Chapter 7

Angular Spectrum Representation

The angular spectrum representation is a mathematical technique to describe optical fields in homogeneous media. Optical fields are described as a superposition of plane waves and evanescent waves which are physically intuitive solutions of Maxwell's equations. The angular spectrum representation is found to be a very powerful method for the description of laser beam propagation and light focusing. Furthermore, in the paraxial limit, the angular spectrum representation becomes identical with the framework of Fourier optics which extends its importance even further.

In this chapter we will consider purely monochromatic fields of angular frequency ω, which can be represented as complex fields $\underline{\mathbf{E}}(\mathbf{r})$ according to (c.f. Chapter 2)

$$\mathbf{E}(\mathbf{r},t) \;=\; \mathrm{Re}\{\underline{\mathbf{E}}(\mathbf{r})\,\mathrm{e}^{-\mathrm{i}\omega t}\}\;. \tag{7.1}$$

In situations where the field is not time-harmonic, we simply replace the complex field $\underline{\mathbf{E}}(\mathbf{r})$ by the spectrum $\hat{\mathbf{E}}(\mathbf{r},\omega)$ and obtain the time-dependent field by Fourier transformation (see Section 2.2).

By the angular spectrum representation we understand the series expansion of an arbitrary field in terms of plane (and evanescent) waves with variable amplitudes and propagation directions (see Section 2.1.3). Assume we know the electric field $\underline{\mathbf{E}}(\mathbf{r})$ at any point $\mathbf{r} = (x,y,z)$ in space. For example, $\underline{\mathbf{E}}(\mathbf{r})$ can be the solution of an optical scattering problem, as shown in Fig. 7.1, for which $\mathbf{E} = \mathbf{E}_{\mathrm{inc}} + \mathbf{E}_{\mathrm{scatt}}$. In the angular spectrum picture, we draw an arbitrary axis z and consider the field $\underline{\mathbf{E}}$

in a plane $z = \text{const.}$ transverse to the chosen axis. In this plane we can evaluate the two-dimensional Fourier transform of the complex field $\underline{\mathbf{E}}(\mathbf{r}) = \underline{\mathbf{E}}(x, y, z)$ as

$$\hat{\mathbf{E}}(k_x, k_y;\, z) \;=\; \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \underline{\mathbf{E}}(x, y, z)\; \mathrm{e}^{-\mathrm{i}\,[k_x x + k_y y]}\, \mathrm{d}x\, \mathrm{d}y\;, \tag{7.2}$$

where x, y are the Cartesian transverse coordinates and k_x, k_y the corresponding spatial frequencies or reciprocal coordinates. Similarly, the inverse Fourier transform reads as

$$\underline{\mathbf{E}}(x, y, z) \;=\; \iint_{-\infty}^{\infty} \hat{\mathbf{E}}(k_x, k_y;\, z)\; \mathrm{e}^{\mathrm{i}\,[k_x x + k_y y]}\; \mathrm{d}k_x\, \mathrm{d}k_y\;. \tag{7.3}$$

Notice that in the notation of Eqs. (7.2) and (7.3) the field $\underline{\mathbf{E}} = (\underline{E}_x, \underline{E}_y, \underline{E}_z)$ and its Fourier transform $\hat{\mathbf{E}} = (\hat{E}_x, \hat{E}_y, \hat{E}_z)$ represent vectors. Thus, the Fourier integrals hold separately for each vector component.

So far we have made no requirements about the field $\underline{\mathbf{E}}$, but we will assume that in the transverse plane the medium is homogeneous, isotropic, linear and source-free. Then, a time-harmonic, optical field with angular frequency ω has to satisfy

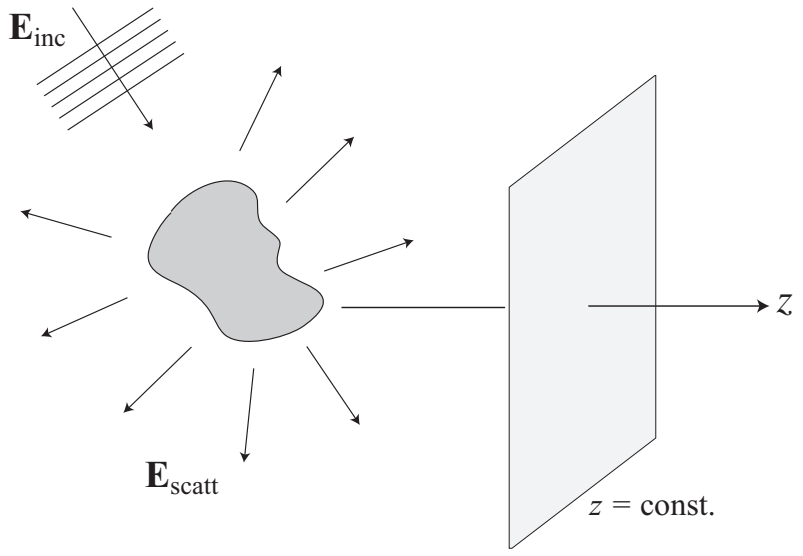

Figure 7.1: In the angular spectrum representation the fields are evaluated in planes ($z = \text{const.}$) perpendicular to an arbitrarily chosen axis z.

the vector Helmholtz equation (c.f. Equation 2.12)

$$(\nabla^2 + k^2)\underline{\mathbf{E}}(\mathbf{r}) \;=\; 0\;, \tag{7.4}$$

where k is determined by $k = (\omega/c)\,n$ and $n = \sqrt{\mu\varepsilon}$ is the index of refraction. Inserting the Fourier representation of $\mathbf{E}(\mathbf{r})$ (Eq. 7.3) into the Helmholtz equation and defining

$$k_z \;\equiv\; \sqrt{(k^2 - k_x^2 - k_y^2)} \qquad \text{with} \quad \mathrm{Im}\{k_z\} \geq 0\;, \tag{7.5}$$

we find that the Fourier spectrum $\hat{\mathbf{E}}$ evolves along the z-axis as

$$\hat{\mathbf{E}}(k_x, k_y;\, z) \;=\; \hat{\mathbf{E}}(k_x, k_y;\, 0)\; \mathrm{e}^{\pm \mathrm{i} k_z z}\;. \tag{7.6}$$

