antireflection coatings possible

By extending the design described in the previous section by another coating layer, the triple-layer design is obtained, with "λ/4-λ/4-λ/4" and "λ/4-λ/2-λ/4" being of special importance.

In case of the λ/4-λ/4-λ/4-coating, the reflectance disappears at λ_0, if the following condition is fulfilled (see Appendix A.5):

$$\frac{n_1 n_3}{n_2} = \sqrt{n_g}.$$

This means there are even more possibilities to combine different coating materials as compared to the two-layer design (see Fig. 2.30 for two examples).

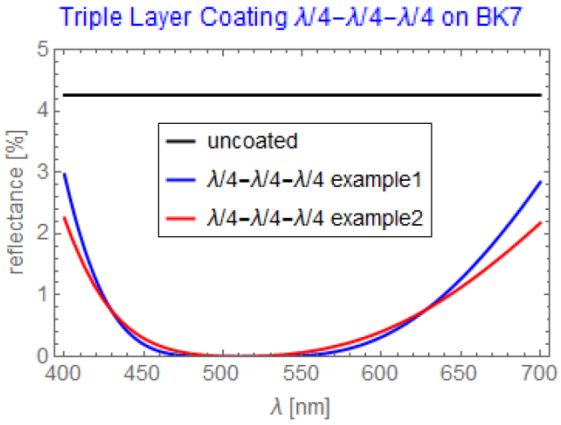


Fig. 2.30: Two "λ/4-λ/4-λ/4"-coatings on BK7 (n_g=1.52). Example 1: n_1=1.38, n_2=2.1 and n_3=1.88, reference wavelength λ_0=510nm; example 2: n_1=1.38, n_2=1.9, n_3=1.7.

For a λ/4-λ/2-λ/4-coating, the amplitude condition is:

$$\frac{n_3}{n_1} = \sqrt{n_g},$$

which is the same as for λ/4-λ/4 double-layer coatings. However, the additional λ/2-coating has a spectrum-broadening effect. Walter Geffcken had this λ/4-λ/2-λ/4-coating type patented as early as 1940 at SCHOTT in Jena. This λ/4-λ/2-λ/4-coating allows for achieving a good broadband antireflection for all glasses. Fig. 2.31 shows two examples on BK7.

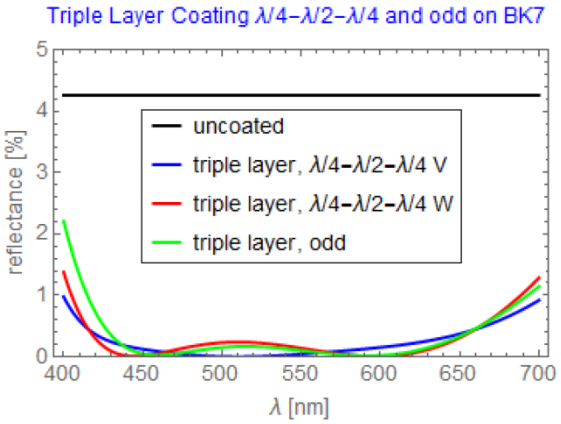


Fig. 2.31: Different triple-layer coatings on BK7 (n_g=1.52). The red and the blue curve depict examples of λ/4-λ/2-λ/4-coatings. Data of the V-shaped blue curve: n_1=1.38, n_2=2.15 and n_3=1.7, reference wavelength λ_0=510nm; data of the slightly W-shaped red curve: n_1=1.38, n_2=2.15, n_3=1.62. Data of the green curve (acc. to Thetford (1969)): n_1=1.38, n_2=2.1, n_3=1.8, d_1=567.2nm/(4n_1), d_2=212.3nm/(4n_2) and d_3=731.4nm/(4n_3).

The green curve in Fig. 2.31 refers to a design that was obtained by numerical optimization: its layer thicknesses clearly deviate from λ/4 or λ/2.

Fig. 2.32 shows another variant: a λ/4-λ/2-λ/2-coating on a high-index glass with n_g=1.85.

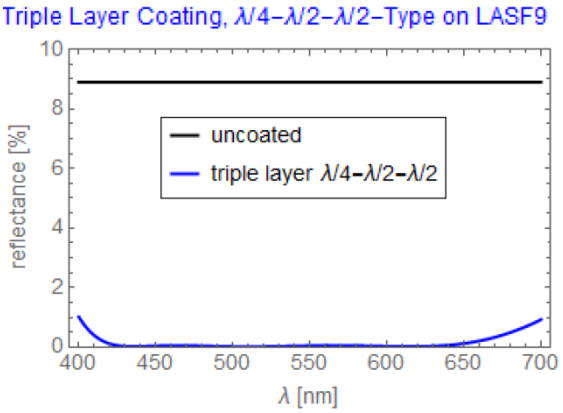


Fig. 2.32: A λ/4-λ/2-λ/2 triple-layer coating on high-index glass LASF9 (n_g=1.85): n_1=1.38, n_2=2.15, n_3=1.7. Reference wavelength λ_0=510nm.

These data demonstrate that with triple-layer coatings, a reflectance less than 0.4% can be achieved over the entire visible spectrum. Equally important is the performance of anti-reflection coatings for incidence angles other than 0°. Fig. 2.33 shows the reflectance profiles of a λ/4-λ/2-λ/4-coating for different incidence angles up to 45°.

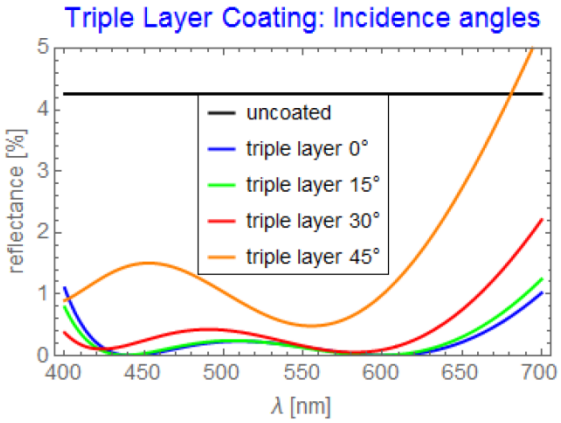


Fig. 2.33: Reflectance profiles of a λ/4-λ/2-λ/4 triple-layer coating (reference wavelength λ_0=510nm) with n_1=1.38, n_2=2.15 and n_3=1.62 on BK7 (n_g=1.52) for different angles of incidence.

In the section "Single-layer coatings" we saw that thickness variations of the coatings lead to color shifts. Multi-layer coatings are typically even more sensitive in this regard. Fig. 2.34 shows the impact of thickness deviations of a triple layer. Too large a thickness leads to increased reflectance values for blue light, whereas layers that are too thin reflect red light more strongly.

Fig. 2.35 shows the reddish color change of a surface with a coating that is thinner than required as compared to the greenish sheen of a coating of correct thickness.

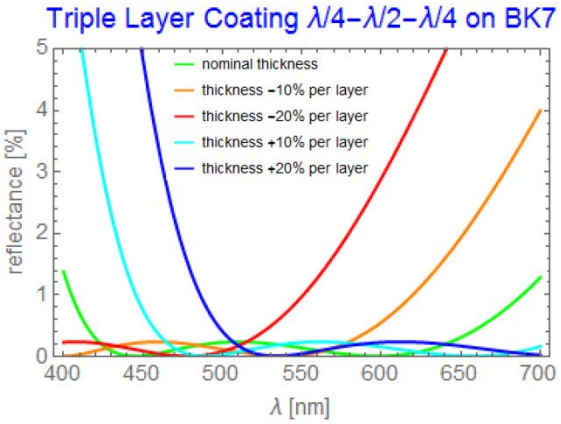


Fig. 2.34: Reflectance profiles of a triple-layer coating for layer thicknesses deviating from the nominal values.

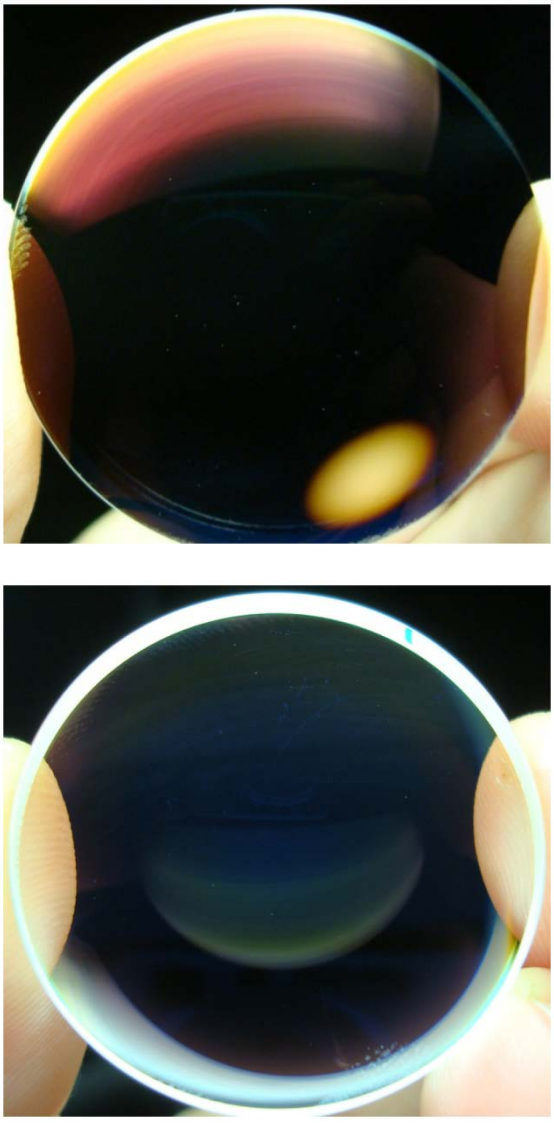


Fig. 2.35: Top: A too thin coating appears reddish; bottom: coating of correct thickness (slightly greenish, since reflectance has „W-profile".)

Since the 1990s there have been coating plants with planetary gearing for moving substrates on planetary tracks. Those special coating plants allow one to apply homogeneous coatings even to lenses of strong curvature (cf. Fig. 2.36).

Fig. 2.36: Vacuum coating plant with planetary gearing to apply very homogeneous coating thicknesses.

Complex multi-layer coatings for high-quality camera lenses

Fig. 2.37 shows a synopsis of the reflectance curves of single-layer, double-layer, and triple-layer coatings on BK7 (n_g=1.52).

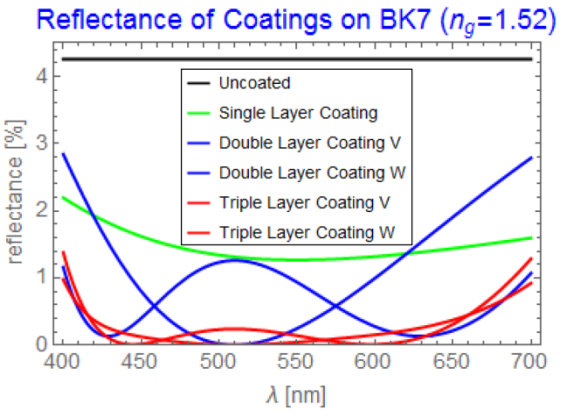


Fig. 2.37: Synopsis of the coatings described so far on low-index glass (n_g=1.52).

The corresponding figure for high-index glass (n_g=1.85, see Fig. 2.38), however, shows a stronger suppression of reflectivity even for the single-layer coatings. When adding a few further coating layers, an excellent broadband antireflection effect is achieved.

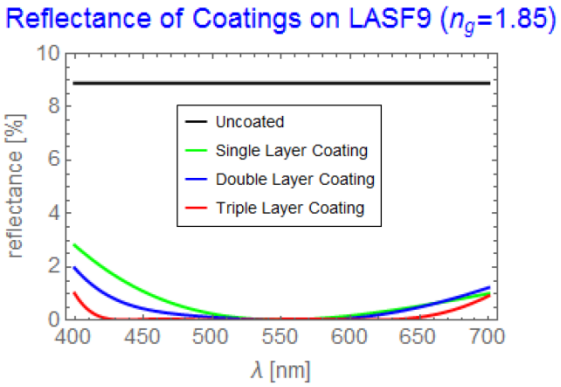


Fig. 2.38: Synopsis of coatings designs on high-index glass with n_g=1.85.

The following figures show the profiles of the same coating designs over a larger wavelength range from 250nm up to 1500nm.

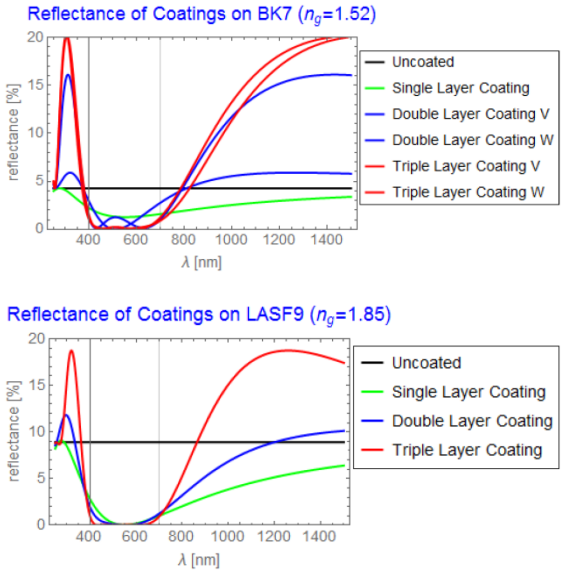


Fig. 2.39: The same coatings shown over a wavelength range that includes UV and IR.

