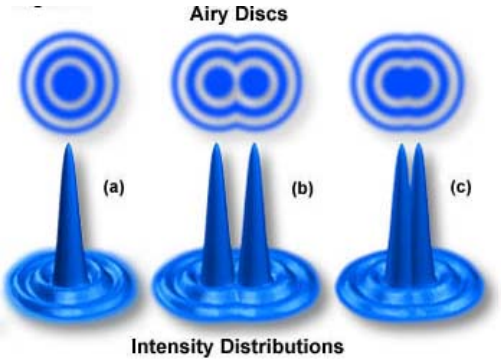
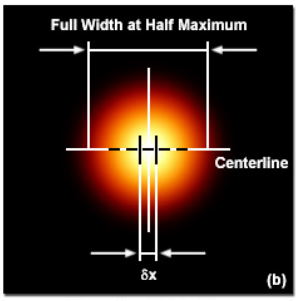
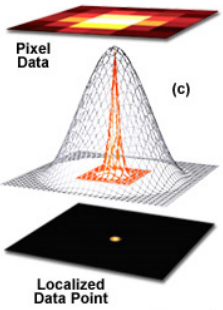
Chapter 4

Optical Microscopy

Instrumentation, Resolution Limits, Aberrations,
Practical Resolution, Modulation Transfer Function





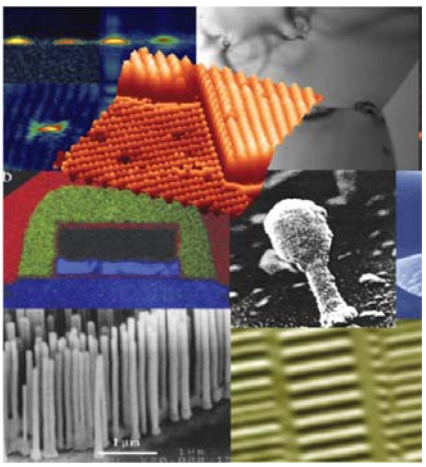


Contents - Chapter 4: Optical Microscopy – Fundamentals

4.1 Introduction

Microscopy = High resolution **2D** or **3D** real space ***spatial imaging*** , i.e., ***mapping*** of materials properties on a ***magnified scale*****.**

⇨ This yields the ***microscopic landscape*** of properties such as morphology, mass density, chemical composition, reflectivity, absorption, luminescence efficiency, or any other property, depending on which type of response signal and probe sample-interaction is recorded as imaging signal. Most common: Imaging of structure, morphology or composition of materials.



⇨ Microscopic imaging is possible in principle for all material characterization methods described in the previous chapter. However, the achievable spatial resolution can be quite different due to the different used probes.

In this chapter:

» The **principle approaches** and methods of microscopy are introduced and the **image formation processes** described,

» **Resolution limits** and conditions for best resolution are derived for optical imaging microscopy and transmission electron microscopy,

Examples

- **Optical microscopy (HM, PCM)**:

  Visible light: Resolution down to 200 nm (≈ λ/2).

- **Electron Microscopy**: (TEM, SEM)

  Electron beams: Small wavelengths λ << Å

  TEM: Resolution of atom rows, limited by optics,

  SEM: Resolution limited by interaction volume,

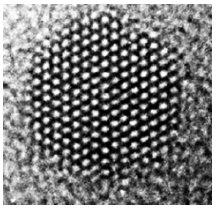
- **Field ion microscopy**: (FIM, FEM)

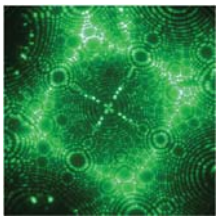
  Imaging using **electrons or ions**:
  Direct imaging of single atoms at surface steps,

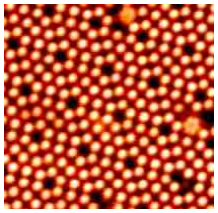
  Limited to tip-shaped samples with radius<100 nm,

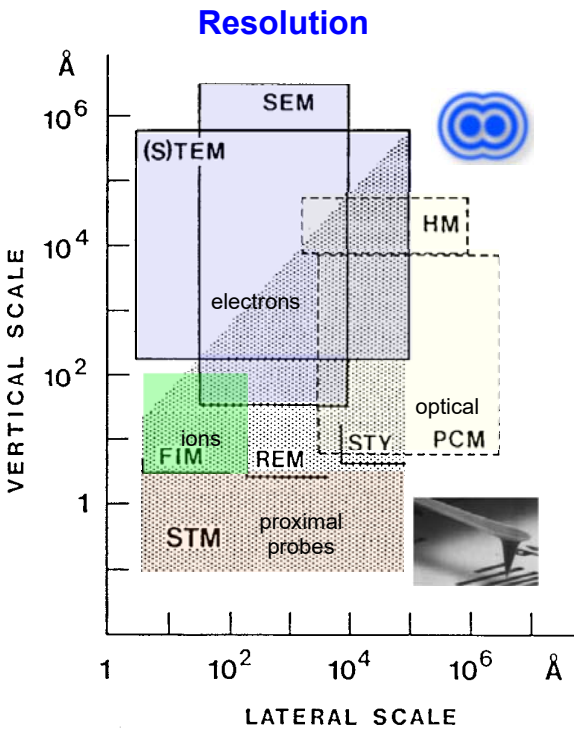
- **Scanning proximal probe microscopy**:

  Proximal probes with very small tip radius and localized (near-field) interaction
  Atomic resolution possible for STM & AFM.









Type of probe determines:

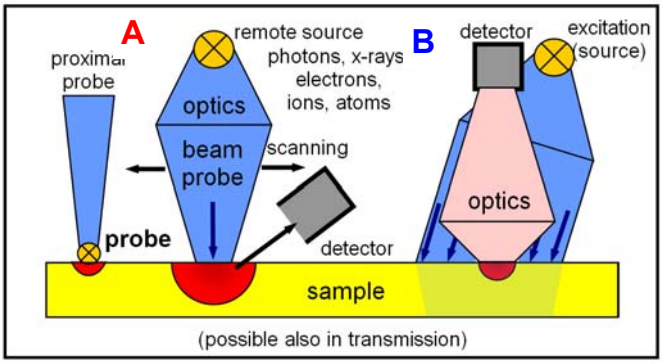
(a) Resolution, (b) probe-sample interaction (c) information type,
(d) required technical instrumentation.

4.2 How to Perform Imaging with High Resolution ?

A. *Scanning Microscopy* = Local excitation + non-local detection

- ⇨ Excitation of a small sample volume by a tightly ***focused beam probe*** produced by ***demagnifying*** **optics**.
- ⇨ Global detection of the response signal from the sample by an integrating detector
- ⇨ **2D images** created by raster scanning the probe over the sample surface

Examples: Scanning electron microscopy(SEM), scanning optical (SOM), scanning probe (SPM), scanning x-ray microscopy (SXRM), …..



B. *Imaging Microscopy* = Non-local excitation + Local detection.

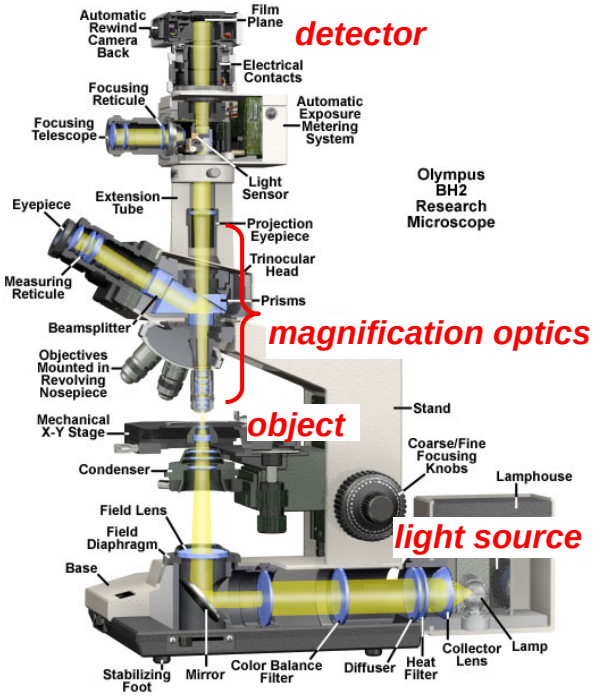
- ⇨ Excitation of a large sample volume but detection only the signal from a small volume that is projected by ***magnifying optics*** onto the detector.
- ⇨ **2D images** created usually by ***parallel*** *detection* of many individual spots using a two dimensional pixel detector such as a photoplate, human eye, CCD camera, channel plate, etc., for which one pixel corresponds to the signal from one small sample spot.

Examples: Transmission electron microscopy (TEM), optical imaging microscopy (OM) …

C. Confocal Microscopy = local excitation *and* local detection (Combination of both).

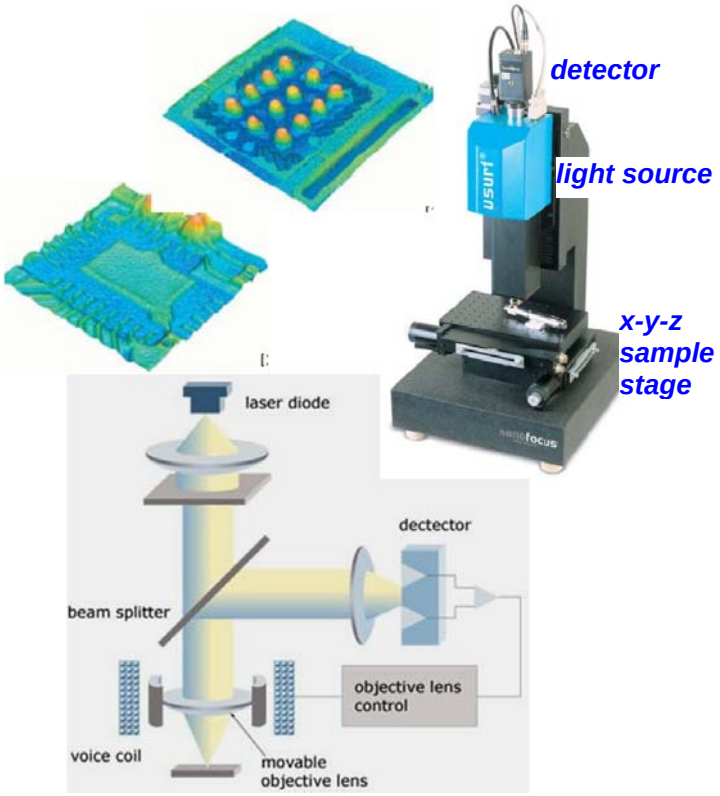
Example: Confocal scanning optical microscopy, ….

Optical Imaging Microscope



⇨ *Magnifying optics* required for image formation, magnification determined by focal length

Laser Scanning Microscope



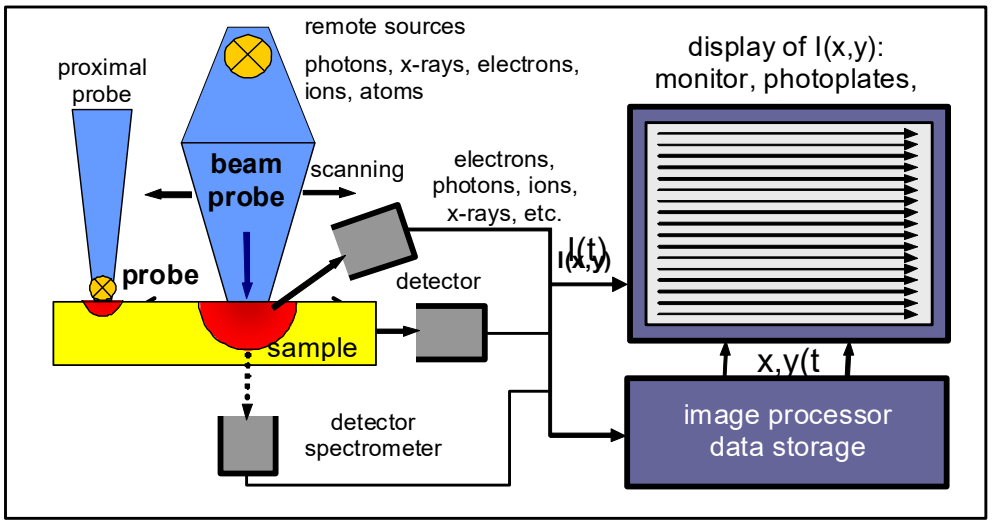
⇨ *Demagnifying* optics for probe formation, Magnification set by scan size

Note: The same type of variants exist for electron and x-ray microscopy

4.3 Scanning Microscopy

In scanning microscopy, an image is created by raster scanning a small probe that locally interacts/excites the sample and the detection of the global response signal I(t) as a function of time, where at a given time *t* the probe is at a different spot (x,y)= f(t) on the sample.

⇨ **Each image point** is recorded ***sequentially*** at a different time linked to a certain spot position !



Magnification = simply the ratio between display and scan size: $M = w_{display} / w_{scan}$

The smallest useful scan size is limited by the probe size, i.e., $w_{scan} > w_{probe}$.

Resolution is determined only by the probe size *as well as* interaction/excitation volume.

The probe size depends on the quality of the demagnifying optics and wavelength of the probe, and the interaction volume on the probe energy and interaction strength.

Characteristic Features of Scanning Microscopy

☺ *No magnifying optics* are required - only *simpler focusing optics* for probe formation, Thus, no demanding high resolution imaging optics are required.

☺ Thus, applicable to *wide range of techniques* because no imaging optics are required.

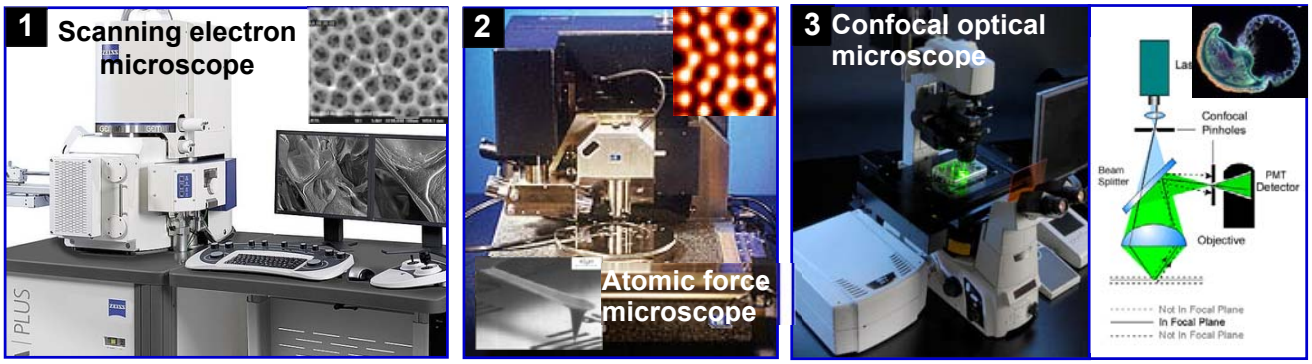
☺ Easy to combine with *spectroscopic* techniques. Thus, probing of many properties possible.

☺ *Ultimate resolution* = determined by the probe size and the interaction volume. For beam probes, the probe size is *limited by diffraction* and the *lens aberrations* of the focusing optics. The *excitation volume* depends on the probe pararmeters. In confocal fluorescenece microscopy (STED) it can be narrowed below the diffraction limit using *non-linear optical effects*.

☹ *Sequential* point-by-point recording is slow and limits the signal integration times per pixel.

😐 Overall, only a *virtual sample image* is obtained that represents measured signal versus (x,y). The image must be displayed on a monitor screen, photo plate, printer, etc. .

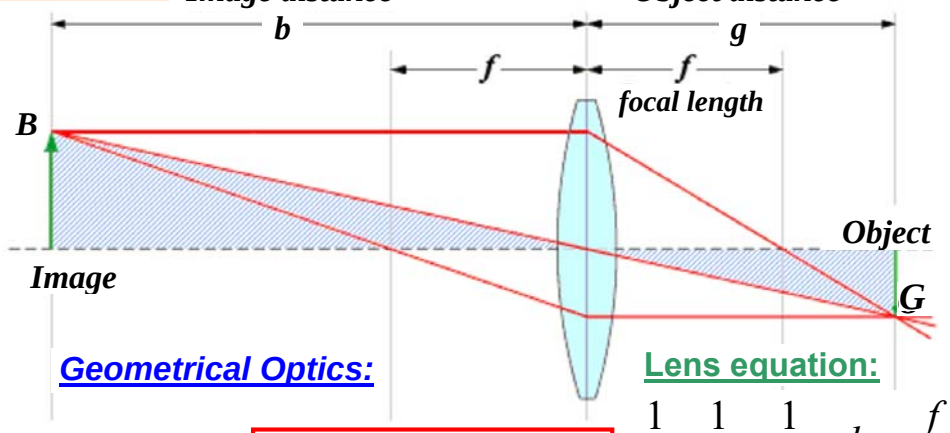
Variants: Scanning Electron Microscope (SEM), Scanning Tunneling (STM) and Scanning Force Microscope (AFM), Scanning Near Field (SNOM) and Scanning Confocal Optical Microscope, …



4.4 Imaging Microscopy

Local detection is achieved by creating an ***enlarged image*** of each sample point using ***magnifying optics*** of lenses or curved mirrors. The enlarged image is viewed or recorded using a 2D detector.

Magnification



Geometrical Optics:

Lens equation:

$$\frac{1}{b}+\frac{1}{g}=\frac{1}{f} \qquad b=\frac{f\cdot g}{f-g}$$

$$g \to f \Rightarrow M \to \infty$$

Magnification:

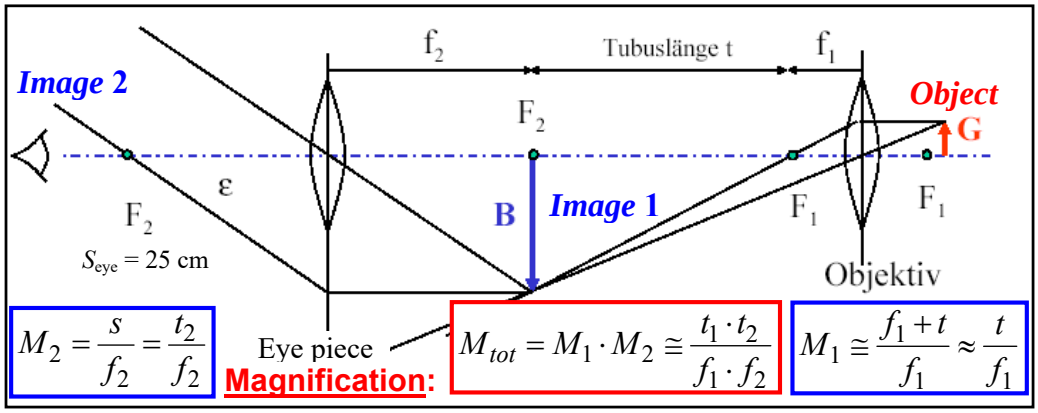
$$M=\frac{B}{G}=\frac{b}{g}=\frac{f}{g-f}$$

Magnification: $M = B / G$ is determined by the lens equation, i.e., object and image distances ***b/g***

⇨ At any time, **each point of the image** = signal collected from ***one point*** **of the sample!**

Magnification of a two lens optical microscope

Given by lens strength f_i and tube length t



$$M_2 = \frac{s}{f_2} = \frac{t_2}{f_2}$$

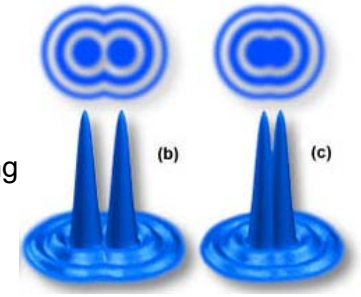
Magnification:

$$M_{tot} = M_1 \cdot M_2 \cong \frac{t_1 \cdot t_2}{f_1 \cdot f_2}$$

$$M_1 \cong \frac{f_1 + t}{f_1} \approx \frac{t}{f_1}$$

⇨ The **tube length** t of optical microscopes is usually fixed (standardized to t =160 mm). Thus, different objective lenses with different f_i have to be used to change the magnification.

Basic features of imaging microscopy:

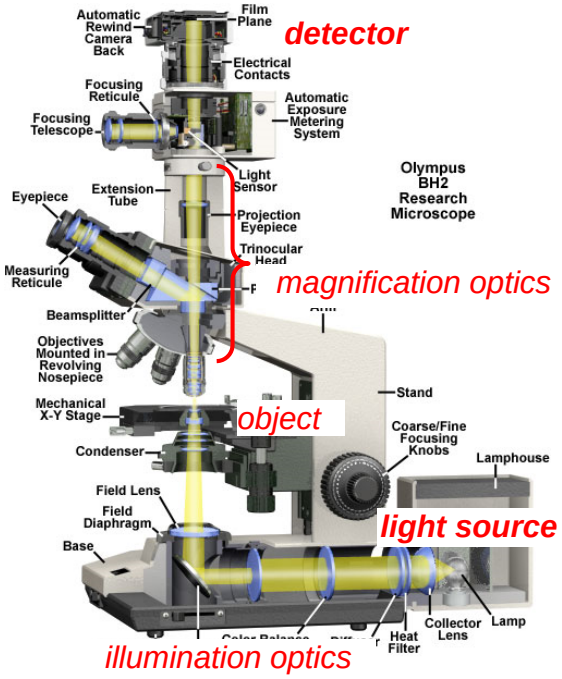


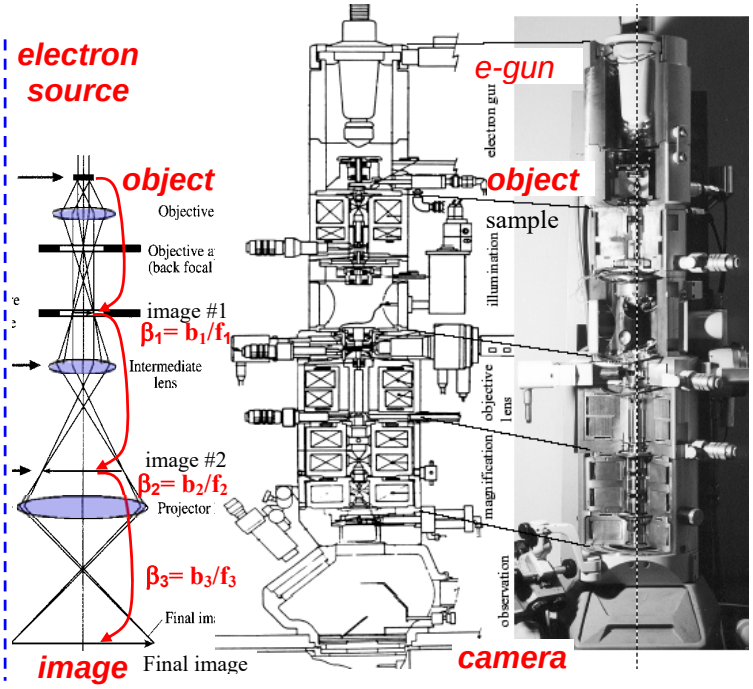
- ***Resolution***: Due to *diffraction* and lens imperfections, the image of each point of the object is smeared out on the detector. Thus, below a certain distance r , two object points can no longer be resolved. The resolution depends on the imaging wavelength λ and the quality of the optical imaging system but it is ***fundamentally limited*** to $r > ½ \lambda$.
- ***Useful magnification***: Although by using strong lenses with small f and a large tube length t the *magnification* $M_{mag} \approx t / f$ can be made be arbitrarily large, due to the resolution limit the ***"useful" magnification is finite***

Instrumentation

For the different probes very different kinds of instrumentation, i.e., beam sources, optical elements, detectors, environmental conditions, etc. is required but the basic opitical system is identical.

Example: Optical Microscope ***versus*** **Transmission Electron Microscope**





Different elements: (1) Beam sources: Light verus electrons, (2) Lenses: Glass versus magnetic/ electrostatic lenses, (3) Environment: Air versus vacuum column, (4) Photon versus electron detectors, (5) High voltage

Key features of imaging microscopy

☺ A *real* physical image of the sample is formed within the image plane.

