Super-resolution imaging : breaking the diffraction limit | 2

(All kudos to Principles of Nanophotonics, by Novotny and Hecht)

2.1 Resolution Limit of Optical Instruments

In microscopy, the verb "*resolve*" can mean both "*to distinguish two very close objects*" and "*to measure correctly the physical size of a very small object*." The **resolution limit** is thus the smallest distance in the object plane that can be faithfully estimated by conjugating the object plane to the image plane.

Consider a small rectangular object of size Δx illuminated and placed at the focal plane of a microscope objective, and attempt to *resolve* it. From the plane wave decomposition, the plane wave spectrum after the object is a sinc function :

$$\tilde{\mathbf{E}}(\alpha,0) = E_0\,\Delta x\,\mathrm{sinc}(\alpha\Delta x)$$

Most of the plane waves in this decomposition are contained in the central lobe, meaning most of the spectrum corresponds to spatial frequencies below $\alpha_0 = \frac{\pi}{\Delta x}$, for a total width typically on the order of $\Delta\alpha = \frac{2\pi}{\Delta x}$. We retrieve the usual Fourier transform relation

$$\Delta x \Delta\alpha \geq 1$$

This illustrates that a tiny object under illumination eventually scatters waves over a broad spatial frequency bandwidth. In far-field optics, the upper bound for $\Delta\alpha$ is given by twice the wavenumber $k = (\omega/c)n = (2\pi/\lambda)n$ of the object medium because we discard spatial frequencies associated with evanescent-wave components. Any scattered field containing waves of higher spatial frequencies will be filtered, and the image formed after propagation will be altered. This means that the resolution limit cannot be better than :

$$\mathrm{Min}[\Delta x] = \frac{1}{2\Delta\alpha}\frac{\lambda}{4\pi n}$$

Resolution limit of a microscope

This upper bound is a theoretical one : for real instruments, the resoltion limit is lower because of the finite size of the imaging systems. A microscope objective is defined by its numerical aperture N.A. = $\sin\theta_{\mathrm{obj}}$. In geometrical optics, rays collected by the objective pupil are inclined at most by an angle θ_{obj} relative to the optical axis. This translates into the framework of Fourier optics using the correspondence between propagation direction for a ray and wavevector component of a plane wave. The key concept behind this correspondance is the fact that the back focal plane of a microscope objective (or of any lens!) is called a Fourier plane :

The Back Focal Plane as a Fourier Plane

We consider a monochromatic wave of wavelength λ and wavenumber $k = 2\pi/\lambda$, and a complex field $U_0(x)$ in the object plane ($z = 0$). For the sake of cimplicity of the argument, propagation is treated in the scalar and paraxial approximation (and not with the full electromagnetism treatment). THis approach is enough to demonstrate our point.

A thin lens of focal length f imparts a quadratic phase to the field :

$$U_L(x) = U_0(x)\,\exp\left(-i\frac{k}{2f}x^2\right). \tag{2.1}$$

At a distance z after the lens, Fresnel diffraction gives :

$$U(x',z) = \frac{e^{ikz}}{i\lambda z}\int_{-\infty}^{\infty} U_L(x)\,\exp\left[i\frac{k}{2z}(x'-x)^2\right]dx. \tag{2.2}$$

We evaluate the field in the back focal plane by setting $z = f$. Substituting $U_L(x)$ yields :

$$U(x',f) = \frac{e^{ikf}}{i\lambda f}\int_{-\infty}^{\infty} U_0(x)\,\exp\left[-i\frac{k}{2f}x^2 + i\frac{k}{2f}(x'-x)^2\right]dx. \tag{2.3}$$

We expand the quadratic term :

$$(x'-x)^2 - x^2 = x'^2 - 2x'x. \tag{2.4}$$

The x^2 terms from the lens and from propagation cancel exactly. Thus we obtain :

$$U(x',f) = \frac{e^{ikf}}{i\lambda f}\exp\left(i\frac{k}{2f}x'^2\right)\int_{-\infty}^{\infty} U_0(x)\,\exp\left(-i\frac{k}{f}xx'\right)dx. \tag{2.5}$$