The strongly increasing reflectance values from the visible to the ultraviolet and infrared wavelength ranges (Fig. 2.39) make clear why an extension of antireflection from the visible spectral range to the UV or IR range leads to coating layer systems that are considerably more complex.

Some materials of individual coating layers are only suitable for certain manufacturing processes. They may not lend themselves, for example, to thermal evaporation. Such materials may still be considered in the early stages of coating design processes to be replaced later on by more robust materials with equivalent reflection reduction properties. As a result, one obtains coatings consisting of a larger number of layers and a smaller number of different materials. For example, Thelen (1969), Thelen (1988), and Furman, Tikhonravov (1992) have given systematical methods in this regard in the theory of thin films. Dobrowolski (1997) provides an overview of numerical optimization methods. A chronology of antireflection coating optimization methods can be found in the work presented by Schallenberg (2006).

AR coatings with seven or more layers may result in reflectance profiles that are as low as 0.1 to 0.2 percent over a large portion of the visible spectrum (see Fig. 2.40). As a consequence, the intensity of glass-glass reflections is reduced by a factor of roughly 1000 as compared to uncoated lenses. At the same time, the reflectance for larger incidence angles is significantly reduced.

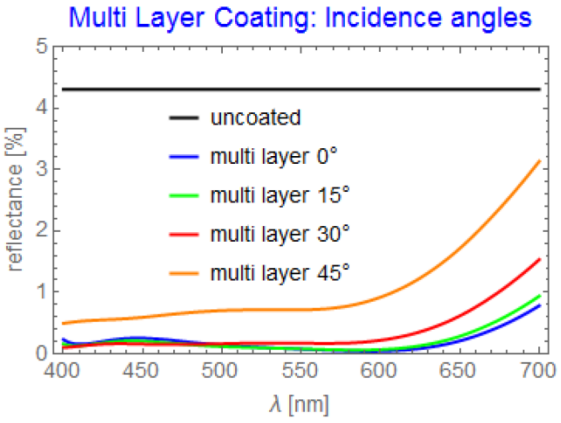


Fig. 2.40: Reflectance curves of a coating consisting of 7 layers on BK7 (n_g=1.52) at different angles of incidence.

Such complex coating systems often have an alternating "baumkuchen" (layer-cake-like) structure. This means they consist of an alternating sequence of low-index layers (e.g. magnesium fluoride) and high-index materials (e.g. hard oxides). The thicknesses of the individual coating layers are numerically optimized and usually differ significantly from each other (for examples of such multilayer coating designs see Cushing (2011) or Macleod (2001)). In coating simulation programs, the propagation of the electrical field components of light waves is computed layer by layer by means of the Fresnel equations (see Appendix A.3). Coating designs are optimized, typically, towards a desired spectral reflectance specification, by variation of the layer thicknesses and refractive indices (materials). Multiple reflections and material absorption in the coatings are taken into account. Production-related boundary conditions can be defined for the thicknesses of single layers and material types.

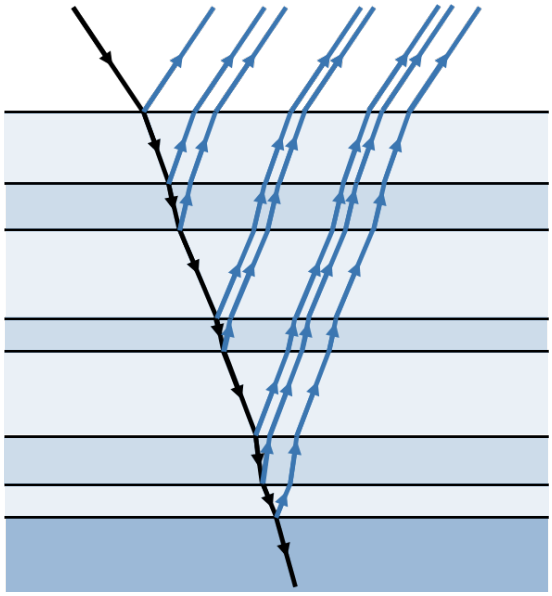


Fig. 2.41: Multi-layer coating with typical "baumkuchen" structure. The individual coating layers are typcially of varying thickness. Only single reflections are shown here.

Apart from the layer-cake structure type (alternating high-index and low-index layers), the multi-layer coating may be designed in such a way that materials with increasing refractive indices are used. The lowest-index material is in contact with air, the highest-index does not exceed the refractive index value of glass. Lord Rayleigh has shown theoretically that reflectance may be reduced to zero when a continuous gradient coating is utilized (Rayleigh (1880)). For infrared glasses with very high refractive indices in the range of 3.5 to 4, it suffices to place a few layers of conventional coating materials in such a graduated arrangement to achieve considerably improved transmission. Even a double-layer coating with magnesium fluoride and a heavy oxide will do the trick.

For glass types with lower refractive indices in the visible spectrum, this is more difficult to achieve. This is, again, due to the fact that coating materials having at the same time the required indices of refraction and the necessary resiliency are not readily available. Recently, though, GRIN (graded refractive index) materials have been developed that were inspired by nano-structured moth eyes (cornea of the night moth). With nano-porous MgF_2 layers, index values ranging from von n = 1.33 at the glass substrate to n=1.16 (top layer) have been realized (Bruynooghe, Tonova, Sundermann, Koch, Schulz (2014)).

So far, we have presented the functional principles underlying antireflection coatings and given some design examples. When we direct our attention to the quality of coatings, production-related aspects play an important role. Among those the following figures prominently: cleanliness of the coating process, realization of very small production tolerances, durability, robustness (see fig. 2.42) and lifetime of the coating under thermal and mechanical stress as well as humidity.

Today, there are many different production processes for coatings. Comprehensive monographs dealing with production processes are, for example, Baumeister (2004) Pulker (1999), Macleod (2001), Knittl (1976) and Anders (1967). The book by Bliedtner, Gräfe (2010) contains a DVD with a film on coating processes.

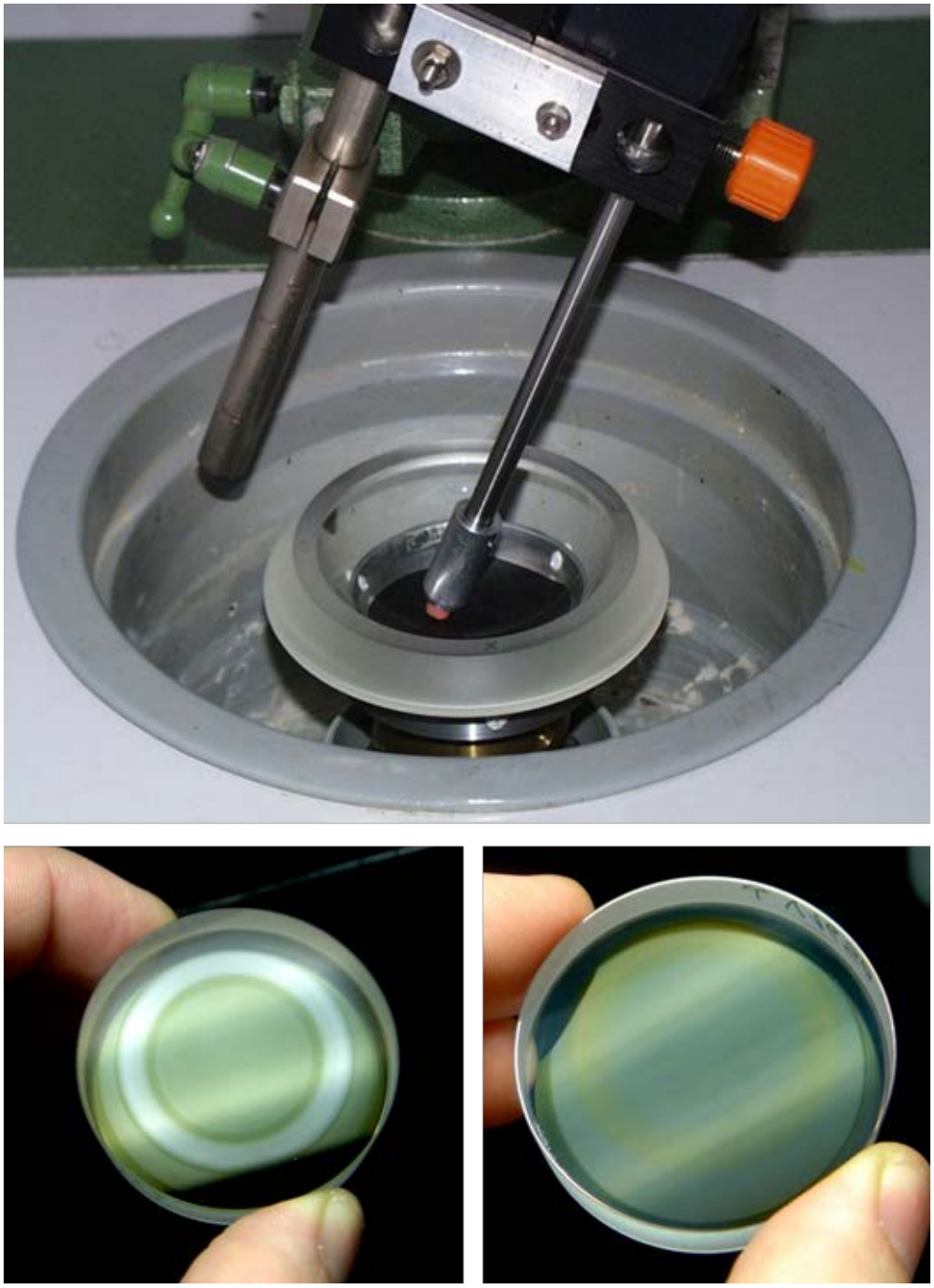


Fig. 2.42: Top: In order to test robustness of the surface it is chafed with an eraser under pressure. Bottom: abrasion of different test coatings after 500 cycles. (Image by courtesy of Thomas Rittmann.)

Simulation and optimization of ghost images in optical systems

Good coatings are a basic prerequisite for the design of low-reflection systems. Conversely, coating of individual optical elements alone need not suffice to guarantee the absence of nuisance reflections. The quality of the reflectivity properties of a lens can, however, be significantly improved, if the evaluation of stray light and ghost images is integrated systematically into the lens design process. Whenever a ghost analysis reveals any dominant reflections, the optical design has to be modified.

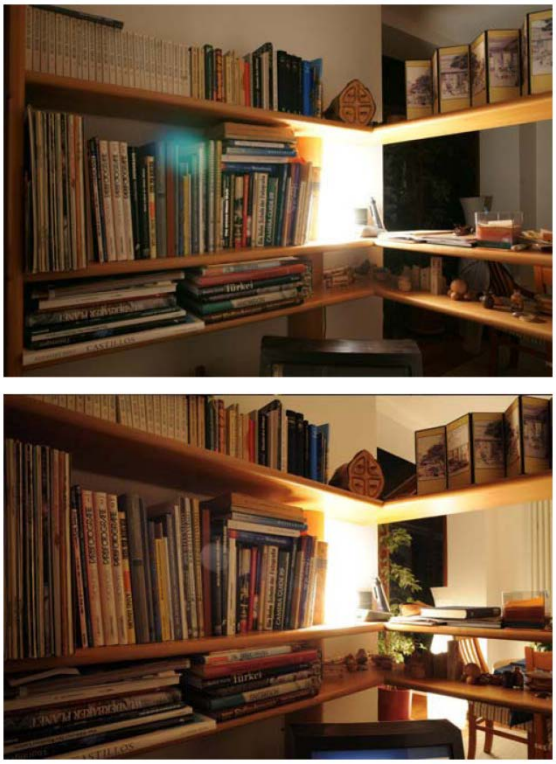


Fig. 3.1: Picture of the same scene taken with two different lenses. The top image (taken with Lens A) displays a striking green reflection, whereas there is virtually no stray reflection with Lens B in the bottom picture. (Images by courtesy of Dr. H. Nasse.)

The reflectivity properties of different lenses may vary considerably (even across lenses of similar design). This is exemplified in Fig. 3.1 showing images of a scene taken with two lenses: lens A produces a strong green reflection, while lens B is inconspicuous.

We present a simplified analysis: We assume that a light ray traversing an optical system and hitting the image sensor is reflected from there with a reflectance of 5%. From the sensor it reverts into the optical system and eventually returns onto the image plane after reflection at a multilayer-AR-coated optical surface (R = approx. 0.2%). The twice-reflected ray then has a relative intensity of 0.002 x 0.05 = 0.0001. The intensity ratio for the "useful ray" vs. the reflected ray is therefore $1:0.0001=10^{-4}$ or $10^{4}:1$, or approximately 13 exposure values.