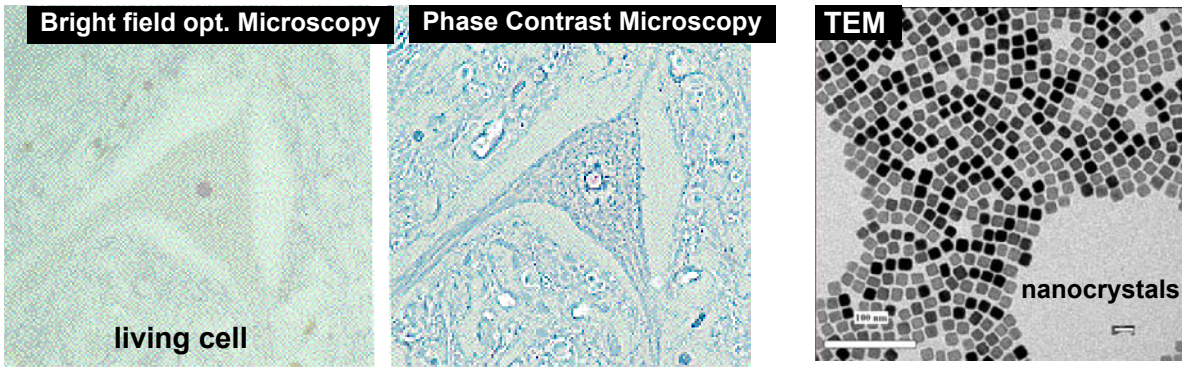
☺ Thus, all sample points are simultaneously imaged at once at the same time: This enables *real time imaging* of the whole sample at high speeds to reveal fast dynamical changes occurring within samples as a function of time.

☺ Long integration times per point are possible that yield high signal-to-noise ratios.

☹ The *resolution* is ***fundamentally limited*** by diffraction (~wavelength) .

Variants of imaging microscopy: Light microscopy (OM) using visible/IR/UV light, transmission electron microscopy (TEM), low energy reflection electron microscopy (LEEM), photo emission electron microscopy (PEEM), field ion microscopy (FIM), etc.

Examples: Left: Optical microscopy with different contrast. Right: TEM image



Different variants of imaging microscopy are distinguished according to (i) the kind of waves used for illumination and (ii) how the contrast is formed in the image.

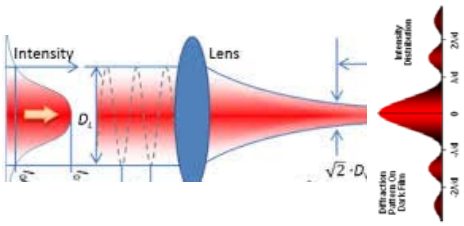
4.5 Resolution of Imaging Microscopy

Generally, the lateral resolution of any microscopy system is limited by two factors:

1. **Diffraction effects** due to the *finite entrance aperture* of the optical system
   = *fundamental* & ultimate resolution limit (= Abbe resolution limit,1873)

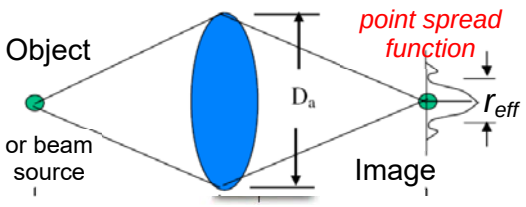
2. **Image distortions/smearing** caused by *lens imperfections* (aberrations)
   = "*practical*" resolution limit (always > diffraction limit)

⇨ Both effects limit the resolution ***independently*** of the magnification of the optical system.



For ***beam scanning microscopy***, both effects lead to a ***broadening of the probe spot size*** on the sample, with a certain minimal diameter that cannot be further reduced.

For ***imaging microscopy*** this leads to a ***broadening of the image*** of each object point, i.e., the smallest resolvable distance between to points is limited.



In sum, the total ***effective resolution*** r_{eff} is given by the square root summation:

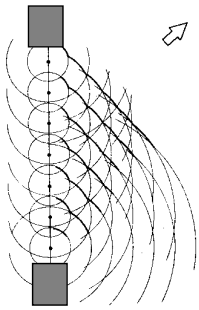
$$r_{eff} = \sqrt{r_{diff}^2 + r_{aber}^2}$$

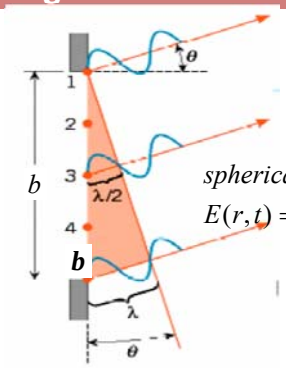
4.6 Abbe Resolution Limit (Diffraction) (see, e.g., E. Hecht: "Optics")

In any optical system, the optical beam is confined by ***apertures*** with finite diameter and size.

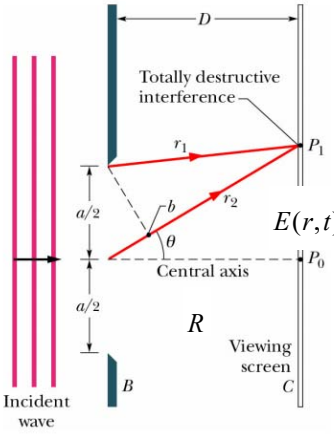
At these apertures, the beams are ***diffracted***. The resulting **diffraction patterns** can be derived using the **Huygens principle**, where the electric field and thus, the intensity at a point *P* away from the aperture is calculated by ***summation over spherical waves*** emitted with wave vector k=2π/λ from each point within the aperture of area *A*.

4.6.1 Diffraction by a Single Slit with a Width *b*





spherical waves : $E(r,t) = \varepsilon / r \cdot \sin(\omega t - kr)$



$E(r,t) = \frac{\varepsilon}{r}\sin(\omega t - kr)$

Emitted spherical wave : $E(r,t) = \varepsilon / r \cdot \sin(\omega t - kr)$

Total field at point *P*:

$$E(R,\theta,t) = \int_A E(r,t)dA = \int_A (\varepsilon / r)\cdot \sin(\omega t - kr)dA \cong \varepsilon / R \int_{-b/2}^{+b/2} \sin(\omega t - k(R - y\sin\theta))dy$$

Fraunhofer far-field diffraction
parallaxial approximation (r>>w))

$$r = R - y\sin\theta + (y^2/2b)\sin^2\theta + .. \cong R - y\sin\theta \quad \text{for} \quad y << R$$

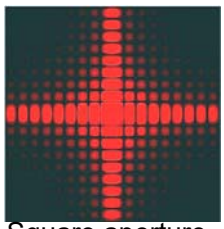
***E*-Field distribution:**

$$\Rightarrow E(R,\theta,t) = 2\frac{\varepsilon}{R}\sin(\omega t)\frac{\sin(1/2\cdot kb\sin\theta)}{kb\sin\theta}$$

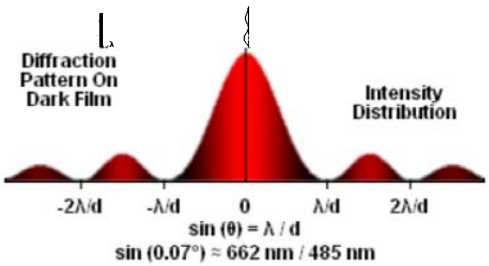
Far-field intensity distribution $I(\theta)$ for a single slit

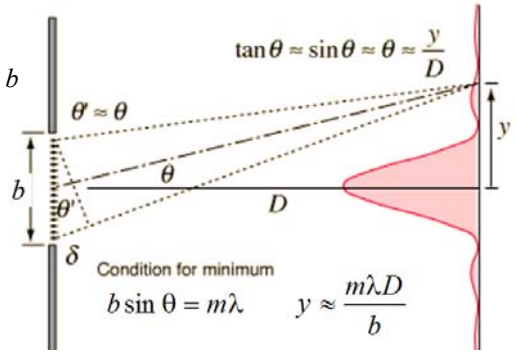
$$I(\theta) = c\cdot\varepsilon_0\left\langle E^2(\theta,t)\right\rangle_t = c\cdot\varepsilon_0\left\langle \sin^2(\omega t)\right\rangle_t \cdot 2E_0^2 \sin^2(\tfrac{1}{2}\cdot kb\sin\theta)/(kb\sin\theta)^2$$

$$I(\theta) = \langle E(t)^2\rangle = I_0\cdot\sin^2\beta/\beta^2$$ with $\beta = \frac{\pi\cdot b\cdot\sin\theta}{\lambda}$, θ = diffraction angle



Square aperture





Basic Properties:

1. **Central intensity maximum I_0** with series of side minma and maxima. $I_0 = c\cdot\varepsilon_0\cdot E_0^2 = c\cdot\varepsilon_0\cdot(\varepsilon/R)^2$
2. **Intensity minima** occur at: $\sin\beta = 0 \rightarrow \beta = m\cdot\pi \rightarrow \sin\theta = \lambda\cdot m/b$, $m = 1,2,3....$
3. **Phase difference** of out-most beams for these minima: $\Delta\phi = m\cdot\lambda$ *note* : $\lambda = \lambda_o/n$ (wavelength in medium)
4. **Broadening of the image** = Separation between the central maximum and the first side minimum: $\sin\theta_1 = \lambda/b$
5. **Stronger diffraction**, *i.e.*, larger angles and broader image **the smaller the slit width *b* .**

4.6.2 Diffraction by a Circular Aperture (Airy disk)

Diffraction pattern of a circular aperture (=Airy disk) with diameter D and radius R_0:
Derived in a similar way by *summation over all spherical waves* emitted from the disk area

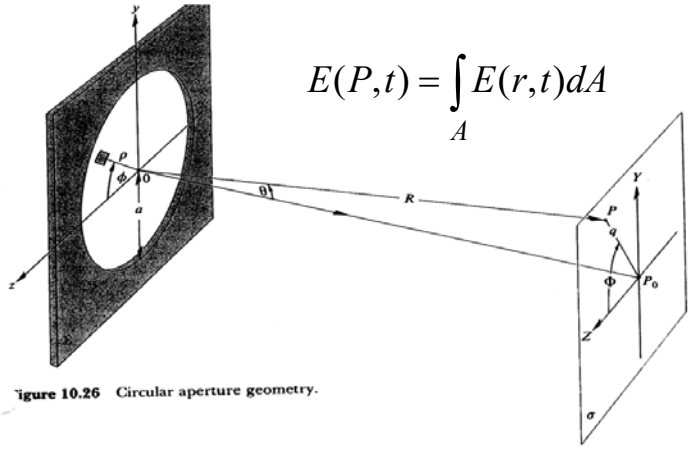
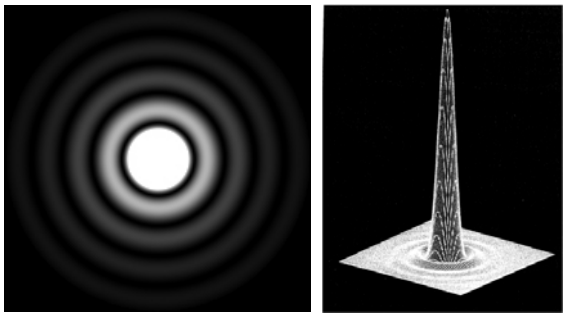


Figure 10.26 Circular aperture geometry.

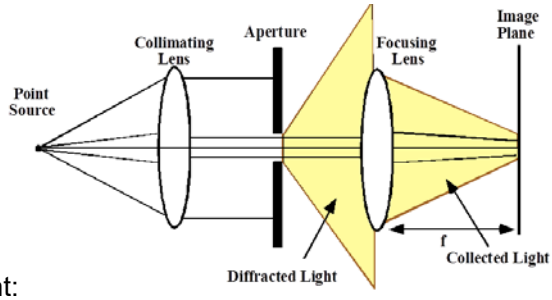


= Central intensity maximum with concentric satellite rings
"Airy intensity distribution"

The Airy intensity distribution

is described by **first order Bessel function $J_1(x)$:**

$$I(\theta) = I_0 \cdot 4 \cdot J_1^2(\beta)/\beta^2$$



where I_0 = primary intensity,
$J_1(\beta)$ = 1st order Bessel function with the argument:
$\beta = 2\pi nR \cdot sin\theta/\lambda$, where R = aperture radius,
λ = wavelength, n = refractive index, and θ = diffraction angle (observation angle).

Properties of the Airy intensity distribution:

Intensity-Minima are given by *zeros* of the Bessel function $J_1(\beta) = 0$ occurring at

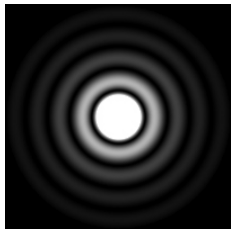
$\beta = 3.83, 7.01, 10.17, \ldots = 2\pi nR \cdot sin\theta/\lambda$ → $\theta_{minma} = m_i\lambda/D$ with

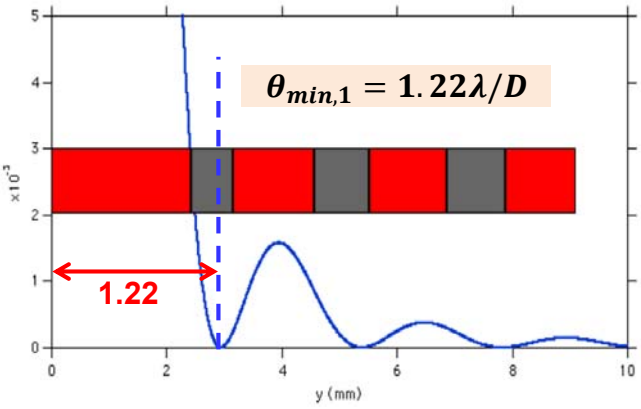
$m_i = 1.22, 2.23, 3.24, \ldots$ $(D = 2R)$

Radius of the 1st diffraction ring (Airy disk)
focused by a lens with a focal distance f onto a screen

$$r_{Airy} = 1.22\lambda f/D$$

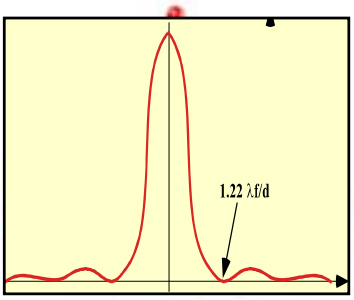
~ smallest achievable focal spot of a laser beam with diameter D and wavelength λ





For each minimum, the path difference between outer beams is: $\Delta\phi_{min,i} = m_i\lambda$ Eq. (1)

Ring	Circular aperture m	I_{max}	I_{total}	Single slit m	I_{max}
Central maximum	0	1	1	0	1
First dark	1.220			1.000	
Second bright	1.635	0.01750	0.084	1.430	0.0472
Second dark	2.233			2.000	
Third bright	2.679	0.00416	0.033	2.459	0.0165
Third dark	3.238			3.000	
Fourth bright	3.699	0.00160	0.018	3.471	0.0083
Fourth dark	4.241			4.000	
Fifth bright	4.710	0.00078	0.011	4.477	0.0050
Fifth dark	5.243			5.000	

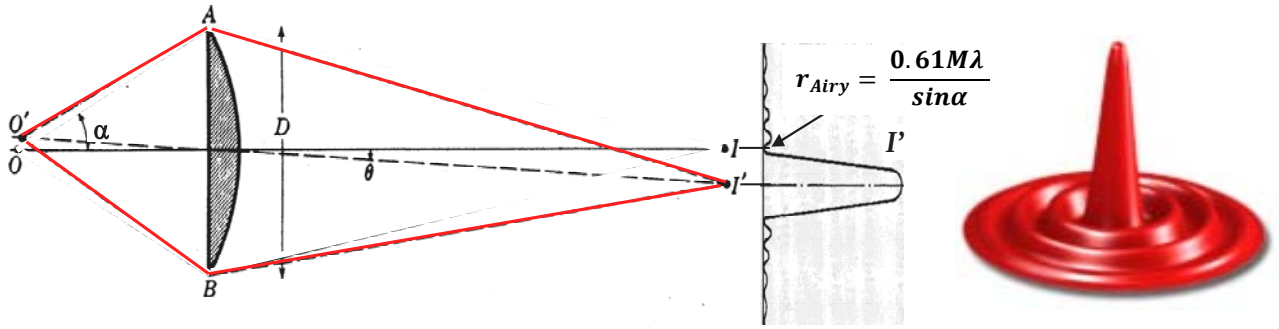


Note: In a medium with refractive index n, the wavelength $\lambda = \lambda_0 / n$ is shorter, accordingly: $\theta_{min} = \theta_{min,vac}/n$
i.e., *the intensity minima are more closely spaced , i.e., the diffraction broadening reduced !*

4.6.3 Diffraction broadening in the Image: Point Spread Function

For a microscope, the resolution is generally described by the so-called ***point spread function*** that is the intensity distribution of the ***image of an ideal point source*** produced by an optical system.

Its shape characterizes the performance of a microscope system and ***includes all broadening factors*** caused both by *diffraction* at the entrance aperture *as well as* by *lens aberrations*.



For a ***perfect diffraction-limited optical system*** *without* any lens aberrations, the lateral point spread function, i.e., radial intensity distribution *I*(r) of a point source is the ***Airy diffraction pattern***, which in the parallaxial approximation (small angles) is:

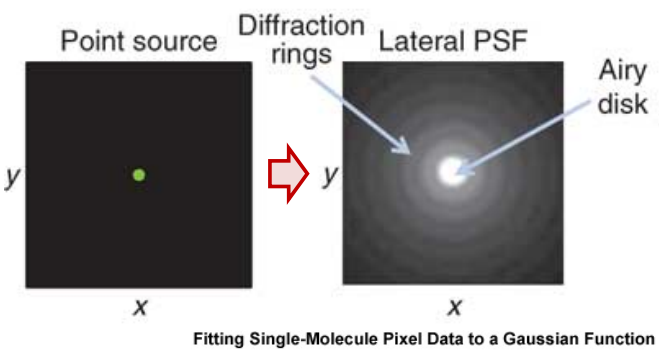
$$I_{Airy}(r) = I_0 \left[\frac{2 \cdot J_1(2\pi \cdot sin\alpha \cdot \frac{r}{M\lambda})}{2\pi \cdot sin\alpha \cdot \frac{r}{M\lambda}}\right]^2$$

» **Radius of the 1st Airy disk**: $R_{Airy} = 0.61\, M \frac{\lambda}{sin\alpha}$

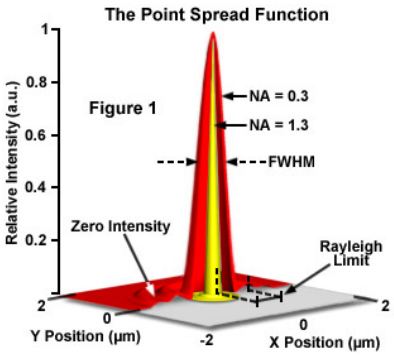
where again J_1 is the first order Bessel function, α is the aperture angle, n the refractive index of the medium between the object and lens, M the magnification and λ the wavelength of the light.

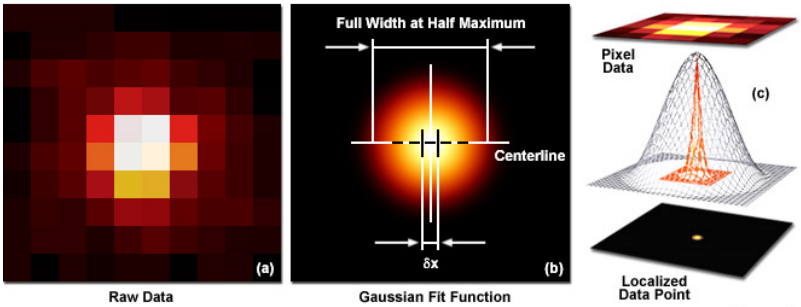
Experimental Determination of the point spread function

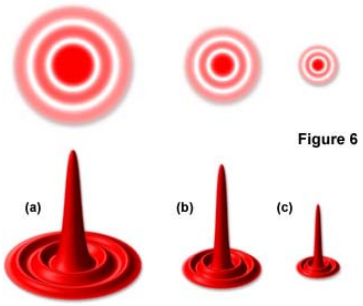
The point spread function of an optical system can be *measured* directly by recording the image of a point source, for example of a single fluorescent molecule or a small nanocrystal or quantum dot.



Fitting Single-Molecule Pixel Data to a Gaussian Function





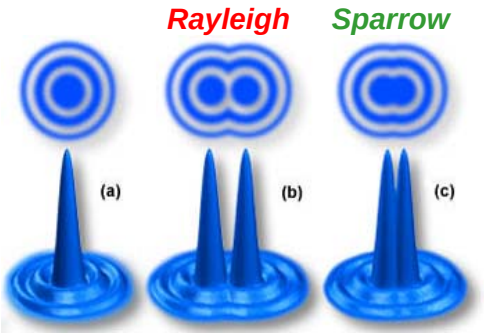


The width of the central maximum in the image, i.e., the ***radius of the first dark ring*** determines the ultimate microscope resolution !

4.6.4 Diffraction Limit of the Lateral Resolution (Abbe Limit)

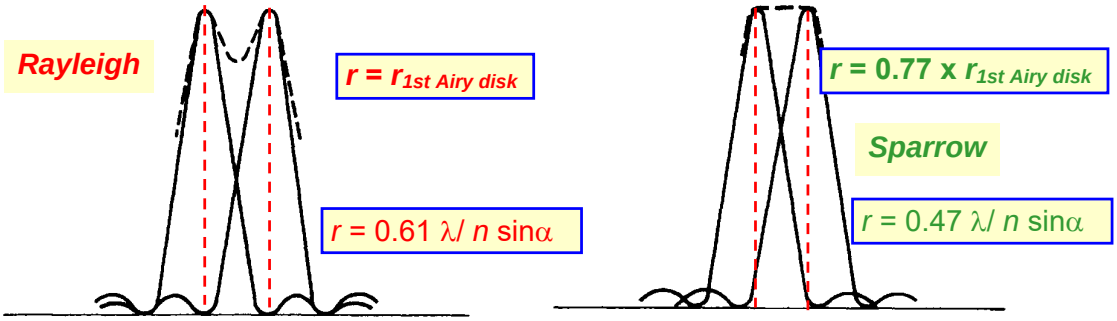
The finite diameter of the objective lens of a microscope acts as a finite size circular aperture that creates an Airy diffraction pattern in the image plane for every point of the object.

When two object points are very close to each other, the point spread function (Airy patterns) of their images overlap so that eventually the two points can no longer be resolved.



(a) Rayleigh Criterion

Two point sources can be just resolved when the central ***maximum*** of the diffraction pattern of one point coincides with the ***minimum*** of the neighbouring point.



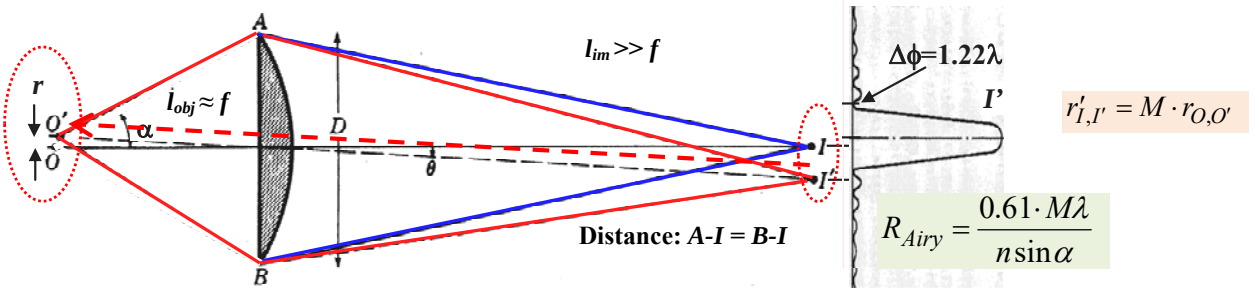
(b) Sparrow Criterion (= Ultimate resolution limit)

Two point sources can be just resolved when there starts to exists an intensity dip between the image of the points.

Abbe resolution limit derived by the Rayleigh criterion

In a **microscope**, the diameter *D* of the objective is finite and thus acts as a finite size circular aperture that creates an Airy diffraction pattern in the image plane for each point of the object.