The $\pm$ sign specifies that we have two solutions that need to be superimposed: the $+$ sign refers to a wave propagating into the half-space $z > 0$ whereas the $-$ sign denotes a wave propagating into $z < 0$. Equation (7.6) tells us that the Fourier spectrum of $\mathbf{E}$ in an arbitrary *image plane* located at $z = \mathrm{const.}$ can be calculated by multiplying the spectrum in the *object plane* at $z = 0$ by the factor $\exp(\pm \mathrm{i} k_z z)$. This factor is called the *propagator* in reciprocal space. In Eq. (7.5) we defined that the square root leading to k_z renders a result with positive imaginary part. This ensures that the solutions remain finite for $z \to \pm\infty$. Inserting the result of Eq. (7.6) into Eq. (7.3) we finally find for arbitrary z

$$\underline{\mathbf{E}}(x, y, z) \;=\; \iint_{-\infty}^{\infty} \hat{\mathbf{E}}(k_x, k_y;\, 0)\; \mathrm{e}^{\mathrm{i}\,[k_x x + k_y y \pm k_z z]}\; \mathrm{d}k_x\, \mathrm{d}k_y \tag{7.7}$$

which is known as the *angular spectrum representation*. In a similar way, we can also represent the magnetic field $\underline{\mathbf{H}}$ by an angular spectrum as

$$\underline{\mathbf{H}}(x, y, z) \;=\; \iint_{-\infty}^{\infty} \hat{\mathbf{H}}(k_x, k_y;\, 0)\; \mathrm{e}^{\mathrm{i}\,[k_x x + k_y y \pm k_z z]}\; \mathrm{d}k_x\, \mathrm{d}k_y\;. \tag{7.8}$$

By using Maxwell's equation $\underline{\mathbf{H}} = (\mathrm{i}\omega\mu\mu_0)^{-1}\,(\nabla \times \underline{\mathbf{E}})$ we find the following relationship between the Fourier spectra $\hat{\mathbf{E}}$ and $\hat{\mathbf{H}}$

$$\begin{aligned}
\hat{H}_x &= Z_{\mu\varepsilon}^{-1}\,[(k_y/k)\,\hat{E}_z - (k_z/k)\,\hat{E}_y]\;, \\
\hat{H}_y &= Z_{\mu\varepsilon}^{-1}\,[(k_z/k)\,\hat{E}_x - (k_x/k)\,\hat{E}_z]\;, \\
\hat{H}_z &= Z_{\mu\varepsilon}^{-1}\,[(k_x/k)\,\hat{E}_y - (k_y/k)\,\hat{E}_x]\;,
\end{aligned} \tag{7.9}$$

where $Z_{\mu\varepsilon} = \sqrt{(\mu_0\mu)/(\varepsilon_0\varepsilon)}$ is the wave impedance of the medium. Although the angular spectra of $\underline{\mathbf{E}}$ and $\underline{\mathbf{H}}$ fulfill Helmholtz equation they are not yet rigorous solutions of Maxwell's equations. We still have to require that the fields are divergence free, i.e. $\nabla\cdot\underline{\mathbf{E}} = 0$ and $\nabla\cdot\underline{\mathbf{H}} = 0$. These conditions restrict the k-vector to directions perpendicular to the spectral amplitudes ($\mathbf{k}\cdot\hat{\mathbf{E}} = \mathbf{k}\cdot\hat{\mathbf{H}} = 0$).

For the case of a purely dielectric medium with no losses the index of refraction n is a real and positive quantity. The wavenumber k_z is then either real or imaginary and turns the factor $\exp(\pm\mathrm{i}\,k_z\,z)$ into an oscillatory or exponentially decaying function. For a certain $(k_x,\, k_y)$ pair we then find two different characteristic solutions: plane waves with $k_x^2 + k_y^2 \leq k^2$ and evanescent waves with $k_x^2 + k_y^2 > k^2$ (see Section 2.1.2).

7.1 Propagation and Focusing of Fields

We have established that, in a homogeneous space, the spatial spectrum $\hat{\mathbf{E}}$ of an optical field $\mathbf{E}$ in a plane $z =$ const. (image plane) is uniquely defined by the spatial spectrum in a different plane $z = 0$ (object plane) according to the linear relationship

$$\hat{\mathbf{E}}(k_x, k_y;\, z) \;=\; \hat{H}(k_x, k_y;\, z)\;\hat{\mathbf{E}}(k_x, k_y;\, 0)\;, \tag{7.10}$$

where $\hat{H}$ is the so-called *propagator in reciprocal space*

$$\hat{H}(k_x, k_y;\, z) \;=\; \mathrm{e}^{\pm\mathrm{i}\,k_z\,z} \tag{7.11}$$

also referred to as the *optical transfer function* (OTF) in free space. Remember that the longitudinal wavenumber is a function of the transverse wavenumber, i.e. $k_z = [k^2 - (k_x^2 + k_y^2)]^{1/2}$, where $k = n\,k_0 = n\,\omega/c = n\,2\pi/\lambda$. The $\pm$ sign indicates that the field can propagate in positive and/or negative z direction. Equation (7.10) can be interpreted in terms of linear response theory: $\hat{\mathbf{E}}(k_x, k_y;\, 0)$ is the input, $\hat{H}$ is a filter function, and $\hat{\mathbf{E}}(k_x, k_y;\, z)$ is the output. The filter function describes the propagation of an arbitrary spectrum through space. $\hat{H}$ can also be regarded as the response function because it describes the field at z due to a point source at $z = 0$. In this sense, it is directly related to the Green's function $\overleftrightarrow{\mathbf{G}}_0$.

The filter $\hat{H}$ is an oscillating function for $(k_x^2 + k_y^2) < k^2$ and an exponentially decreasing function for $(k_x^2 + k_y^2) > k^2$. Thus, if the image plane is sufficiently separated from the object plane, the contribution of the decaying parts (evanescent waves) is zero and the integration can be reduced to the circular area $(k_x^2 + k_y^2) \leq k^2$. In other words, the image at z is a *low pass filtered* representation of the original field at $z = 0$. The spatial frequencies $(k_x^2 + k_y^2) > k^2$ of the original field are filtered out during propagation and the information on high spatial variations gets lost. Hence, there is always a loss of information on propagating from near- to far-field and only structures with lateral dimensions larger than

$$\Delta x \;\approx\; \frac{1}{2k} = \frac{\lambda}{4\pi n} \tag{7.12}$$

can be imaged with sufficient accuracy. Here, n is the index of refraction. This equation is qualitative and we will provide a more detailed discussion later. In general, higher resolution can be obtained by a higher index of refraction of the embodying system (substrate, lenses, etc.) or by shorter wavelengths. Theoretically, resolutions down to a few nanometers can be achieved by using far-ultraviolet radiation or X-rays.