When extending this single-ray analysis to a whole bundle of rays emanating from an object point and hitting the image plane, we obtain a realistic estimate of the maximum possible relative intensity due to one single reflection: If all reflected rays of the bundle were focused in one point in the image plane, the relative intensity of this reflection would actually amount to approximately 13 exposure values. However, if the reflected light converges to a point lying in front of or behind the image plane, the reflections in the image plane are very much weaker, because the light is distributed over a larger surface area. This can be controlled during development of a lens: One tries to avoid ghost images that are focused in the image plane.

In modern optical designs, dozens of such image sensor/glass reflections and hundreds of glass/glass reflections exist. However, it is generally not the case that all of the reflected light reaches the image plane; a number of light rays leaves the optical path and is absorbed by mechanical lens components or the camera body.

As was elaborated in the first section of this article, the distribution of reflection intensities is often very inhomogeneous. This is due to significant optical aberrations along the

ghost ray paths. Those may result for instance, in local dot-like or crescent-shaped light concentrations. In addition, strong reflections can arise even outside the nominal ray path owing to large incidence angles and/or due to total internal reflection. In view of the complexity of reflection phenomena, fairly comprehensive optical simulations are required, which we will describe in more detail in the following sections.

The number of double reflections in the image plane and their intensity

As we have seen, light splits up at every refractive surface into a transmitted part and a reflected part. When either part reaches some other refractive surface, it will, again, be partly reflected and transmitted, and so on. In this way, a substantial number of ghost ray paths is created. Fig. 3.2 shows multiple reflection paths up to four-fold reflections for a simple lens. Only paths with an even number of reflections (2x, 4x, 6x,…) are eventually headed toward the image plane and may disturb the picture. Any paths with an odd number of reflections (1x, 3x, 5x,…) are eventually reverted back into object space and are thus irrelevant as far as image quality is concerned.

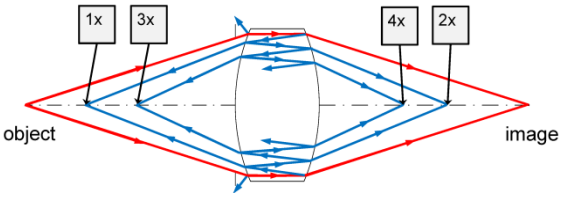


Fig. 3.2: With a refractive optical system (i.e. one consisting exclusively of surfaces whose function per design is to refract and not to reflect light) – a thick lens in this example – only light ray paths with an even number of reflections can reach the image space.

For a camera lens with two lens elements, i.e. four optical interfaces, there are glass-glass reflections between the following surface pairs (see Fig. 3.3):

4-3, 4-2, 4-1,

3-2, 3-1,

2-1.

This amounts to a total of 3+2+1=6 combinations of reflections. Generally, a lens with n optical interfaces will have

$$N_{2R,G-G} = 1 + 2 + \ldots + (n-1) = \frac{n(n-1)}{2}$$

glass-glass reflections.

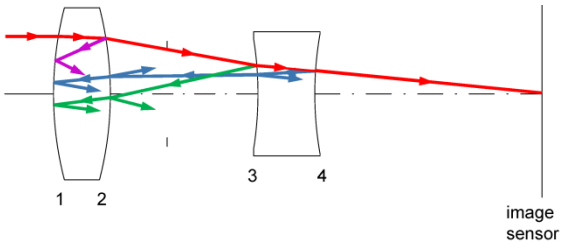


Fig. 3.3: Glass-glass reflections of a camera lens with two lens elements.

Furthermore, light reflected in the image plane, i.e. at the image sensor or film, may return into the optical system and be reflected at each of the n optical surfaces back towards the image plane (see Fig. 3.4). This results in n additional (image sensor / glass) reflections.

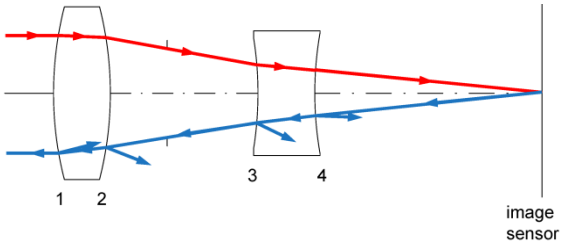


Fig. 3.4: Image sensor / glass reflections for a camera lens with two lens elements.

The total number $N_{2R,total}$ of light paths with two reflections (between any pair of optical surfaces including the image sensor) is therefore

$$N_{2R,total} = 1 + 2 + \ldots + n = \frac{n(n+1)}{2}.$$

The dominant term in this expression is $n^2/2$. Hence, to a good approximation, the number of reflections rises quadratically with the number of surfaces. Fig. 3.5 shows the number of reflections against the number of optical surfaces; some values that correspond to the number of optical surfaces of some ZEISS camera lenses are specially labelled. While a Tessar lens with 6 air-glass surfaces has 15 glass-glass ghost ray paths, some zoom lenses exhibit around 1000 ghost ray paths.

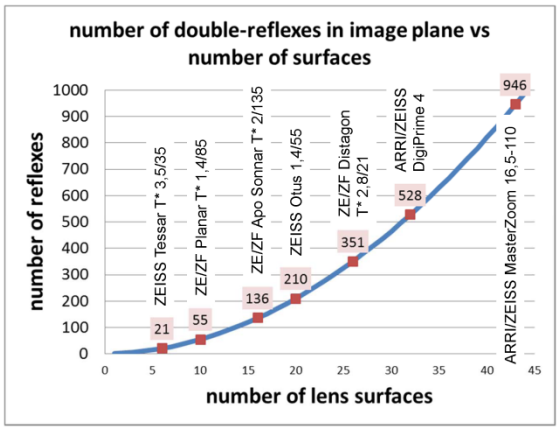


Fig. 3.5: Number of reflections as a function of the number of refractive lens surfaces. The number of surfaces (air-glass) of some ZEISS camera lenses is indicated.

For any lens, the number of reflections involving the image sensor is much lower than that of reflections involving glass surfaces in the lens only. Albeit, the intensity of reflections involving the image sensor are, on average, greater by one order of magnitude and therefore require particular attention. Image sensors often have reflectance values of about 5%, with some variations among the different models and manufacturers. Film, too, reflects about 5% of the incoming light (see Fig. 3.6).

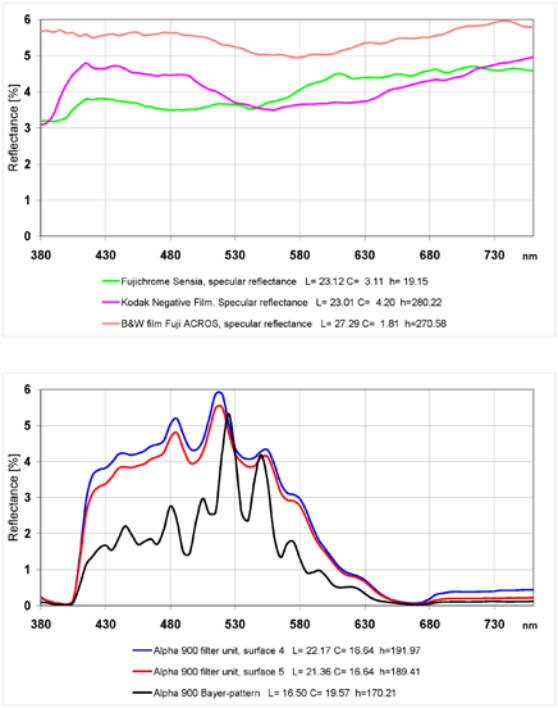


Fig. 3.6: Top: reflectance of different photographic films; bottom: reflectance curves of the image sensor of a digital camera.

By way of illustration, if a coated glass surface has a reflectance of 0.2%, any reflections involving the image sensor will be greater in magnitude by a factor of about 25 as compared to the glass-glass reflections (5% x 0.2% vs. 0.2% x 0.2%).

number of reflections	uncoated lens	single-layer coating	multi-layer coating
0	1	1	1
1	$5.00 \cdot 10^{-2}$	10^{-2}	$2.0 \cdot 10^{-3}$
2	$2.50 \cdot 10^{-3}$	10^{-4}	$4.0 \cdot 10^{-6}$
3	$1.25 \cdot 10^{-4}$	10^{-6}	$8.0 \cdot 10^{-9}$
4	$6.25 \cdot 10^{-6}$	10^{-8}	$1.6 \cdot 10^{-11}$
5	$3.13 \cdot 10^{-7}$	10^{-10}	$3.2 \cdot 10^{-14}$
6	$1.56 \cdot 10^{-8}$	10^{-12}	$6.4 \cdot 10^{-17}$

Fig. 3.7: Relative intensities of reflections for multiple glass-glass reflections with a reflectance per refractive surface of 5% (uncoated), 1% (single-layer coating) and 0.2% (multi-layer coating).

number of reflections	uncoated lens	single layer coating	multi layer coating
0	1	1	1
1	0.05	0.05	0.05
2	$2.5 \cdot 10^{-3}$	$5 \cdot 10^{-4}$	$1 \cdot 10^{-4}$
3	$1.25 \cdot 10^{-4}$	$2.5 \cdot 10^{-5}$	$5 \cdot 10^{-6}$
4*	$6.25 \cdot 10^{-6}$	$2.5 \cdot 10^{-7}$	$1 \cdot 10^{-8}$
4**	$6.25 \cdot 10^{-6}$	$5 \cdot 10^{-8}$	$4 \cdot 10^{-10}$

*: 2 reflections at image sensor; **: 1 refl. Image

Fig 3.8: Relative intensities of reflections for multiple image sensor/glass reflections with an assumed reflectance of the image sensor of 5%.

Twofold and fourfold reflections

According to Fig. 3.7, the intensity of light that is reflected at four coated surfaces is more than four orders of magnitude lower than the intensity of twice reflected light. In other words, the intensity of light reaching the sensor after a fourfold reflection is much weaker than that due to a double reflection. Fourfold reflections thus result in a quasi-continuous straylight distribution as opposed to the structured ghost images that are due to twofold reflections.

On the other hand, the number of light paths with four reflections is much higher, roughly proportional to the fourth power of the number of surfaces. The exact expression is

$$N_{4R} = \frac{1}{24}(5n^4 - 10n^3 + 7n^2 - 2n).$$

The intensity of reflection multiplied with the number of reflections gives an approximate measure for the (average global) straylight level. Fig. 3.9 summarizes the number of reflections N_{2R} and N_{4R}, the reflection intensity r^2 and r^4 as well as the average straylight level $r^2 \cdot N_{2R}$ and $r^4 \cdot N_{4R}$, for uncoated surfaces (assumed reflectance value r=5%), single-layer coated surfaces (r=1% per surface), and multi-layer coated surfaces (r=0.2% per surface).

number of surfaces	N_{2R}	r^2	$r^2 \cdot N_{2R}$	N_{4R}	r^4	$r^4 \cdot N_{4R}$	$(r^4 \cdot N_{4R})/(r^2 \cdot N_{2R})$	
10	55	$2.5 \cdot 10^{-3}$	0.1375	1695	$6.25 \cdot 10^{-6}$	0.0106	0.0770	uncoated
20	210	$2.5 \cdot 10^{-3}$	0.525	30115	$6.25 \cdot 10^{-6}$	0.1882	0.3585	
30	465	$2.5 \cdot 10^{-3}$	1.1625	157760	$6.25 \cdot 10^{-6}$	0.9860	0.8482	
40	820	$2.5 \cdot 10^{-3}$	2.05	507130	$6.25 \cdot 10^{-6}$	3.1696	1.5461	
50	1275	$2.5 \cdot 10^{-3}$	3.1875	1250725	$6.25 \cdot 10^{-6}$	7.8170	2.4524	
10	55	$1 \cdot 10^{-4}$	0.0055	1695	$1 \cdot 10^{-8}$	0.00002	0.0031	single-layer coating
20	210	$1 \cdot 10^{-4}$	0.021	30115	$1 \cdot 10^{-8}$	0.0003	0.0143	
30	465	$1 \cdot 10^{-4}$	0.0465	157760	$1 \cdot 10^{-8}$	0.0016	0.0339	
40	820	$1 \cdot 10^{-4}$	0.082	507130	$1 \cdot 10^{-8}$	0.0051	0.0618	
50	1275	$1 \cdot 10^{-4}$	0.1275	1250725	$1 \cdot 10^{-8}$	0.0125	0.0981	
10	55	$4 \cdot 10^{-6}$	0.00022	1695	$1.6 \cdot 10^{-11}$	$2.71 \cdot 10^{-8}$	0.0001	multi-layer coating
20	210	$4 \cdot 10^{-6}$	0.00084	30115	$1.6 \cdot 10^{-11}$	$4.82 \cdot 10^{-7}$	0.0006	
30	465	$4 \cdot 10^{-6}$	0.00186	157760	$1.6 \cdot 10^{-11}$	$2.52 \cdot 10^{-6}$	0.0014	
40	820	$4 \cdot 10^{-6}$	0.00328	507130	$1.6 \cdot 10^{-11}$	$8.11 \cdot 10^{-6}$	0.0025	
50	1275	$4 \cdot 10^{-6}$	0.0051	1250725	$1.6 \cdot 10^{-11}$	$2.00 \cdot 10^{-5}$	0.0039	

Fig. 3.9: Number of reflections, intensity of reflection and their product (approximate average straylight level) for uncoated, single-layer coated and multi-layer coated surfaces.