As shown on the previous slides, the radius of the 1st Airy disk of the image of a point source located in the object place at point O is $R_{Airy} = 0.61M\lambda / n\sin\alpha$ where $\lambda_{Airy} = \lambda_{vac}/n$ is the wavelength of light in the medium with refractive index n between the lens and the objective, and M the magnification.



⇨ Tracing back this position I' to the object plane, the image point of the Airy minimum corresponds to an object at position *O'* that according to the Raleigh criterion can still be resolved.

As seen in the figure, point *O'* is separated vertically from point *O* by the distance $r = \tan\theta_{min} \cdot l_{obj}$ where $l_{obj} \sim f$ is its distance from the objective lens.

Since for $\lambda << D$, $\tan\theta_{min} \approx \theta_{min}$ it follows that $r_{min} = 1.22\,\lambda/n \cdot l_{obj}/D_{len} = 0.61\lambda/n \cdot f/R_{lens}$.

Because $f/R \sim 1/\sin\alpha$ the ***diffraction limited resolution*** turns out to be: $r_{min} = 0.61\,\lambda / n\sin\alpha$

Resolution and Numerical Aperture

The **diffraction limited resolution** $r_{dif} = 0.61\ \lambda / n \sin\alpha$

increases with (i) decreasing wavelength λ, and
(ii) decreasing acceptance angle α

Wavelength (nm)	Resolution (nm)
360 – 450 nm (violet)	190 – 250 nm
450 – 500 nm (blue)	250 – 300 nm
500 – 570 nm (green)	300 – 350 nm
620 – 750 nm (red)	380 – 460 nm

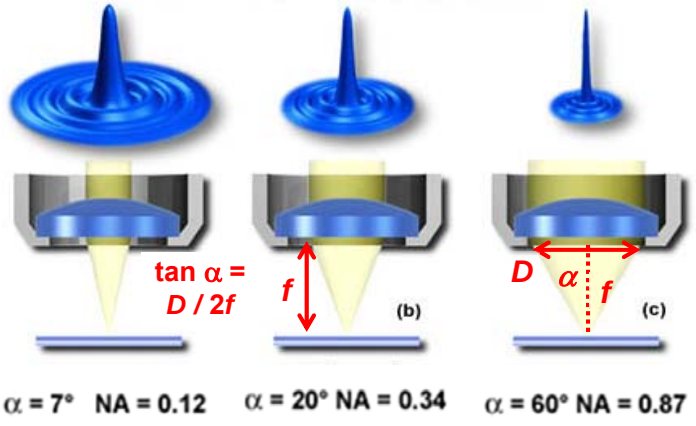
The denominator
NA = $n \sin\alpha$ is called ***"Numerical Aperture"***

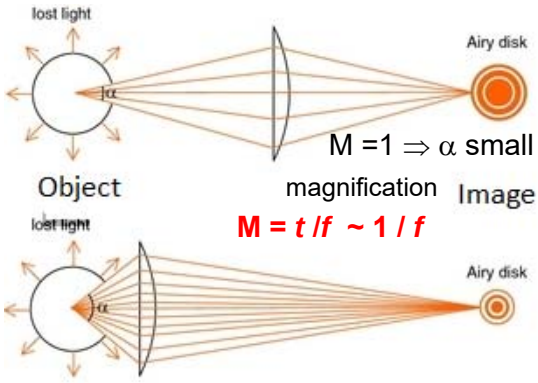
It combines the aperture angle α of the entrance lens and the refractive index *n* of the medium between sample and lens and thus, characterizes the properties of a given optical lens system.

The numerical aperture NA can be increased by (a) increasing the diameter of the objective lens and/or (b) decreasing the focal length *f* of the lens as shown by the figures below:

⇨ *High NA* = small working distance WD » small *f*-number » high magnification [$M \sim t / f$]

⇨ But: For n =1, sinα is limited to < 1 ⇒ NA < 1.



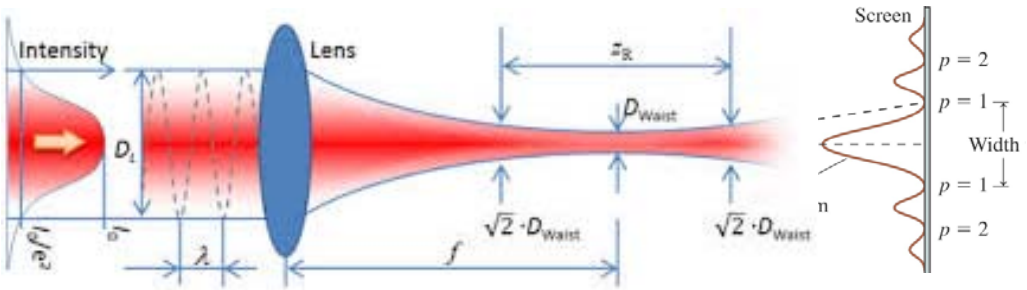


M =4 ⇒ α, NA large

4.6.5 Resolution of Scanning Microscopy: Minimal Spot Size

In scanning (laser) microscopy, imaging is performed by scanning a tightly focused laser spot over the sample surface and recording the reflected or transmitted light intensity as a function of spot position.

The resolution is given by minimal spot size, which is again limited by the diffraction at the focusing lens, or more specific, for a quasi parallel laser beam focused by the beam diameter (=beam waist).

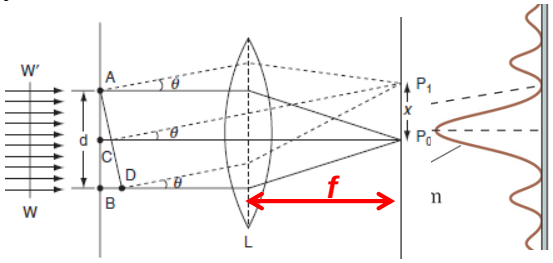


The diffraction angle of the 1st Airy disk caused by the diffraction is again given by $\theta_{min,1} = 1.22 \cdot \lambda/Dn$, which is appears at a distance x from the optical axis with $x = f \tan\theta$. When the lens diameter is large, θ_{min} is small and thus, $x = f \tan\theta \approx f \bullet \theta \approx f \bullet$ **NA.** Accordingly,

the diameter of the Airy disk (=spot size) is:

$$d_{spot} = 1.22 \cdot \lambda\ f/D$$

where the parameter f/D is called
"f" number of the lens.



⇨ Note that for a laser with *Gaussian beam profile*, the laser does not uniformly fill the aperture, for which reason the spot size is slightly increased to $d_{spot} = 4/\pi \cdot \lambda\ f/D$. Example: λ = 680nm, D = 5 mm, f = 5mm » d_{spot} = 1.9 µm

4.7 Modulation Transfer Function: Resolution in *Fourier* Space

4.7.1 Resolutoin for a Periodic Line Grating

For a 2*b* periodic line grating with line width *b*, the angles φ_i of the diffraction maxima are given by:

$\sin \varphi_i = i \cdot \lambda / (2\, n\, b)$ with i … diffraction order, *n* … refractive index, λ … wavelength

If the grating is imaged with an objective lens with a capture angle α, only diffracted beams with $\varphi_i < \alpha$ can contribute to the image formation.

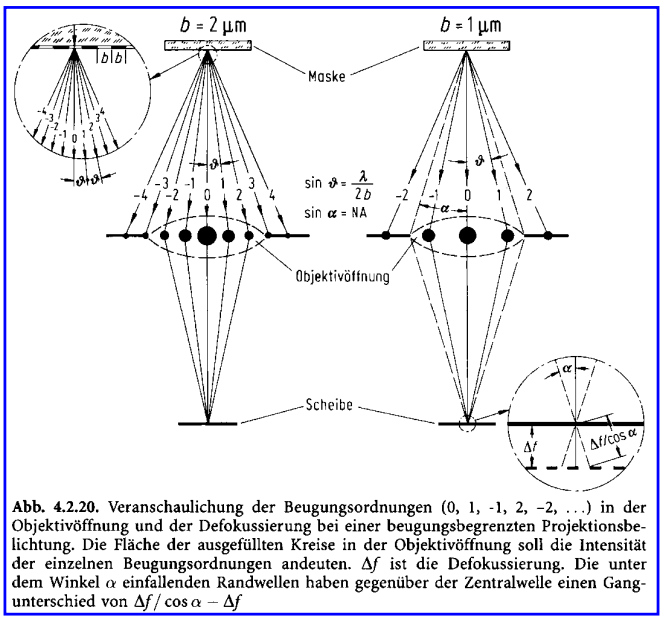


Abb. 4.2.20. Veranschaulichung der Beugungsordnungen (0, 1, -1, 2, -2, …) in der Objektivöffnung und der Defokussierung bei einer beugungsbegrenzten Projektionsbelichtung. Die Fläche der ausgefüllten Kreise in der Objektivöffnung soll die Intensität der einzelnen Beugungsordnungen andeuten. Δf ist die Defokussierung. Die unter dem Winkel α einfallenden Randwellen haben gegenüber der Zentralwelle einen Gangunterschied von $\Delta f / \cos\alpha - \Delta f$

Resolution Criterion:

The grating can only be resolved if at least the 1st order diffracted beams are still collected by the lens ($\varphi_1 \leq \alpha$).

Condition: $\sin \varphi_1 = \lambda / (2\, n\, b_{min}) =! \sin \alpha$

→ $R = b_{min} = 0.5 \cdot \lambda / (n \cdot \sin \alpha)$

or $R = 0.5 \cdot \lambda / NA$

with $NA = (n \cdot \sin \alpha)$

(numerical aperture of the lens)

⇨ Thus, the objective acts as a low pass filter in diffraction space !

⇨ **Off-axis illumination** increases the resolution because more higher order diffraction maxima go through the objective.

4.7.2 Fourier Decomposition of the Sample Contrast

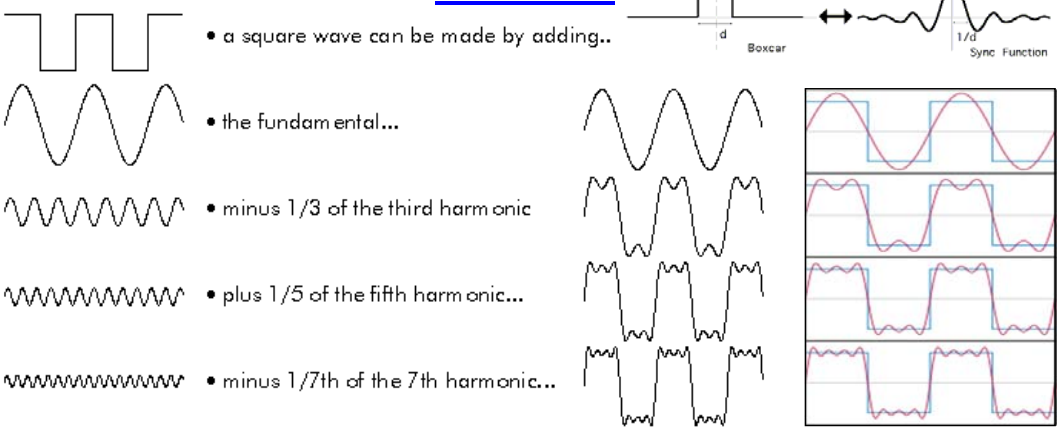
Any real space object with intensity contrast $C_s(x,y)$ can be thought to be composed up of a sum of sinus and cosine functions of various wavelengths λ or wave vectors $k = 2\pi/\lambda$, with the **amplitude** of each Fourier component $A(k)$ or $A(\lambda)$

(For periodic functions, the integral can be replaced by a summation)

$$C(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(k_x,k_y)\, e^{-i(k_x x + k_y y)}\, dk_x\, dk_y$$

$$C(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} A(n)\sin(k_n x) + B(n)\cos(k_n x) \quad \text{where } k_n = \frac{2\pi n}{\lambda}$$

Example:



$$|A(n_{odd})| = 1/n$$

$$A(n_{even}) = 0$$

Box function:

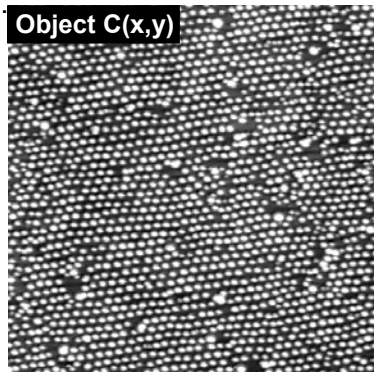
⇨ An exact 1:1 Fourier representation of $C(x)$ requires an ***infinite number* of Fourier components**.

⇨ In a ***real image formation process***, however, the higher order spatial frequencies are cut off, i.e, $A(k)=0$ for $k > k_c$. Thus, only a *finite Fourier series* contributes to the image and as a result, the sample contrast $C_s(x,y)$ is no longer exactly reproduced, i.e., $C_{image}(x,y) \neq C_{sample}(x,y)$.

Fourier Representation of the Sample

The Fourier space representation of the **intensity contrast $C_s(x,y)$** of a real space object is consists of the sum of plane waves with an infinite number of spatial frequencies **$k = 2\pi/\lambda$** . Each plane wave is weighed by a different amplitude **$A(k_x,k_y)$** (=Fourier component), i.e.

$$C(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(k_x,k_y)e^{-i(k_x x + k_y y)}dk_x d_y$$



$k_x = 2\pi/\lambda_x$ and $k_y = 2\pi/\lambda_y$ are the **wave vectors** in *x* and *y* direction of each plane wave, which are inversely proportional to the corresponding **spatial wavelength $\lambda_{x,y}$.**

$A(k_x,k_y)$ is the **amplitude** of the plane wave component with wave vector (k_x,k_y). Thus: **$C(x,y) = FFT(A(k_x,k_y))$**.

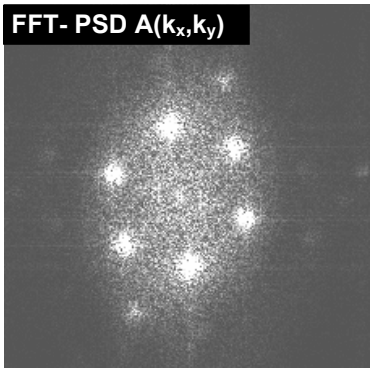
The **Fourier amplitudes** **$A(k_x,k_y)$** (=contribution of each frequency or wavelength to C(x,y)) can be calculated by the ***inverse*** **Fourier transformation (iFFT)** of the object contrast function **$C(x,y)$**.

Thus: **$A(k_x,k_y) = iFFT(C(x,y))$**

$$A(k_x,k_y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C(x,y)e^{+i(k_x x + k_y y)}dxdy$$

From the knowledge of the Fourier amplitudes **$A(k_x,k_y)$** the sample structure **$C(x,y)$** can be reconstructed and vice versa.

FFT ⇄ **i-FFT**



4.7.3 Modulation Transfer Function

(a) Ideal optical imaging system

For an ideal ***perfect*** imaging process, each spatial frequency contained in the object would be *exactly* reproduced, i.e., $A_{image}(k_{ximage},k_{yimage}) = A_{sample}(k_x,k_y)$.

Accordingly, the sample contrast

$C(x,y) = FFT(A(k_x,k_y))$ would be ***exactly*** reproduced in the image *without any loss of information*.

(a) Real optical imaging system

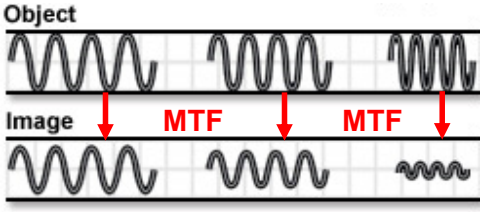
In a ***real*** imaging system, different spatial frequencies are reproduced differently well because high spatial frequencies (small period structures) are generally reproduced less well than low spatial frequencies. This means that the amplitudes $A(k)$ of Fourier components with high spatial frequencies contained in the object are reduced by a certain factor in the image.

⇨ The ability of a lens system to reproduce different spatial frequencies of the object in the image is characterized by the **modulation transfer function MFT(*f*) or MFT(*k*).**

Definition of the modulation transfer function:

$$MTF(k) = \frac{\text{image contrast modulation}}{\text{object contrast modulation}} = \frac{A_{im}(k)}{A_{obj}(k)}$$

at a given spatial frequency *k*.



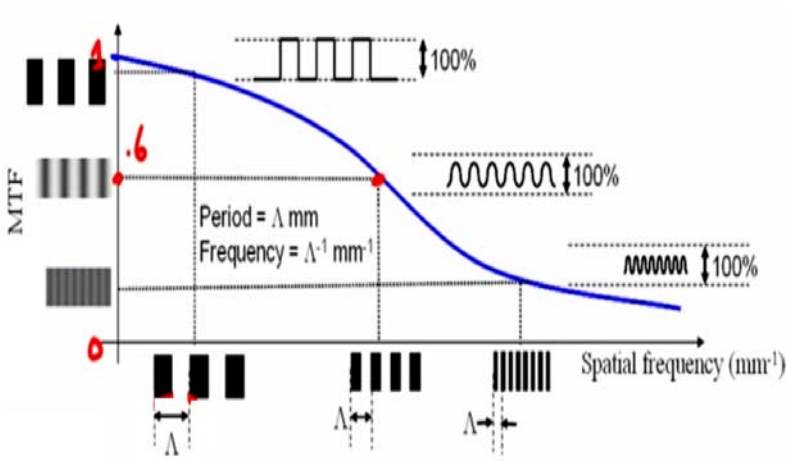
Properties of the Modulation Transfer Function

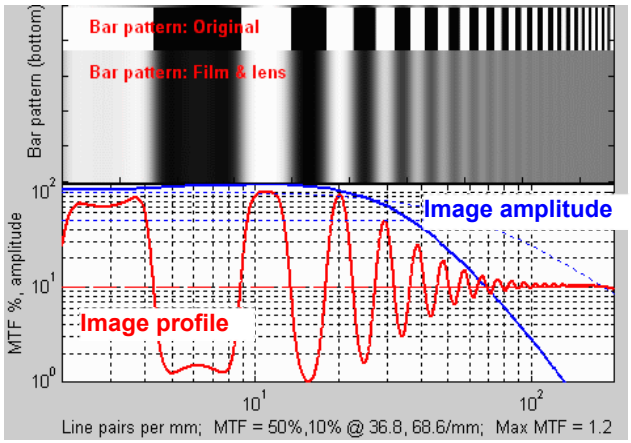
For a perfect imaging MTF = 1 at all frequencies, whereas for a real imaging system the MTF decreases with increasing wave vector, i.e., with decreasing periodicities contained in the sample.

As a result, there exists a certain **maximum *cut-off spatial frequency* k_c** above which the MTF is zero, meaning that all high frequency information is eliminated in the image formation.

⇨ **A lens system acts like a low pass Fourier filter such that the contributions of the higher spatial frequencies in the image are removed.**

Real modulation transfer MTF(*k*) plotted versus spatial frequency $k = 1/\lambda$



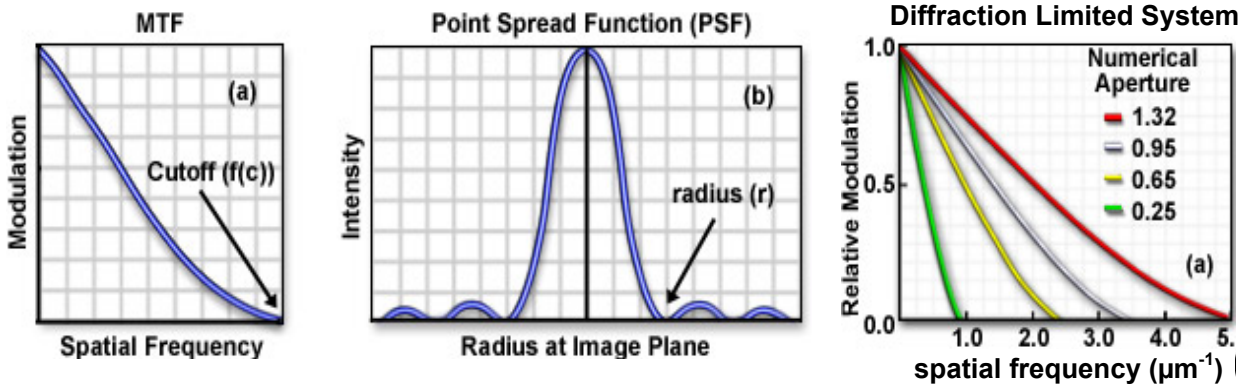


Modulation Transfer Function MTF(*k*) of a Diffraction Limited System

The modulation transfer function is the inverse FFT of the ***point spread function*** (=*image of a point source*).

For a **diffracting limited optical system** the MTF can be thus directly calculated from the

Airy intensity distribution as: $MTF(k) = 2(\phi - \cos\phi \sin\phi)/\pi$ **where** $\phi = \cos^{-1}(k\,\lambda_{illu}/2NA)$



- At *low spatial frequencies*, MTF(k) = $A_{im} / A_{obj} \sim 1$. Thus, the image contrast is equal to that of the sample. At *higher frequencies*, MTF decreases and falls to zero at the cut-off frequency k_c.
- For a diffraction limited system, the cut off frequency k_c is given by: $k_c = f_c = 2.1\ NA / \lambda$
  Thus, the highest spatial frequency is determined by the Sparrow limit.
- The corresponding radius of the first dark concentric ring surrounding the central peak of a point spread function is given by: $r = 0.61\ \lambda / NA$
- Objectives with low NA produce a wider point spread functions and thus, lower resolution.

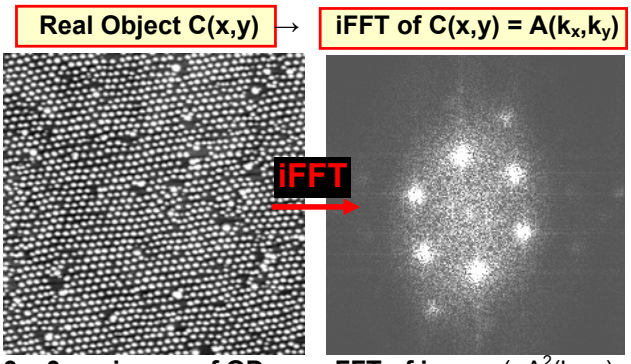
4.7.4 Calculation of the Image using the Modulation Transfer Function

Once the modulation transfer function MTF(k) is known, the Fourier component amplitudes of the image, i.e., $A_{image}(k_x,k_y)$ can be calculated using: $A_{im}(k) = MTF(k) \cdot A_{obj}(k)$

From the resulting $A_{image}(k_x,k_y)$, the **real space image $S_{image}(x,y)$** can be calculated by back transformation of $A_{image}(k_x,k_y)$ into real space:

$$S_{im}(x) = iFFT[A_{im}(k)] = iFFT[MTF(k) \cdot A_{obj}(k)]$$

Example: Influence of the modulation transfer function on the image

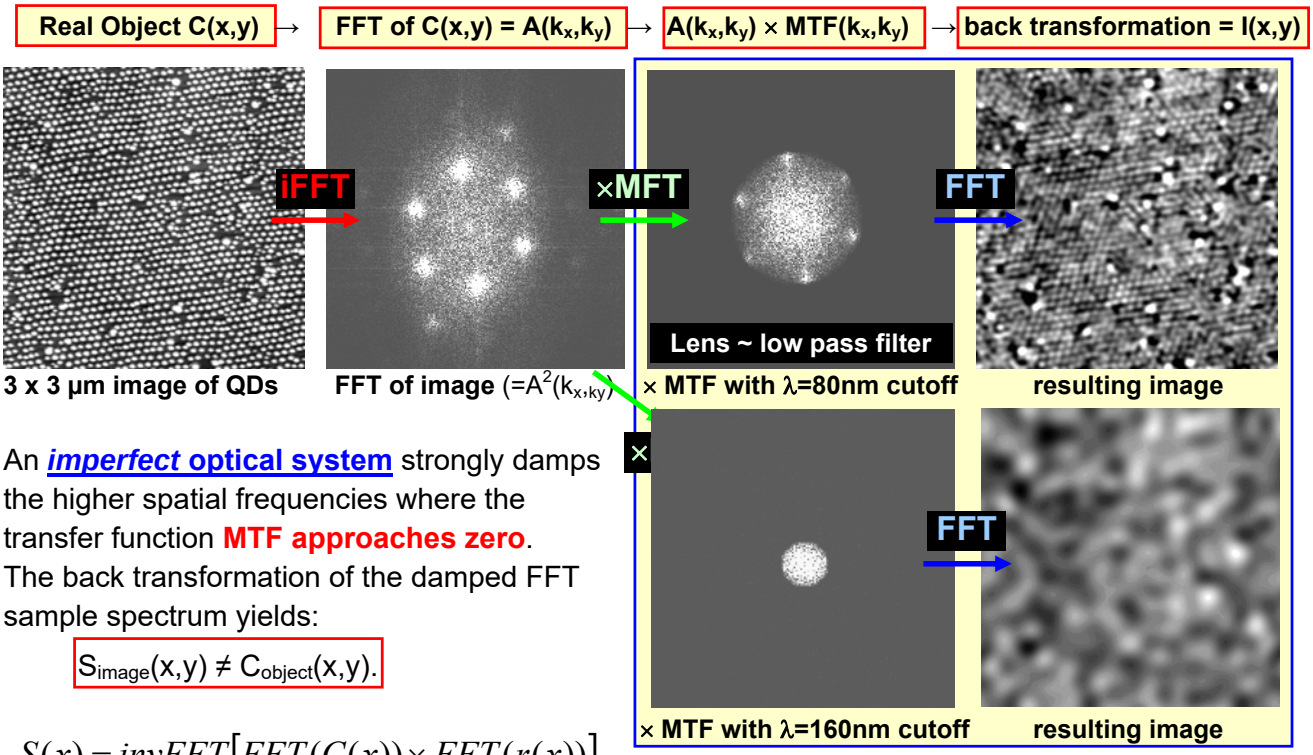


The contrast modulation of the object at a spatial frequency k is given by the corresponding **Fourier amplitude** $A(k_{x,ky})$ of the object contrast $C(x,y)$.