The integral is, up to normalization, the Fourier transform of the object-plane field. We define the spatial frequency :

$$\alpha = \frac{k}{2\pi f}x' = \frac{x'}{\lambda f}. \tag{2.6}$$

Then :

$$U(x',f) \propto \tilde{U}_0(\alpha)\,\exp\left(i\frac{k}{2f}x'^2\right). \tag{2.7}$$

Up to a quadratic phase factor and normalization, the field in the back focal plane of a thin lens is given by :

$$U(x',f) \propto \mathcal{F}\{U_0(x)\}\left(\alpha = \frac{x'}{\lambda f}\right). \tag{2.8}$$

Thus, the back focal plane is the *Fourier plane* of the field in the object plane : the position x' in the back focal plane corresponds to the spatial frequency α, or equivalently to the transverse wave vector $k_x = k \sin\theta$.

Resolution limit of a microscope

The physical lateral size of a back focal plane is limited.We see from the previous derivation that it directly translates into an upper bound in the spatial frequencies that can propagate into the system :

$$\alpha_{\max} = (2\pi)\Phi_{\mathrm{BFP}}/2\lambda f$$

In addition, because it is conjugated at infinity in the object plane, the physical size of the BFP also limits the aperture of the system : the BFP is the microscope's entrance pupil. The edge of the BFP defines what correspond to the numerical aperture of the system.

$$\mathrm{NA} = n\sin\theta = n\frac{\alpha_{\max}}{k} = n\frac{\Phi_{BFP}}{2f}$$

The actual filtering by a microscope makes the diffraction limit condition tighter than what suggests free space propagation, and can be read :

$$\mathrm{Min}[\Delta x] = 1/(2\alpha_{\max}) = \frac{\lambda}{4\pi NA}$$

In electromagnetism, only waves with a wavevector forming an angle less than θ_{obj} with the optical axis are effectively collected. Equivalently, the microscope can only collect waves with spatial frequencies below $\alpha_{\lim} = \frac{2\pi}{\lambda}\sin\theta_{\mathrm{obj}}$.

Thus, diffracted plane waves are efficiently collected by the objective pupil only if :

$$\alpha_0 < \alpha_{\max} \longrightarrow \Delta x > \frac{\lambda}{2\sin\theta_{\mathrm{obj}}}$$

The microscope resolution limit thus corresponds to a criterion for collecting spatial frequencies diffracted by an object. Two cases arise :

– If Δ_x is large enough (object is sufficiently big), most spatial frequencies diffracted by the object are collected by the microscope. The image formed in the image plane will contain all the information from the object and faithfully reconstruct it : aside from magnification, the object size can be determined.

– If Δ_x is too small, some spatial frequencies will not be collected, and the microscope acts as a low-pass filter. The collected spectrum is truncated : information necessary to faithfully reconstruct the object image is missing. The reconstructed image, aside from magnification, appears larger than the actual object. The typical size of this image is determined by the instrument resolution $\Delta_{x_{\text{res}}}$.

Resolution can be improved by maximizing the numerical aperture, up to 1 in air, corresponding to a half-collection angle of $\pi/2$.

Other standard but arbitrary definitions of the resolution limit

For a circular pupil, **Rayleigh's criterion for the resolution limit** is :

$$\Delta_{x_{\text{res}}} = \frac{1.22\,\lambda}{2\text{N.A.}} = \frac{1.22\,\lambda}{2\sin\theta_{\text{obj}}}$$

2.2 Near-field optical microscopy

Collecting propagating waves is a fundamental aspect of optical resolution, and is thus one way of defining the diffraction limit.

In near-field optical microscopy, an electromagnetic field generated by a localized source interacts with a sample placed very close to it. After this interaction, the resulting field propagates to the distant (far-field) detection region, where it is recorded. A central question is how information associated with structures much smaller than the wavelength can be encoded in the measured radiation. In particular, since evanescent waves do not reach the far-field, one may ask how near-field information becomes accessible at all.