Let us now determine how the fields themselves evolve. For this purpose we denote the transverse coordinates in the object plane at $z = 0$ as (x', y') and in the image plane at $z =$ const. as (x, y). The fields in the image plane are described by the angular spectrum Eq. (7.7). We just have to express the Fourier spectrum $\hat{\mathbf{E}}(k_x, k_y\,; 0)$ in terms of the fields in the object plane. Similarly to Eq. (7.2) this Fourier spectrum can be represented as

$$\hat{\mathbf{E}}(k_x, k_y; 0) \;=\; \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \underline{\mathbf{E}}(x', y', 0)\, \mathrm{e}^{-\mathrm{i}\,[k_x x' + k_y y']}\, \mathrm{d}x'\, \mathrm{d}y' \;. \tag{7.13}$$

After inserting into Eq. (7.7) we find the following expression for the field $\underline{\mathbf{E}}$ in the image plane $z =$ const.

$$\begin{aligned} \underline{\mathbf{E}}(x, y, z) &= \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \underline{\mathbf{E}}(x', y';\, 0) \iint_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\,[k_x (x - x') + k_y (y - y') \pm k_z z]}\, \mathrm{d}x'\, \mathrm{d}y' \mathrm{d}k_x\, \mathrm{d}k_y \\ &= \underline{\mathbf{E}}(x, y; 0) \;*\; H(x, y; z) \;. \end{aligned} \tag{7.14}$$

This equation describes an invariant filter with the following impulse response

(*propagator in direct space*)

$$H(x,\,y;\,z) \;=\; \iint_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\,[k_x x + k_y y \pm k_z z]}\,\mathrm{d}k_x\,\mathrm{d}k_y\;. \tag{7.15}$$

H is simply the inverse Fourier transform of the propagator in reciprocal space $\hat{H}$ (7.11). The field at $z =$ const. is represented by the convolution of H with the field at $z = 0$.

7.1.1 Paraxial Approximation

In many optical problems the light fields propagate along a certain direction z and spread out only slowly in the transverse direction. Examples are laser beam propagation or optical waveguide applications. In these examples the wavevectors $\mathbf{k} = (k_x, k_y, k_z)$ in the angular spectrum representation are almost parallel to the z-axis and the transverse wavenumbers (k_x, k_y) are small compared to k. We can then expand the square root of Eq. (7.5) in a series as

$$k_z \;=\; k\,\sqrt{1-(k_x^2+k_y^2)/k^2} \;\approx\; k \;-\; \frac{(k_x^2+k_y^2)}{2\,k}\;. \tag{7.16}$$

This approximation is called the paraxial approximation and it considerably simplifies the analytical integration of the Fourier integrals. In the following we shall apply the paraxial approximation to find a description for weakly focused laser beams.

7.1.2 Gaussian Beams

We consider a fundamental laser beam with a linearly polarized, Gaussian field distribution in the beam waist

$$\underline{\mathbf{E}}(x',\,y',\,0) \;=\; \mathbf{E}_0\,\mathrm{e}^{-(x'^2+y'^2)/w_0^2}\;, \tag{7.17}$$

where $\mathbf{E}_0$ is a constant field vector in the transverse $(x,\,y)$ plane. We have chosen $z = 0$ at the beam waist. The parameter w_0 denotes the beam waist radius. We can calculate the spatial Fourier spectrum at $z = 0$ as[1]

[1] $\int_{-\infty}^{\infty} \exp(-ax^2 + \mathrm{i}bx)\,\mathrm{d}x = \sqrt{\pi/a}\,\exp(-b^2/4a)$ and $\int_{-\infty}^{\infty} x\exp(-ax^2 + \mathrm{i}bx)\,\mathrm{d}x = \mathrm{i}b\sqrt{\pi}\exp(-b^2/4a)/\,(2a^{3/2})$

$$\begin{aligned}\hat{\mathbf{E}}(k_x, k_y; 0) &= \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} \mathbf{E}_0 \, \mathrm{e}^{-(x'^2+y'^2)/w_0^2} \, \mathrm{e}^{-\mathrm{i}\,[k_x x' + k_y y']} \, \mathrm{d}x' \, \mathrm{d}y' \\ &= \mathbf{E}_0 \, \frac{w_0^2}{4\pi} \, \mathrm{e}^{-(k_x^2+k_y^2)\, w_0^2/4}, \end{aligned} \tag{7.18}$$

which is again a Gaussian function. We now insert this spectrum into the angular spectrum representation Eq. (7.7) and replace k_z by its paraxial expression in Eq. (7.16)

$$\underline{\mathbf{E}}(x, y, z) = \mathbf{E}_0 \, \frac{w_0^2}{4\,\pi} \, \mathrm{e}^{\mathrm{i}kz} \iint_{-\infty}^{\infty} \mathrm{e}^{-(k_x^2+k_y^2)(\frac{w_0^2}{4}+\frac{\mathrm{i}z}{2k})} \, \mathrm{e}^{\mathrm{i}\,[k_x x + k_y y]} \, \mathrm{d}k_x \, \mathrm{d}k_y \, , \tag{7.19}$$

This equation can be integrated and gives as a result the familiar paraxial representation of a Gaussian beam

$$\underline{\mathbf{E}}(x, y, z) = \frac{\mathbf{E}_0 \, \mathrm{e}^{\mathrm{i}kz}}{(1 \, + \, 2\,\mathrm{i}\,z/kw_0^2)} \, \mathrm{e}^{-\frac{(x^2+y^2)}{w_0^2} \frac{1}{(1+2\,\mathrm{i}\,z/kw_0^2)}} \, . \tag{7.20}$$

To get a better feeling for a paraxial Gaussian beam we set $\rho^2 = x^2 + y^2$, define a new parameter z_0 as

$$z_0 = \frac{k\, w_0^2}{2} \, , \tag{7.21}$$

and rewrite Eq. (7.20) as

$$\underline{\mathbf{E}}(\rho, z) = \mathbf{E}_0 \, \frac{w_0}{w(z)} \, \mathrm{e}^{-\frac{\rho^2}{w^2(z)}} \, \mathrm{e}^{\mathrm{i}\,[kz - \eta(z) + k\rho^2/2R(z)]} \tag{7.22}$$

with the following abbreviations

$$\begin{aligned} w(z) &= w_0(1 + z^2/z_0^2)^{1/2} && \text{beam radius} \\ R(z) &= z(1 + z_0^2/z^2) && \text{wavefront radius} \\ \eta(z) &= \arctan z/z_0 && \text{phase correction} \end{aligned} \tag{7.23}$$