This amplitude is obtained by the Fourier transformation of the object contrast distribution $C(x,y)$ or $I(x,y)$ that represents the real space structure of the object

A ***perfect* optical system** with infinitely high resolution would have a modulation transfer function of **MTF = 1** (unity) **for *all* spatial frequencies**. Only then, it follows that $I_{image}(x,y) = I_{object}(x,y)$!!!

Image formation calculated using the modulation transfer function



An ***imperfect* optical system** strongly damps the higher spatial frequencies where the transfer function **MTF approaches zero**. The back transformation of the damped FFT sample spectrum yields:

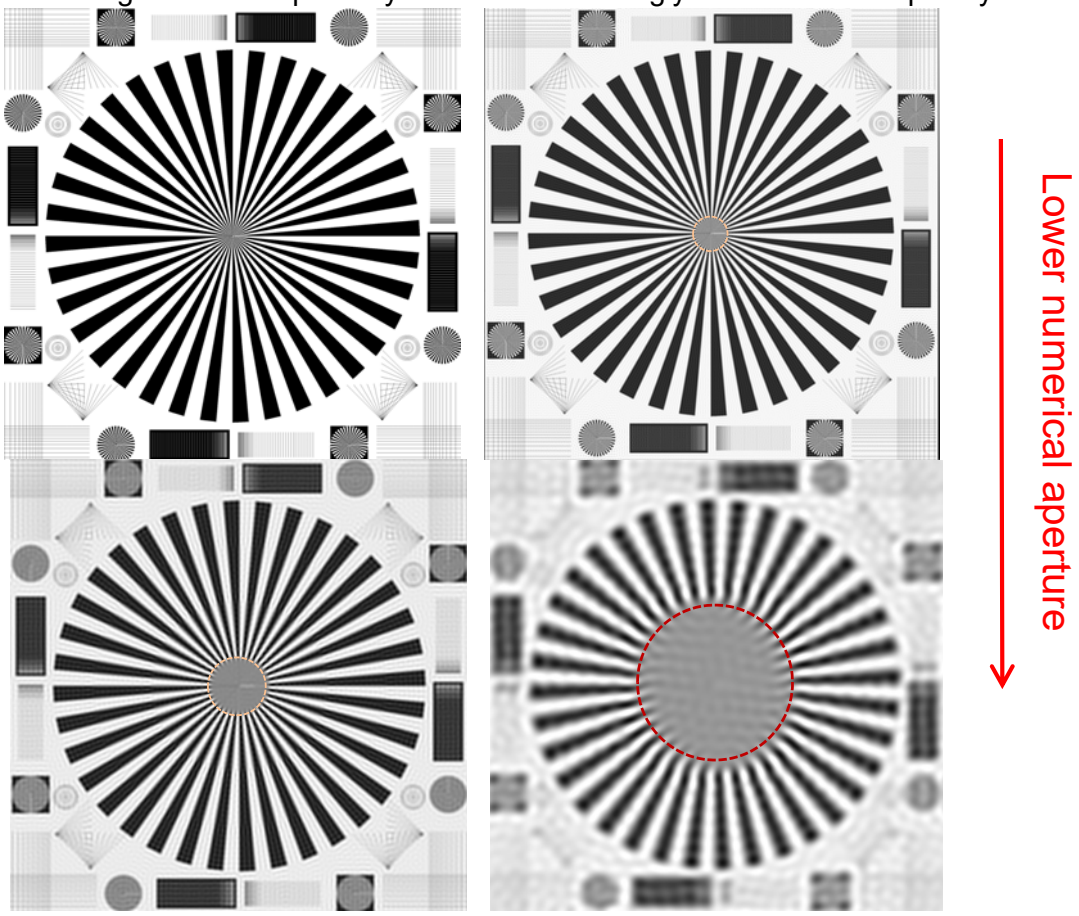
$S_{image}(x,y) \neq C_{object}(x,y)$.

$$S(x) = invFFT\left[FFT(C(x)) \times FFT(r(x))\right]$$

The **modulation transfer function includes** both image broadening effects due to **diffraction *and* Lens imperfections**. A lens system without aberrations is termed diffraction limited.

Other Example: Testpattern for Quantification of the Image Resolution

"Siemens Star" imaged with an optical system with increasingly lower cut-off frequency



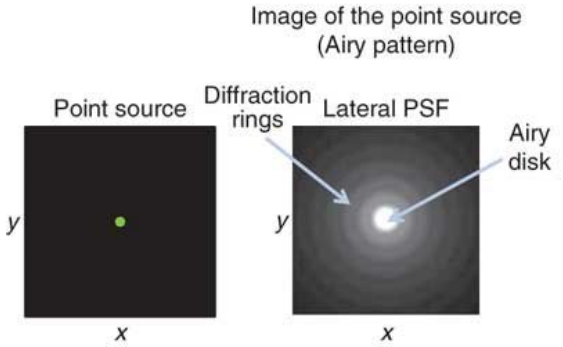
4.7.5 Measurement of the Modulation Transfer Function

The **modulation transfer function** and **point spread function** of any optical system including a microscope can be ***determined*** using a **point-like light source** such as single fluorescing molecules, quantum dots or nanocrystals, or illuminated point like apertures.

Procedures:

(a) The intensity distribution of the image of the point source corresponds to the **point spread function** of the microscope.

The **Fourier transform** of this intensity profile of the point source image yields the **modulation transfer function MTF(*k*)**

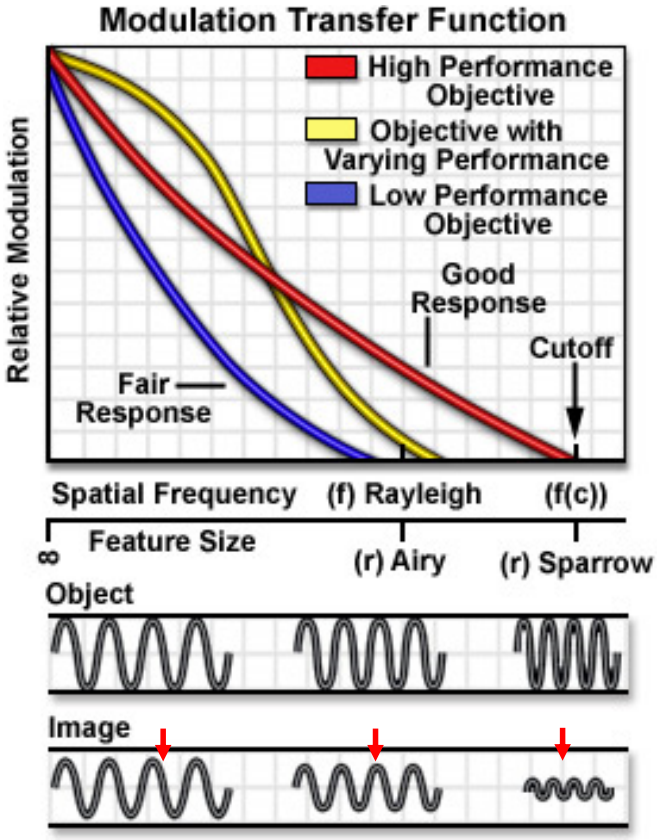


(b) The MTF can be also ***directly measured***:
It is the diffraction pattern formed in the ***back focal plane*** of the objective system.

Alternatively, the intensity profile of a sharp shadow edge can be also used to determine the MTF function, again by calculating the Fourier transform of the intensity profile in the image.

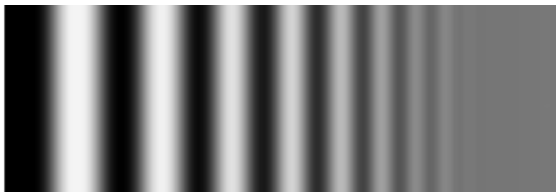
In general, for an ideal point source the **modulation transfer function** corresponds to the ***diffraction pattern*** formed in the ***back focal plane*** of the objective system, and the point spread function to the intensity distribution in the image plane.

Thus, the MTF and PSF are strictly interrelated by the Fourier transformation.



Properties of the MTF:

- ❖ continuously decreases with increasing spatial frequency until it reaches zero at the cut-off frequency k_c (or f_c)
- ❖ all frequencies higher than k_c are completely blocked by the optics.
- ❖ high performance objectives with large NA have a higher cut-off frequency and higher MFT values at lower frequencies.



4.8 Axial Resolution and Depth of Field

The point spread function is actually a ***3D intensity distribution*** in (x,y) as well as *z* (axial) direction.

Axial Airy intensity distribution:

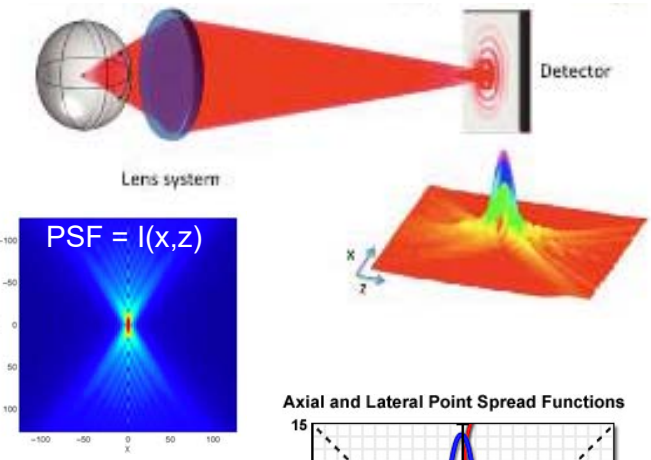
$$I_{Airy}(x=0,z) = C \cdot \frac{NA^4}{M^2}\left[\frac{\sin(\pi NA^2 \cdot z/2n * M^2\lambda)}{\pi NA^2 \cdot z/2n * M^2\lambda}\right]^2$$

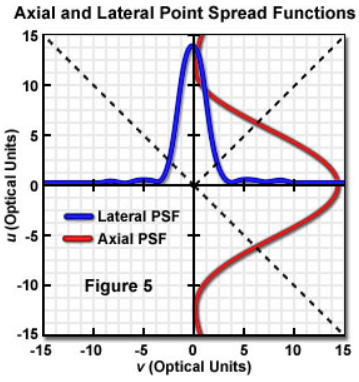
This means that the image of a point source is not only broadened in the lateral, but also *in the axial direction*, with an axial width (=1st zero of *I*(*z*)

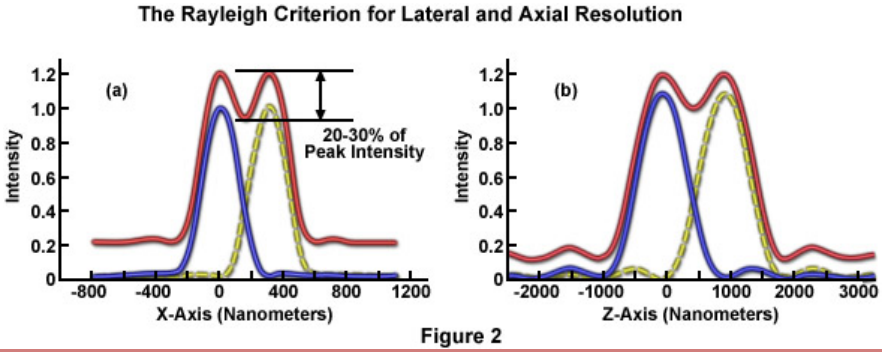
$$\Delta z_{Airy} = 2n * \cdot M^2\lambda / NA^2$$

Thus, the ***axial resolution***, defined as minimal z-separation of two points that can be discriminated in the microscope when focusing through the object is limited to:

$$\Delta z_{res} = 2n * \lambda / NA^2$$

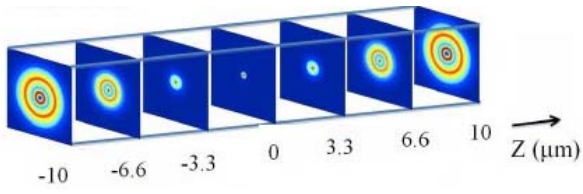






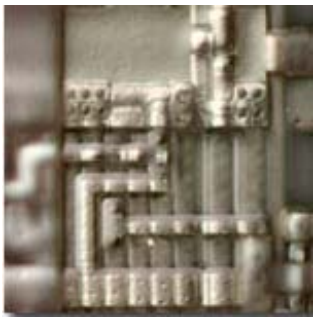
Depth-of-Field / Focus

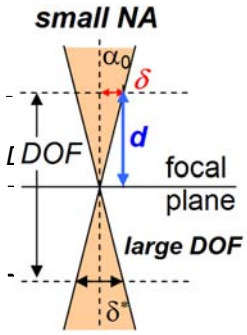
Due to the broadening of the PSF in the axial direction an object part that is offset from the ideal focal plane appears broadened in the image plane.

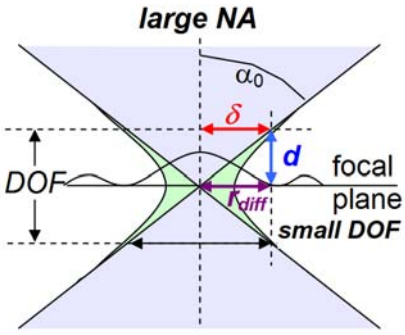


The **Depth-of-field (DOF)** is thus defined as the vertical distance or sample height over which thea microscopy image appears still sharp.

⇨ Like the lateral resolution, the DOF is determined by acceptance angle α_0 (*NA*) of the objective lens.







Derivation of the DOF using the following resolution criterion:

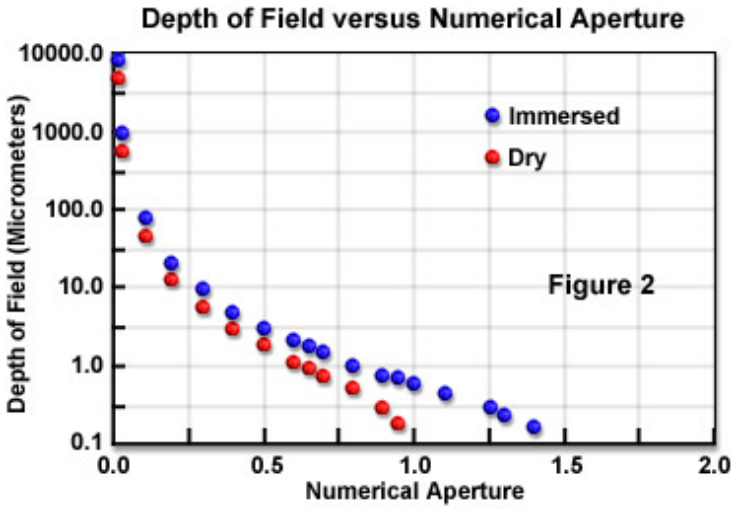
Criterion: DOF = Depth of sample across which the **broadening** $\delta(z)$ due to the spreading of the light cone by α_0 is equal to the lateral resolution r_{diff} in the focal plane, i.e., $\delta(z) = r_{diff}$

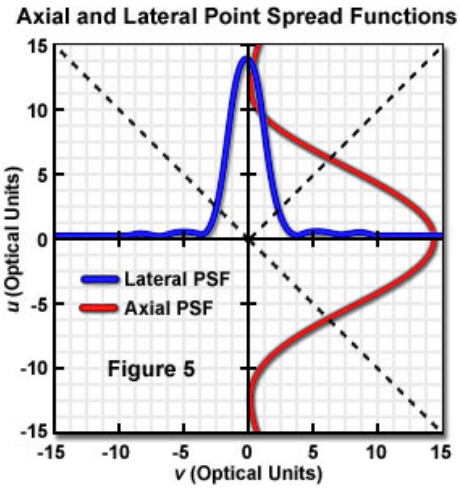
From above figure: $\delta = d \cdot \tan\alpha$ & $r_{diff} = 0.61\,\lambda / n \sin\alpha \Rightarrow d = 0.61\,\lambda / (\tan\alpha \cdot n \sin\alpha)$

$\Rightarrow$ $DOF = 2d = 1.22\,\lambda \cdot (n^2 - NA^2)^{1/2} / NA^2$ For small α: $DOF \approx n \cdot \lambda / NA^2$

» Increasing the numerical aperture NA *decreases* the depth-of-field of the image !

$DOF = 2d = 1.22\,\lambda \cdot (n^2 - NA^2)^{1/2} / NA^2$ For small α: $DOF \approx n \cdot \lambda / NA^2$





Conclusions:

⇨ For large sample thicknesses or large topography variations, the image is not sharp over the whole microscopy image.

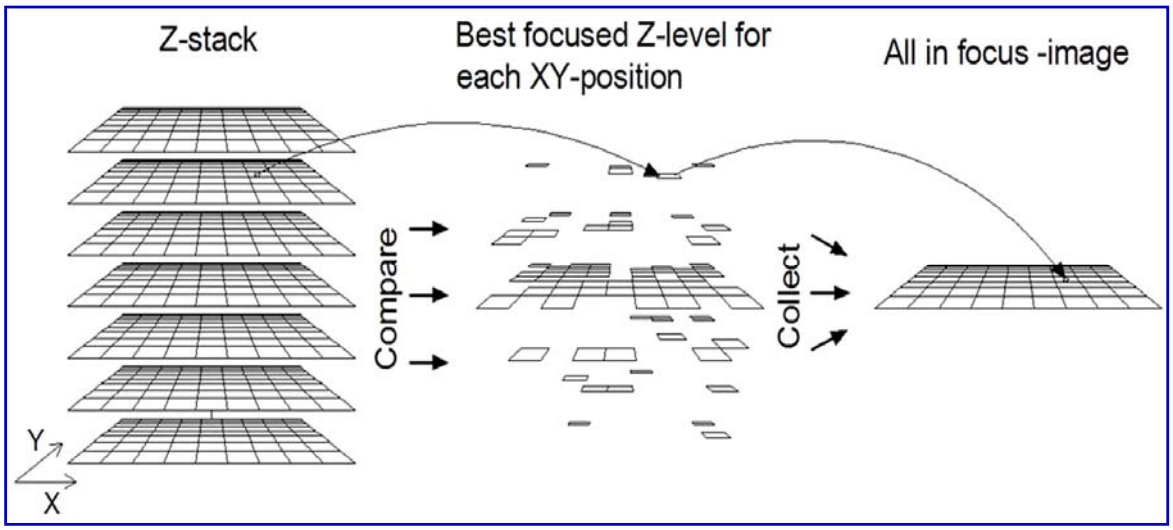
⇨ A **high lateral resolution** and a **large DOF cannot be achieved at the same time**.

⇨ A **large DOF** mean a **low axial resolution** because at large DOF the microscope cannot discriminate between objects located at different z-positions. This ability is called axial resolution, which is important for 3D imaging of the objects in x,y,z direction.

4.9 Z-Series: A Method to Overcome the DOF Limitation

The limit of image sharpness due to the finite DOF can be overcome by recording a whole series of images (**= "z-stack"**) with incremental vertical shifts of the focal plane.

From each image, the **sharp parts** are selected by an image processing software, and then a total ***virtual image*** is constructed only from the sharp regions for the whole sample series.

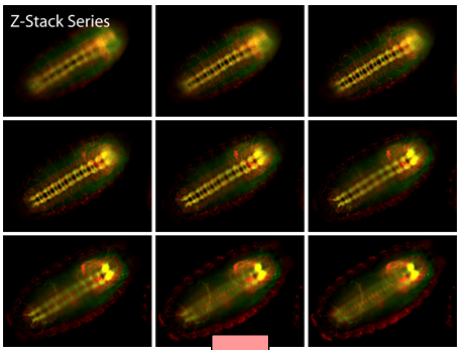


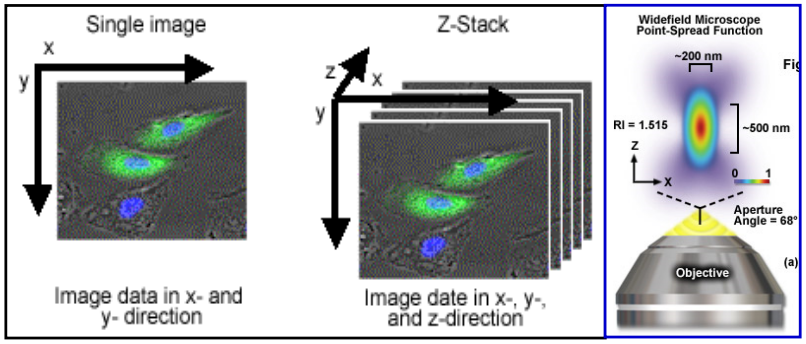
Selection criterion:
The sharpest z-element is the one with the highest local contrast modulation,
i.e., the element containing the highest spatial frequencies.

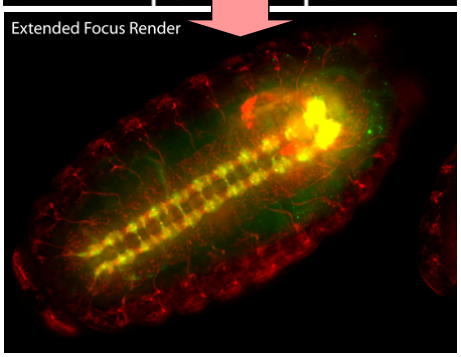
Example: Z-Series in Optical Imaging Microscopy

When a z-image series is recorded with precisely-controlled z-displacement increments Δz, the x,y coordinates of the sharp parts of the images can be correlated with their location in z-direction.
In this way, a ***complete 3D sample image*** is obtained.

Examples for z-series 3D image reconstruction





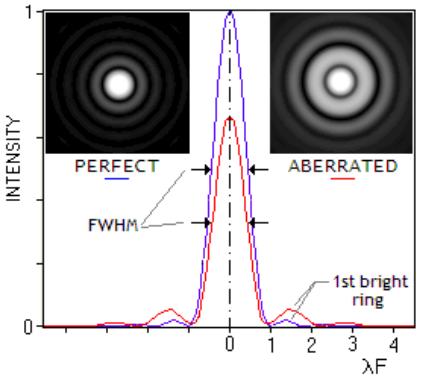


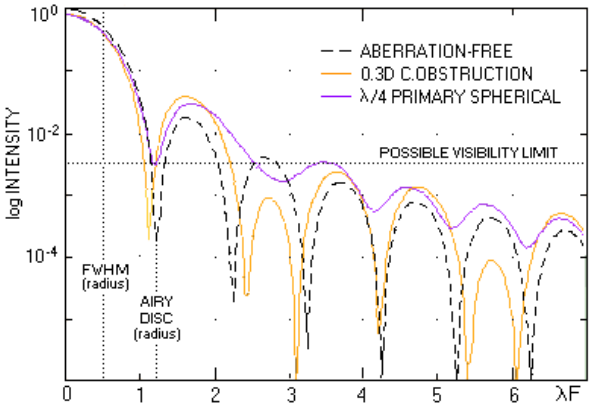
Lateral resolution: $\Delta x = 0.66\,\lambda / NA \sim \lambda / 2$ **Vertical resolution**: $\Delta z = DOF \sim n \cdot \lambda / NA^2 \sim \lambda$

4.10 Resolution Limits due to Lens Aberrations

For imperfect optical systems, the PSF point spread function is ***additionally broadened***.