To address this conceptually, we keep the discussion general and do not assume a specific illumination profile (such as a near-field probe or a tightly focused beam), nor do we model probe–sample interactions. These effects can be treated more rigorously, but here we focus on the essential physics.

We consider three planes (see Fig. 4.20) :

1. the source plane at $z = -z_0$,
2. the sample plane at $z = 0$, and
3. the detection plane at $z = z_\infty$.

The source plane could represent, for example, the end facet of a near-field probe or the focal plane of a confocal microscope. The sample plane separates two media of refractive indices n_1 and n_2, respectively.

Angular Spectrum Representation of the Source Field

Using the angular spectrum formalism, the electric field in the source plane can be expressed as

$$E_{\text{source}}(x,y;-z_0) = \iint_{-\infty}^{\infty} \tilde{E}_{\text{source}}(k_x,k_y;-z_0)\, e^{i(k_x x + k_y y)}\, dk_x\, dk_y. \tag{2.9}$$

Propagation from $z = -z_0$ to the sample plane $z = 0$ is described by the usual propagator (see Eq. (3.2)) :

$$E_{\text{source}}(x,y;0) = \iint \tilde{E}_{\text{source}}(k_x,k_y;-z_0)\, e^{i(k_x x + k_y y + k_{z1} z_0)}\, dk_x\, dk_y. \tag{2.10}$$

Here k_{z1} is the longitudinal wavenumber in the medium of index n_1.

Because the source is very close to the sample ($z_0 \ll \lambda$), the incoming field at the surface **contains both propagating and evanescent components**. As indicated schematically in Fig. 4.21, evanescent waves with large transverse wavenumber $k_{\|} = \sqrt{k_x^2 + k_y^2}$ experience strong exponential attenuation, with the suppression increasing as $k_{\|}$ becomes larger.

Interaction with a Thin Sample

For conceptual clarity, we assume that the sample is an extremely thin object described by a transmission function $T(x,y)$. Under this assumption, its effect on the incident field is modeled by the simple relation

$$E_{\text{sample}}(x,y;0) = T(x,y)\, E_{\text{source}}(x,y;0). \tag{2.11}$$

This approximation neglects back-action of the sample on the source field (e.g. probe scattering).

Since the product of two functions in real space corresponds to a convolution in Fourier space, the spectrum of the transmitted field becomes

$$\tilde{E}_{\text{sample}}(\kappa_x,\kappa_y;0) = \iint \tilde{T}(\kappa_x - k_x,\kappa_y - k_y)\, \tilde{E}_{\text{source}}(k_x,k_y;0)\, dk_x\, dk_y \tag{2.12}$$

$$= \iint \tilde{T}(\kappa_x - k_x,\kappa_y - k_y)\, \tilde{E}_{\text{source}}(k_x,k_y;-z_0)\, e^{ik_{z1} z_0} dk_x\, dk_y, \tag{2.13}$$

where $\tilde{T}$ is the Fourier transform of the transmission function.

Equations (2.9)–(2.13) form the basis of understanding how near-field information is encoded in the spatial spectrum and how it can be transferred toward the far-field detection plane.

We now propagate the sample field E_{sample} to the detector located in the far field at $z = z_\infty$. As seen previously, the far field corresponds to the spatial spectrum in the source plane.

Here, however, we are interested in the spatial spectrum in the *detector* plane. Thus we propagate $\tilde{E}_{\text{sample}}$ as

$$E_{\text{detector}}(x,y;z_\infty) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{E}_{\text{sample}}(\kappa_x,\kappa_y;0)\, e^{i(\kappa_x x+\kappa_y y)} e^{i\kappa_z z_\infty}\, d\kappa_x\, d\kappa_y. \tag{2.14}$$

Because of the propagator $\exp(i\kappa_z z_\infty)$, only plane–wave components satisfying

$$|\boldsymbol{\kappa}_{\|}| \leq k_3 = \frac{\omega}{c}\, n_3 \tag{2.15}$$

reach the detector, where the transverse wavenumber is defined as $\boldsymbol{\kappa}_{\|} = (\kappa_x^2 + \kappa_y^2)^{1/2}$.