Fig. 3.10 shows the intensity ratios of fourfold vs twofold reflections against the number of surfaces. Intensity ratio = (intensity of fourfold reflection * number of fourfold reflections) divided by (intensity of twofold reflection * number of twofold reflections), i.e. $(r^4 \cdot N_{4R})/(r^2 \cdot N_{2R})$.

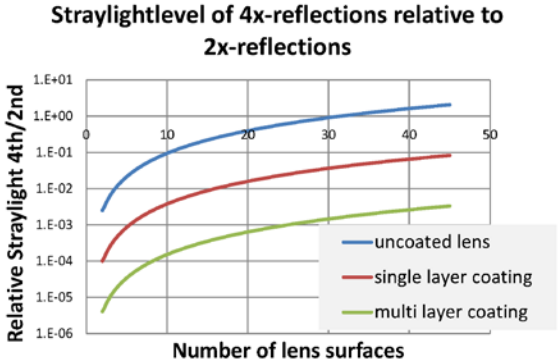


Fig. 3.10: Ratio of the straylight levels of fourfold vs. twofold reflections $(r^4 \cdot N_{4R})/(r^2 \cdot N_{2R})$.

The more surfaces a system has and the higher the reflectance values are, the greater is the influence of fourfold reflections on the average straylight level. We have determined the number of surfaces for which the straylight level due to fourfold reflections is equal in magnitude to that of twofold reflections: 31 surfaces for an uncoated optical system, 156 surfaces for single-layer coatings, and 775 surfaces for multi-layer coatings. In a (fairly large) optical system with 31 refractive surfaces, double reflections outweigh fourfold reflections, in terms of straylight intensity, by a factor of 25 for single-layer coatings and 625 for multi-layer coatings. Therefore we may (fortunately) neglect fourfold reflections in straylight simulations, otherwise computational complexity (which is already substantial) would become prohibitively large.

Irradiance of non-focused reflections

So far, in our discussion of reflection intensities we have only considered the reflectance values of the surfaces involved. The size of any given ghost image and its distribution in the image plane are relevant factors that impact its intensity. The more distant the focus of the ghost ray path is from the image plane, the weaker is the reflection. The irradiance of any given ghost image is inversely proportional to its surface area (see Fig. 3.11).

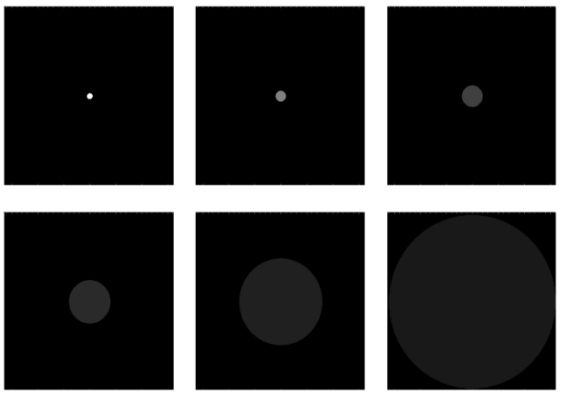


Fig. 3.11: The irradiance of reflections decreases, the farther away they are focused from the image plane. The pictures show the respective brightness value (proportional to $\log_2$(surface area)) with a defocusing is successively doubled over the image sequence.

Ghost ray paths in camera lenses

Not all the light that is reflected at glass surfaces gets back onto the image plane. Large part of the reflected light impinges, after reflection of the lenses, on mounting parts, the iris stop or, in the image space, on areas outside the image sensor. For many camera lenses, on average only about 25% of the light reflected from an optical surface reaches the image sensor, when a light source is present in the center of the image field. If the light source is, however, located at the periphery of the field of view, only about 5% of the reflected light reaches the image sensor.

The values stated here are only approximate averages taken over all optical surfaces of the lens: There may, however, be individual optical surfaces, for which the reflected light reaches the image plane in its entirety, and other optical surfaces, for which no light at all reaches the image sensor. Figures 3.12, 3.13, 3.14 show ghost ray paths for light that comes from three distant light sources, located in the center and at the periphery of the field, respectively.

Fig. 3.12 shows the ghost ray paths with reflections at surfaces 9 and 7. In this case, the light from all source positions reaches the image plane after the two reflections. The focus of the reflected light lies in front of the image plane, which means it reaches the image plane out-of-focus and is thus weakened in intensity.

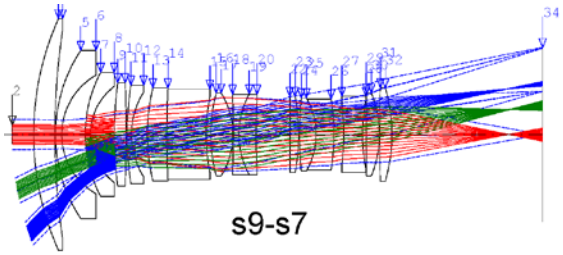


Fig. 3.12: Ghost ray paths from the optical surfaces 9 and 7 and nominal light paths (dashed) for light that comes from three distant light sources (red, green and blue).

Fig. 3.13 shows the ghost ray paths for reflections at the adjacent surfaces 9 and 8: Only the light from the light source in the center of the field of view reaches the image plane. This light is spread over large areas of the image plane and is therefore very weak. Any light emanating from light sources located at the field periphery is intercepted at mechanical mounting parts and does not reach the image plane.

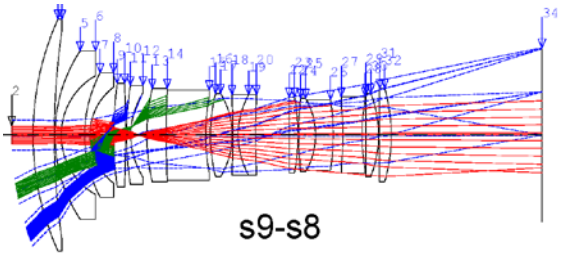


Fig. 3.13: Ghost ray paths with reflections at surfaces 9 and 8.

The combination of surfaces 28 and 13 is similar to the previous case: The reflected light from the light source in the center of field is strongly weakened, and no light from the peripheral light sources reaches the image plane (Fig. 3.14).

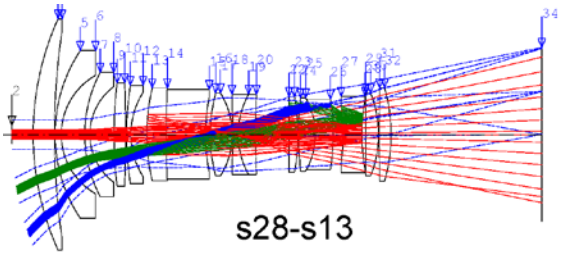


Fig. 3.14: Ghost ray paths between surfaces 28 and 13.

For most SLR cameras, the angles at which chief rays impinge on the image plane increase towards the periphery of the field, i.e. the chief rays are “non-telecentric” (see Blahnik (2014)). As a consequence, sensor-glass reflections can occur only in a certain region near the center of field: With increasing distance from the center, an ever smaller part of the light cone impinging on the image is reflected back into the lens. Beyond a certain image height, rays are not reflected back into the image plane (see Fig. 3.15).

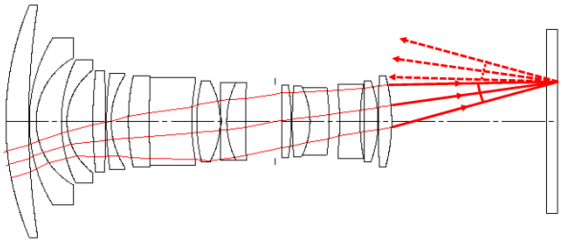


Fig. 3.15: In many lenses, the light rays heading towards the periphery of the field impinge non-telecentrically on the image sensor. Beyond a certain image height, there are no more sensor-glass reflections.

Apart from optical elements, mechanical components like mounting parts, the iris stop or lens bevels may also cause straylight in the image plane. During the pre-development phase and during prototyping, critical surface and light path combinations can be identified. For those surfaces straylight distributions are computed with the aid of opto-mechanical simulation models. Such calculations are very time-consuming, because the components are modeled as CAD data and often consist of thousands of surfaces. A detailed exposition of straylight modeling of optical systems is provided by Fest (2013).

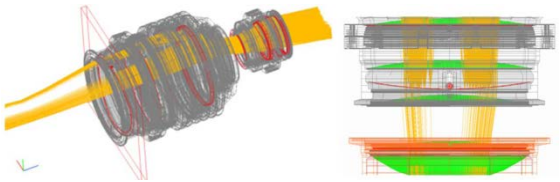


Fig. 3.16: Simulation of ghost ray paths by reflection and scattering at mechanical components modeled as CAD data.

These straylight calculations show possible weak points and give hints for the optimization of the geometries of mounting parts. Generally, high-quality lacquers help minimize straylight emerging from the marginal lens areas (see Fig. 3.17).

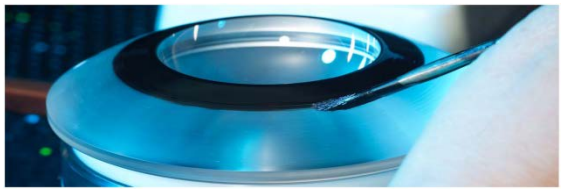


Fig. 3.17: A substantial amount of light escapes from the optical domain via lens edges and mounting parts. High-quality lacquers at the lens edges minimize straylight from the marginal lens areas.

In addition to reducing reflections, antireflection coatings have an influence on the transmission properties of lenses. Let R denote the reflected portion of the light per surface and n the number of refractive lens surfaces. Then the portion of light transmitted to the image plane is

$$T = (1 - R)^n.$$

Fig. 3.18 shows the relative transmission in f-stop units as a function of the number of optical surfaces. When we took the pictures shown in the first chapter, with the uncoated demonstration lens Distagon 2,8/21 ZE (26 glass-air transitions), we had to adjust exposure time as follows: In order to capture the bright areas in the scene with the same brightness as with the corresponding T*-coated lens, exposure had to be increased by (an equivalent of) two f-stops. This is in good accordance with the data shown in Fig. 3.18.

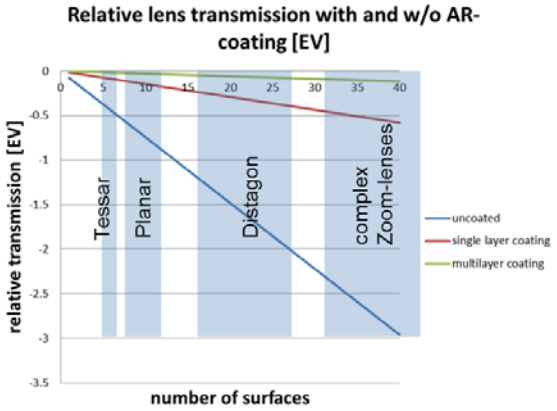


Fig. 3.18: Transmission loss in terms of exposure values (f-stops) for uncoated, single-layer and multi-layer coated lenses. For different lens types, typical approximate numbers of air-glass surfaces are indicated.

Simulation, analysis and optimization of ghost images in camera lenses

The simulation of ghost images supports optical design during the development stage, since it reveals optical surfaces that contribute significantly to reflections. It also serves to identify the origin of reflections that one may observe during prototyping (see Fig. 3.19). Ghost simulations require extensive ray tracing to determine the intensity distributions to a sufficient degree of accuracy.

In order to save computing time, one may attempt to perform reflection analyses via paraxial calculation that involve (reciprocal) linear expressions. In doing so one has to accept strong simplifications. Such calculations were performed even before computers existed: Goldberg (1925) analyzed the stray-light behavior of various lens types theoretically and practically. His calculations yielded worst-case estimates for small apertures near the center of field. However, such calculations no longer hold for larger apertures or for points at the periphery of the field. Many potential problems such as reflections that are due to large incidence angles, total reflection or ghost ray paths beyond the nominal optical path as well as the light concentration in caustics cannot be identified by paraxial calculations. Also, paraxial calculations tend to overestimate problems in cases when a lot of reflected light escapes from the optical system, e.g. towards mounting parts. The only viable way to achieve valid estimates requires time-consuming simulations of the real rays paths.