This is due to the fact that *imperfect lenses* do not focus all beams exactly on the same image spot in the image. This leads to a smearing of the image and thus, the resolution is decreased.





The resulting ***effective* resolution** is described by the ***RMS sum*** of diffraction broadening r_{diff} *and* aberration broadening r_{aber} :

$$r_{eff} = (r^2_{dif} + r^2_{aber})^{1/2}$$

The overall performance of an ojective lens is then characterized by an **effective numerical *aperture*** NA_{eff} : which is always smaller than the geometrical numerical aperture !

$$NA_{eff} = n \sin\alpha \cdot r_{dif} / r_{eff}$$

There are **three main types of lens errors** (aberrations) that degrade the spatial resolution:

(1) Spherical Aberration, (2) Chromatic Aberration and (3) Astigmatism.

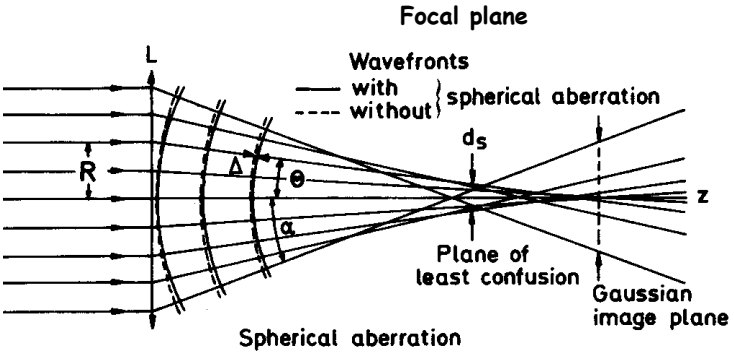
4.10.1 Spherical Aberration

Spherical aberration arises from the spherical shape of optical lenses.

A spherically shaped lens acts differently on central and far off-axis rays:

» Off-axis beams are bend more strongly towards the optical axis.

» Thus, parallel beams are not focused on the same point on the optical axis.



Result: Broadening of the spot size.

Spot radius r_{sph} **in the focal plane:** $r_{sph} = C_s \cdot \alpha^n$ C_s … spherical aberration constant.

α - exponent: $n \sim 3$

Radius in the plane of least confusion: $r^*_{sph} = r_{sph} / 4$

Spherical aberration decreases with *decreasing* lens curvature, i.e., *increasing* focal length.

Example: Calculation of spherical aberration for a simple *half-spherical* lens

Focal length *f* for a thin spherical lens: $\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$. Half-spherical lens: $\boxed{f_0 = \frac{R}{n-1}}$

(1) **Incidence angle** γ on curved lens surface: $\sin\gamma = y/R$

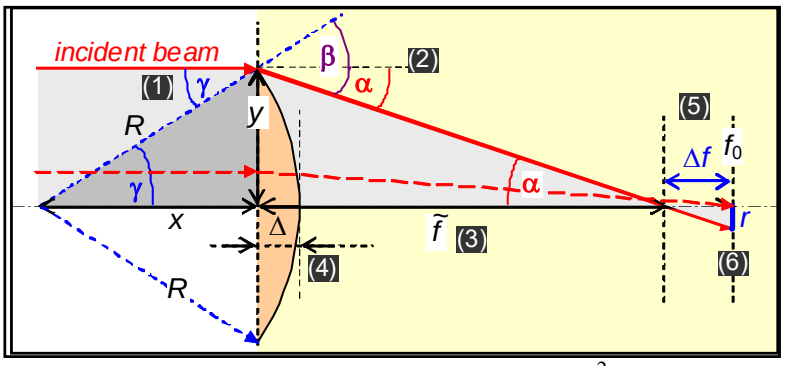
(2) **Refraction** of beam at curved surface (Snell´s law): $\sin\beta = n \cdot \sin\gamma$ » Focusing angle $\alpha = (\beta - \gamma)$

(3) **Focal distance of outer rays**: $\tilde{f}(y) = y/\tan\alpha \approx f_0\left[1 - k(y/R)^2\right]$ for $\alpha < 10°$, k = *constant*

(4) **Additional shift** Δ of origin: $\Delta(y) = R - x = R(1-\cos\gamma) \approx 1/2R \cdot (y/R)^2$ for $\alpha < 10°$

⇨ **Effective focal distance of outer rays**:

$f(y) = \tilde{f}(y) - \Delta(y)$

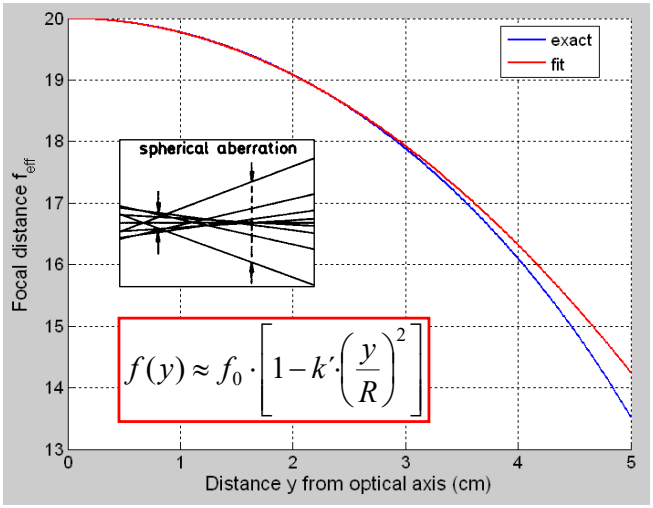


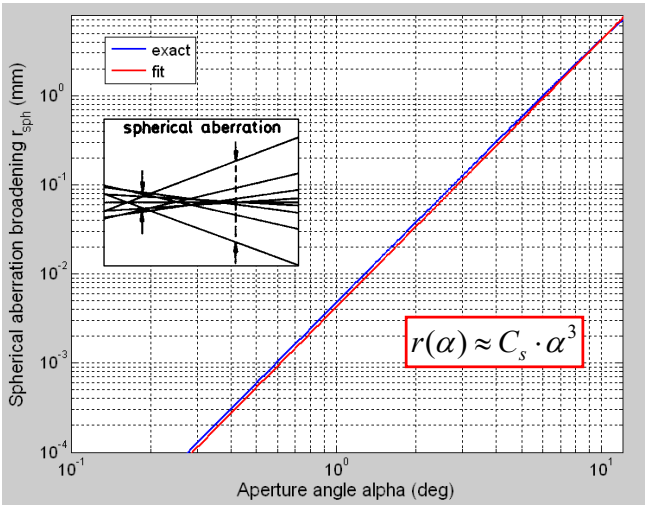
(5) **Change of focal distance**: $\Delta f(y) = f_0 - \tilde{f} - \Delta \approx f_0 \cdot k' \cdot \left(\frac{y}{R}\right)^2 = f_0(n-1)^2 \cdot k' \cdot \left(\frac{y}{f_0}\right)^2 \approx f_0(n-1)^2 \cdot k' \cdot \alpha^2$

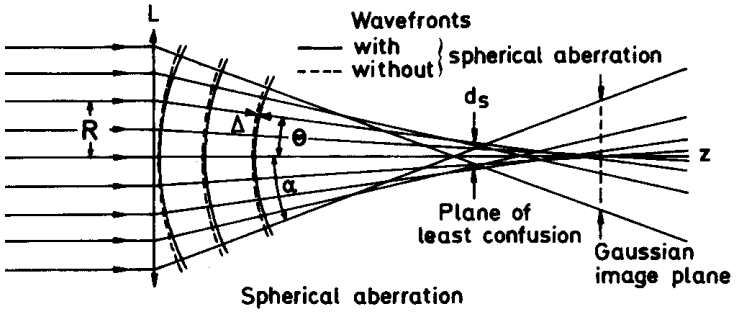
(6) **Broadening**: $r(\alpha) = \Delta f(\alpha) \cdot \tan\alpha \approx f_0(n-1)^2 \cdot k' \cdot \alpha^3$ $\boxed{r(\alpha) \approx C_s \cdot \alpha^3}$

Focal Distance and Broadening as a Function of Distance and Aperture Angle

Example: Curvature *R* = 10 cm, refractive index *n* = 1.5 , corresponding focal length f_0 = 20 cm

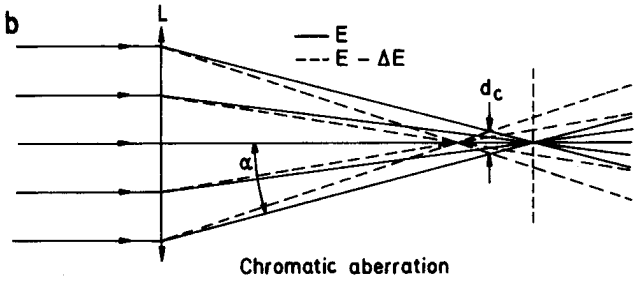






4.10.2 Chromatic Aberration

Light with different wavelengths or energies are refracted differently by the lens due to the **dispersion n(λ) of the refractive index** of the lens material.



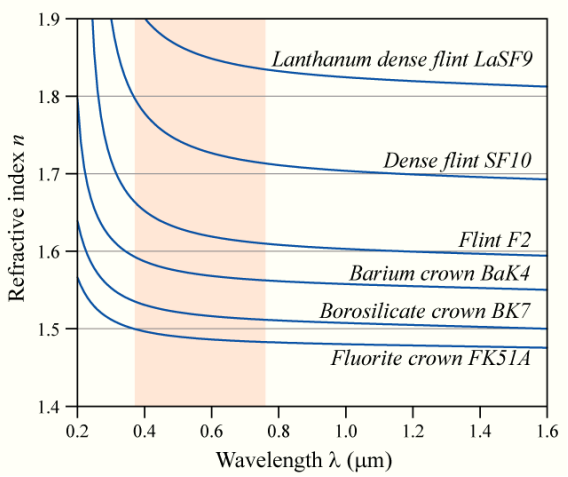
➢ Broadening of the spot size (spot radius r_{chr}) according to

$\Delta f \sim \Delta\lambda$:

$$r_{chr} = C_c \frac{\Delta\lambda}{\lambda}\alpha^m = C_c \frac{\Delta E}{E_0}\alpha^m$$

C_c … chromatic aberration constant: ~ dn /dλ or ~ dE /dλ or ~ df /dλ, $m \sim 1$

Example: Dispersion $n(\lambda)$ of the refractive index for typical lens



Chromatic aberration

⇨ strongly depends on the material properties

⇨ increases with increasing wave-length spread of the used light

4.10.3 Astigmatism

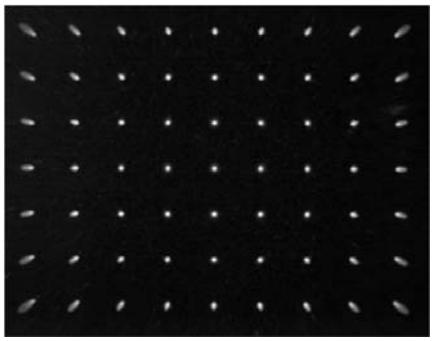
= Non-uniform action of the lens in two orthogonal directions.

Main origin:

Asymmetries of the lenses, misalignments of the optical system, electric stray fields due to charging in electron microscopy.

➢ Broadening of the spot size/radius r_{ast} according to

$$r_{ast} = \alpha\Delta f$$



In all, the

⇨ **total effective resolution** is given by:

$$r_{eff} = (r^2_{dif} + r^2_{sph} + r^2_{chr} + r^2_{ast} + \ldots)^{1/2}$$

i.e. by the sum of all broadeing factors.

4.11 Maximizing the Microscope Resolution

For *any* microscope, the **total effective resolution r_{tot}** is given by the ***sum*** *of diffraction and aberration broadening* according to:

Total effective resolution:

$$r_{tot} = \sqrt{r_{dif}^2 + r_{sph}^2 + r_{chr}^2 + r_{ast}^2} = 0.61 \cdot \lambda / NA_{eff}$$

which is characterized by the *effective* numerical aperture NA_{eff}

⇨ For highest resolution ***all*** **terms must be minimized** !

However: Diffraction broadening and aberration broadening *counteract* each other !

$$r_{tot} = \sqrt{\left(\frac{0.61\lambda}{n \sin\alpha}\right)^2 + \left(C_{sph}\alpha^3\right)^2 + \left(C_{chr}\frac{\Delta\lambda}{\lambda}\alpha\right)^2 + \ldots}$$

- *Diffraction broadening is proportional to* $\sim 1 / \sin\alpha$
- *Aberration broadening is proportional to* $\sim C \cdot \alpha^n$

⇨ With increasing aperture α, diffraction broadening decreases but lens aberration increase.

Consequences:

⇨ The resolution cannot be simply increased by increasing the lens diameter:
Instead a ***compromise*** must be made that minimizes the ***sum*** of all broadening factors.

⇨ This means that there exists an ***optimum aperture angle*** α_{opt} for which the *highest resolution* is practically achieved.

⇨ This optimization is particularly important for electron microscopy, where the lens aberrations are very large and dominate the achievable resolution.

4.11.1 Reducing Diffraction Broadening

Numerical apertures $NA = n \cdot \sin\alpha$: Means that α as well as *n* should be maximized

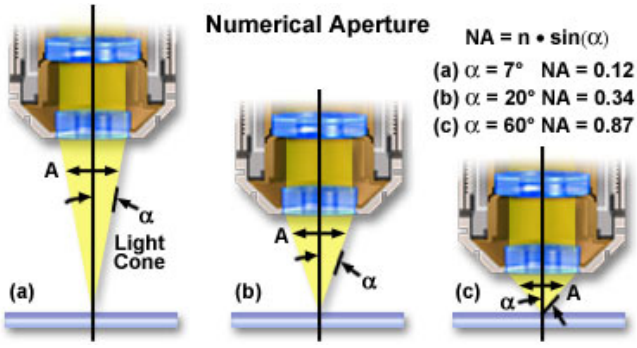
(a) Increasing the aperture angle

$\alpha = \tan(D/2f)$

requires lenses with large diameter *D* and/or smaller focal length *f*.

⇨ For microscopes with fixed tube length and with fixed objective diameter *D*, α can be increased only by **decreasing the focal length *f*** of the objective,

⇨ Reducing *f* means also reducing the working distance *w* ~ *f*



$$r_{diff} = 0.61 \cdot \lambda / n \sin\alpha$$

Consequence: Because at the same time the magnification *M* = *t*/*f*, increasing *f* simultaneously increases the magnification:

⇨ High NA (and high resolution) is obtained only at high magnification !

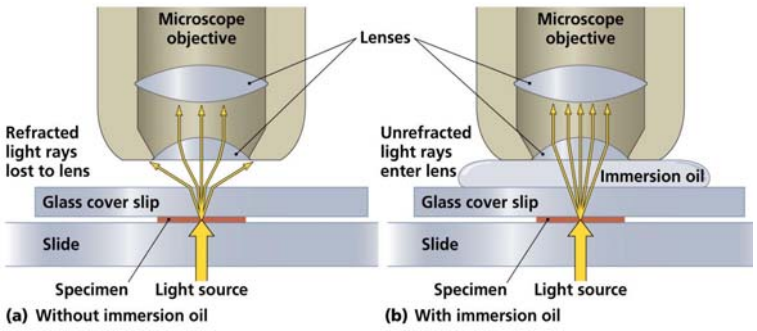
⇨ A high NA also gives a high brightness of the image, which is proportional to √(NA)

Resolution & Numerical Aperture for λ = 550nm (green)

Magnification	N.A.	Resolution (µm)
4x	0.10	1.375
10x	0.25	0.61
20x	0.40	0.37
40x	0.65	0.29
60x	0.75	0.29
N.A. = Numerical Aperture		

(b) Increasing the refractive index n: Immersion Microscopy

Introducing a ***liquid medium*** with refractive index $n > 1$ between the sample and objective reduces the diffraction angles and thus, increases α_{eff}.

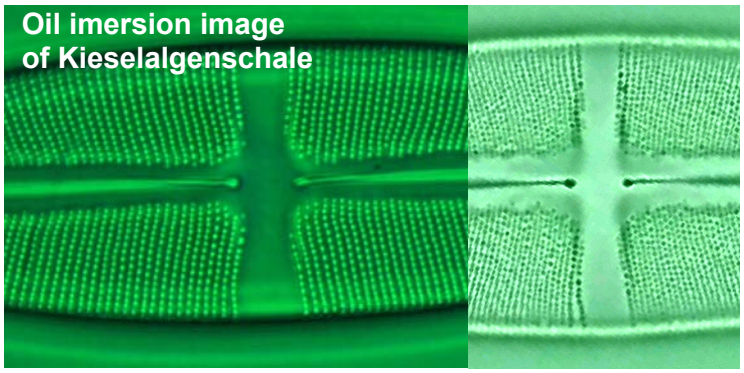


This is called ***immersion microscopy***.

Numerical apertures NA > 1 can be obtained, which reduces diffraction broadening and improves resolution

$$r_{diff} = 0.61 \cdot \lambda / n \sin\alpha$$

air: $n = 1$ **, water:** $n = 1.33$ **, glycerine:** $n = 1.47$
synthetic oils: $n = 1.51$, **Bromonaphthalene:** $n = 1.659$, **Methylene iodide:** $n = 1.74$,



4.11.2 Correction of Lens Aberrations $r_{eff} = \sqrt{r_{diff}^2 + r_{aber}^2} = 0.61 \cdot \lambda / NA_{eff}$

⇨ To achieve a diffraction limited resolution, all lens errors must be corrected !

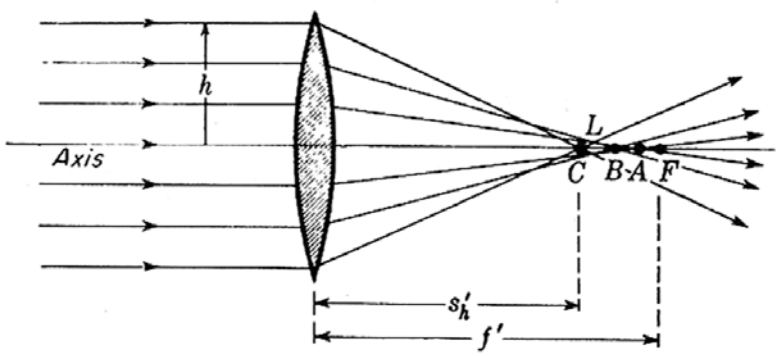
Non-perfect lens system: Performance defined by *effective numerical aperture* $NA_{eff} < n \sin\alpha$.
This value is *smaller* than $n \sin\alpha$ and accounts for additional broadenings caused by lens aberrations.

⇨ Objectives with better corrections yield higher *effective* NA_{eff} values close to $NA_{ideal} = n \sin\alpha$

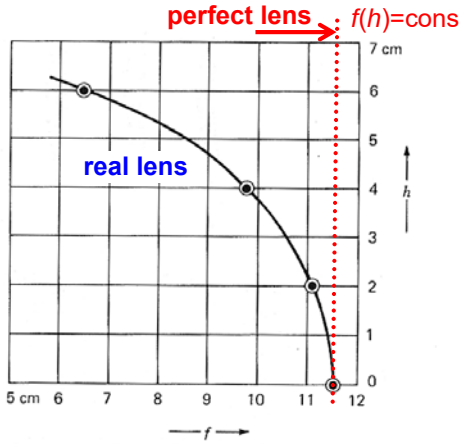
⇨ Combination of several lenses in one objective allows to compensate of most lens aberrations.

(i) Spherical aberration correction: **Cause:** Spherical curvature of lenses results in stronger bending of outer rays towards the optical axis. Strong effect for "thick" lenses

Broadening: $r_{sph} \sim C_s \cdot \alpha^n$



see: Jenkins & White: Fundamental of Optics

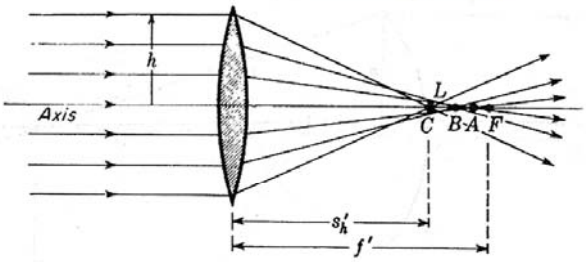


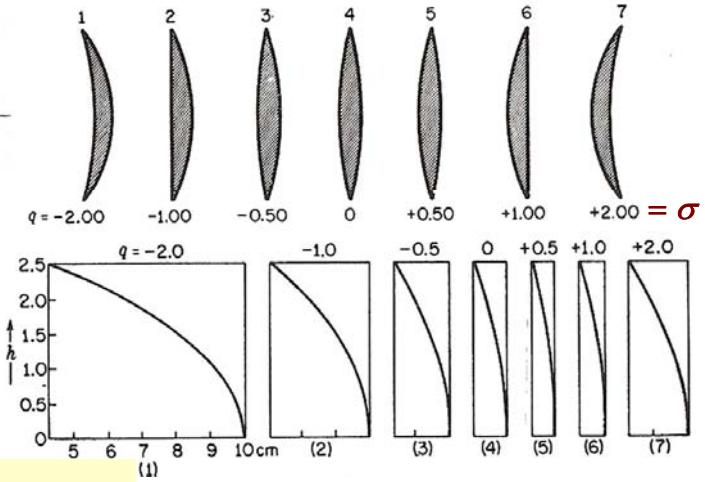
Focal point *f* versus distance *h* of ray from the optical axis

Solution #1: Use of several lenses with *smaller curvature* and smaller aberration

Solution #2: Use of *asymmetric lenses* with two different curvatures on each side to minimize spherical aberration

Lens shape factor: $\sigma = \dfrac{R_2 + R_1}{R_2 - R_1}$ determines the change of $f(h)$ ~ spherical aberration





Spherical aberration is *minimized* when:

$$\sigma = -\frac{2(n_{lens}^2 - 1)}{n_{lens} + 2} \cdot \left(\frac{s_{obj} - s_{im}}{s_{obj} + s_{im}} \right) = -\frac{2(n_{lens}^2 - 1)}{n_{lens} + 2} \cdot p$$

where s_{obj} and s'_{im} are the object and image distances, respectively.