Taking into account the finite collection angle of a lens with numerical aperture NA, we obtain the restricted condition

$$|\boldsymbol{\kappa}_{\|}| \leq k_3\, \text{NA}. \tag{2.16}$$

This seems at first to be merely a restatement of the diffraction limit. To better understand the consequences, let us rewrite the spectrum of the source field as

$$\tilde{E}_{\text{source}}(k_x,k_y;0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{E}_{\text{source}}(\tilde{k}_x,\tilde{k}_y;0)\, \delta(\tilde{k}_x - k_x)\, \delta(\tilde{k}_y - k_y)\, d\tilde{k}_x\, d\tilde{k}_y, \tag{2.17}$$

which, as illustrated in Fig. 4.21, simply expresses the field as a sum of discrete spatial frequencies.

Thus the source field can be regarded as an infinite set of partial source fields with discrete spatial-frequency pairs. For each pair of frequencies $\pm(\tilde{k}_x,\tilde{k}_y)$ we evaluate the interaction with the sample and its far-field contribution at the detector. Pairs are used because they correspond to equal-amplitude plane (or evanescent) waves forming a stationary standing-wave pattern in the sample plane. After evaluating each pair independently, we then sum all contributions.

Recall that we convolved $\tilde{T}(k'_x,k'_y)$ with $\tilde{E}_{\text{source}}(k_x,k_y;0)$. A source field containing a single pair of spatial frequencies $\mathbf{k}_{\|} = \pm(k_x,k_y)$ simply shifts the transverse wavevectors of the sample :

$$\boldsymbol{\kappa}_{\|} = \mathbf{k}_{\|} \pm \mathbf{k}'_{\|}, \tag{2.18}$$

i.e. it translates the sample spectrum $\tilde{T}$ by $\pm\mathbf{k}_{\|}$. A normally incident plane wave (represented by $\delta(\mathbf{k}_{\|})$) does not shift the spectrum.

Plane waves with maximal transverse wavevector (represented by $\delta(\mathbf{k}_{\|} \pm k)$) propagate parallel to the surface and shift the spectrum by $\pm k$, thereby bringing previously inaccessible spatial frequencies of $T(k'_x,k'_y)$ into the propagating region $|\boldsymbol{\kappa}_{\|}| < k$.

Evanescent components ($\delta(\mathbf{k}_{\|} \pm 2k)$) shift $\tilde{T}$ by $\pm 2k$ and allow spatial frequencies up to $k'_{\|} = 3k$ to enter the cone of propagating waves..

Thus high spatial frequencies of the sample are combined with high spatial frequencies of the probe field so that the difference wavevector corresponds to a propagating mode reaching the far field.

This is analogous to creation of long-wavelength Moiré patterns when two high-frequency gratings multiply. We conclude that a confined source field with a large spatial-frequency bandwidth allows high spatial frequencies of the sample to be accessed in the far field. The better the confinement, the better the achievable resolution.

To estimate the highest detectable spatial frequencies, we use Eqs. (2.16) and (2.18) :

$$\left| k'_{\| \max} \pm k_{\|, \max} \right| = \frac{2\pi}{\lambda}\,\mathrm{NA}. \tag{2.19}$$

For a confined source field with characteristic lateral extent L (aperture diameter, tip diameter, etc.), the maximum transverse wavevector is approximately $k_{\|, \max} \approx \pi/L$. Thus

$$k'_{\| \max} \simeq \left| \frac{\pi}{L} \mp \frac{2\pi}{\lambda}\,\mathrm{NA} \right|. \tag{2.20}$$

In the limit $L \ll \lambda$, the second term can be neglected, and the source confinement alone determines the highest detectable spatial frequency. One must remember, however, that the detection bandwidth remains limited to a disk of radius k_3, and that high spatial frequencies are always mixed with low ones—making image reconstruction nontrivial.