The transverse size of the beam is usually defined by the value of $\rho = \sqrt{x^2 + y^2}$ for which the electric field amplitude has decreased to a value of $1/\mathrm{e}$ of its center value

$$|\underline{\mathbf{E}}(x, y, z)| \, / \, |\underline{\mathbf{E}}(0, 0, z)| = 1/\mathrm{e} \, . \tag{7.24}$$

It can be shown that the surface defined by this equation is a hyperboloid whose asymptotes enclose an angle

$$\theta = \frac{2}{k w_0} \tag{7.25}$$

with the z-axis. From this equation we can directly find the correspondence between the numerical aperture (NA $= n \sin\theta$) and the beam angle as NA $\approx 2n/kw_0$. Here we used the fact that in the paraxial approximation, θ is restricted to small beam angles. Another property of the paraxial Gaussian beam is that close to the focus, the beam stays roughly collimated over a distance $2z_0$. z_0 is called the *Rayleigh range* and denotes the distance from the beam waist to where the beam radius has increased by a factor of $\sqrt{2}$. It is important to notice that along the z-axis ($\rho = 0$) the phases of the beam deviate from those of a plane wave. If at $z \to -\infty$ the beam was in phase with a reference plane wave, then at $z \to +\infty$ the beam will be exactly out of phase with the reference wave. This phase shift is called *Gouy phase shift*. The $180°$ phase change happens gradually as the beam propagates through its focus. The phase variation is described by the factor $\eta(z)$ in Eq. (7.23). The tighter the focus the faster the phase variation will be.

A qualitative picture of a paraxial Gaussian beam and some of its characteristics are shown in Fig. 7.2. It has to be emphasized that once the paraxial approximation is introduced, the field $\mathbf{E}$ fulfills no longer Maxwell's equations. The error becomes larger the smaller the beam waist radius w_0 is. Another important aspect of Gaussian beams is that they do not exist, no matter how rigorous the theory that

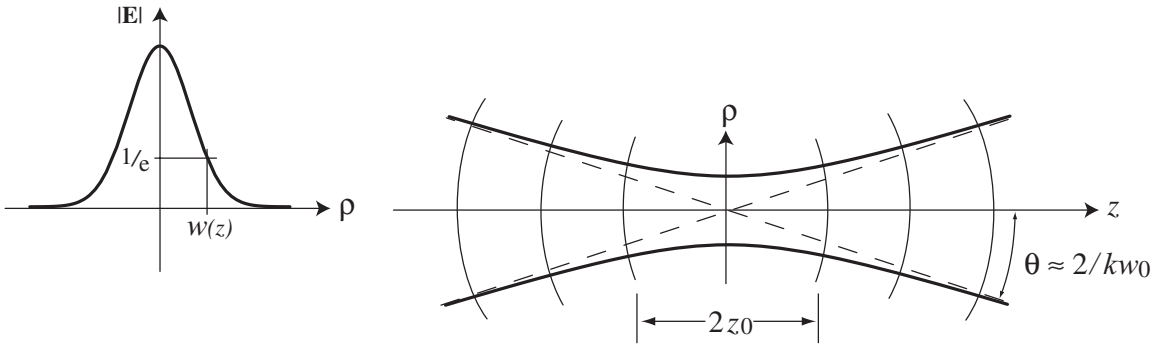

Figure 7.2: Illustration and main characteristics of a paraxial Gaussian beam. The beam has a Gaussian field distribution in the transverse plane. The surfaces of constant field strength form a hyperboloid along the z-axis.

describes them! The reason is that a Gaussian beam profile demands a Gaussian spectrum. However, the Gaussian spectrum is infinite and contains evanescent components that are not available in a realistic situation. Thus, a Gaussian beam must always be regarded as an approximation. The tighter the focus, the broader the Gaussian spectrum and the more contradictory the Gaussian beam profile will be. The angular spectrum representation can be used to derive a rigorous description of focussed fields (e.g. Novotny, Principles of Nano-Optics).

7.2 Far-field Approximation

In this section we will derive the important result that Fourier Optics and Geometrical Optics naturally emerge from the angular spectrum representation.

Consider a particular (localized) field distribution in the plane $z = 0$. The angular spectrum representation tells us how this field propagates and how it is mapped onto other planes $z = z_0$. Here, we ask what the field will be in a very remote plane. Vice versa, we can ask what field will result when we focus a particular far-field onto an image plane. Let us start with the familiar angular spectrum representation of an optical field

$$\underline{\mathbf{E}}(x, y, z) \;=\; \iint_{-\infty}^{\infty} \hat{\mathbf{E}}(k_x, k_y;\, 0)\; \mathrm{e}^{\mathrm{i}\,[k_x x + k_y y \pm k_z z]}\; \mathrm{d}k_x\, \mathrm{d}k_y \;. \tag{7.26}$$

We are interested in the asymptotic far-zone approximation of this field, i.e. in the evaluation of the field in a point $\mathbf{r} = \mathbf{r}_\infty$ at infinite distance from the object plane. The dimensionless unit vector s in the direction of $\mathbf{r}_\infty$ is given by

$$\mathbf{s} = (s_x, s_y, s_z) = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right) \;, \tag{7.27}$$

where $r = (x^2 + y^2 + z^2)^{1/2}$ is the distance of $\mathbf{r}_\infty$ from the origin. To calculate the far-field $\mathbf{E}_\infty$ we require that $r \to \infty$ and rewrite Eq. (7.26) as

$$\underline{\mathbf{E}}_\infty(s_x, s_y) \;=\; \lim_{kr\to\infty} \iint\limits_{(k_x^2+k_y^2)\le k^2} \hat{\mathbf{E}}(k_x, k_y;\, 0)\; \mathrm{e}^{\mathrm{i}\,kr\,[\frac{k_x}{k}s_x + \frac{k_y}{k}s_y \pm \frac{k_z}{k}s_z]}\; \mathrm{d}k_x\, \mathrm{d}k_y \;, \tag{7.28}$$

where $s_z = \sqrt{1 - (s_x^2 + s_y^2)}$. Because of their exponential decay, evanescent waves do not contribute to the fields at infinity. We therefore rejected their contribution and

reduced the integration range to $(k_x^2 + k_y^2) \leq k^2$. The asymptotic behavior of the double integral as $kr \to \infty$ can be evaluated by the method of *stationary phase*. A clear outline of this method can be found in other textbooks (e.g. Born & Wolf, Principles of Optics). Without going into details, the result of Eq. (7.28) can be expressed as

$$\underline{\mathbf{E}}_\infty(s_x, s_y) \;=\; -2\pi \mathrm{i} k\, s_z\, \hat{\mathbf{E}}(k s_x, k s_y;\, 0)\, \frac{\mathrm{e}^{\mathrm{i}kr}}{r}\;. \tag{7.29}$$