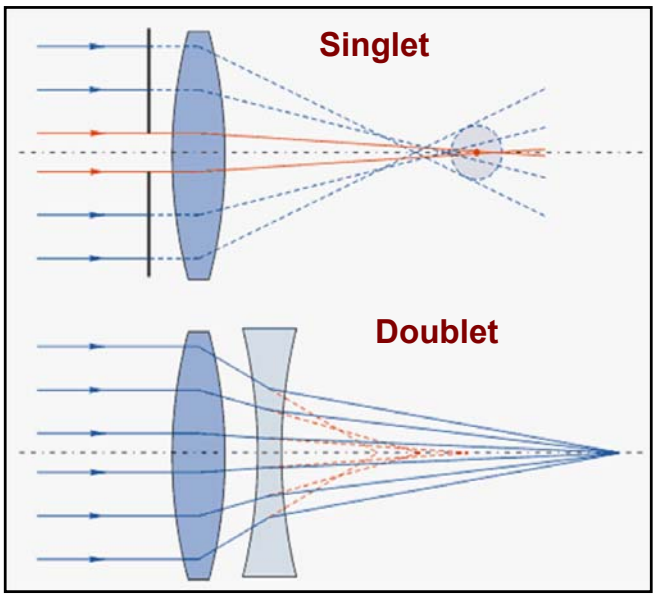
$$p = \left(\frac{s_{object} - s_{image}}{s_{object} + s_{image}} \right)$$

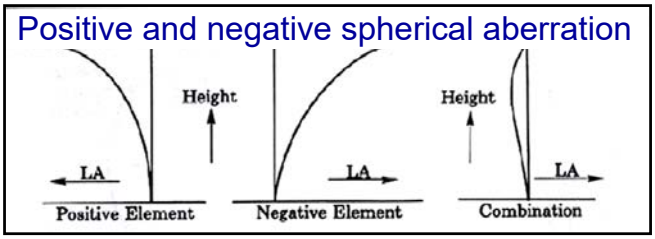
For an object at infinite ($s_{obj} = \infty$, $p = -1$, $n_L = 1.50$),

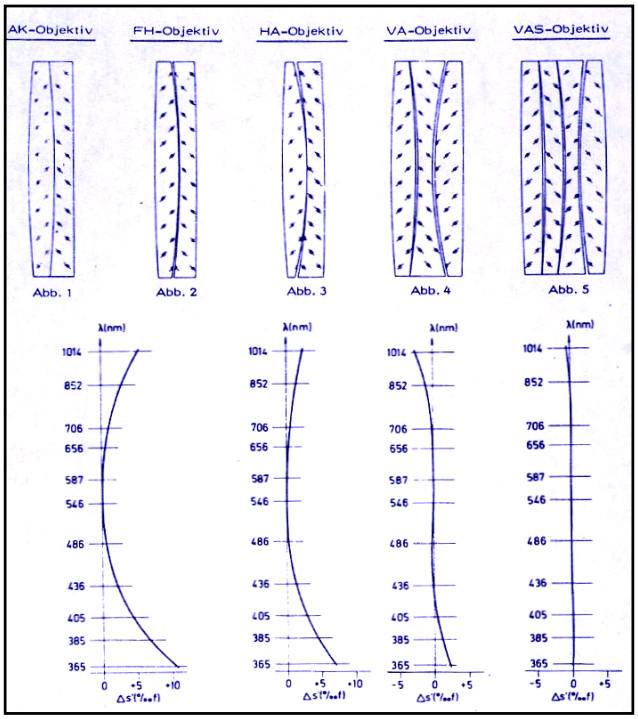
Optimum lens shape: yields $\sigma_{opt} \sim 0.7$

Solution #3: Combination of positive (convex) and negative (concave) lenses

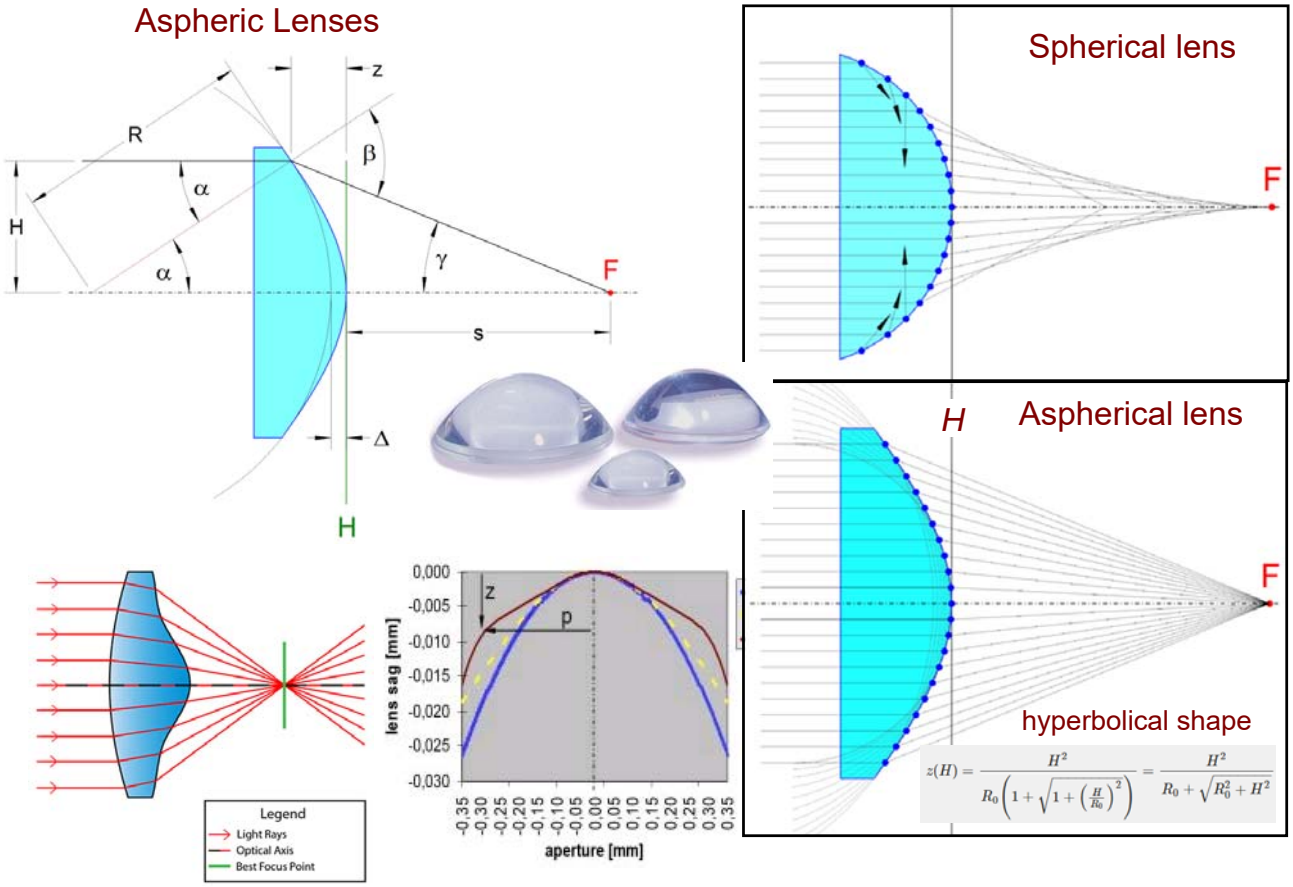
with ***positive*** and **negative** spherical aberration constants (**dublets** and triplets)







Solution #4: Use of more expensive *nonspherical lenses*

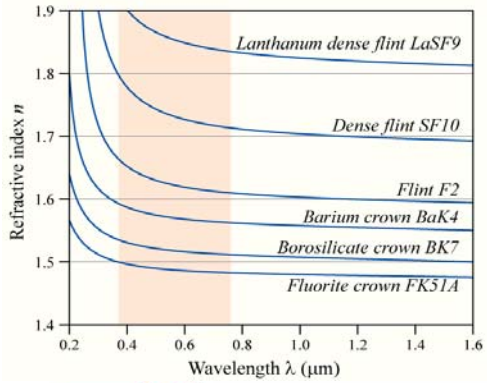


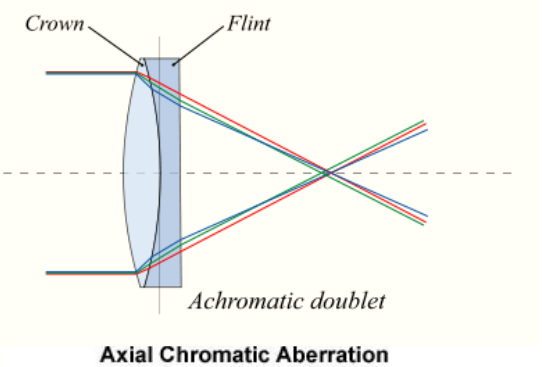
(ii) Correction of chromatic aberration

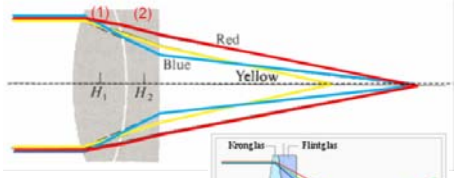
Cause:
Wavelength dependence, i.e., dispersion of the refractive index $n(\lambda)$

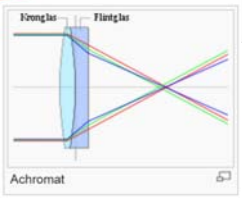
Solution:

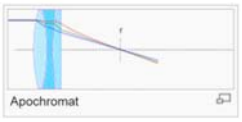
Lens ***doublets*** or ***triplets*** of convex and concave lenses consisting of glass materials with different dispersion (crown, flint glasses and fluorspar).

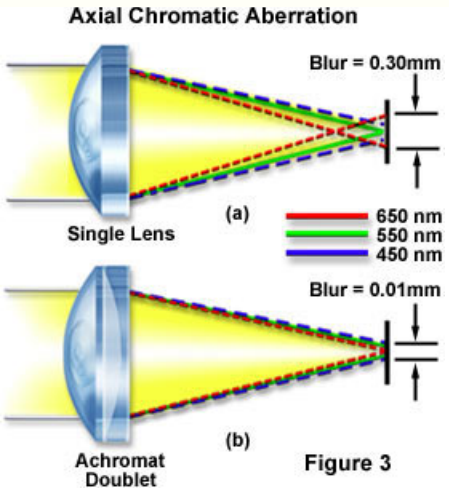












(iii) Correction of astigmatism

Cause: Asymmetric lenses or misaligned/unparallel optical elements

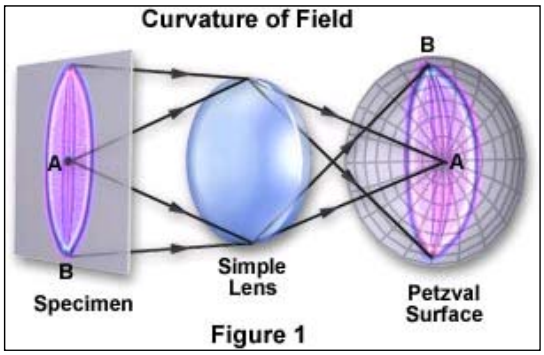
Solution: Improved lens symmetry, precise alignment.

(iv) Curvature of field-of-view

For spherical lenses, the image is not focused on a plane but rather on a curved surface.

Cause: Curvature of lens surfaces.

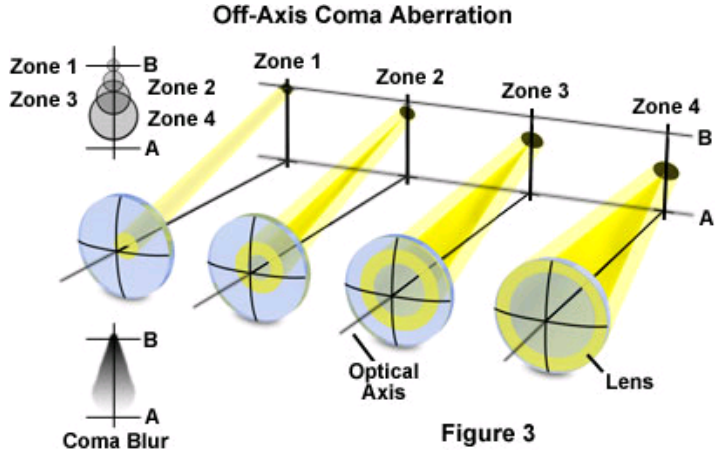
Solution: Correction by additional lens elements.

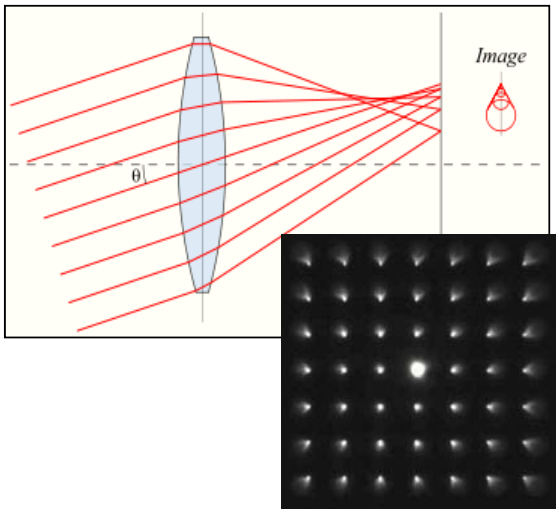


(v) Coma = smearing of outer part of image

Cause: Inclined beams have different focal position for outer beams through the lens.

Solution: Correction by additional lens elements.

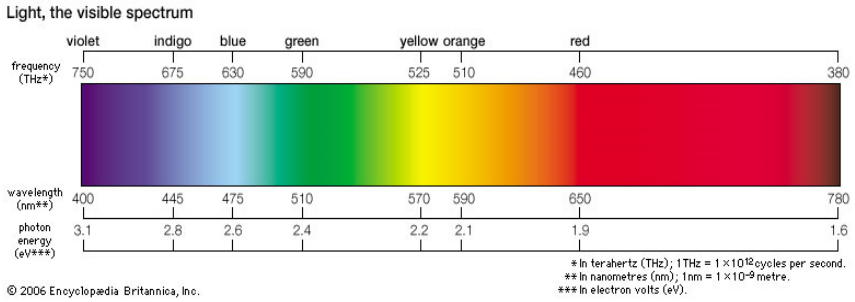




4.12 Application to Light Microscopy

In optical microscopy, visible (or near-infrared) light is used for imaging of the sample.

Wavelength: $\lambda[nm] = c/\upsilon = 1240/E[eV]$ **Energy:** $E[eV] = h \cdot \upsilon = 1240/\lambda[nm]$



Color	Energy	Wavelength
violet	3.2 – 2.8	380–450 nm
blue	2.8 – 2.5	450–495 nm
green	2.5 – 2.2	495–570 nm
yellow	2.2 – 2.1	570–590 nm
orange	2.1 – 1.9	590–620 nm
red	1.9 – 1.6	620–750 nm

Resolution limits: Since for VIS & NIR optical systems, effective aberration corrections exist that nearly fully eliminate aberration broadening, ***diffraction*** is the main remaining ***limitation*** for the spatial ***resolution*** of VIS & NIR microscopy.

⇨ The best resolution is about ½ λ of the illumination wavelength λ using aberration corrected objectives with high NA ~ 1

$r_{res} = r_{diff} = 0.61\ \lambda / NA$

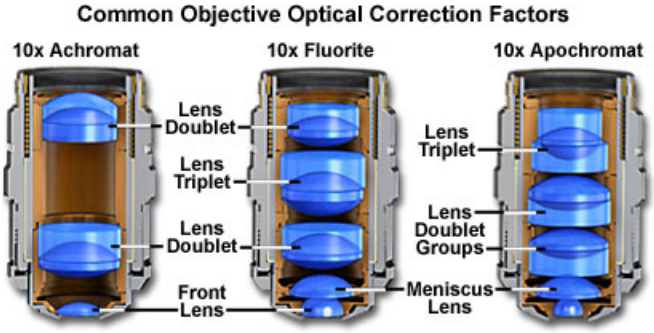
⇨ Note: Recently, sub λ/2 resolution scanning fluorescence microscopy has been developed.

Wavelength (nm)	Resolution (nm)
360 – 450 nm (violet)	190 – 250 nm
450 – 500 nm (blue)	250 – 300 nm
500 – 570 nm (green)	300 – 350 nm
620 – 750 nm (red)	380 – 460 nm

Types and Specifications of Objective Lenses

Achromats: Objectives corrected for axial chromatic aberration at two wavelengths (486 nm blue and 656 nm red) and corrected for spherical aberration in the color green (546 nm; see table).

Fluorites or semic-apochromats: Additional lenses for spherical corrections at two wavelengths.

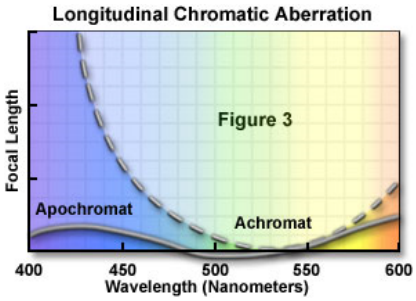


Apochromats: Best corrected objectives, correction of spherical aberrations for wide a wavelength region.

Objective Correction for Optical Aberration

Objective Type	Spherical Aberration	Chromatic Aberration	Field Curvature
Achromat	1 Color	2 Colors	No
Plan Achromat	1 Color	2 Colors	Yes
Fluorite	2-3 Colors	2-3 Colors	No
Plan Fluorite	3-4 Colors	2-4 Colors	Yes
Plan Apochromat	3-4 Colors	4-5 Colors	Yes

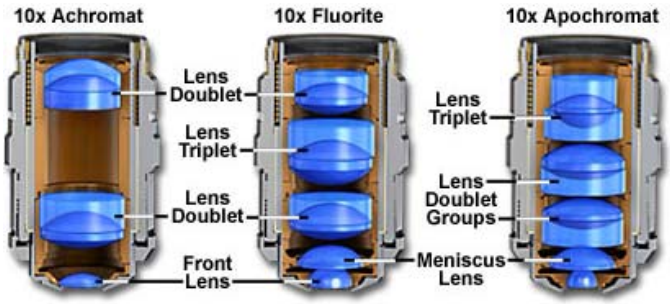
see: www.micro.magnet.fsu.edu



Plan: Lenses with field of view curvature correction are termed with additional "plan" prefix.

Specifications of Objective Lenses

- Type of Objective (Degree of correction)
- Magnification (number)
- Numerical Aperture
- Tube length (mm)
- Cover Glass correction (thickness)



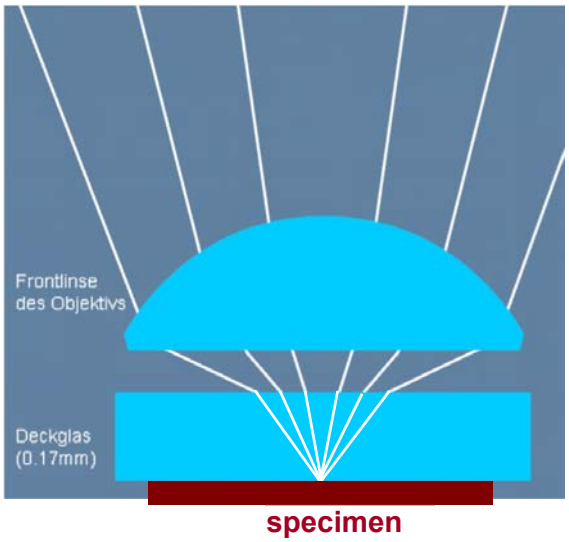
Resolution and Numerical Aperture for λ = 550nm (green)

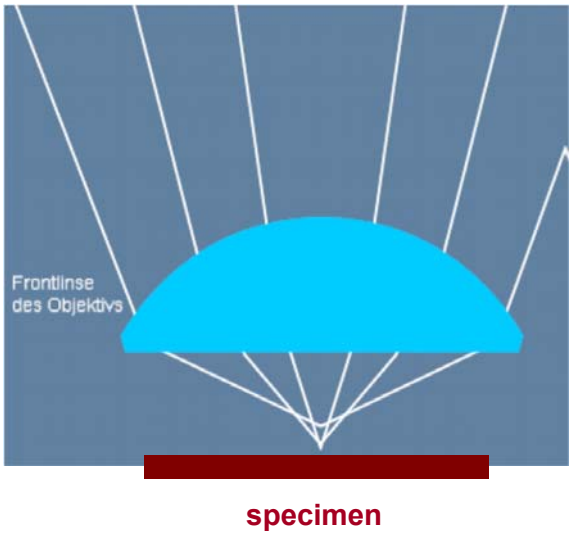
	Objective Type					
	Plan Achromat		Plan Fluorite		Plan Apochromat	
Magnification	N.A.	Resolution (µm)	N.A.	Resolution (µm)	N.A.	Resolution (µm)
4x	0.10	2.75	0.13	2.12	0.20	1.375
10x	0.25	1.10	0.30	0.92	0.45	0.61
20x	0.40	0.69	0.50	0.55	0.75	0.37
40x	0.65	0.42	0.75	0.37	0.95	0.29
60x	0.75	0.37	0.85	0.32	0.95	0.29
N.A. = Numerical Aperture						

Violet: λ = 420nm (25% improvement) www.micro.magnet.fsu.edu

» Objectives with better corrections yield significantly higher resolution !

Cover Glass Correction (common for biological specimen)





Comparison of Image Quality: Achrochromat

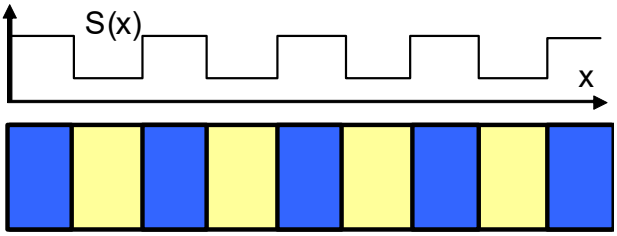
Planapochromat

4.13 Contrast in Microscopy

Contrast is produced by local variations in sample properties, which gives rise to changes in the response signal S(x,y) that is detected and forms the microscopy image.

Contrast is defined as relative difference in signal ΔS from adjacent spots on the sample normalized to the average value:

$$C_{im}(x) = \frac{S(x_1) \quad S(x_2)}{\overline{S}}$$



Contrast is crucial for high resolution microscopy, because if the contrast is smaller than the noise level, the actual resolution is zero ! Thus, contrast needs to be maximized for best imaging conditions and highest resolution.

Contrast depends on many factors:

- How strongly the local sample properties and vary across the sample, i.e., what kind of micro- or nanostructure is present in the sample,
- The type and strength of the probe-sample interaction,
- The probe energy or wavelength,
- The type of detected response signal and what detection scheme is used (e.g., spectroscopic detection, angle resolved detection, energy or wavelength filtering, etc.)

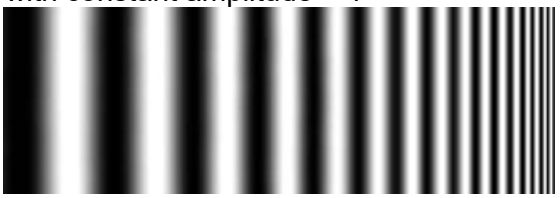
⇨ Contrast depends not only on the sample properties but also on the ***imaging conditions*** !!

4.13.1 Contrast and Resolution

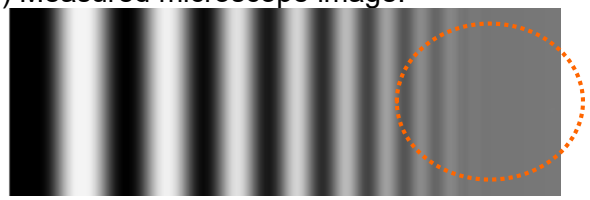
Contrast has a significant effect on resolution. This is because small sample features can be ***resolved only*** when the signal difference ΔS in the image is larger than a certain threshold value (=visibility criterion). Due to the **broadening of the image** by diffraction or lens aberrations, when the feature size decreases the signal difference ΔS or signal amplitude ***continuously decreases*** and drops to zero when the feature size becomes smaller than the resolution limit.

Example: Stripes with decreasing line spacing

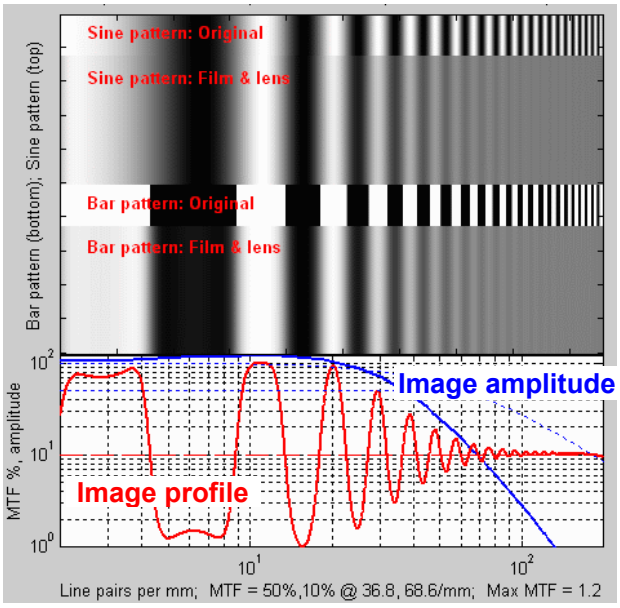
(a) Object of sinusoidal contrast line grating with constant amplitude = 1



(b) Measured microscope image:



⇨ The image contrast decreases when the line spacing decreases !



4.13.2 Visibility and Noise in Practical Microscopy Resolution

Visibility defines how well different features in an image can be distinguished.

