

5. LIGHT MICROSCOPY.

- We use Fourier optics to describe *coherent image formation*, imaging obtained by illuminating the specimen with spatially coherent light.
- We define resolution, contrast, and phase-sensitive methods to enhance contrast.

5.1. Abbe's theory of imaging

- A convergent lens produces at its *back* focal plane the Fourier transform of the field distribution at its *front* focal plane [1, 2]. One way to describe an imaging system (e.g. a microscope) is in terms of a system of two lenses that perform two successive Fourier transforms.

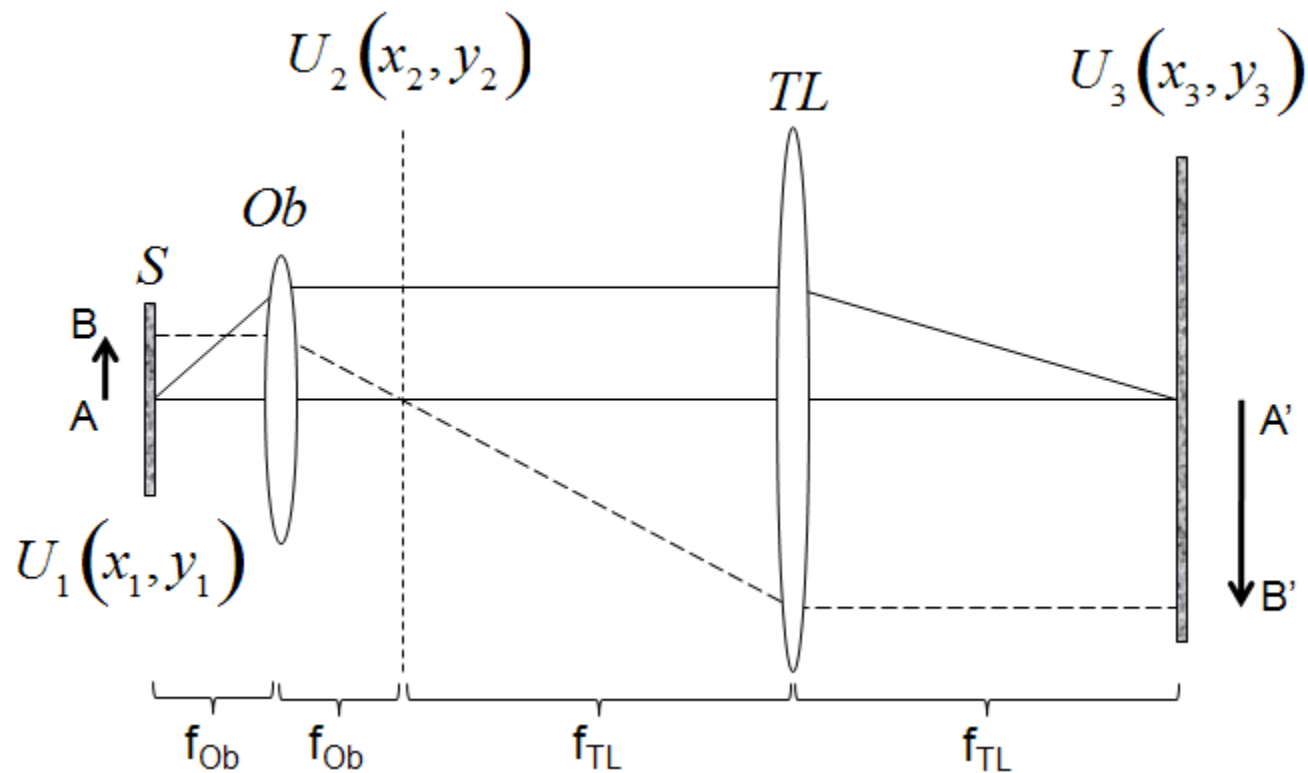


Figure 1. Coherent image formation in a microscope: Ob , objective lens, TL , tube lens.

- Figure 1 shows the geometrical optics image formation through this microscope. A system where the back focal plane of the first lens (objective, Ob) overlaps with the front focal plane of the second (tube lens, TL) is called *telecentric*.
- U_3 is the Fourier transform of U_2 , which is the Fourier transform of U_1 .

- It can be easily shown that applying two forward Fourier transforms recovers the original function, up to a reversal in the coordinates.

$$F \left\{ F \left[f(x, y) \right] \right\} = f(-x, -y). \quad (0.1)$$

- Thus the image A'B' is inverted with respect to the object AB. Because the two lenses have different focal distances, the image field is also scaled by a factor M, called *transverse magnification*.
- We calculate this magnification by evaluating the field U_3 as a function of U_1 ,

$$U_3(k_{x2}, k_{y2}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_2(x_2, y_2) \cdot e^{-i(k_{x2} \cdot x_2 + k_{y2} \cdot y_2)} dx_2 dy_2 \quad (0.2)$$

$$U_2(k_{x1}, k_{y1}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_1(x_1, y_1) \cdot e^{-i(k_{x1} \cdot x_1 + k_{y1} \cdot y_1)} dx_1 dy_1$$

- In Eq. 2, the spatial frequencies are defined as

$$k_{x1} = \frac{2\pi x_2}{\lambda f_{Ob}}; k_{y2} = \frac{2\pi y_2}{\lambda f_{Ob}} \quad (0.3)$$

$$k_{x2} = \frac{2\pi x_3}{\lambda f_{TL}}; k_{y2} = \frac{2\pi y_3}{\lambda f_{TL}}$$

- Plugging Eq. 2b in Eq. 2a, one finds the final expression that relates the image and the object

$$U_3(x_3, y_3) = U_1\left(-\frac{x_1}{M}, -\frac{y_1}{M}\right), \quad (0.4)$$

- where the magnification is given by

$$M = \frac{f_{TL}}{f_{Ob}}. \quad (0.5)$$

- Potentially, the ratio f_{TL}/f_{Ob} can be made arbitrarily large (for instance, cascading many imaging systems).

- This does not mean that the microscope is able to resolve arbitrarily small objects. We have already encountered the limited resolution in extracting the structure of inhomogeneous objects via scattering experiments (Section 2.5). The microscope obeys the limits. Unlike magnification, resolution is fundamentally limited by the laws of physics.