This equation tells us that the far-fields are entirely defined by the Fourier spectrum of the fields $\hat{\mathbf{E}}(k_x, k_y;\, 0)$ in the object plane if we replace $k_x \to k s_x$ and $k_y \to k s_y$. This simply means that the unit vector s fulfills

$$\mathbf{s} = (s_x, s_y, s_z) = \left(\frac{k_x}{k}, \frac{k_y}{k}, \frac{k_z}{k}\right)\;, \tag{7.30}$$

which implies that only one plane wave with the wavevector $\mathbf{k} = (k_x, k_y, k_z)$ of the angular spectrum at $z = 0$ contributes to the far-field at a point located in the direction of the unit vector s (see Fig. 7.3). The effect of all other plane waves is cancelled by destructive interference. This beautiful result allows us to treat the field in the far-zone as a collection of rays with each ray being characterized by a particular plane wave of the original angular spectrum representation (Geometrical optics). Combining Eqs. (7.29) and (7.30) we can express the Fourier spectrum $\hat{\mathbf{E}}$ in terms of the far-field as

$$\hat{\mathbf{E}}(k_x, k_y;\, 0) \;=\; \frac{\mathrm{i}r\, \mathrm{e}^{-\mathrm{i}kr}}{2\pi k_z}\, \underline{\mathbf{E}}_\infty(\frac{k_x}{k}, \frac{k_y}{k})\,, \tag{7.31}$$

keeping in mind that the vector s is entirely defined by k_x, k_y. This expression can be substituted into the angular spectrum representation (Eq. 7.26) as

$$\underline{\mathbf{E}}(x, y, z) \;=\; \frac{\mathrm{i}r\, \mathrm{e}^{-\mathrm{i}kr}}{2\pi} \iint\limits_{(k_x^2 + k_y^2) \leq k^2} \underline{\mathbf{E}}_\infty(\frac{k_x}{k}, \frac{k_y}{k})\; \mathrm{e}^{\mathrm{i}\,[k_x x + k_y y \pm k_z z]}\; \frac{1}{k_z}\, \mathrm{d}k_x\, \mathrm{d}k_y \tag{7.32}$$

Thus, as long as evanescent fields are not part of our system then the field $\underline{\mathbf{E}}$ and its far-field $\underline{\mathbf{E}}_\infty$ form essentially a Fourier transform pair at $z = 0$. The only deviation is given by the factor $1/k_z$. In the approximation $k_z \approx k$, the two fields form a perfect Fourier transform pair. This is the limit of *Fourier Optics*.

As an example consider the diffraction at a rectangular aperture with sides $2L_x$ and $2L_y$ in an infinitely thin conducting screen which we choose to be our object

plane ($z=0$). A plane wave illuminates the aperture at normal incidence from the back. For simplicity we assume that the field in the object plane has a constant field amplitude $\mathbf{E}_0$ whereas the screen blocks all the field outside of the aperture. The Fourier spectrum at $z=0$ is then

$$\begin{aligned}\hat{\mathbf{E}}(k_x, k_y; 0) &= \frac{\mathbf{E}_0}{4\pi^2}\int_{-L_y}^{+L_y}\int_{-L_x}^{+L_x} \mathrm{e}^{-\mathrm{i}\,[k_x x' + k_y y']}\,\mathrm{d}x'\,\mathrm{d}y' \\ &= \mathbf{E}_0\,\frac{L_x L_y}{\pi^2}\,\frac{\sin(k_x\,L_x)}{k_x\,L_x}\,\frac{\sin(k_y\,L_y)}{k_y\,L_y}\,, \end{aligned} \tag{7.33}$$

With Eq. (7.29) we now determine the far-field as

$$\underline{\mathbf{E}}_\infty(s_x, s_y) = -\mathrm{i}ks_z\mathbf{E}_0\,\frac{2L_x L_y}{\pi}\,\frac{\sin(ks_x\,L_x)}{ks_x\,L_x}\,\frac{\sin(ks_y\,L_y)}{ks_y\,L_y}\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\,, \tag{7.34}$$

which, in the paraxial limit $k_z \approx k$, agrees with Fraunhofer diffraction.

Equation (7.29) is an important result. It links the near-fields of an object with the corresponding far-fields. While in the near-field a rigorous description of fields

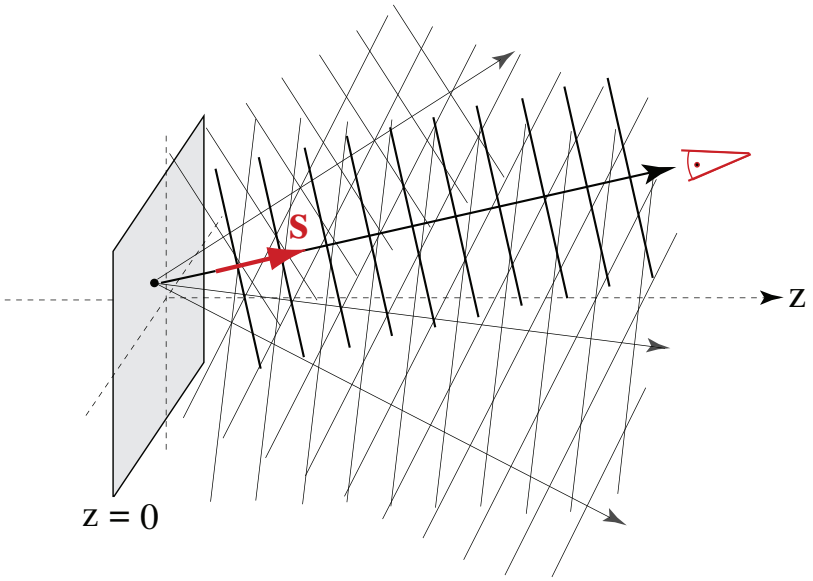

Figure 7.3: Illustration of the far-field approximation. According to the angular spectrum representation, a point in the source plane $z = 0$ emits plane waves in all possible directions. However, a distant detector ($kr \gg 1$) measures only the plane wave that propagates towards it (in direction of unit vector s). The fields of all other plane waves are canceled by destructive interference.

is necessary, the far-fields are well approximated by the laws of Geometrical Optics.