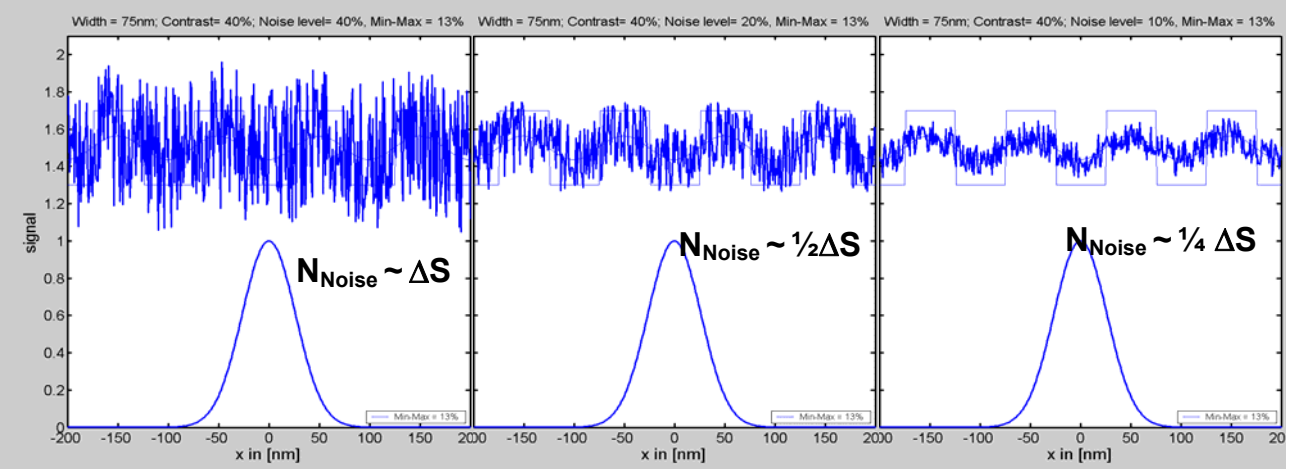
Visibility criterion: Adjacent features in an image can only be discerned if the **signal contrast ΔS** is ~ five times larger than the **noise level N.**

$$\Delta S > 5\sqrt{n}$$

For random noise, the noise level is typically assumed to be proportional to the root mean square deviation of the signal from its average value, i.e., $N = \sigma_{RMS} = \sqrt{n}$

As a result, the **signal-to-noise ratio S/N** strongly influences the **practically achievable microscope resolution**. This means that the theoretically resolution limit can be reached only for (i) ***high contrast specimen*** and/or (ii) ***sufficiently low noise levels*** in the images.

Example: Image signal S(x) for a periodic sample C(x) images with a Gaussian response function for different superimposed *uncorrelated random noise* levels *N*.



Sources of noise: Fluctuations of illumination intensity, of sample emission, detector noise, amplifier noise, mechanical and acoustic noise, etc. ….

Noise is particular a ***problem for scanning microscopy*** because the measurement time per pixel or image point is very small (= total recording time / number of pixels (10^6)).

$$S(t) = \bar{S} + N(t)$$

Noise reduction can be achieved by:

» *improved hardware* (low noise detectors, amplifiers, vibration damping, electromagnetic shielding)
» *increasing the illumination intensity* when the absolute noise level is constant.
» *time averaging* or signal integration.

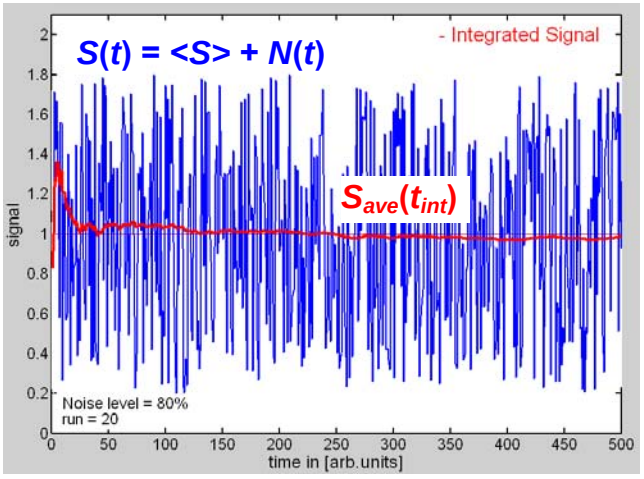
Time averaging and integration:

For ***uncorrelated statistical noise***, the *root mean square (RMS) noise* ***level σ decreases*** with increasing averaging or integration time.

This is because uncorrelated random noise averages out, leaving a constant signal value after a sufficiently long integration time.

$$S_{ave}(\tau) = t^{-1}\int_t \left(\bar{S} + N(t)\right)dt$$

$$= \bar{S} + t^{-1}\int_\tau N(t)dt \rightarrow \bar{S}$$

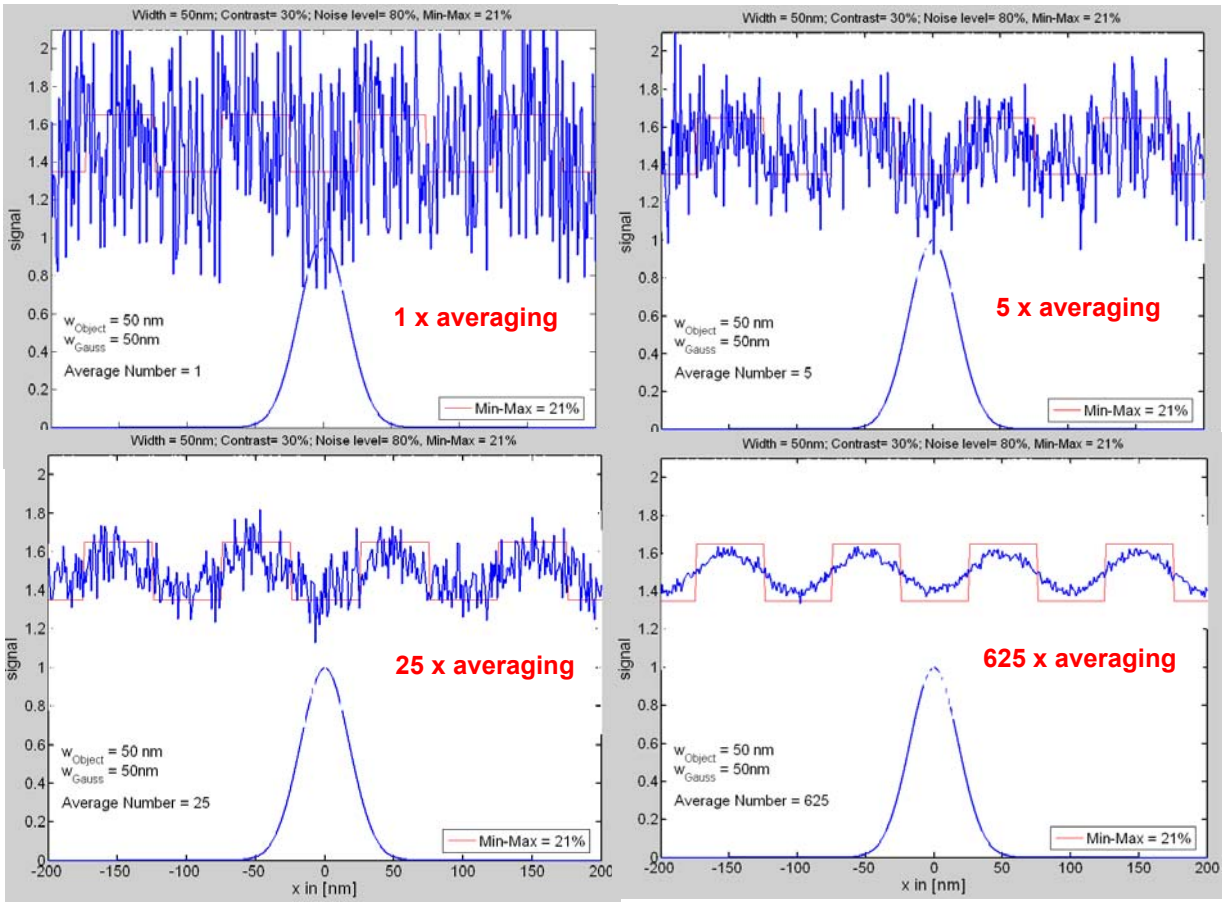


Resulting noise level σ_0^{RMS} of average signal:

When the signal is measured *n*-times giving a set of values S_i the RMS noise of the average signal is given by the an the standard deviation of the S_i values:

$$\sigma_0^{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(S_i - \bar{S}\right)^2}$$

Example how the signal-to-noise ratio improves by time averaging:



Example: Noise Reduction in Scanning Electron Microscopy

According to statistics, the **RMS noise** $\sigma_{n,ave}$ of an *n*-times averaged signal decreases according to:

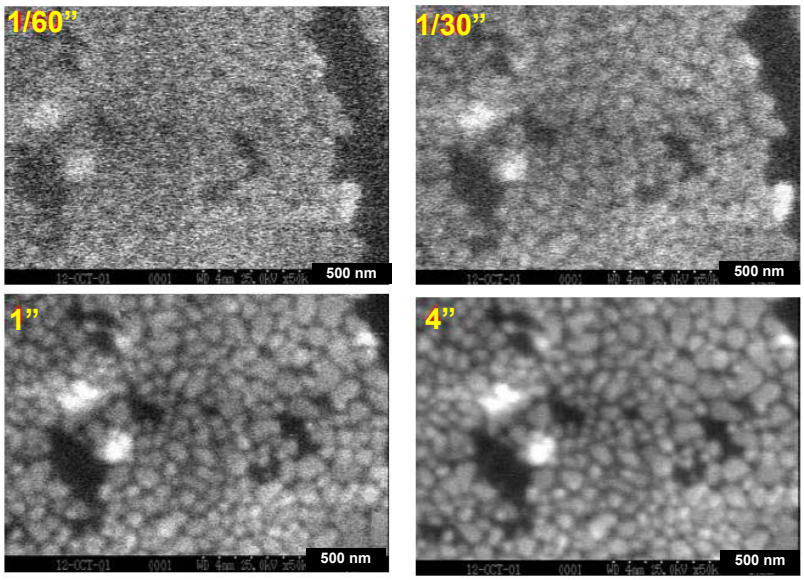
$$\sigma_{n,ave} = \sigma / \sqrt{n}$$

This means that the noise $\sigma_{n,ave}$ of the average signal decreases by the factor $1/\sqrt{n}$ with increasing number *n* of averaging steps, which is equivalent to increasing the integration time τ.

Thus, the **signal-to-noise ratio** $S/N = S/\sigma$ improves as

$$S / N = S / \sigma_n = (S / \sigma_0^{RMS}) \sqrt{n}$$

Example: SEM images obtained using different integration times



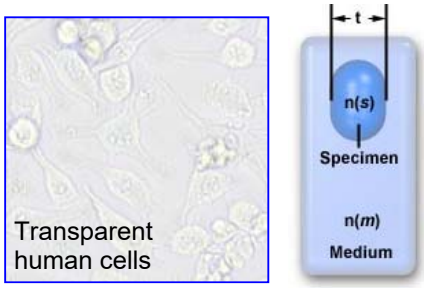
SEM images of 50-200 nm Au crystallites on a carbon template recorded with different integration times (in seconds).

$$\frac{Signal}{Noise} = \frac{S}{N} = \frac{\bar{n}}{\sqrt{\bar{n}}} = \sqrt{\bar{n}}$$

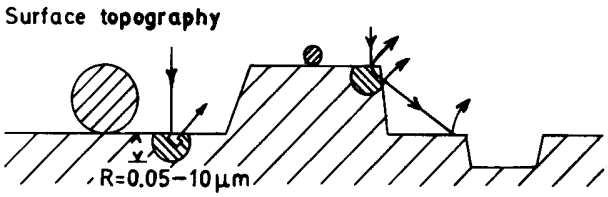
4.14 Contrast Mechanisms

Generally, **many different mechanisms** contribute to the contrast in the response signal:

- **Absorption and thickness contrast** due to variation of the transmission according to $T(x) \sim e^{-\alpha(x)d(x)}$
- **Reflection/backscattering contrast** of the probe due to variation of the reflectivity and backscattering yield
- **Phase/interference contrast** due to constructive or destructive interference of beams experiencing different phase shifts when transmitted or reflected from the sample,



- **Topographic contrast** due to local tilts, edges, shadowing, multiple scattering, etc.,
- **Diffraction & angular scattering contrast** due to sideway deflection by diffraction or by diffuse scattering at surface roughness, topographic features or local microstructure,



- **Spectroscopic / color contrast** due to energy/wavelength differences in the transmitted beam or in the excited response signal such as, e.g., fluorescence, luminescence, cathodoluminescence, characteristic x-ray emission, fluorescence markers, etc. (see Chap. 2).

Contrast formation is often ***very complex***. This applies in particular for electron microscopy where many different types of interaction processs occur simultaneously. In addition, the contrast mechanisms can ***vary across the sample surface*** when the topography is complex.

Contrast for the Different Microscopy Methods

Optical microscopy:

- Contrast due to local difference in optical constants $n(x)$ and $\alpha(x)$, color contrast,
- Contrast in fluorescence yield, topographic contrast (reflection mode), ..

X-rays:

- Z-contrast due to different x-ray absorption.
- Strong diffraction contrast due to local variation in lattice structure and lattice constants, strain contrast,

Electrons:

- Z-contrast due to dependence of electron scattering on the atomic Z-number.
- Absorption contrast and contrast due to variation of secondary electron generation yield and diffuse scattering
- Topographic contrast due to shadowing and edge effects (SEM).
- Diffraction and phase contrast (TEM) due to differences in the lattice structure and lattice constants.

Example: Scanning Electron Microscopy

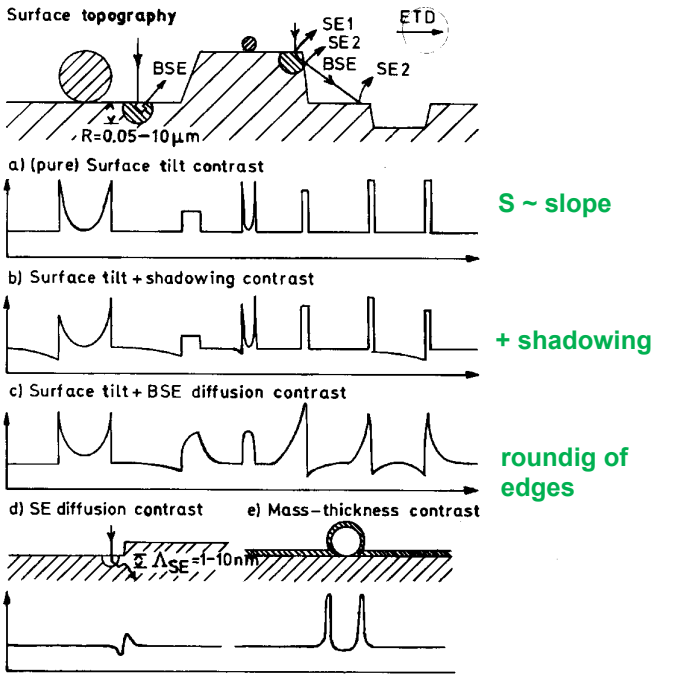


Fig. 6.1. Contributions to topographic contrast demonstrated schematically by surface contours (top) and linescans of SE signals; (**a**) surface tilt contrast, (**b**) shadowing contrast, (**c**) BSE diffusion contrast, (**d**) SE diffusion contrast and (**e**) mass-thickness contrast

4.15 Contrast Enhancement Techniques

For many sample specimen, the relative differences in the transmitted light or for electron microscopy is often rather small and thus, the image ***contrast weak***.

In order to fulfill the visibility criterion and actually discern different features, therefore ***contrast enhancement techniques*** are employed, by which the magnitude of the contrast in the microscopy images is increased.

Contrast enhancement techniques for imaging (and scanning) microscopy

1. **Sample *modification*:** Staining (dyes), selective etching, fluorescence markers.
2. ***Grazing incidence* illumination** and/or **detection** (=angular filtering) to enhance topography contrast due to increased shadowing and scattering at step edges.
3. ***Spectral filtering*** of detected signal to take advantage of differences in $n(E,\lambda)$, $\alpha(E,\lambda)$, $I(E,\lambda)$ of different parts of the sample and to obtain enhanced material selectivity.
4. **Modification of the optical path** in the microscope using additional beam blocking or phase shifting elements. This results in **dark field microscopy, phase contrast microscopy, differential interference contrast microscopy,** etc.

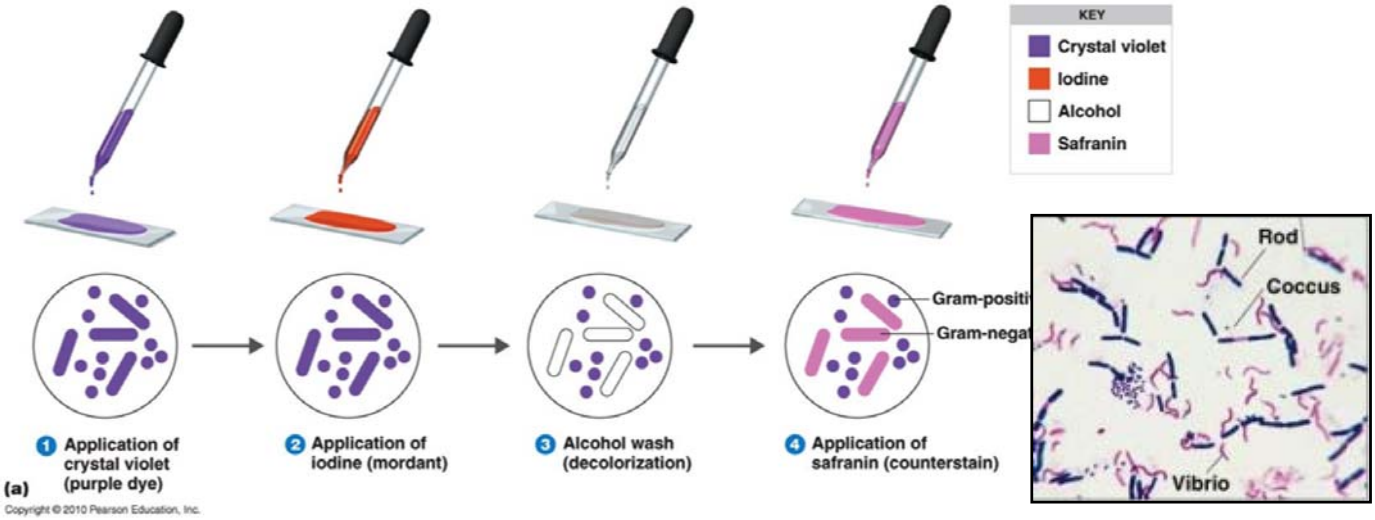
Note: In general, the image contrast is not solely determined by the sample properties, but also depends strongly on the light path and optical configuration of the imaging optics which can be modified and optimized for a given application.

4.15.1 Staining and Sample Modification

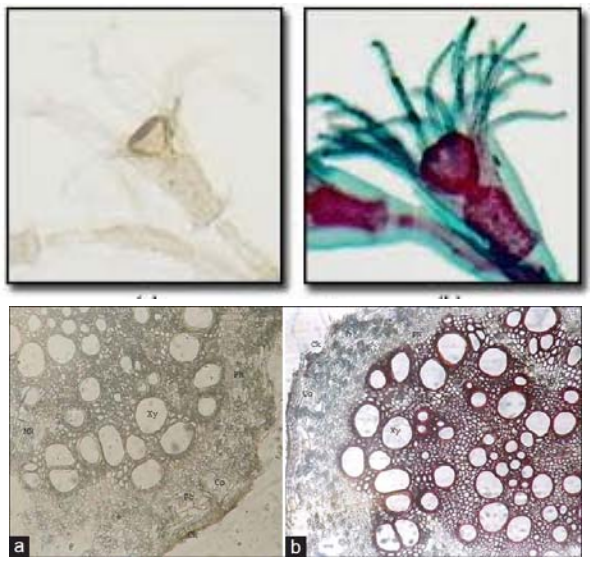
⇨ Local modification of the absorption or emission of biological or organic specimen by selective absorption of color dye molecules that attach to specific sites of the sample structure.

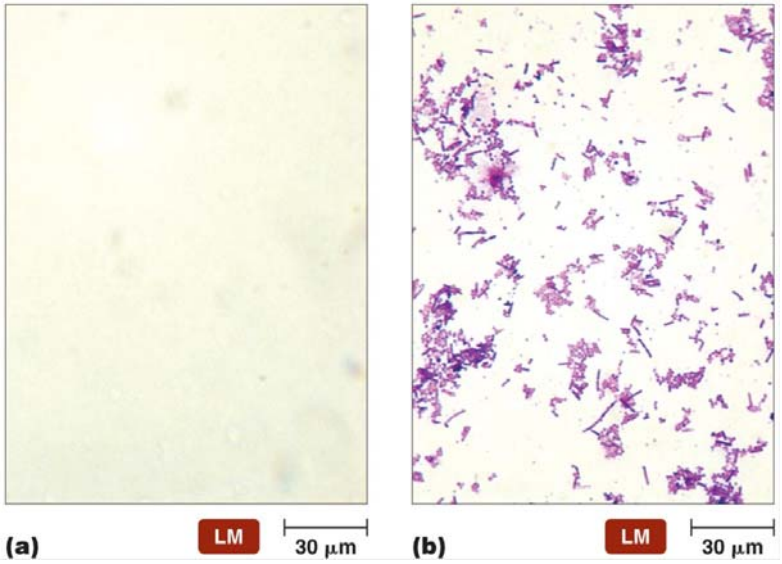
⇨ For metallurgical specimen often etching or staining is applied that selectively removes material according to its composition or local orientation, ...

Staining:



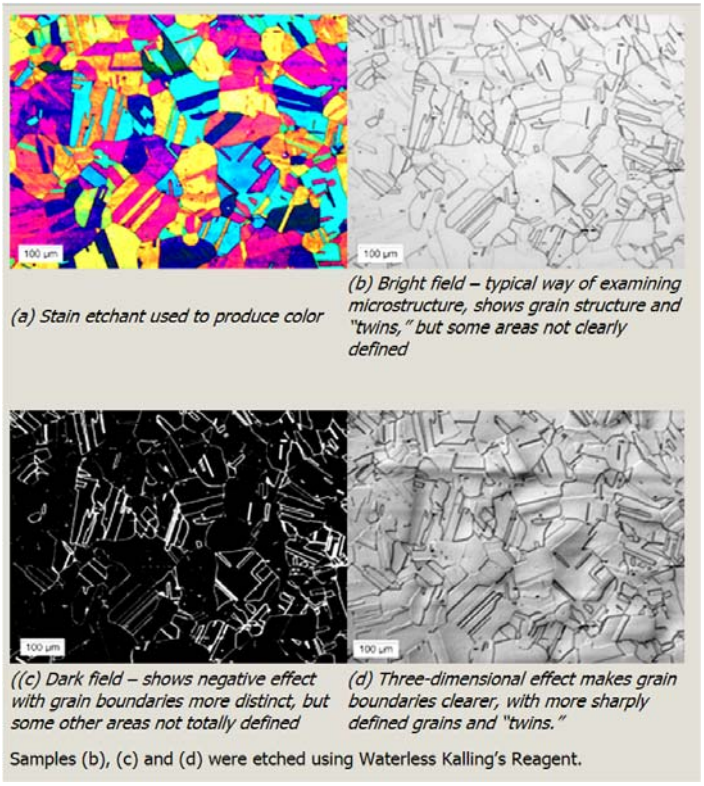
Stained versus unstained organic specimen imaged by optical microscopy





Example: Etched Metallurgical Specimen

The etchants used in metallographic examination are solutions of acid and chemicals and are applied to attack selectively a polished surface, thus permitting microstructural examination



(a) Stain etchant used to produce color

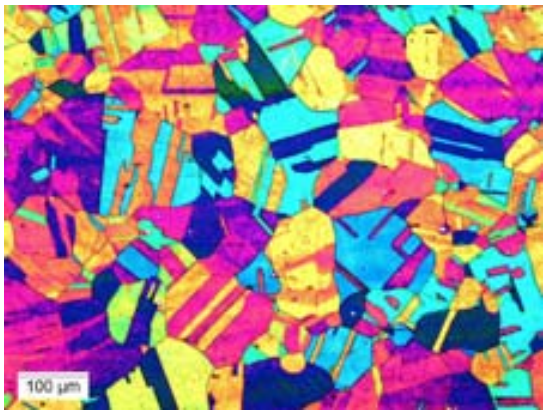
(b) Bright field – typical way of examining microstructure, shows grain structure and "twins," but some areas not clearly defined

((c) Dark field – shows negative effect with grain boundaries more distinct, but some other areas not totally defined

(d) Three-dimensional effect makes grain boundaries clearer, with more sharply defined grains and "twins."