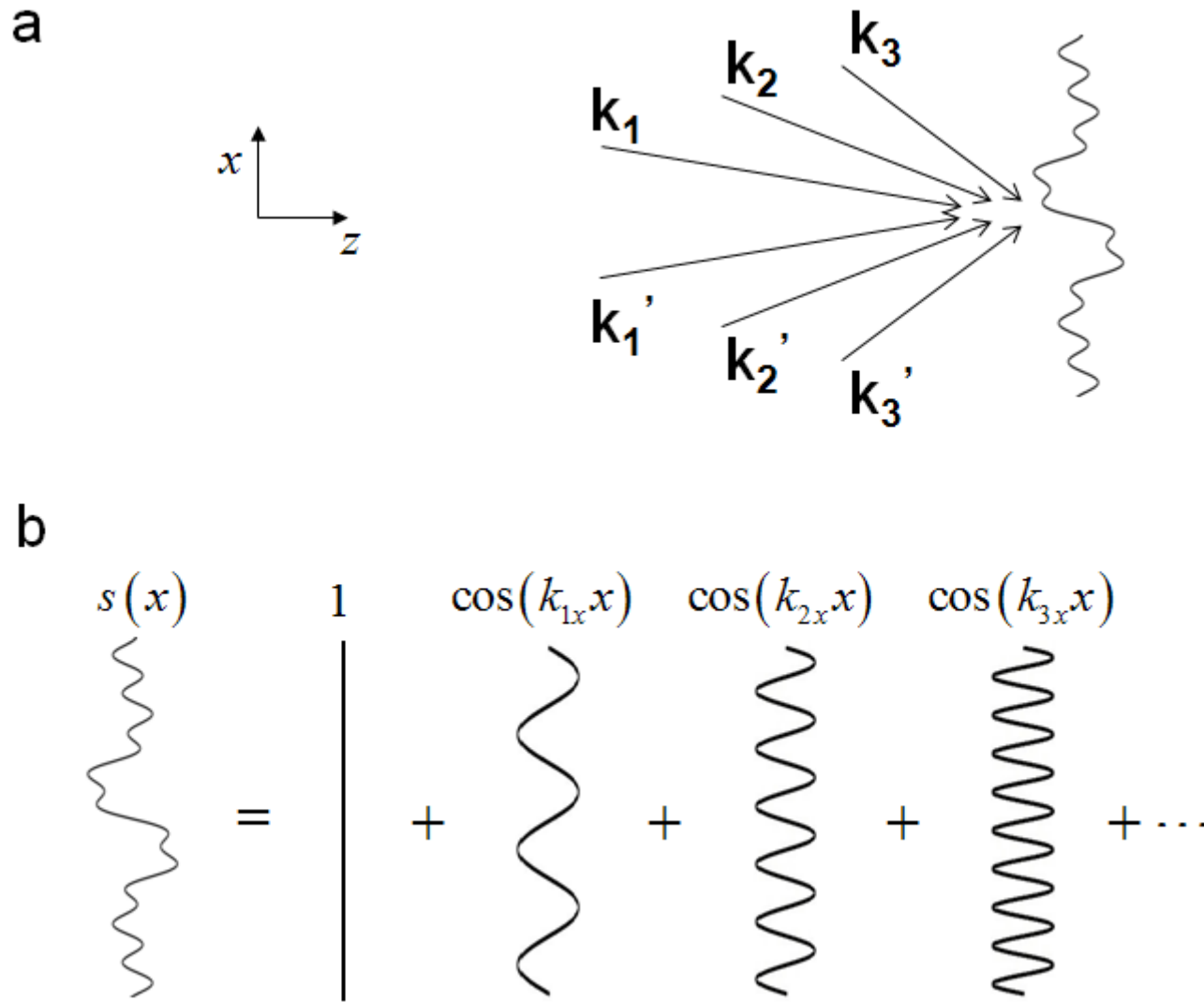


Figure 2. a) Abbe's concept of imaging as an interference phenomenon: the pairs of wave vectors $\mathbf{k}_{1,2,3}-\mathbf{k}'_{1,2,3}$ generate standing waves of different frequencies along the x -axis. b) frequency decomposition of the resulting field.

- Figure 2 provides a physical explanation for the image formation, originally formulated by Abbe in 1873.
- Abbe's theory in his own words: "The microscope image is the interference effect of a diffraction phenomenon" [3].
- Thus, a given image field is formed by the interference between plane waves propagating along different directions (Fig. 2a). The resulting field can therefore be decomposed into sinusoids of various frequencies and phase shifts (Fig. 2b).
- The same picture applies at the sample plane, where each spatial frequency generates pairs of plane waves (diffraction orders) propagating symmetrically with respect to the optical axis (Fig. 3). As this frequency increases, the respective diffraction angle reaches the point where it exceeds the maximum allowed by the objective.