7.3 Fresnel and Fraunhofer Diffraction

Diffraction refers to the observation that light rays break away from their geometrical paths, which is to say, that the wave nature of radiation becomes relevant. In this section we will discuss two important regimes of diffraction theory, Fresnel and Fraunhofer diffraction.

The far-field approximation derived in the previous chapter has its limitations. It has been assumed that the observation point is at infinite distance from the source plane. However, how far do we have to go to be approximately at infinity? The best is to look at this question geometrically and to consider the situation illustrated in Fig. 7.4. R is the observation distance defined as the distance from the origin in the source plane (e.g. center of an aperture) to the observer. On the other hand, r is the true distance between source point $(x', y', 0)$ and the observation point (x, y, z).

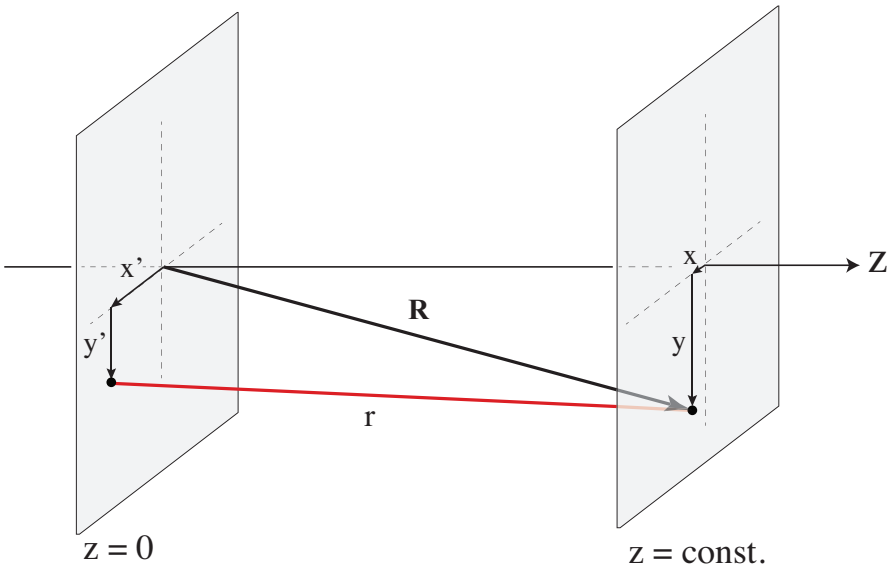

Figure 7.4: Coordinates used in discussion of Fresnel and Fraunhofer diffraction.

The square of r calculated as

$$r^2 = (x - x')^2 + (y - y')^2 + z^2 \;=\; R^2 \left[1 - \frac{2(xx' + yy')}{R^2} + \frac{x'^2 + y'^2}{R^2}\right] . \qquad (7.35)$$

After taking the square root on both sides and invoking the paraxial approximation we obtain

$$r(x', y') \;=\; R \,-\, [x'(x/R) + y'(y/R)] \,+\, \frac{x'^2 + y'^2}{2\,R} \,+\, \ldots . \qquad (7.36)$$

To determine the field at the observation point, we have to sum up the waves emanating from different locations in the source plane (x', y'). This yields integrals of the form

$$\int_{z=0} A(x', y') \, \frac{\exp[-\mathrm{i}\; kr(x', y')]}{r(x', y')} \, dx' \, dy' \, , \qquad (7.37)$$

where $A(x', y')$ is some amplitude function. The "summing up" of elementary spherical waves is referred to as *Huygens' principle*. Because of the large distance between source and observer we can safely replace $r(x', y')$ in the denominator by R or z. However, we cannot apply this approximation to the exponent since we would eliminate the effects of interference. Thus, we need to retain at least one of the additional terms in the expansion of r in (7.36).

Let us denote the maximum extent of the sources at $z = 0$ as D, that is $D/2 = \mathrm{Max}\{\sqrt{x'^2 + y'^2}\}$. If the observation distance is sufficiently large ($R \gg D$), we can neglect the last term in Eq. (7.36) and we end up with *Fraunhofer diffraction*. On the other hand, if the last term is not negligible, we speak of *Fresnel diffraction*. Fresnel diffraction becomes considerably more complicated because the exponent depends on the square of the source plane coordinates.

The transition from Fresnel to Fraunhofer diffraction happens at a distance z that roughly corresponds to the Rayleigh range z_0 of a Gaussian beam (c.f. Equation 7.21), that is, the distance where the beam transitions into a spherical wave. Expressing the beam waist as $w_0 = D/2$ we obtain

$$z_0 \;=\; \frac{1}{8}\, k\, D^2 \, . \qquad (7.38)$$

As an example, let us consider a laser beam with beam diameter of $D = 3\,\mathrm{mm}$ and wavelength $\lambda = 532\,\mathrm{nm}$ (green). It turns out that $z_0 = 13\,\mathrm{m}$, which is quite a distance to reach the far field!

7.4 The Point-Spread Function

The point-spread function is a measure of the resolving power of an imaging system. The narrower the point-spread function the better the resolution will be. As the name implies, the point-spread function defines the spread of a point source. If we have a radiating point source then the image of that source will appear to have a finite size. This broadening is a direct consequence of spatial filtering. A point in space is characterized by a delta function that has an infinite spectrum of spatial frequencies k_x, k_y. On propagation from the source to the image, high-frequency components are filtered out. Usually the entire spectrum $(k_x^2+k_y^2) > k^2$ associated