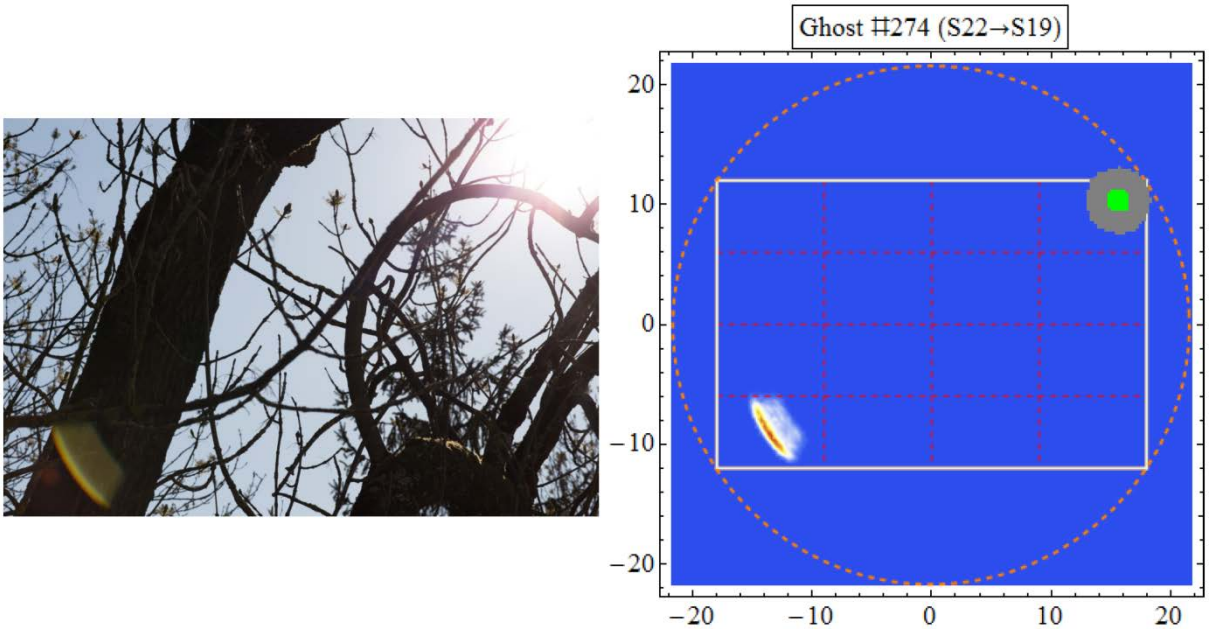


Fig. 3.19: Identification of the origin of a real reflection (bottom left) by means of a simulation. From the simulation data of all surface combinations it can be concluded which surface combination has led to the observed reflection (here: surfaces 22 and 19).

Fig. 3.20 shows an analysis of all ghost images of a lens with 31 refractive surfaces. Each square shows the image intensity distribution due to the reflections of one glass/glass or glass/sensor combination. The total of 32 surfaces (including the image sensor) results in 31*32/2=496 combinations that are shown in the matrix.

It can be seen that the image sensor/glass reflections (displayed in the last row of the matrix) have a higher intensity, since the image sensor reflects much more intensely than the glass surfaces.

For further literature on analysis and modeling of ghost images in camera lenses, see Gross, Blechinger, Achtner (2008), chapter 39, and Abd El-Maksoud, Sasian (2011).

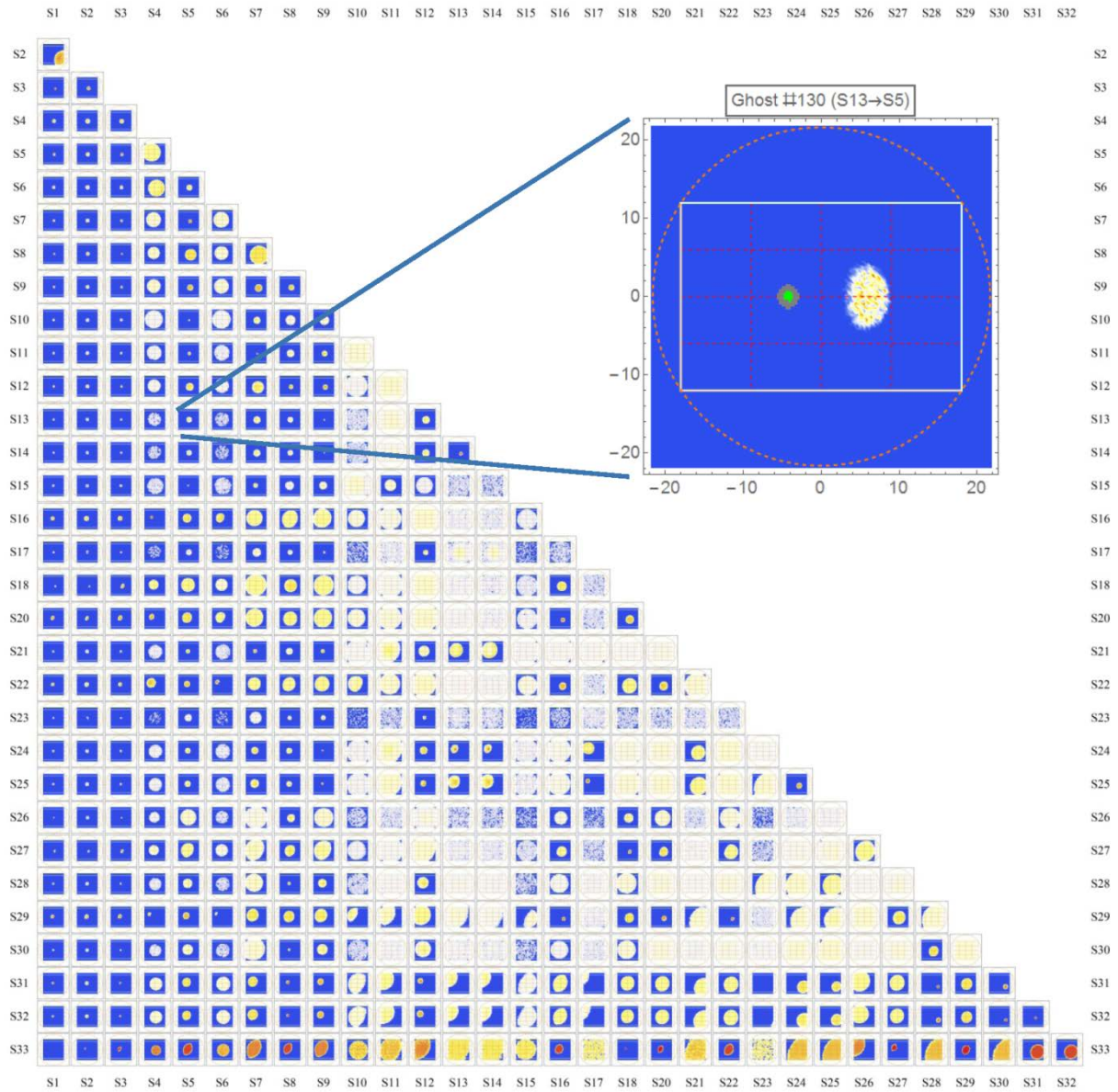


Fig. 3.20: Distribution of the intensity of individual reflections in the image plane of all glass/glass and image sensor/glass combination in the lens. The bottom row of the matrix shows the image sensor/glass reflections that have a clearly higher individual intensity than the other glass/glass reflections.

Computational costs of ghost image analysis

We consider a zoom lens with 26 refractive surfaces corresponding to n·(n+1)/2 = 351 surface combinations. We assume that we analyze this lens with the following set of parameters:

- 5 different values of object distance
- 7 focal lengths
- 3 positions of the light source
- 3 wavelengths

The number of ghost ray paths to be calculated for a ray from the light source is equal to the product of these numbers, i.e. (351·5·7·3·3) = 110,565.

To determine the distribution of intensity for each of these ghost ray paths with the required precision, the path of a sufficient number of rays through this lens has to be calculated. For ray tracing, the ray intersection points and angles at each surface are to be determined successively for all optical surfaces of the system. In order to do so, Snell's law of refraction,

$$n \sin i = n' \sin i',$$

is applied to each refractive surface. Fig. 3.21 shows the equation set for the calculation at a spherical surface in the tangential plane.

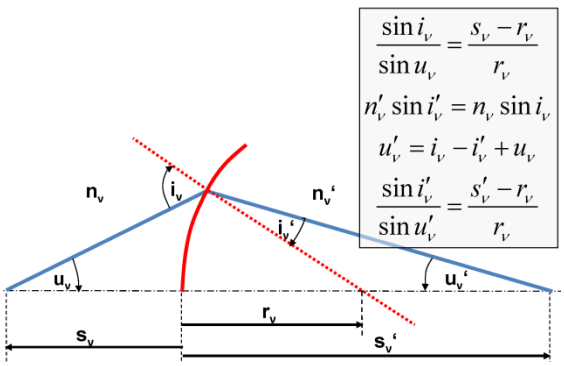


Fig. 3.21: Formula set for ray tracing at a spherical surface in the tangential plane.

If we choose 300 x 300 (i.e. about 100,000) rays emanating from the light source and passing into the camera lens, we obtain a total of 10 billion rays to trace for an analysis of the entire lens! This huge number of 10 billion rays takes about one day with the computing power available today. For a long time in the past, ghost simulations of such dimensions were not feasible; as recently as at the beginning of the 1980s, the computation times of the available computers used to be 100,000 times longer than today. That is this computational problem would have taken about 300 years with the computers available in the 1980s.

Although the theories of image formation were formulated for a large range of applications in the 1950s, only a few concrete examples of application could actually be calculated at that time. Today it is possible to calculate realistic ghost images of complex camera lenses taking into account the residual reflectance of antireflection coatings (see Fig. 3.22).

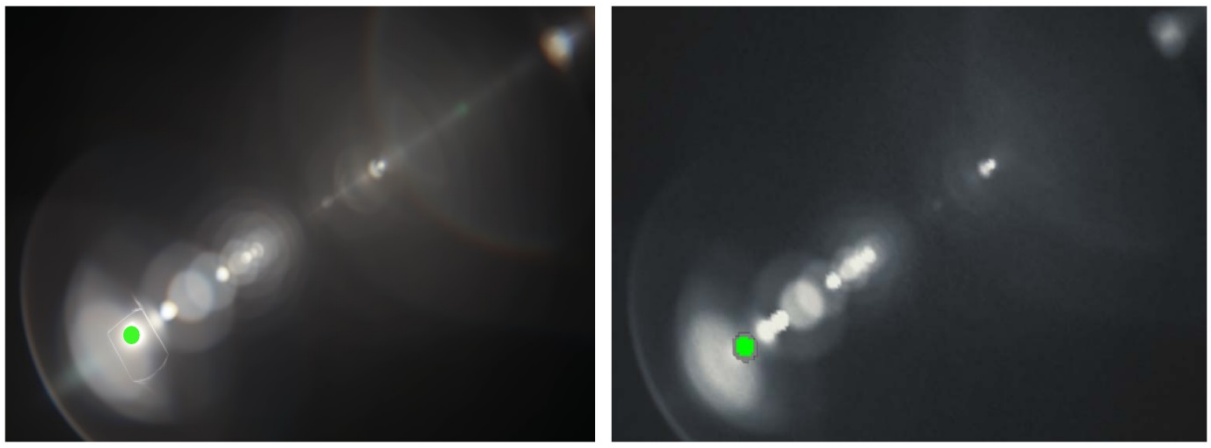


Fig. 3.22: The ghost image in a real photograph (left) and the image simulated for the same light source position (right) match well. The green dot marks the position of the light source.

But it was only after the rapid development of the computing power in the past decades that the number of ray tracings that is required for numerical image calculations was made possible. We are going to describe the historical development of computers for optical calculations in the next section.

About the history of computers at ZEISS

The introduction of mainframe computers into the optics industry in Germany started in the 1950s. In 1953, Leitz in Wetzlar installed a Z5 computer made by Zuse (Vollrath (2009)). ZEISS in Jena developed and built two OPREMA machines (**Op**tik**re**chen-**ma**schinen, machines for optical calculations, see Fig. 3.23) and used them, starting in 1955, for optical calculations. In 1958, Konrad Zuse, the famous computer pioneer, installed a Z22 computer at the West German ZEISS plant in Oberkochen.

Fig. 3.23: First computing machine at ZEISS: The "OPREMA" was developed at ZEISS in Jena and put into operation for optical calculations in 1955.

To ensure reliably running operations, the OPREMA machine at Jena consisted of two computers, which occupied one large room each (see Fig. 3.24). Each computer consisted of 16,626 relays. Programming and numerical inputs were done by plugging cable connections at the plug board.

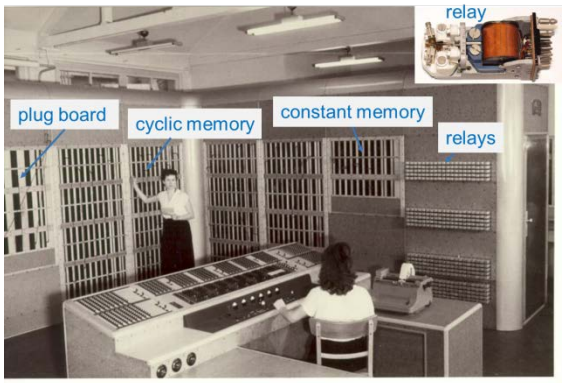


Fig. 3.24: Setup of the OPREMA. One relay (top right), i.e. 1 bit of information, was larger than a USB stick of today.

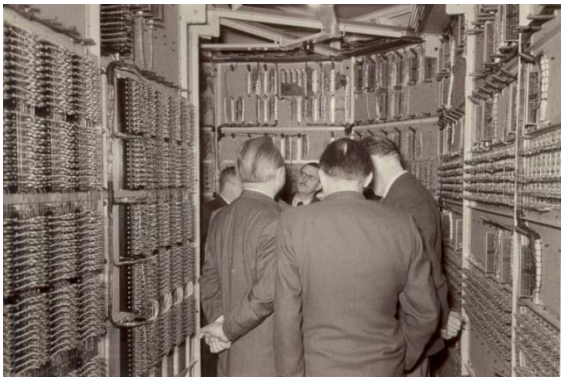


Fig. 3.25: "Inside" the OPREMA. Entrance through the door at right-hand image boarder of Fig. 3.24.