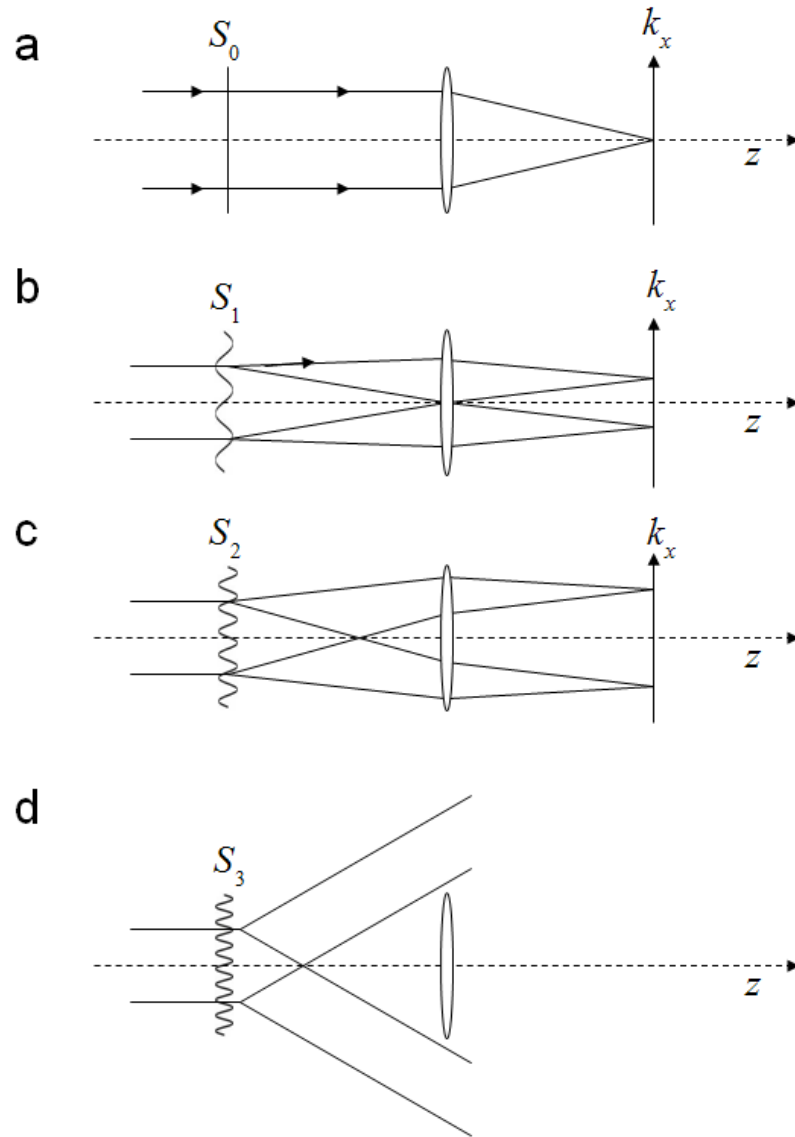


Figure 3. Low-pass filtering effect by the microscope objective.

- This framework allowed Abbe to derive his famous formula for the resolution limit, which we will now discuss.

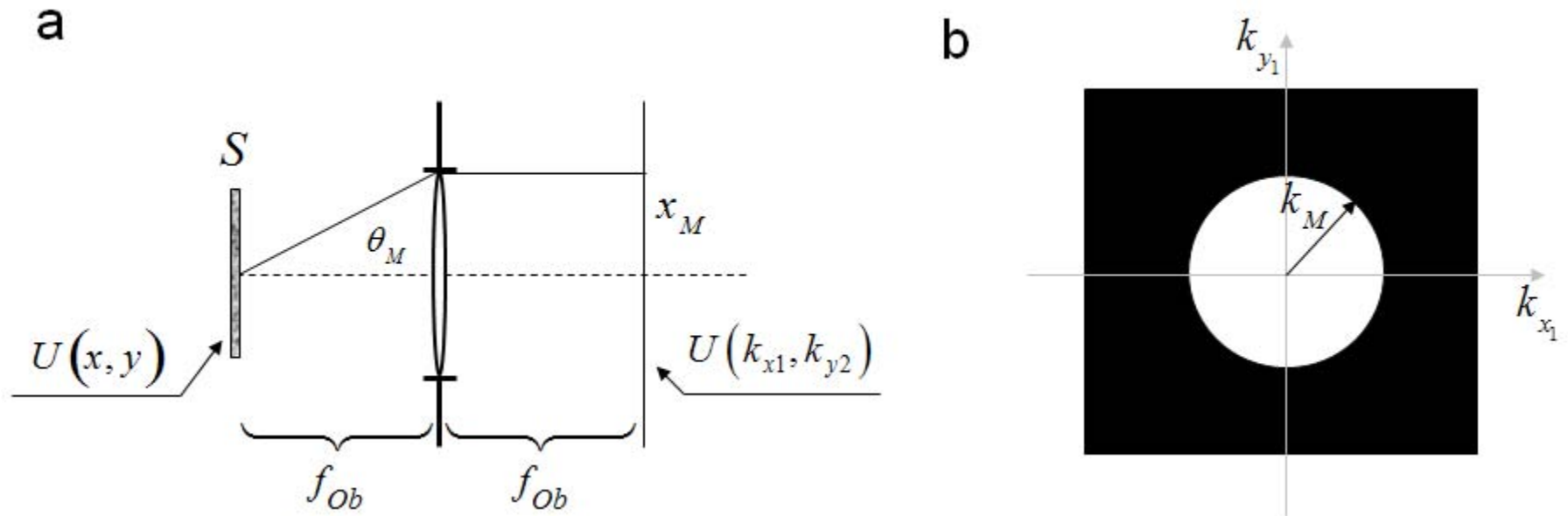


Figure 4. Frequency cut-off in a light microscope: a) maximum angle subtended by the entrance pupil from the specimen; b) entrance pupil.

- Figure 4 illustrates how the apertures present in the microscope objective limit the maximum angle associated with the light scattered by the specimen.

- The effect of the objective is that of a *low-pass filter*, with the cut-off frequency in 1D given by

$$\begin{aligned}
 k_M &= \frac{2\pi}{\lambda f_{ob}} \cdot x_M \\
 &= \frac{2\pi}{\lambda} \cdot \theta_M
 \end{aligned}
 , \tag{0.6}$$

- θ_M is the maximum angle subtended by the *entrance pupil* from the specimen.
- *Qualitatively*, this low-pass filter has the effect of smoothing-down the details in the sample field, i.e. limiting the spatial resolution of the instrument.
- *Quantitatively*, need to find the relationship between the ideal (infinite resolution) sample field $U_1(x_1, y_1)$ and the smooth (image) field, $\underline{U}_3(x_3, y_3)$.
- This can be done in two equivalent ways: 1) express the sample field as the Fourier transform of the field at the *entrance pupil*; 2) express the image field as the Fourier transform of the *exit pupil*.

- Following the first path, the image field can be expressed in terms of the Fourier transform of $U_2(x_2, y_2)$. Rewriting Eq. 2a, we obtain

$$\underline{U}_3(x_3, y_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \underline{U}_2(k_{x2}, k_{y2}) \cdot e^{i(k_{x2} \cdot x_3 + k_{y2} \cdot y_3)} dk_{x2} dk_{y2} \quad (0.7)$$

- $\underline{U}_2(k_{x2}, k_{y2})$ is the frequency domain field that is *truncated* by the exit pupil function.
- The pupil function, $P(k_{x2}, k_{y2})$, is the transfer function of the microscope,

$$\underline{U}_2(k_{x2}, k_{y2}) = U_2(k_{x2}, k_{y2}) \cdot P(k_{x2}, k_{y2}), \quad (0.8)$$

- U_2 is the unrestricted (of infinite support) Fourier transform of the sample field U_1 .
- Combining Eqs. 7 and 8, we obtain the image field, \underline{U}_3 , as the Fourier transform of a *product* between the ideal field, U_2 , and the pupil function, P .