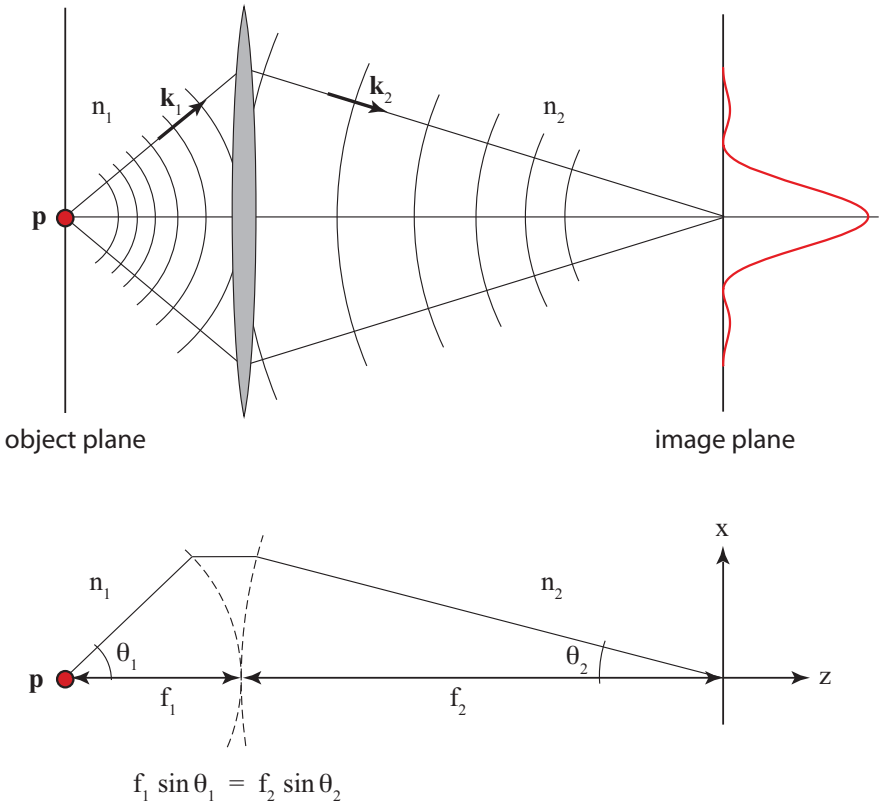

Figure 7.5: Calculation of the point-spread-function (PSF). The fields of a point source are projected onto an image plane. Because of the loss of evanescent waves and the finite angular collection angle of the imaging system, the point appears as a function with finite width.

with the evanescent waves is lost. Furthermore, not all plane wave components can be collected, which leads to a further reduction in bandwidth. The reduced spectrum is not able to accurately reconstruct the original point source and the image of the point will have a finite size.

The smallest radiating electromagnetic unit is a dipole. As shown in Fig. 7.5, to calculate the point-spread function (PSF) we have to trace the dipole's field through an imaging system that focuses it onto an image plane. We will choose the origin of coordinates $(x, y, z) = (0, 0, 0)$ at the focus and use the angular spectrum representation of Eq. (7.32) to calculate the fields in the image plane. It is convenient to represent Eq. (7.32) in spherical coordinates by using the substitutions

$$k_x \;=\; k_2\,\sin\theta_2\,\cos\phi, \qquad k_y \;=\; k_2\,\sin\theta_2\,\sin\phi, \qquad k_z \;=\; k_2\,\cos\theta_2\;. \tag{7.39}$$

Furthermore, due to the symmetry of our problem it is favorable to express the transverse coordinates (x, y) of the field point as

$$x \;=\; \rho\cos\varphi \qquad y \;=\; \rho\sin\varphi\;. \tag{7.40}$$

Finally, we note that the integration in Eq. (7.32) runs over a plane, which is not a constant-coordinate surface in spherical coordinates. We therefore transform the planar integration surface into a spherical one using

$$\frac{1}{k_z}\,\mathrm{d}k_x\,\mathrm{d}k_y \;=\; k_2\,\sin\theta_2\,\mathrm{d}\theta_2\,\mathrm{d}\phi\,, \tag{7.41}$$

which is illustrated in Fig. 7.6.

Taken all together, the focal field represented by Eq. (7.32) can be written as

$$\underline{\mathbf{E}}(\rho,\varphi,z) = \frac{\mathrm{i}k_2 f_2\,\mathrm{e}^{-\mathrm{i}k_2 f_2}}{2\pi}\int_0^{\mathrm{Max}[\theta_2]}\int_0^{2\pi}\underline{\mathbf{E}}_\infty(\theta_2,\phi)\;\mathrm{e}^{\mathrm{i}k_2 z\cos\theta_2}\;\mathrm{e}^{\mathrm{i}k_2\rho\sin\theta_2\cos(\phi-\varphi)}\sin\theta_2\,\mathrm{d}\phi\,\mathrm{d}\theta_2 \tag{7.42}$$

Here, we have replaced the distance r_∞ between the focal point and the surface of the reference sphere of the lens by the focal length f_2. We have also limited the integration over θ_2 to the finite range $[0\,..\,\mathrm{Max}[\theta_2]]$ because any lens will have

a finite size. Furthermore, since all fields propagate in the positive z-direction we retained only the $+$ sign in the exponent of Eq. (7.32).

To evaluate Eq. (7.42) we need to insert the field $\mathbf{E}_\infty$ of the dipole after it has been refracted by the lens. To simplify the analysis we will ignore the vectorial nature of dipole and its fields. Furthermore, we will assume that the imaging system can be treated in the paraxial approximation, or small angle limit, for which $\sin\theta_1 \approx \theta_1$ and $\sin\theta_2 \approx \theta_2$. Using the far-field term of Eq. (6.45) and ignoring the angular dependence (scalar point source), the dipole field before refraction at the lens is

$$E_1 \;=\; -\frac{p\,k_1^2}{4\pi\varepsilon_0\varepsilon_1}\;\frac{\exp(\mathrm{i}k_1 f_1)}{f_1}\;. \tag{7.43}$$

We trace this field through the lens and then insert it as E_∞ into Eq. (7.42) above. In essence, the dipole field is a field that uniformly illuminates the focusing lens, that is $E_\infty(\theta_2,\phi) \approx \mathrm{const.}$. Using the mathematical relation

$$\int_0^{2\pi} \mathrm{e}^{\mathrm{i}x\cos(\phi-\varphi)}\,\mathrm{d}\phi \;=\; 2\pi\,J_0(x)\;, \tag{7.44}$$

we can carry out the integration over ϕ analytically. Here, J_0 is the 0th-order Bessel function. The final expression for the focal field now contains a single integration

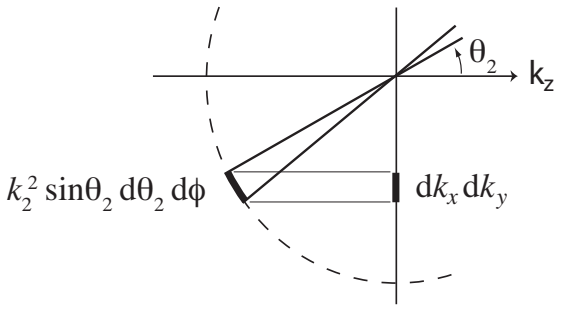

Figure 7.6: Illustration of the substitution $(1/k_z)\,\mathrm{d}k_x\,\mathrm{d}k_y = k\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$. The factor $1/k_z = 1/(k\cos\theta)$ ensures that the differential areas on the plane and the sphere stay equal.