It took the OPREMA 0.8 seconds for one multiplication (Winkler (2013)). This sounds ridiculous from today's point of view, but it was a great breakthrough at that time and meant a considerable reduction of the computation time. It was an advantage that calculation errors could be excluded for the first time in history, the detection and tracing of which used to be laborious. Fig. 3.26 shows some performance data of the main-frames of the first generation at ZEISS in Jena.

Computer Type	OPREMA	ZRA 1	ICL 1905	ODRA 1305	EC 1040	workstation
Generation	1	2	3	3	3	current
Manufacturerer	ZEISS JENA	ZEISS JENA	ICL/England	ELWRO/Polen	Robotron (RGW)	
Year of manufacture	1954	1960	1968	1974	1977	2014
Technical principle	relais	electronic tube	transistor	integrated circuits	integrated circuits	integrated circuits
Number of central processing units	2	3	2	2	1	12
Processing speed (average number of operations / second)	2	150	150000	250000	380000	10^{10}
Input	manual	punch card, punched tape	punch card, punched tape	punch card, punched tape, magnetic tape	punch card, punched tape, magnetic tape	keyboard
Output	typewriter	numeric printer	printer, punch card, punched tape	printer, punch card, punched tape	printer, punch card, punched tape	monitor, printer
Main storage technology	relais	magnetic drum	matrix memory	matrix memory	matrix memory	integrated circuits
Capacity of main storage	1 k words	4 k words	64 k words	128 k words	512 kByte	64 GB
External memory	-	-	12 MB	18 MB	14 MB	1 TB
Cost per hour of operation (German Mark)	60,-	150,-	320,-	650,-	900,-	4,-

Fig. 3.26: Performance data of the first computing machines at ZEISS in Jena.

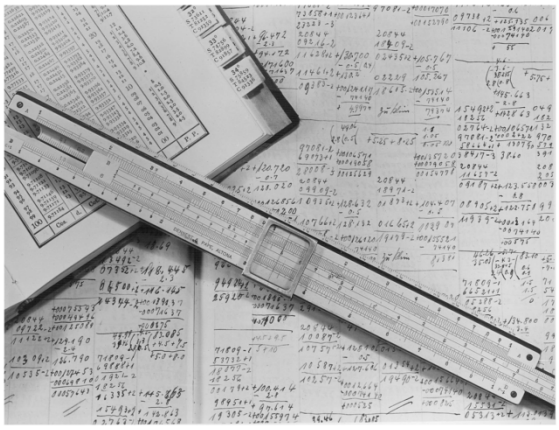


Fig. 3.27: Optical calculations using paper, pencil and logarithm tables.

Up to that time, trained optics calculation staff used to need about one working week to trace about 400 rays, using log tables and, from the 1940s onwards, also small hand-operated 4-species desktop calculators (electrically driven as of 1952) as well as handwritten calculations (see Fig. 3.27).

This number of ray tracings was just sufficient to evaluate the optical performance status of a Tessar lens consisting of seven refractive surfaces.

By the end of the 1960s, there was a leap in computing speed thanks to the use of transistors and integrated circuits. Since the 1970s, the number of feasible ray tracings has been rising by more than a factor of 20 per decade. Today, in 2016, it is possible to calculate 100 million ray transitions at spherical surfaces in one second using ten CPUs in parallel. The calculation of aspherical (i.e. non-spherical) surfaces takes about ten times longer than that of spherical surfaces, since the ray intersection point with the aspherical surface and the angle of refraction to the surface normal have to be determined numerically using iteration methods. Fig. 3.28 shows the historical development of the computing speed for optical calculations since 1900.

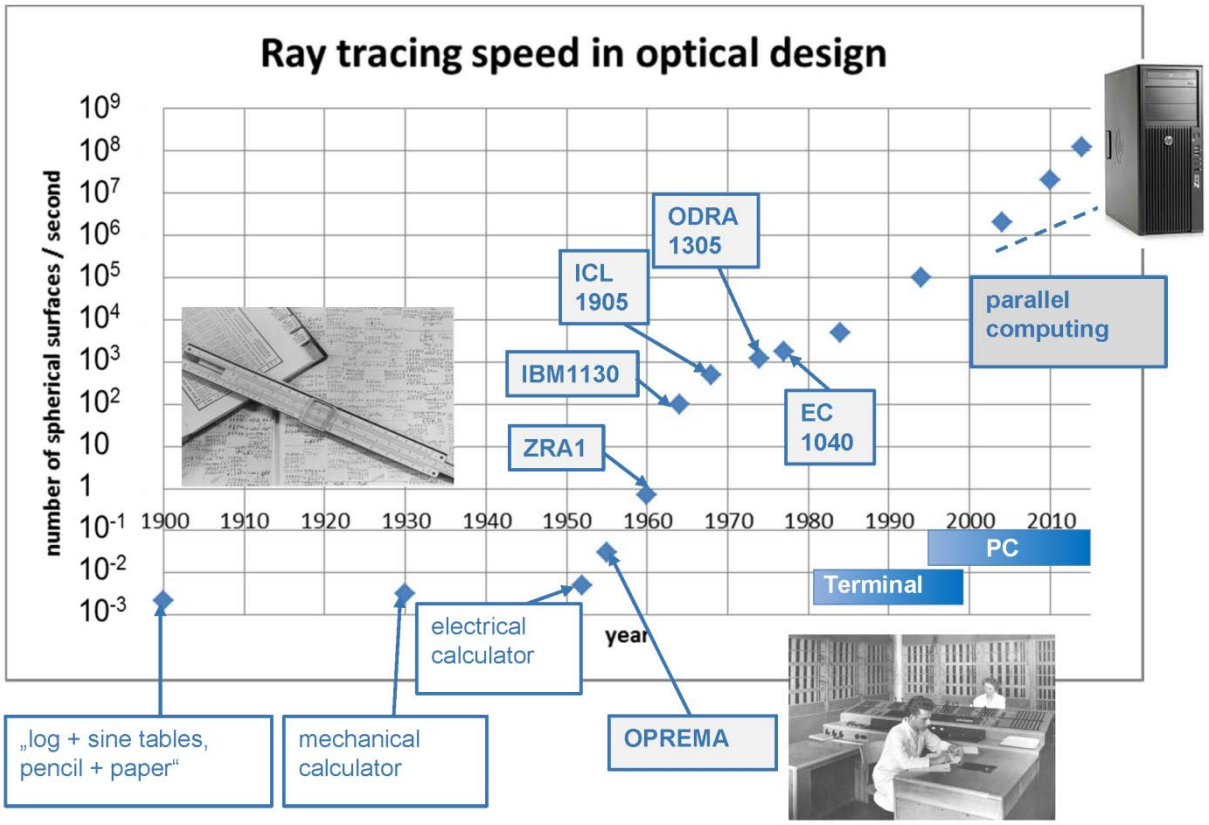


Fig. 3.28: Number of ray tracings per second at a spherical surface, from 1900 until today (2016).

Acknowledgements

The authors would like to thank Dr. Wolfgang Wimmer for historical information from the ZEISS archive about T- and T*-coating as well as about the development of computing machines; Dr. Harry Schlemmer for information about the history of coatings at ZEISS; Dr. Diana Tonova, Dr. Thomas Koch and Thomas Rittmann for information about the design and production of optical coatings; Andreas Bogenschütz for the preparation of product photos; Hans-Joachim Frasch for information about the speed of optical calculations; and Prof. Dr. Thomas Schneider for valuable and detailed discussions on different topics of this work. Finally, we would particularly like to thank Dr. Hubert Nasse for providing us with extensive laboratory data and photo material. We are grateful for our joint efforts over the past years that have resulted in major progress in ghost image analysis.

References

Abbe, E. (1873). *Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung*, Archiv für mikroskopische Anatomie, Vol. 9, Issue 1, pp. 413-418.

Abd El-Maksoud, R. H., Sasian, J. (2011). *Modeling and analyzing ghost images for incoherent optical systems*. Applied Optics, Vol. 50, Issue 15, pp. 2305-2315.

Adams, A. (1980). *The Print*. Republication of the book of 1950 by Little, Brown and Company, New York, Boston.

Anders, H. (1967). *Thin Films in Optics*, Focal, New York.

ARRI (2011). ALEXA Anamorphic De-Squeeze white paper, https://www.arri.com/camera/alexa/downloads/.

Baumeister, P. W. (2004). *Optical Coating Technology*, SPIE, Bellingham.

Berek, M. (1926). *Über Kohärenz und Konsonanz des Lichtes*. Parts I to IV, Z. f. Physik **36** (1926), pp. 675-688, pp. 824-838, **37** (1926), pp. 387-394; **40** (1927), pp. 420-450.

Blahnik, V. (2014). *About the irradiance and apertures of camera lenses.* http://www.zeiss.com/content/dam/Photography/new/pdf/en/cln_archiv/cln_en_web_special_irradiance_and_apertures.pdf

Bliedtner, J., Gräfe, G. (2010). *Optiktechnologie*. Carl Hanser Fachbuchverlag.

Bruynooghe, S., Tonova, D., Sundermann, M., Koch, T., Schulz, U. (2014). *Antireflection coatings combining interference multilayers and a nanoporous MgF_2 top layer prepared by glancing angle deposition*, Surface & Coatings Technology 267 (2015), pp. 40–44.

Cittert, P. H. van (1934). *Die wahrscheinlichste Schwingungsverteilung in einer von einer Lichtquelle direkt oder mittels einer Linse beleuchteten Ebene*, Physica, Haag **1**, pp. 201-210.

Cushing, D. (2011). *Enhanced Optical Filter Design*, SPIE, Bellingham.

Czapski, S. (1904). *Grundsätze der Theorie der optischen Instrumente nach Abbe*, published by Otto Eppenstein, Johann Ambrosius Barth, Leipzig.

Darmont, A. (2012). *High Dynamic Range Imaging – Sensors and Architectures*. SPIE Press, Bellingham, Washington USA.

Debevec, P. E., Malik, J. (1997). *Recovering High Dynamic Range Radiance Maps from Photographs*, Siggraph Conference, July 1997.

Dobrowolski, J. A. (1997). *Numerical Methods for Optical Thin Films*, Optics & Photonics News, June 1997.

Dörband, B., Müller, H., Gross, H. (2012). *Handbook of Optical Systems. Metrology of Optical Components and Systems*. Wiley-VCH.

DxOMark database of measured data, among others of a variety of digital cameras: http://www.dxomark.com

Fährmann, T. (2015). *Master Anamorphics Flare Sets Showreel*, short film made with ARRI/ZEISS Master Anamorphics and Flare Sets as well as ARRI ALEXA.

Fest, E. (2013). *Straylight Analysis and Control*, SPIE Press.

Flügge, J. (1960). *25 Jahre reflexmindernde Vergütung*, ZEISS-Mitteilungen, Vol. 2, Issue 3.

Fresnel, M. A. (1823). *Mémoire sur la loi des modifications que la réflexion imprime a la lumière polarisée*, Mémoires de l'Acadêmie des sciences de l'Institut de France, Paris.

Furman, Sh. A., Tikhonravov, A. V. (1992). *Basics of Optics of Multilayer Systems*. Editions Frontieres, Paris.

Gänswein, B. (1997). *Dünne Schichten für die Optik und Kristallherstellung in Oberkochen von 1946 bis heute*, Vakuum in Forschung und Praxis, Nr. 3, pp. 209-211, Wiley-VCH.

Geffcken, W. (1940). *Überzug aus mindestens drei Schichten von verschiedener Brechungszahl für einen nichtmetallischen Gegenstand zur Verminderung von dessen Oberflächenreflexion*, Deutsches Reichspatent 758 767.

Goldberg, E. (1925). *Der Aufbau des Photographischen Bildes, Teil 1: Helligkeitsdetails*. Enzyklopädie der Photographie und Kinematographie, No. 99, publisher: Wilhelm Knapp, Halle.

Gonzalez, R. C., Woods, R. E. (2008). *Digital Image Processing*. Prentice Hall, Upper Saddle River, NJ.

Grant, B. G. (2011). *Field Guide to Radiometry.* SPIE Press, Bellingham, Washington USA.

Gross, H., Blechinger, F., Achtner, B. (2008). *Handbook of Optical Systems. Volume 4; Survey of Optical Instruments.* Wiley-VCH.

Hertz, H. (1887). *Ueber sehr schnelle electrische Schwingungen*, Annalen der Physik, **267** (7), pp. 421-448.

Hopkins, H. H. (1953). *On the diffraction theory of optical images*, Proc. Roy. Soc. (London) A 217, pp. 408-432.

Hullin, M. B., Eisemann, E., Seidel, H.-P., Lee, S. (2011). *Physically-Based Real-Time Lens Flare Rendering*. In ACM Transactions on Graphics (Proceedings of SIGGRAPH 2011).

Huygens, C. (1690, completed 1678). *Traité de la lumière [Treatise about light],* Leyden.