- The image field \underline{U}_3 can be written as the convolution between U_1 and the Fourier transform of P ,

$$\begin{aligned}\underline{U}_3(x_3, y_3) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_3(x_3', y_3') \cdot g(x_3 - x_3', y_3 - y_3') dx_3' dy_3' \quad (0.9) \\ &= U_1(x_3 / M, y_3 / M) * g(x_3, y_3).\end{aligned}$$

- g is the Green's function or PSF of the instrument, and is defined as

$$g(x_3, y_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(k_{x2}, k_{y2}) \cdot e^{i(k_{x2} \cdot x_3 + k_{y2} \cdot y_3)} dk_{x2} dk_{y2}. \quad (0.10)$$

- If the input (sample) field U_1 is a point, expressed by a δ -function, the imaging system blurs it to a spot that in the imaging plane is of the form $\delta(x_1, y_1) * g(x_1, y_1) = g(x_1, y_1)$. g is called sometimes the *impulse response* of the coherent imaging instrument. A point in the sample plane is “smeared” by the instrument into a spot whose size is given by the width of g .

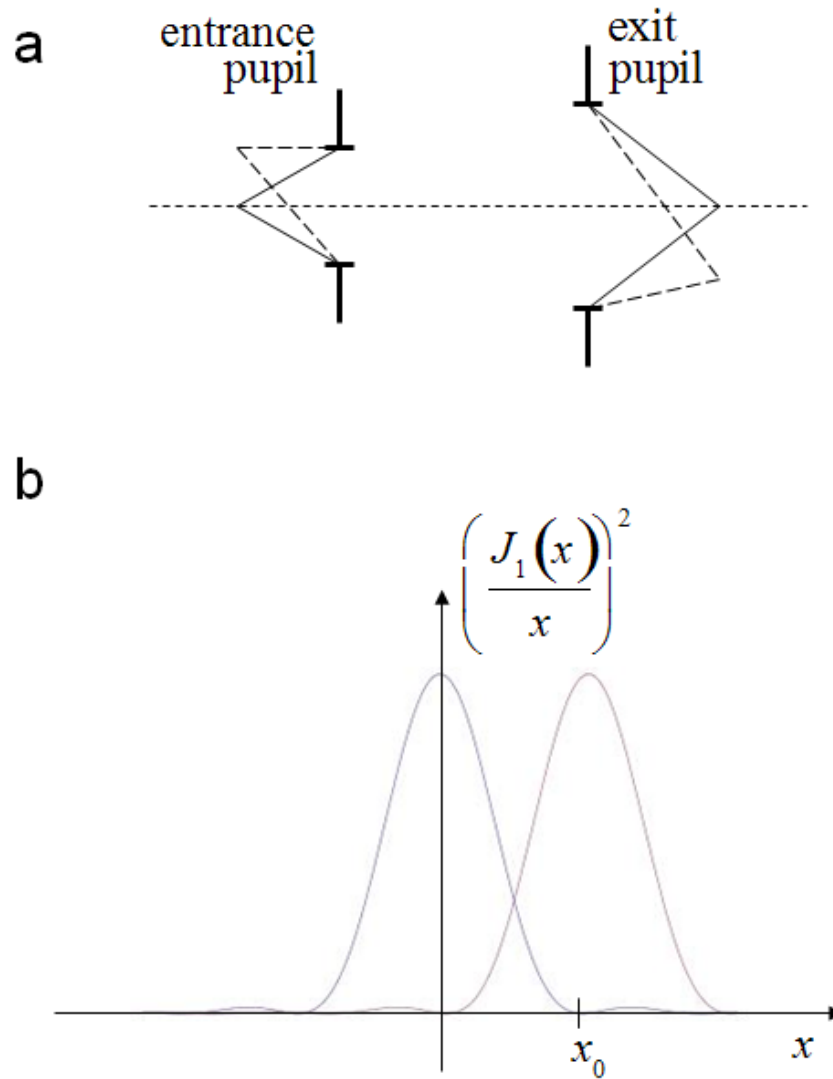


Figure 5. a) Entrance and exit pupils of an imaging system. b) Two points considered resolved by the Rayleigh criterion.

- The impulse response g is merely the field diffracted by the exit pupil (Fig. 5a).

- To obtain an expression for the impulse response g , we need to know the pupil function P . Most commonly, the pupil function is a disk, defined as

$$P(k_{x2}, k_{y2}) = \begin{cases} 1, & \text{if } k_{x2}^2 + k_{y2}^2 \leq k_M^2 \\ 0, & \text{otherwise} \end{cases}, \quad (0.11)$$

$$= \Pi(k / 2k_M)$$

- Π denotes a “rectangular” function in polar coordinates, with $k = \sqrt{k_{x2}^2 + k_{y2}^2}$.
- The Fourier transform of P yields g of the form

$$g(\rho) = \frac{J_1(k_M \rho)}{k_M \rho}$$

$$\rho^2 = x^2 + y^2 \quad (0.12)$$

$$k_M = 2\pi \frac{r_M}{\lambda f_{ob}},$$

- J_1 is the Bessel function of first kind and order, r_M is the radius of the exit pupil.

- The intensity profile, $|g(\rho)|^2$, is depicted in Fig. 5b. Gaskill [1] and others refer to the function $2J_1(x)/x$ as “Sombrero” function, due to its 2D surface plot resembling a Mexican hat.
- The Rayleigh criterion for resolution postulates that two points are considered *resolved* if the maxima associated with their diffraction patterns are separated by at least the first root of the function (normalized coordinate x_0 in Fig. 5). In other words, the maximum of one function overlaps with the root of the second function. This root occurs at $x_0 = 1.22\pi$.
- The resolution, $\rho^{(0)}$, is obtained by solving

$$2\pi \frac{r_M}{\lambda f_{ob}} \cdot \rho_0 = 1.22\pi \quad (0.13)$$

- r_M / f_{Ob} represents the maximum half-angle subtended by the entrance pupil (Fig. 4); this quantity is referred to as the *numerical aperture* of the objective, *NA*.
- We obtain the well known result for resolution defined by the Rayleigh criterion,

$$\rho_0 = 0.61 \frac{\lambda}{NA}. \quad (0.14)$$

- Resolution can be increased by using higher *NA* values or shorter wavelengths. Without the use of immersion liquids, the numerical aperture is limited to $NA = 1$; thus, at best, the microscope can resolve features of the order of $\lambda/2$.
- Recently, research has brought the concept of limited resolution into question. It has been shown that if instead of the *linear* interaction between light and the specimen presented here, a *nonlinear* mechanism is employed, the resolving power of microscopes can be extended, virtually indefinitely [4].