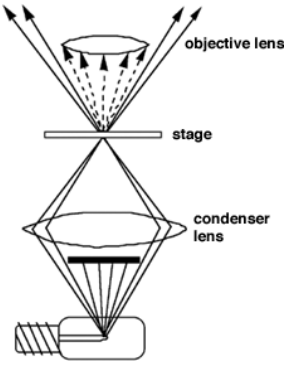
Samples (b), (c) and (d) were etched using Waterless Kalling's Reagent.

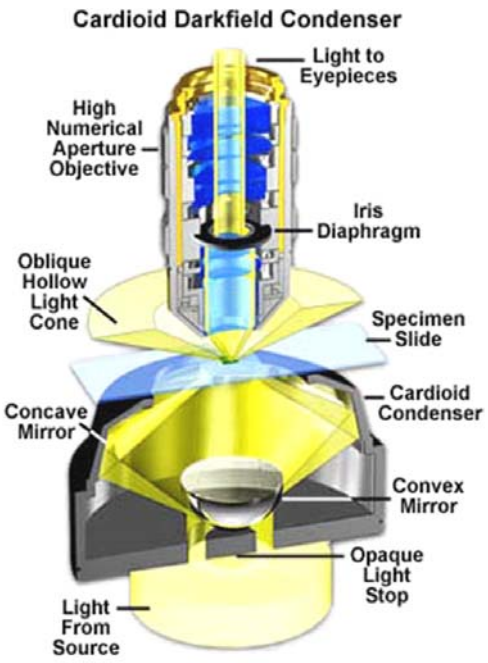
Microstructure of a stainless Type 330 sample @ 100 x magnification, prepared using a tint etch consisting of a solution of 40 ml hydrochloric acid (HCL) + distilled water (H_2O) + one gram potassium meta bisulfite ($K_2S_2O_5$) + 4 grams ammonium biflouride (NH_4F–HF) at room temperature.



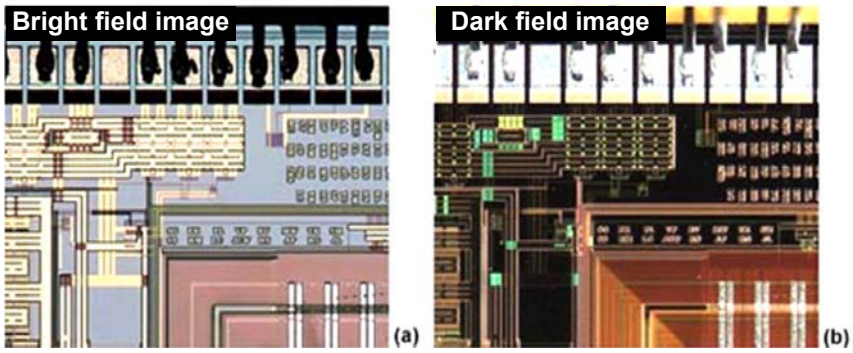
4.15.2 Dark-Field Microscopy (=enhances edge contrast)

Working principle: ***Off-axis illumination*** of the sample and ***blocking*** the central light passing through the specimen without angular scattering or deflection. The image is then formed *exclusively* by ***scattered light***. Without scattered light, the image appears dark (=*dark-field image*).



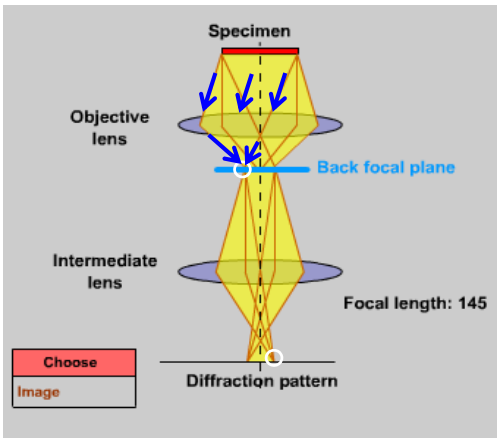


In **dark-field imaging**, the specimen is illuminated with a cone of oblique rays. Without angular scattering in the sample these rays pass outside of the objective lens and not contribute to the image. The specimen therefore appears dark (= "dark-field microscopy"). Only at features such as steps or interfaces where light is scattered the sample appears bright, i.e., the contrast is drastically enhanced. The same principle is used in dark-field transmission electron microscopy.

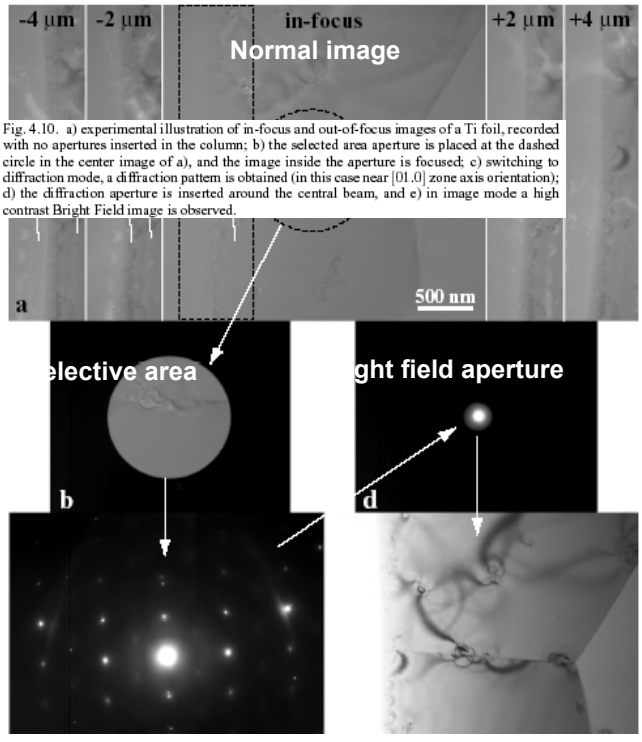


Bright-field / Dark-field Imaging in TEM

In TEM, bright-field or dark-field contrast is obtained by blocking part of the beams using apertures inserted in the back focal plane where the diffraction pattern of the sample is formed.



Example: TEM imaging of a thin Ti foil



Dark field: Central intensity maximum is blocked and only angularly deflected electrons are selected for image formation.

Thus, the sample appears dark, except for parts with enhanced angular deflection.

Bright field: Only central intensity maximum is used for image formation.

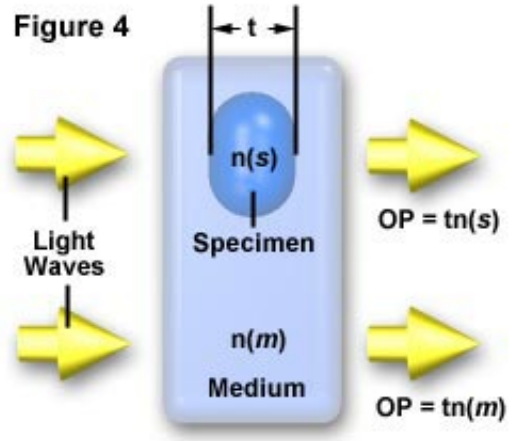
At imperfections, electrons are sideways deflected and thus, they appear dark.

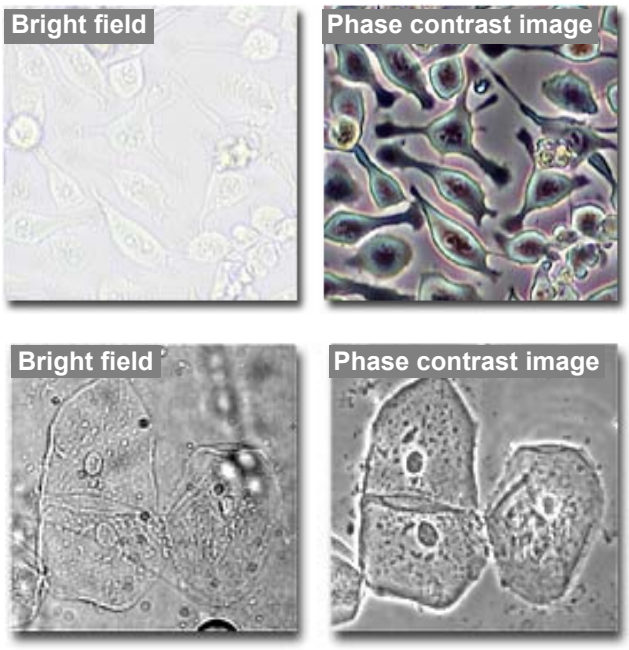
4.15.3 Phase Contrast Microscopy (=enhances refractive index contrast)

For transparent (e.g., organic) samples, absorption differences are usually very small. Although, the absorption contrast is weak the refractive index can still vary significantly. This creates phase changes in the transmitted optical beams due to the ***optical path differences***.

In **phase contrast microscopy**, these optical path or phase differences in the transmitted beams are converted into amplitude, *i.e.*, intensity differences. Thus, the contrast in transparent samples is strongly enhanced.

Optical Path Difference in Phase Objects





Example: Transparent, colorless human cells imaged with optical microscopy under differing contrast modes. Left: Bright field, Right: Phase contrast.

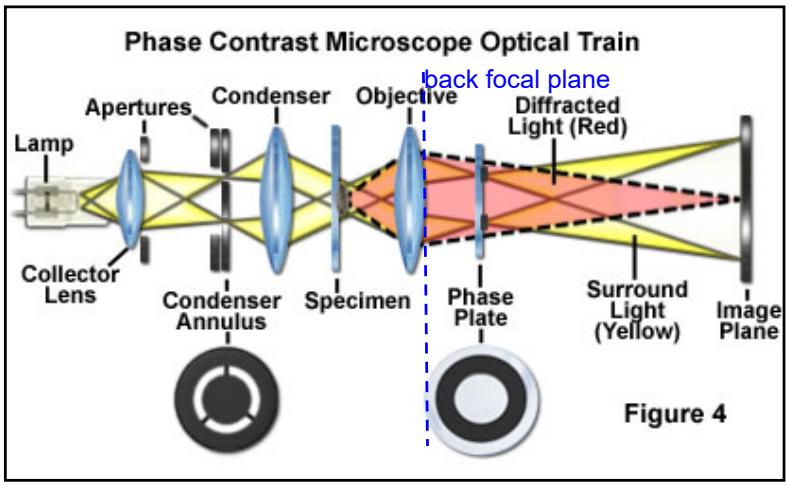
Working principle (for transmission mode)

1. The conversion of phase differences into intensity differences is achieved by separating the directly transmitted unscattered light (yellow) from the diffracted light (red in the figure) from the sample. These two components are projected onto different locations in the objective back focal plane due to parallel beam off-axis illumination coming from a condenser annulus aperture.

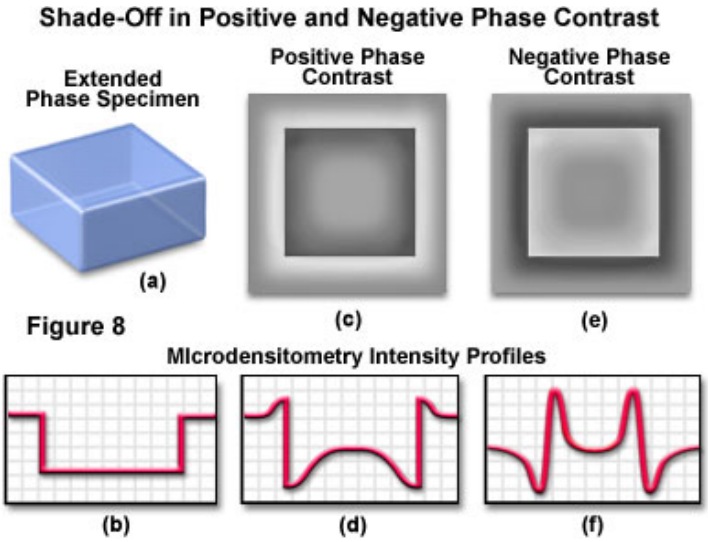
2. The **directly transmitted light** appears as a ring in the back focal plane of the objective. There a ring-shaped ***phase plate*** of etched or semi-transparent glass is inserted, inducing a phase shift of the directly transmitted light.

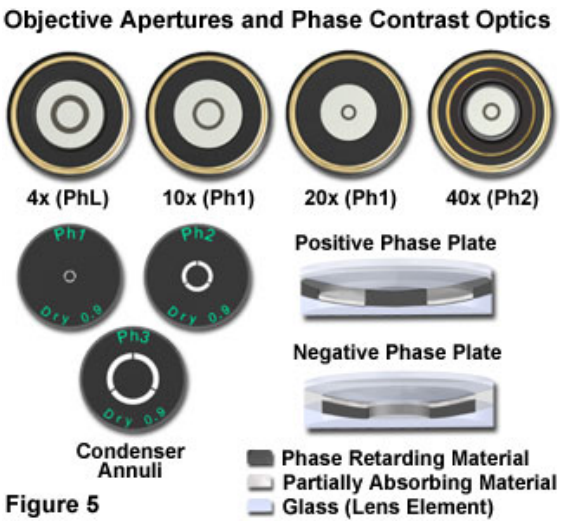
   The **diffracted light** is spread all over the back focal plane and is thus not much phase shifted.

   Thus, the phase plate induces a **phase shift** between the direct and the diffracted light rays.



3. When the different light rays are combined in the image plane, the **phase difference** between these rays leads to strong **interference effects** and thus, **large intensity modulations**, which enhances the contrast from the sample. A strong contrast appears at the points in the sample where the refractive index changes.





4.15.4 Differential Interference Contrast Microscopy (DIC/Nomarski Microscopy)

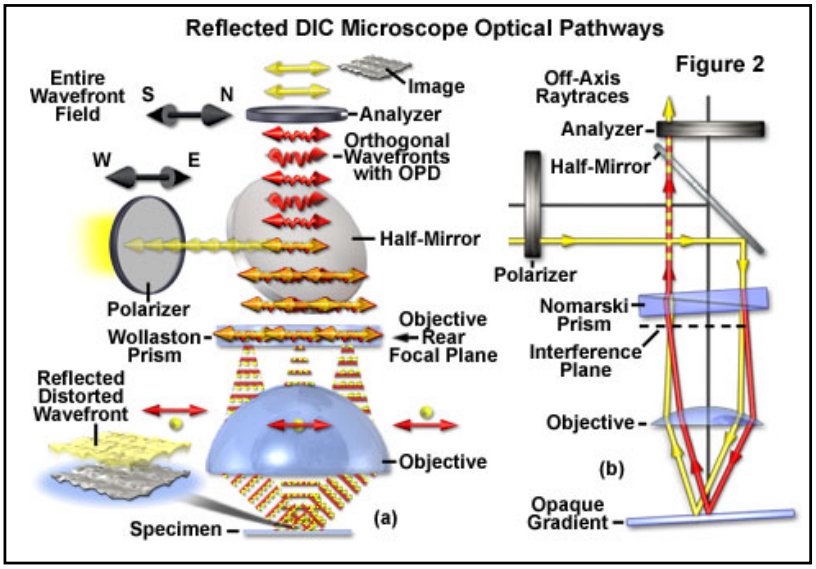
Differential interference contrast microscopy **enhances** the **sensitivity to topography** (reflection mode) or to **thickness** / **refractive index variations** (transmission mode) in the sample.

Working principle:

(1)The illumination light passes through a linear polarizer and projected through a half-mirror on the sample.

(2) A bi-refringent Nomarski prism separates the light into two orthogonally polarized components, which are slightly *sheared* with respect to each other. The objective focuses sheared wave fronts on the surface.

(3) The two shear waves experience varying optical path differences as a function of surface topography and are collected by the objective and are recombined and superimposed by the Nomarski prism.



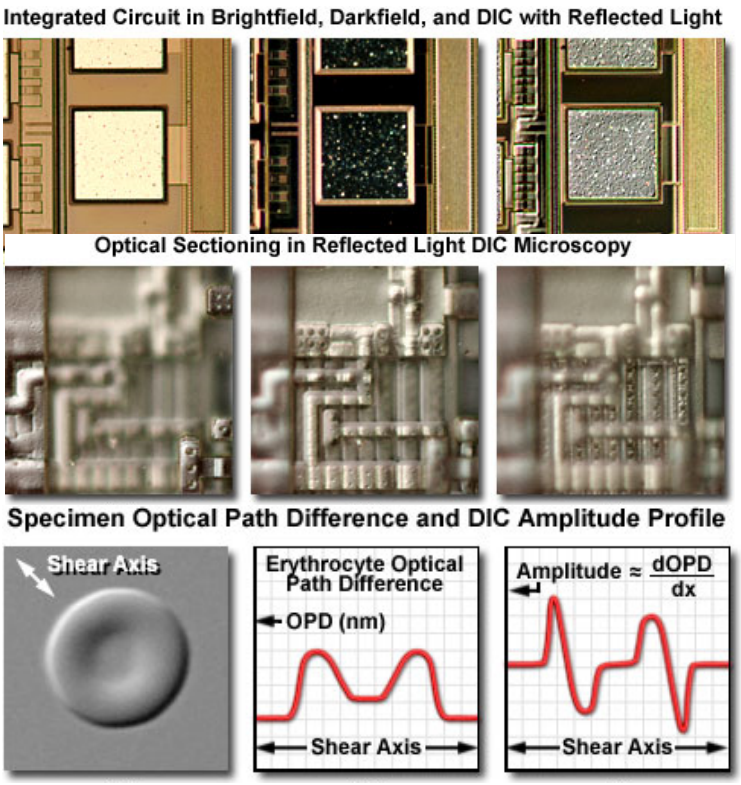
(4) After the Nomarski prism, the waves pass through the half-mirror and the second polarizer analyzer. Components of the orthogonal wave fronts parallel to the analyzer transmission vector to pass through in a common azimuth, and subsequently interfere in the image plane. By adjusting the relative orientation of the polarizer and analyzer axes, the interference contrast in the images can be tuned.

Resulting DIC contrast

Because the intensities in the image are extremely sensitive to the phase gradient between the orthogonally polarized beams, the Nomarski microscope drastically **enhances the resolution in z-direction** down to a few nanometers.

Thus, the topographic contrast in the images is strongly increased and is particularly sensitive to surface steps.

Examples:



Distinction between phase contrast (PCM) and differential interference contrast (DIC):

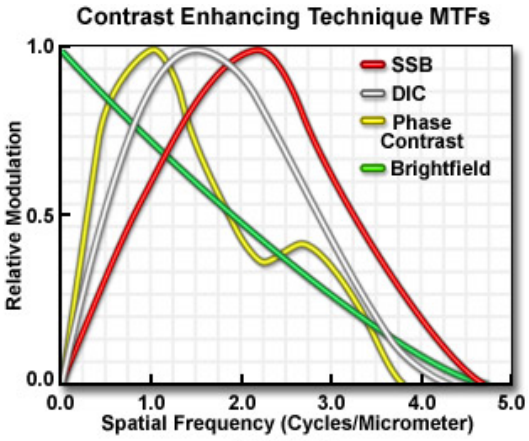
PCM: Intensity ~absolute phase differences

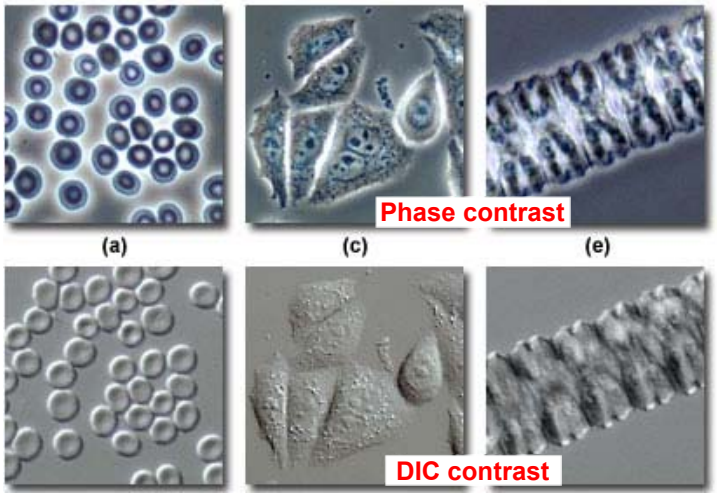
DIC: Intensity ~ *phase gradients* or
~ *topography gradients*

⇨ Thus, DIC images appear as derivative of the bright field images, as if the sample were illuminates from the side.

4.15.5 Comparison of Different Contrast Enhancement Techniques

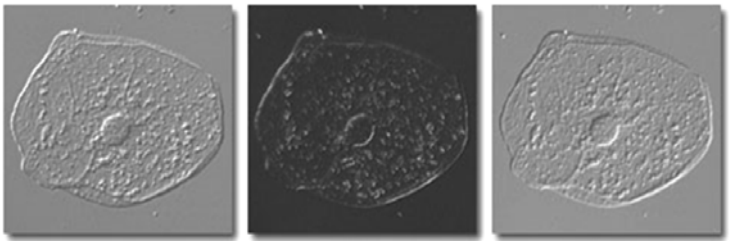
Modulation transfer functions:





Effect of off-axis illumination

Positive and Negative Bias in Differential Interference Contrast



4.16 Summary

- Two main approaches for microscopy exist: **Scanning** and **imaging microscopy**.
  These either consist of local excitation and non-local detector (scanning microscopy), or of non-local excitation and local detection (imaging microscopy).
- The **resolution** in scanning microscopy is limited by the spot size and interaction volume.
- In imaging microscopy, the **resolution is limited by two factors**:
  (a) Diffraction at the objective aperture, which yields: $r_{diff} = 0.61\,\lambda / n \sin\alpha$
  (b) Lens aberrations, introducing broadening $r_{sph} = C_s \cdot \alpha^n$ $n \sim 3$, $C_s \sim f$
  due to spherical & chromatic aberrations $r_{chr} = C_c \cdot dn/n \cdot \alpha^n$ $n \sim 1$, $C_c \sim f$
  Other aberration effects include coma and field of view curvature.
  The **total effective resolution** is thus given by the sum $r_{eff} = (r^2_{dif} + r^2_{aber})^{1/2}$
- The **depth-of-focus** decreases with increasing numerical aperture: $DOF \approx n \cdot \lambda / NA^2$
  By recording *z-stacks* (focus through series), 3D images can be obtained.
- Due to efficient aberrations corrections, the resolution in optical microscopy is mainly diffraction limited and the highest resolution is achieved for ***high numerical apertures*** $NA = n \sin\alpha$ and ***short wavelength*** λ imaging systems.
  For electron and x-ray microscopy the resolution is mainly aberration limited because of the much lower quality of the lenses.

❖ **Point spread function and modulation transfer function**

The image formation process can be described either by the

(i) **convolution of the object with the *point spread function***, or by

(ii) convolution of the Fourier transformation of the object with the ***modulation transfer function***, which describes the contrast modulation amplitude in the image versus that of the sample amplitude as a function of spatial frequency, i.e., periodicity.

Both functions precisely describe the performance of a microscope.

❖ ***Contrast*** is the second crucial parameter for microscopy.
It influences the "*practical*" resolution due to the visibility criterion.

» Therefore, a ***high signal-to-noise ratio*** is essential high resolution imaging.
This can be much improved by proper signal averaging and increase of integration times.

» Many different mechanisms contribute to the contrast such as differences in absorption, reflectivity, topography, diffraction, composition (Z-contrast), strain, phase differences, etc.

» For many specimen, the contrast is weak / insufficient.

» ***Contrast enhancement techniques*** are therefore employed such as staining, selective etching, phase-, interference-contrast, off-axis illumination, etc. .

This modifies the modulation transfer function and enhances it at higher spatial frequencies, thus, increasing the visibility especially of small scale structures closed the resolution limit.