ISO (2009). Photography, Electronics still-picture cameras, *Methods for measuring opto-electronic conversion functions (OECFs).* ISO 14524:2009; Vernier, Geneva, Switzerland.

Jones, L. A., Condit, H. R. (1941). *The Brightness Scale of Exterior Scenes and the Computation of Correct Photographic Exposure*, JOSA, vol. 31, issue 11, pp. 651-678.

Knittl, Z. (1976). *Optics of Thin Films*, Wiley, London.

Laue, M. von (1907). *Die Entropie von partiell kohärenten Strahlenbündeln*, Annalen der Physik, **23**, pp. 1-43.

Leistner, K. (1941). *Zeiss T-Optik*, Photographie und Forschung, Band 3, Heft 6, pp. 188-192.

Lummer, O., Reiche, F. (1910). *Die Lehre von der Bildentstehung im Mikroskop von Ernst Abbe*. Vieweg, Braunschweig.

Macleod, H. A. (2001). *Thin-Film Optical Filters*, Institute of Physics Publishing, Bristol and Philadelphia.

Mann, S., Picard, R. W. (1995). *On being ´undigital´ with digital cameras: extending dynamic range by combining differently exposed pictures*, Proc. IST, pp. 422-428.

Maxwell, J. C. (1865). *A Dynamical Theory of the Electromagnetic Field*, Philosophical Transactions of the Royal Society, **155**, pp. 459-512.

McCann, J. J., Rizzi, A. (2012). *The Art and Science of HDR Imaging*, Wiley.

Musset, A., Thelen, A. (1970). *Multilayer Antireflection Coatings*. Progress in Optics, ed. E. Wolf, North-Holland, Amsterdam, Vol. 18, pp. 201-237.

Nasse, H. (2008). *How to read MTF curves - Part I,* http://www.zeiss.com/content/dam/Photography/new/pdf/en/cln_archiv/cln30_en_web_special_mtf_01.pdf

Newton, I. (1704). *Opticks*, Royal Society, London.

Palmer, J. M., Grant, B. G. (2010). *The Art of Radiometry.* SPIE Press, Bellingham, Washington USA.

Pulker, H. K. (1999). *Coatings on Glass*, Elsevier, Amsterdam.

Lord Rayleigh (1880). *On reflection of vibrations at the confines of two media between which the transition is gradual*, Proc. London Math. Soc. **11,** pp. 51–56.

Reinhard, E., Ward, E., Pattanaik, S., Debevec, P., Heidrich, W., Myszkowski, K. (2010). *High Dynamic Range Imaging*, Elsevier, Morgan Kaufmann.

Richter, R. (1940). *Die Bedeutung der Zeiss-T-Optik für die Photographie und Projektion*. ZEISS-Nachrichten, special edition no. 5.

Sauer, H. (20. Februar 1973). *The development of anti-reflection coating for photographic lenses: From the Zeiss T-coating to multi-layer coating*. ZEISS Press-Information.

Schallenberg, U. B. (2006). *Antireflection design concepts with equivalent layers*. Applied Optics, Vol. 45, Issue 7, pp. 1507-1514.

SCHOTT Glass Data: http://www.schott.com/german/download/index.html

Schrödinger, E. (1920). *Über die Kohärenz in weitgeöffneten Bündeln*. Annalen der Physik, Vol. 366, Issue 1, pp. 69-86.

Smakula, A. (1935). *Verfahren zur Erhöhung der Lichtdurchlässigkeit optischer Teile durch Erniedrigung des Brechungsexponenten an den Grenzflächen dieser Teile*, Deutsches Reichspatent 685 767.

Smakula, A. (1941). *Über die Reflexionsverminderung an Grenzflächen zwischen durchsichtigen Körpern*, Glastechn. Ber. **19**, pp. 377-386.

Smakula, A. (1942). *Grundlagen der Reflexionsverminderung*, Film und Farbe, Schriftenreihe der Reichsfilmkammer, Band 9, Max Hesse Verlag, Berlin, pp. 69-74.

Steinbach, M. (2015). *Development of a method to measure the veiling glare contributions of the lens and the sensor and its influence on the limit of the dynamic range when shooting movies,* Bachelor Thesis, Cologne University of Applied Sciences.

Taylor, D. H. (1904). *Lenses*. United Kingdom patent 29561.

Thetford, A. (1969). *A Method of Designing Three-layer Anti-reflection Coatings*, Optica Acta, Vol. 16, No. 1, pp. 37-43.

Thelen, A. (1969). *Design of multilayer interference filters*, *Physics of Thin Films,* ed. G. Hass and R. E. Thun, Academic Press, New York.

Thelen, A. (1988). *Design of Optical Interference Coatings*, McGraw-Hill, New York.

Vollrath, W. (2009). *Die Entwicklung von Leica Objektiven damals und heute*, in: Max Berek, Schöpfer der ersten Leica Objektive, Pionier der Mikroskopie, published by Knut Kühn-Leitz, Lindemanns Verlag.

Winkler, J. F. H. (2008). *Die Oprema – der Relaisrechner des Zeisswerks Jena*. Lecture at the Friedrich Schiller University, Jena. http://psc.informatik.uni-jena.de/publ/oprema-JW-2008.pdf

Young, T. (1802). *On the Theory of Light and Colours*, Philosophical Transactions of the Royal Society of London **92**, pp.12-48.

Zernike, F. (1938). *The concept of degree of coherence and its application to optical problems*, Physica (The Hague) **5** (8), pp. 785-795.

Appendix

A.1: Reflection and transmission of a plane wave upon transition between two media

The Fresnel coefficients describe the reflected ("r") and transmitted ("t") electrical field components of light waves at the interface separating two media with refractive indices n and n', respectively:

$$r_s(n,n',i,i') = \frac{n\cos i - n'\cos i'}{n\cos i + n'\cos i'},$$

$$r_p(n,n',i,i') = \frac{n'\cos i - n\cos i'}{n'\cos i + n\cos i'},$$

$$t_s(n,n',i,i') = \frac{2n\cos i}{n\cos i + n'\cos i'},$$

$$t_p(n,n',i,i') = \frac{2n\cos i}{n'\cos i + n\cos i'}.$$

Here, i denotes the angle between the ray representing the wave that is incident onto the interface plane and the surface normal. The angle between the ray (or wave normal) and the surface normal that is transmitted into the medium n' is denoted by i'; it is called angle of refraction (cf. Fig. A.1).

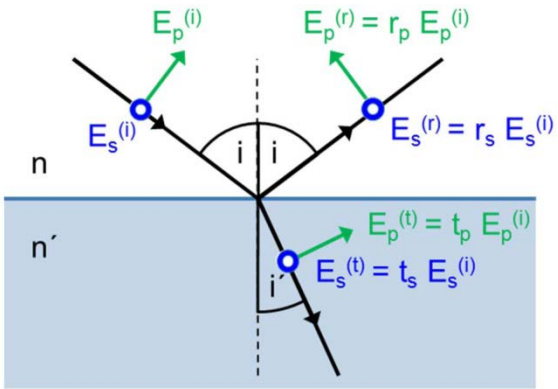


Fig. A.1: Electrical field components upon transition at the interface between two media with refractive indices n und n´.

The subscript s designates the component that is perpendicular to the plane of incidence. The plane of incidence is the plane spanned by the ray vector and the surface normal. The subscript p denotes the component that is parallel to the plane of incidence (see Fig. A.1).

Inside the plane of incidence the light propagates as follows: The angle i'' of the reflected ray and the surface normal is, by the law of reflection, equal to the incidence angle i; the refracted ray angle i´ is given by Snell´s Law:

$$i'' = i \qquad \text{(reflection)},$$

$$n'\sin i' = n\sin i \qquad \text{(refraction)}.$$

Figures A.2 and A.3 show the reflected intensities

$$R_s = r_s^2,$$

$$R_p = r_p^2,$$

as well as the corresponding average

$$R_{mean} = \frac{R_s + R_p}{2}$$

for different combinations of the refractive indices n = 1, 1.38, 1.52, 1.85, and 2.3. It is common practice to express and represent "the" reflectance of coatings, by means of the average $R_{mean,}$ which is then simply denoted by R. This quantity can be interpreted as "mean reflected intensity of unpolarized light."

The corresponding quantities for the transmitted intensities are $T_s = t_s^2$, $T_p = t_p^2$ and $T_{mean} = (T_s + T_p)/2$.

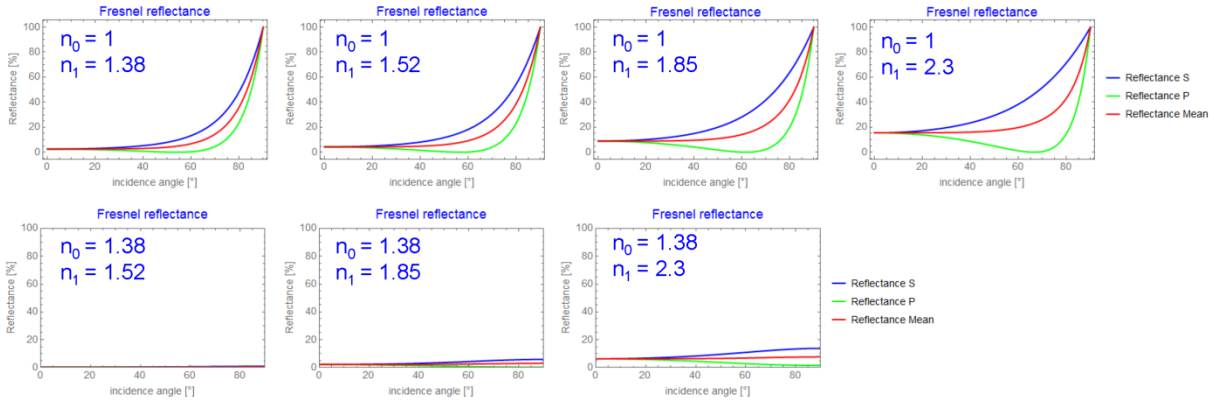


Fig. A.2: Reflected intensities at a single interface for different combinations of refractive indices n_0 und n_1 shown against incidence angles from 0° to 90°.

In camera lenses, the angles of incidence along the nominal optical path rarely exceed 50°. Fig. A.3 presents the reflectance profiles for angles ranging from 0° to 50°.

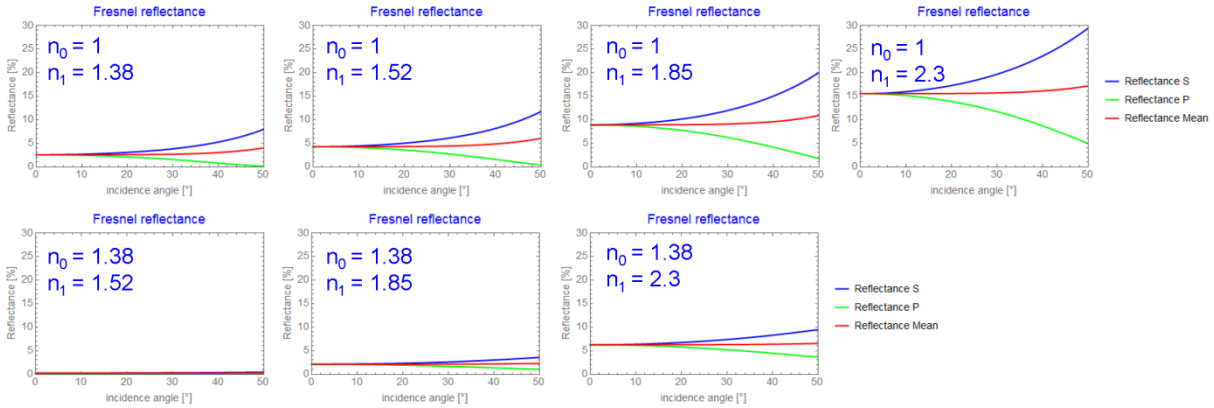


Fig. A.3: Reflectance profiles as in Fig. A.2, but shown for angles of incidence from 0° to 50°.

A.2: Reflectance of single-layer coatings

We define the following abbreviatory notation for the transmission and reflectance coefficients (for the s and p components, respectively) referring to a wave travelling in the medium n_0 and incident at an angle i:

$$r_{0,1} = r(n_0, n_1, \dots),\ t_{0,1} = t(n_0, n_1, \dots),$$

$$r_{1,g} = r(n_1, n_g, \dots),\ t_{1,g} = t(n_1, n_g, \dots).$$

In an analogous manner, $r_{1,0}$ and $t_{1,0}$ describe, respectively, the reflected and transmitted amplitudes of the wave coming out of the medium n_1 at the interface to the medium n_0 (cf. Fig. A.4)

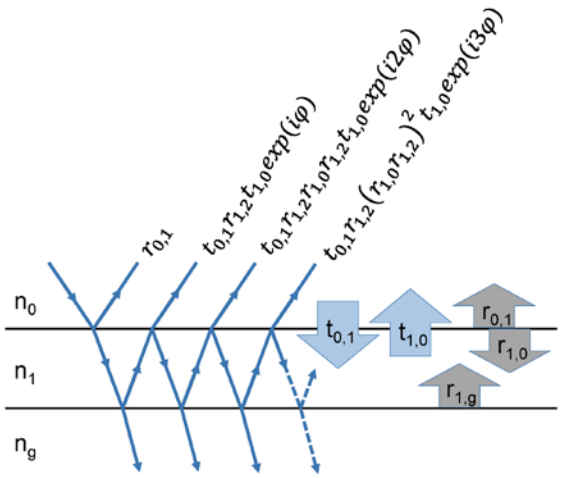


Fig. A.4: Waves reflected at a single-layer coating with refractive index n_1 between two media with refractive indices n_0 and n_g.