5.2. Imaging of Phase Objects

- *Resolution* is a property of the instrument itself, while *contrast* depends on both the instrument and sample.
- We analyze a special class of samples, which do not absorb or scatter light significantly. They only affect the phase of the illuminating field and not its amplitude; these are generally known as *phase objects*.
- Consider a plane wave, e^{ik_0z} , incident on a specimen, characterized by a *complex transmission function* of the form $A_s e^{i\phi_s(x,y)}$. An ideal imaging system generates at the image plane an identical (i.e. phase and amplitude) replica of the sample field, up to a scaling factor defined by the magnification.
- Since A_s is not a function of spatial coordinate (x, y) , the image field amplitude, A_i , is also a constant.

- Because the detector is only responsive to intensities, the measurement at the image plane yields no information about the phase,

$$\begin{aligned} I_i &= \left| A_i \cdot e^{i\phi(x,y)} \right|^2 \\ &= A_i^2 \end{aligned} \quad (0.15)$$

- Equation 15 states that imaging a phase object produces an intensity image that is constant across the plane, i.e. the image has *zero contrast*.
- For this reason, imaging transparent specimens such as live cells is very challenging. Developing clever methods for generating contrast of phase objects has been driving the microscopy field since the beginning, four centuries ago.
- The *numerical* filtering shown in Section 4.5, will not help, because the intensity has essentially no structure at all. There are *optical* techniques that can be employed to enhance contrast.

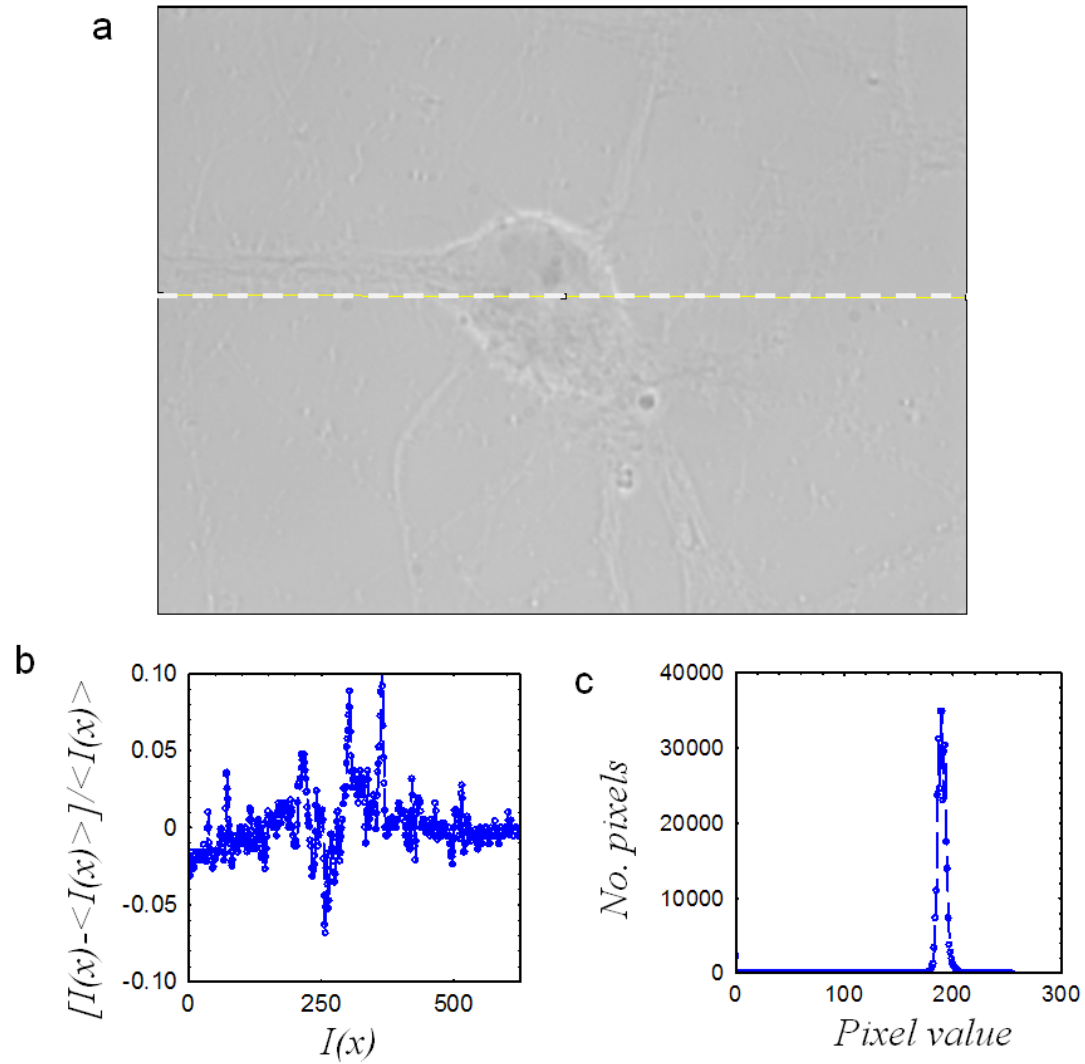


Figure 6. a) Low-contrast image of a neuron. b) Intensity profile along the line shown in a. c) Histogram of intensity distribution in a.

- Consider the intensity profile along one direction for a transparent sample (Fig. 6a). The low contrast is expressed by the small deviation from the mean, of only few %, of the intensity fluctuations, $[I(x) - \langle I(x) \rangle] / \langle I(x) \rangle$. For highest contrast, these normalized fluctuations approach unity.
- Another manifestation of the low contrast is the narrow histogram of the pixel values, which indicates that the intensity at all points is very similar. To achieve high contrast, this distribution must be broadened.
- One straight forward way to increase contrast optically is to simply remove the low-frequency content of the image, i.e. DC component, before the light is detected. For coherent illumination, this *high-pass* operation can be easily accomplished by placing an obstruction on-axis at the Fourier plane of the objective (see Fig. 7).