over the variable θ_2. Skipping the constant prefactors we obtain

$$\underline{E}(\rho,\varphi,z) \propto \int_0^{\mathrm{Max}[\theta_2]} \mathrm{e}^{\mathrm{i}k_2 z\cos\theta_2}\, J_0(k_2\rho\sin\theta_2)\,\sin\theta_2\,\mathrm{d}\phi\,\mathrm{d}\theta_2\ . \tag{7.45}$$

Using $\sin\theta_2 \approx \theta_2$, setting $z=0$ (image plane), and using

$$\int x J_0(x)\mathrm{d}x \;=\; x J_1(x)\ , \tag{7.46}$$

we find the following result for the intensity in the image plane

$$\lim_{\theta_{\max}\ll\pi/2} \left|\underline{\mathbf{E}}(\rho,z=0)\right|^2 \;=\; \frac{\pi^4}{\varepsilon_0^2 n_1 n_2}\,\frac{p^2}{\lambda^6}\,\frac{\mathsf{NA}^4}{M^2}\left[2\,\frac{J_1(2\pi\tilde{\rho})}{(2\pi\tilde{\rho})}\right]^2,\quad \tilde{\rho}=\frac{\mathsf{NA}\,\rho}{M\lambda} \tag{7.47}$$

Here, we have used a normalized coordinate $\tilde{\rho}$, which is expressed in terms of the magnification $M=(n_1/n_2)(f_2/f_1)$ and the numerical aperture

$$\mathsf{NA} \;=\; n_1\,\sin(\mathrm{Max}[\theta_1])\ . \tag{7.48}$$

Note that $\sin(\mathrm{Max}[\theta_1]) = (f_2/f_1)\sin(\mathrm{Max}[\theta_2])$. The result (7.47) is the *point-spread-function*, first derived by Abbe in 1873. Its functional form is given by the term in

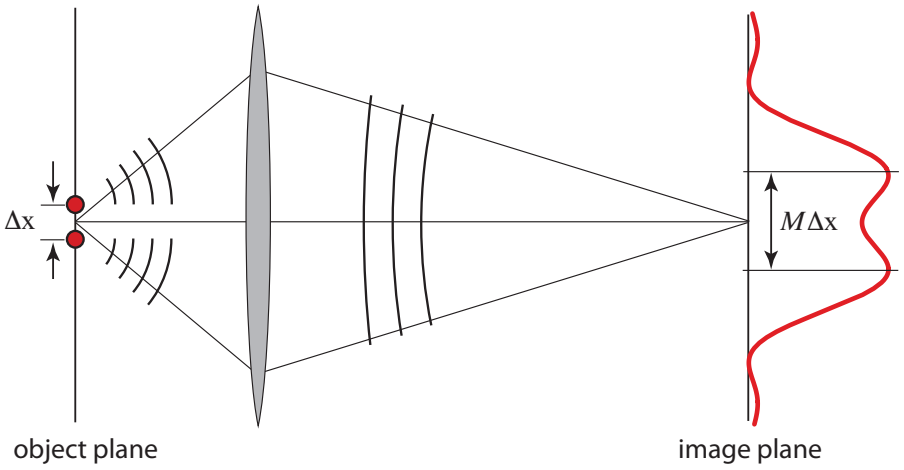

Figure 7.7: Illustration of the resolution limit. Two simultaneously radiating point sources separated by $\Delta r_{||}$ in the object plane generate a combined point-spread function in the image plane. The two point sources are resolved if they can be distinguished based on their image pattern.

brackets which is known as the *Airy function*. It tells us that the image of a point is no longer a point, but a function with a finite width. This width determines the resolution of an imaging system. In essence, two points in the object plane have to be separated by more than the width of the PSF in order to be distinguishable. This is illustrated in Fig. 7.7.

The point-spread function can be measured by using a single quantum emitter, such as a single molecule of quantum dot, as a point emitter. Fig. 7.8 shows such a measurement together with a fit according to Eq. (7.47). The point-spread-function has been recorded by using a NA $= 1.3$ lens to collect the fluorescence photons from a single DiI molecule with center wavelength of $\lambda \approx 580$ nm.

The width of the point-spread function Δx is usually defined as the radial dis-

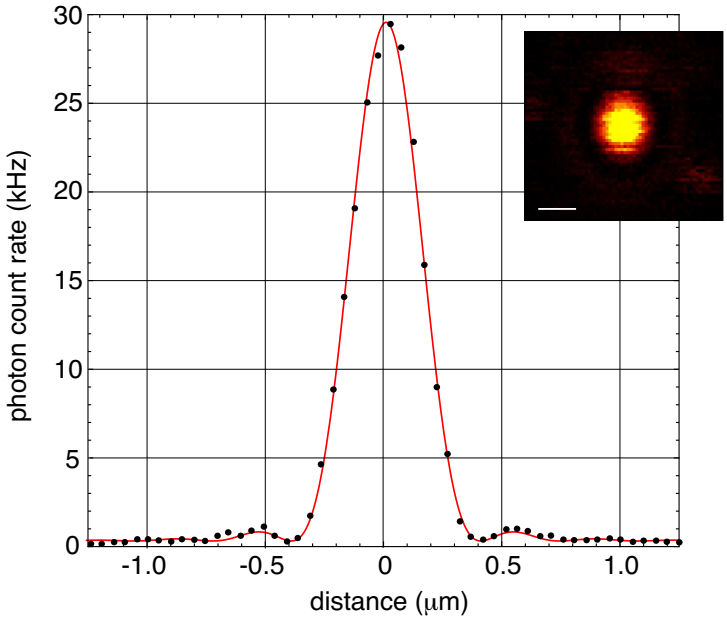

Figure 7.8: Point-spread function measured with a single molecule point source. Fluorescence photons emitted by a DiI molecule are collected with a NA $= 1.3$ objective lens. The center wavelength is $\lambda \approx 580$ nm. The data points correspond to a horizontal line cut through the center of the fluorescence rate image shown in the inset. The solid curve corresponds to the Airy function.

tance for which the value of the paraxial point-spread function becomes zero, or

$$\Delta x \;=\; 0.6098\,\frac{M\,\lambda}{\mathsf{NA}}\;. \tag{7.49}$$

This width is also denoted as the *Airy disk radius*. It depends in a simple manner on the numerical aperture, the wavelength and the magnification of the system.