For a single-layer coating, the following complex-valued amplitudes ρ_j of the individual reflected waves are obtained after multiplication of the reflectance and transmission coefficients (see Fig. A.4):

$$\rho_0 = r_{0,1},$$

$$\rho_1 = t_{0,1}r_{1,g}t_{1,0}exp(i\varphi),$$

$$\rho_2 = t_{0,1}r_{1,g}r_{1,0}r_{1,g}t_{1,0}exp(i2\varphi),$$

$$\rho_3 = t_{0,1}r_{1,g}(r_{1,0}r_{1,g})^2t_{1,0}exp(i3\varphi),$$

$$\dots$$

$$\rho_n = t_{0,1}r_{1,g}(r_{1,0}r_{1,g})^{n-1}t_{1,0}exp(in\varphi).$$

The phase shift caused by the coating of directly neighboring reflected partial waves is:

$$\varphi = \frac{2\pi}{\lambda} 2\, n_1\, d\, cosi'$$

with ray angle i′ within medium n_1 (see section "Single-layer coatings").

Summing up all partial waves yields the expression

$$\rho = \sum_{k=0}^{\infty} \rho_k$$
$$= r_{0,1}$$
$$+ t_{0,1}t_{1,0}r_{1,g}exp(i\varphi)\sum_{k=0}^{\infty}(r_{1,0}r_{1,g})^k\, exp(ik\varphi).$$

This sum has the form of a geometric series:

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

with $x = r_{1,0}r_{1,g}\exp(i\varphi)$. Moreover, with the Fresnel coefficients, the following relation applies for the s and the p components:

$$t(n,n',.)t(n',n,.) - r(n,n',.)r(n',n,.) = 1.$$

Thus we obtain the following result:

$$\rho = \frac{r_{0,1} + r_{1,g}\exp(i\varphi)}{1 - r_{1,0}r_{1,g}\exp(i\varphi)}.$$

For this calculation an idealized plane interface extending to infinity and an idealized infinite coherence length were assumed. Those assumptions are unproblematic since, as a general rule, the reflectance values for antireflection coatings in the visible spectral range are relatively small: The geometric series thus converges very quickly; the term ρ_2 is about 100 times smaller than ρ_0 and ρ_1. As a consequence, the reflected intensity is basically determined by two-beam interference of the first two reflected waves (see Fig. A.5).

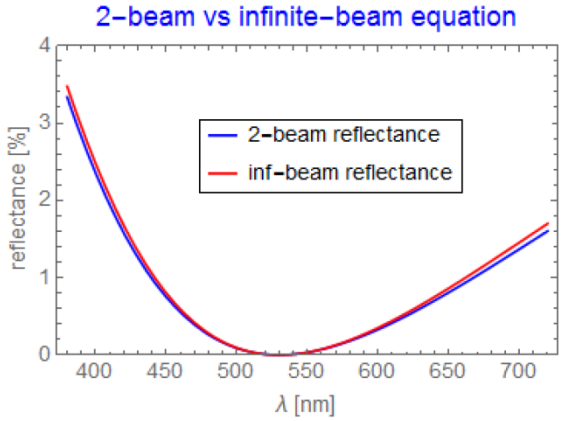


Fig. A.5: Comparison of the reflectance profiles for a single-layer coating with n_0=1, n_c=1.38, n_g=1.9: The red curve is based on the exact ansatz of infinitely many reflections in the coating; the blue curve is based on two-beam interference only, i.e. neglecting any terms involving more than two reflections: $R_{2-beam} = r_{0,1}^2 + \rho_{1,g}^2 + 2r_{0,1}\,\rho_{1,g}\cos\varphi$, where $\rho_{1,g} = t_{0,1}r_{1,g}t_{1,0}$. The two curves basically coincide. It is therefore legitimate to use the two-beam interference equation.

This means that the mechanism of destructive interference of a single-layer coating can be described by two-beam interference to very good approximation. For reflectance calculation with simulation programs, however, there is no benefit to be gained from two-beam interference formulas, instead the general formula set is used.

The reflected intensity is calculated as

$$R = \rho\,\rho^* = \frac{r_{0,1}^2 + r_{1,g}^2 + 2r_{0,1}r_{1,g}\cos\varphi}{1 + r_{1,0}^2 r_{1,g}^2 - 2r_{1,0}r_{1,g}\cos\varphi}.$$

For a stack of coating layers, i.e. consecutive single-layer coatings, the same reflectance and transmission values result, no matter whether the light comes from air and proceeds via the coating into the glass or vice versa, from glass via the coating into air. Of course, this holds true only if the angles in each medium remain unchanged. From the equations above, the following relations may be derived:

$$r_s(n',n,i',i) = -r_s(n,n',i,i'),$$

$$r_p(n',n,i',i) = -r_p(n,n',i,i').$$

Since the expression for the reflected intensity R contains only products of these reflected amplitudes, reversibility (in the sense stated above) follows. This also applies to multi-layer coatings as can be seen by successive application of our reasoning (see Appendix A.3).

If the Fresnel coefficients are evaluated for the special case of perpendicular incidence, one obtains the equations $r = \frac{n-n'}{n+n'}$ and $t = \frac{2n}{n+n'}$, and for the reflected intensity

$$R = \frac{(1+n_1^2)(n_1^2+n_g^2) - 4n_1^2 n_g + (1-n_1^2)(n_1^2-n_g^2)\cos\left(4\pi n_1\frac{d}{\lambda}\right)}{(1+n_1^2)(n_1^2+n_g^2) + 4n_1^2 n_g + (1-n_1^2)(n_1^2-n_g^2)\cos\left(4\pi n_1\frac{d}{\lambda}\right)}.$$

Here the refractive index of the single-layer coating is denoted n_1 and of the glass substrate n_g respectively. Making use of the identity

$$\cos\varphi = 1 - 2\,\sin^2(\varphi/2),$$

one obtains the equation stated by Smakula (1941) and Musset, Thelen (1970):

$$R = \frac{n_1^2(1-n_g)^2 - (1-n_1^2)(n_1^2-n_g^2)\sin^2\left(2\pi n_1\frac{d}{\lambda}\right)}{n_1^2(1+n_g)^2 - (1-n_1^2)(n_1^2-n_g^2)\sin^2\left(2\pi n_1\frac{d}{\lambda}\right)}.$$

A.3: Multi-layer coatings

If we add to a given coating another single layer, the transition is calculated by replacing the reflectance of the bottom interface $r_{1,2}$ by the effective reflected wave of the second coating lying underneath:

$$r_{1,2} \rightarrow \frac{r_{1,2} + r_{2,3}exp(i\varphi_2)}{1 - r_{2,1}r_{2,3}exp(i\varphi_2)}.$$

Performing this substitution in

$$\rho_1 = \frac{r_{0,1} + r_{1,2}exp(i\varphi_1)}{1 - r_{1,0}r_{1,2}exp(i\varphi_1)}$$

yields

$$\rho_2 = \frac{r_{0,1} + \frac{r_{1,2} + r_{2,3}exp(i\varphi_2)}{1 - r_{2,1}r_{2,3}exp(i\varphi_2)}exp(i\varphi_1)}{1 - r_{1,0}\frac{r_{1,2} + r_{2,3}exp(i\varphi_2)}{1 - r_{2,1}r_{2,3}exp(i\varphi_2)}exp(i\varphi_1)}.$$

If further coating layers are present, the following recursive substitution rule has to be used:

$$r_{j-1,j} \rightarrow \frac{r_{j-1,j} + r_{j,j+1}exp(i\varphi_j)}{1 - r_{j,j-1}r_{j,j+1}exp(i\varphi_j)}.$$

By successively applying Snell's law of refraction, i.e.

$$n_{j-1}\sin i_{j-1} = n_j \sin i_j,$$

the cosine terms in the Fresnel coefficients can be expressed as a function of the angle of incidence i_0 on the multi-layer coating:

$$\cos i_j = \sqrt{1 - \sin^2 i_j} = \sqrt{1 - \left(\frac{n_0}{n_j}\right)^2 \sin^2 i_0}.$$

This means that the Fresnel coefficients r_s, r_p, t_s and t_p depend, at each interface, on the three parameters (n_{j-1}, n_j, i_0).

A.4: Reflectance of double-layer coatings

By means of the method described in A.3, one obtains the reflectance of the double-layer coating. For perpendicular incidence and n_0=1 we have:

$$R = \frac{X}{1 + X}$$

where

$$X = \frac{n_g}{4}\left\{\left[\left(\frac{1}{n_g} - 1\right)\cos\varphi_1 \cos\varphi_2 + \left(\frac{n_1}{n_2} - \frac{n_2}{n_1 n_g}\right)\sin\varphi_1 \sin\varphi_2\right]^2 + \left[\left(\frac{1}{n_1} - \frac{n_1}{n_g}\right)\sin\varphi_1 \cos\varphi_2 + \left(\frac{1}{n_2} - \frac{n_2}{n_g}\right)\cos\varphi_1 \sin\varphi_2\right]^2\right\},$$

$\varphi_1 = 2\pi n_1 d_1/\lambda$, and $\varphi_2 = 2\pi n_2 d_2/\lambda$.

For a λ/4-λ/4-coating, all cosine terms disappear at the reference wavelength $\lambda_0 = 4\,n\,d$. In order to obtain vanishing reflectivity, R=0 (that is X=0), the coefficient of the $\sin\phi_1\sin\phi_2$ term needs to vanish. The following amplitude condition results:

$$\left(\frac{n_1}{n_2} - \frac{n_2}{n_1 n_g}\right) = 0$$

or

$$\frac{n_2}{n_1} = \sqrt{n_g}.$$

For the λ/4-λ/2-coating, only the $\sin\phi_1\cos\phi_2$ term remains for λ_0, which yields the amplitude condition

$$\left(\frac{1}{n_1} - \frac{n_1}{n_g}\right) = 0$$

or

$$n_1 = \sqrt{n_g}.$$

As stated before this is the same expression as the one applicable to single-layer coatings, since the λ/2-layer is neutral for λ_0.

A.5: Reflectance of triple-layer coatings

The reflectance of a triple-layer coating, for perpendicular incidence and n_0=1 is given by

$$R = \frac{X}{1+X}$$

where

$$X = \frac{n_g}{4}\,(A^2 + B^2),$$

$$A = \left(\frac{1}{n_g} - 1\right)\cos\varphi_1 \cos\varphi_2 \cos\varphi_3$$

$$-\left(\frac{n_2}{n_1 n_g} - \frac{n_1}{n_2}\right)\sin\varphi_1 \sin\varphi_2 \cos\varphi_3$$

$$-\left(\frac{n_3}{n_2 n_g} - \frac{n_2}{n_3}\right)\cos\varphi_1 \sin\varphi_2 \sin\varphi_3$$

$$-\left(\frac{n_3}{n_1 n_g} - \frac{n_1}{n_3}\right)\sin\varphi_1 \cos\varphi_2 \sin\varphi_3,$$

$$B = \left(\frac{1}{n_3} - \frac{n_3}{n_g}\right)\cos\varphi_1 \cos\varphi_2 \sin\varphi_3$$

$$-\left(\frac{1}{n_2} - \frac{n_2}{n_g}\right)\cos\varphi_1 \sin\varphi_2 \cos\varphi_3$$

$$-\left(\frac{1}{n_1} - \frac{n_1}{n_g}\right)\sin\varphi_1 \cos\varphi_2 \cos\varphi_3$$

$$-\left(\frac{n_2}{n_1 n_3} - \frac{n_1 n_3}{n_2 n_g}\right)\sin\varphi_1 \sin\varphi_2 \sin\varphi_3,$$

$\varphi_1 = 2\pi n_1 d_1/\lambda$, $\varphi_2 = 2\pi n_2 d_2/\lambda$, and $\varphi_3 = 2\pi n_3 d_3/\lambda$.

With a λ/4-λ/4-λ/4-coating, only the last summand in B survives, and the reflectance at λ_0 equals zero, if the following condition is fulfilled:

$$\frac{n_1 n_3}{n_2} = \sqrt{n_g}.$$

For a λ/4-λ/2-λ/4-coating, the only remaining term is $\sin\varphi_1 \cos\varphi_2 \sin\varphi_3 = -1$, and the amplitude condition to be fulfilled in order to have R=0 is:

$$\frac{n_3}{n_1} = \sqrt{n_g}.$$

If the λ/2-coating lies at a different position in the stack of layers, the following holds:

$$\frac{n_3}{n_2} = \sqrt{n_g} \qquad (\lambda/2\text{-}\lambda/4\text{-}\lambda/4),$$

$$\frac{n_2}{n_1} = \sqrt{n_g} \qquad (\lambda/4\text{-}\lambda/4\text{-}\lambda/2).$$


ZEISS