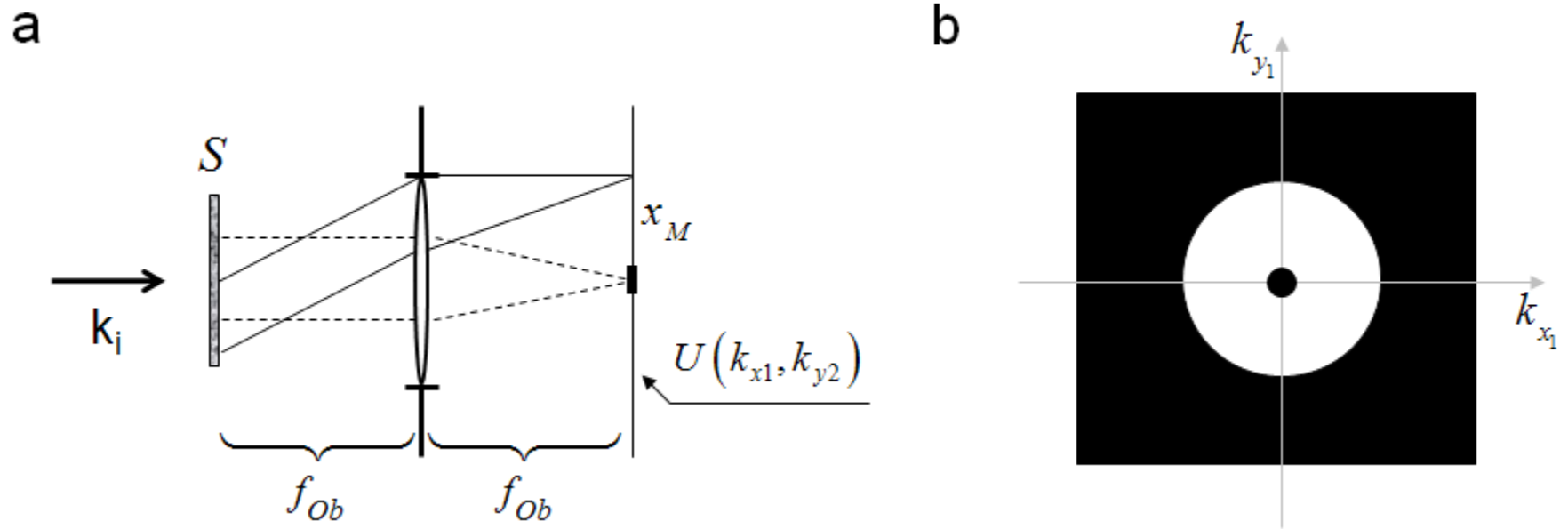


Figure 7. Dark field microscopy: a) the unscattered component is blocked; b) the entrance pupil showing the low frequency obstruction.

- In the absence of the specimen, the incident plane wave is focused on axis and, thus, entirely blocked. This type of “zero-background” imaging is called *dark field microscopy*.

- This is one of the earliest modalities of generating contrast. Originally, this idea was implemented with the illumination beam propagating at an angle higher than allowed by the NA of the objective. This oblique illumination is such that, without a sample, all the light is blocked by the entrance pupil of the system.

5.3. Zernike's phase contrast microscopy.

- Phase contrast microscopy (PCM) represents a major breakthrough in the field of light microscopy. Developed in the 1930s by the Dutch physicist Frits Zernike, for which he received the Nobel Prize in physics in 1953, PCM strikes with its simplicity and yet powerful capability.
- Much of what is known today in cell biology can be traced back to this method as it allows *label-free, noninvasive investigation* of live cells.
- The principle of PCM exploits the early theory of image formation due to Abbe. The image field is the result of the *superposition* of fields originating at the specimen. For coherent illumination, this image field, U , can be conveniently decomposed into its spatial average, U_0 , and fluctuating component, $U_1(x, y)$,

$$\begin{aligned} U(x, y) &= U_0 + [U(x, y) - U_0], \\ &= U_0 + U_1(x, y) \end{aligned} \quad (0.16)$$

- In Eq. 16, the average field, $U_0 = \langle U(x, y) \rangle_{x,y}$, can be expressed as

$$U_0 = \frac{1}{A} \iint U(x, y) dx dy, \quad (0.17)$$

- A is the area of the image.
- This average field can be defined only when the coherence area of the field is larger than the field of view. The summation of complex fields over areas larger than the coherence field is meaningless.
- Taking the Fourier transform of Eq. 16a, we obtain

$$\tilde{U}(k_x, k_y) = \delta(0, 0) + \tilde{U}_1(k_x, k_y). \quad (0.18)$$

- The average field U_0 is the unscattered field, which is focused on axis by the objective, while U_1 corresponds to the scattered component.

- The decomposition in Eq. 16a describes the image field as the interference between the scattered and unscattered components. The resulting image intensity is that of an interferogram,

$$\begin{aligned}
 I(x, y) &= |U(x, y)|^2 \\
 &= |U_0|^2 + |U_1(x, y)|^2 + 2|U_0| \cdot |U_1(x, y)| \cdot \cos[\Delta\phi(x, y)], \quad (0.19)
 \end{aligned}$$

- $\Delta\phi$ is the phase difference between the scattered and unscattered field.
- For the optically thin specimens of interest here, the phase $\Delta\phi$ exhibits small variations, for which the corresponding intensity change is insignificant.
- The intensity is very sensitive to $\Delta\phi$ changes around $\Delta\phi = \pi/2$, or, equivalently, if we replace the cosine term with a sine.

- Realizing that the Taylor expansion around zero gives a quadratic function of the phase in the case of the cosine, $\cos(x) \simeq 1 - x^2/2$, which is negligible for small x values, and a linear dependence for sine, $\sin(x) \simeq x$.
- Zernike understood that by shifting the phase of the unscattered light by $\pi/2$, the image intensity will suddenly exhibit great contrast.
- Let us investigate another variable in generating contrast, namely the ratio between the amplitudes of the two interfering beams.

- We define the *contrast* of this interference pattern as

$$\begin{aligned}
 \gamma(x, y) &= \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \\
 &= \frac{2|U_0| \cdot |U_1(x, y)|}{|U_0|^2 + |U_1(x, y)|^2} \\
 &= \frac{2\beta(x, y)}{1 + \beta(x, y)^2},
 \end{aligned} \tag{0.20}$$

- β is the ratio between the amplitudes of the two fields, $\beta = \frac{|U_1|}{|U_0|}$.
- The contrast γ is a quantitative measure for how intensity across the image varies as a function of $\Delta\phi$.

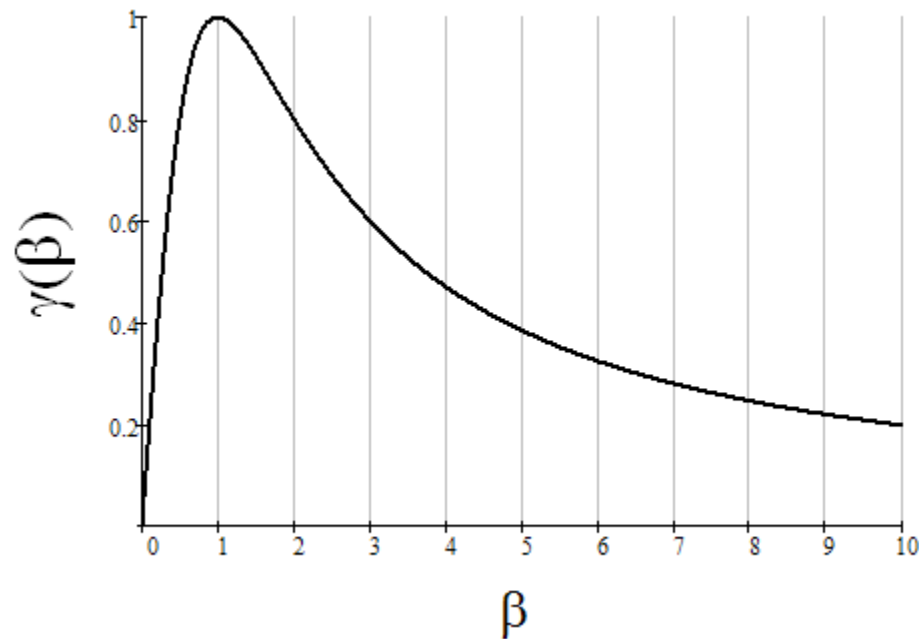


Figure 8. Contrast in the intensity image vs. the ratio between the scattered and unscattered field amplitude, $\beta=|U_1|/|U_0|$ (see Eq. 4.20).

- Figure 8 shows the behavior of γ vs. β . The maximum contrast is achieved when $\beta = 1$, i.e. when the two amplitudes are equal, a well known result in interferometry.

- For transparent samples, $|U_0| \gg |U_1|$, i.e. the unscattered light is much stronger than the scattered light, another reason for low contrast. In addition to the $\pi/2$ phase shift, attenuating $|U_0|^2$ (the unscattered light) is beneficial for improving the contrast.
- Let us now discuss Zernike's optical implementation of these ideas.

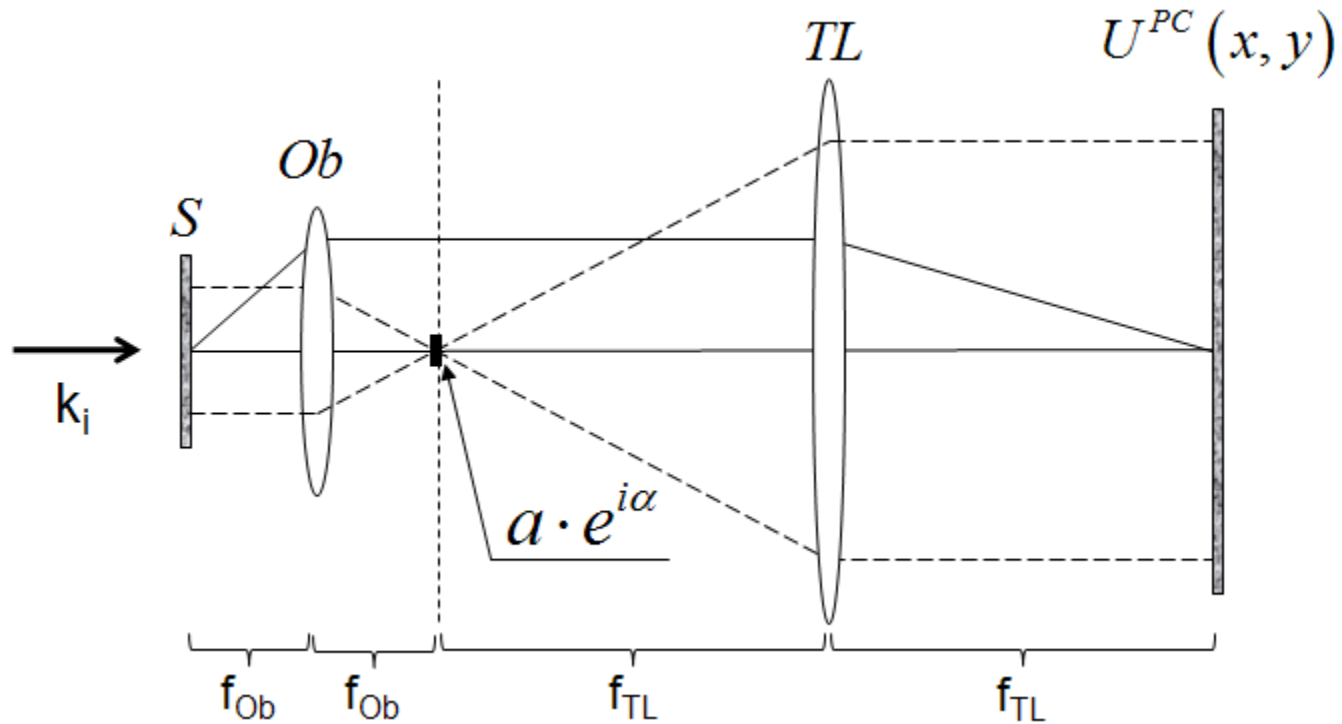


Figure 9. Phase contrast microscope.

- Placing a small metal film that covers the DC component in the Fourier plane of the objective can both *attenuate* and *shift* the phase of the unscattered field (see Fig. 9).

- Consider that the transmission function of this *phase contrast filter* is $a \cdot e^{i\alpha}$, where a describes its attenuation and α the phase shift [5].
- Find the values for a and α that produce optimal contrast.
- Consider that the specimen under investigation is a phase object and that the image field is of unit amplitude, i.e.

$$\begin{aligned}
 U(x, y) &= e^{i\phi(x, y)} \\
 &= 1 + U_1(x, y),
 \end{aligned}
 \tag{0.21}$$

- $U_0 = 1$ is the unscattered field.
- With the phase contrast filter, the new field becomes

$$\begin{aligned}
 U^{PC}(x, y) &= a \cdot e^{i\alpha} + U_1(x, y) \\
 &= a \cdot e^{i\alpha} + e^{i\phi(x, y)} - 1.
 \end{aligned}
 \tag{0.22}$$

- The intensity of the *phase contrast image* has the form

$$\begin{aligned}
 I^{PC}(x, y) &= a^2 + 1 + 1 + 2[a \cos(\alpha + \phi) - \cos \alpha - \cos \phi] \\
 &= a^2 + 2[1 - a \cos \alpha - \cos \phi + a \cos(\alpha + \phi)]
 \end{aligned}
 \tag{0.23}$$

- We can neglect $1 - \cos \phi$ for small ϕ . If we chose $\alpha = \pm \pi/2$ (thus *positive* and *negative* phase contrast), we readily obtain

$$\begin{aligned}
 I^{PC}(x, y) &= a^2 \pm 2a \sin \phi(x, y) \\
 &\simeq a^2 \pm 2a\phi(x, y)
 \end{aligned}
 \tag{0.24}$$

- The $\pm \pi/2$ shift of the unscattered light produces an intensity that is now linear in ϕ . The contrast is further improved for an attenuation of the order $a = 2\phi$.
- Since the attenuation factor is fixed and ϕ is sample dependent, it is impossible to satisfy the condition of maximum contrast for all points simultaneously.

- A practical value of the order of $a = 2 - 3$ is typically used, corresponding to phase shifts of the order of 1-1.5 rad, which is common for live cells.

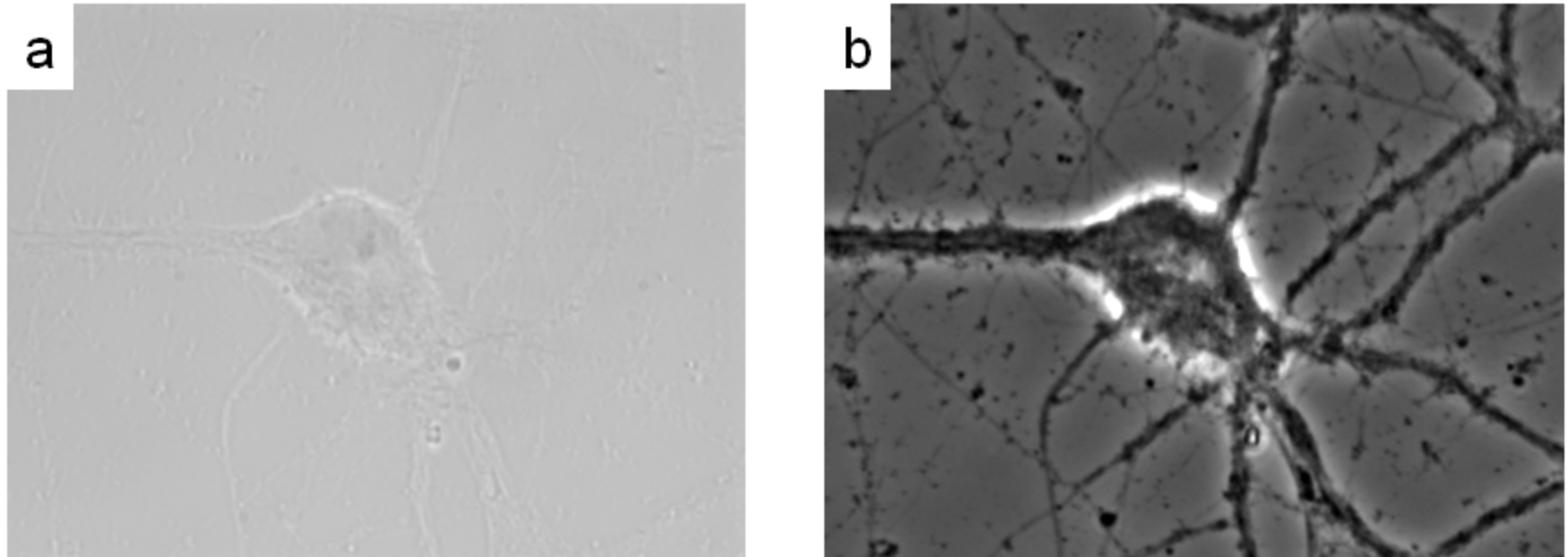


Figure 10. Bright field (a) and phase contrast (b) image of an unstained neuron.

- The powerful capability of PCM is illustrated in Fig. 10. Phase contrast microscopy is significantly more effective in enhancing contrast than the dark field method.

- Instead of removing the unscattered light completely, in phase contrast some DC field is maintained, but is placed in quadrature ($\pi/2$ out of phase) with the scattered component.
- An entertaining analogy can be drawn between phase contrast and martial arts: while carefully blocking the opponent's strikes makes a good approach for defense, managing to turn the opponents energy against him/her is perhaps even better (and energy efficient). In microscopy, the unscattered (direct or DC component) light is, of course, the enemy.
- Figure 10b reveals the well known glowing edges associated with phase contrast, generally known as “halo” effects. This interesting phenomenon happens whenever light from such edges scatter strongly such that portions of this high frequency field end up reaching the phase contrast filter and receive the treatment of the DC field.

- At the image plane, we now have two types of high-frequency fields, phase shifted by $\pi/2$ and not, which generate high contrast fringes (bright or dark halos, depending on the sign of the phase shift filter). The existence of the halo indicates the failure of our assumption for a perfect 2D Fourier transform at the back focal plane of the objective. This is a manifestation of the 3rd dimension of the object.
- The quantitative derivation of the halo intensity distribution requires taking into account a 3D scattering model of the object, beyond the scope of our discussion here.
- Recent advances in QPI have had an impact on the halo artifacts as well. We will discuss in Section 12.2 a new method (Spatial Light Interference Microscopy) that transforms a common phase contrast into a QPI instrument and, in the process, eliminates the halo almost entirely.