Smallest Focal Spot

Tasso R. M. Sales
Rochester Photonics Corporation, 330 Clay Road, Rochester, New York 14623
(Received 19 March 1998)

According to diffraction theory the size of a focused spot is determined by wavelength, numerical aperture, and aperture shape. But even when these parameters are fixed it is possible to achieve super-resolution and reduce the spot size even further. It is shown, however, that the three-dimensional spot size cannot be arbitrarily reduced. The minimum focal spot able to probe high-resolution volumetric information in an optical system is about half of the diffraction-limited spot size. [S0031-9007(98)07535-8]

PACS numbers: 42.25.Fx, 03.50.–z, 41.85.Ct

It is a well-known result that, because of diffraction, the image of a point object is no longer a point but spreads over a certain spatial volume whose exact shape depends on the geometry of the aperture. In the case of a circular aperture, the focal distribution has been well characterized [1]. In particular, on a transverse plane of the focal point one finds the famous Airy disk pattern [1] described by a central spot surrounded by low-intensity diffraction rings or sidelobes. Within the bounds of classical optics and disregarding material or source limitations, the size of a focused beam can be continuously decreased by reducing the wavelength or increasing the numerical aperture of the objective. The remarkable fact, however, is that even when these parameters are fixed, the size of the focused spot can be further reduced by means of super-resolution techniques [2]. What happens is that the resolution improvement is typically limited to a single direction followed by a resolution loss in a complementary direction, together with some reduction of brightness and increase of sidelobe intensity.

There is an interest in super-resolution for both theoretical and practical reasons. From a fundamental point of view, super-resolution offers a strategy to overcome the limits of diffraction and increase resolution without the need to modify significantly the optical apparatus. From a practical perspective, there are numerous applications that benefit from super-resolution including astronomy, image processing, confocal scanning microscopy, optical data storage, and laser printing. Depending on the particular problem, it may be more interesting to obtain a super-resolved diffraction pattern in a specific direction. For example, in optical data storage and laser printing it is desirable that the focal point possess a small transverse dimension (leads to larger data density) but a larger axial response or focal depth (leads to better tolerances in beam focusing). An example of such a focal spot can be generated by an annular aperture [3], which retains a large depth of focus over considerable distances and presents a Bessel amplitude pattern on a transverse plane. It is also possible to obtain axial resolution improvement [4], particularly for confocal scanning microscopy, where the resolution of volumetric image reconstruction is limited by the axial depth response [5].

An unexplored form of super-resolution involves the simultaneous increase of resolution over all cross sections of the focal spot, or three-dimensional super-resolution, such that the focal spot is confined within a spatial volume smaller than the diffraction limit. This situation presents interest for confocal microscopy allowing higher lateral resolution together with better depth discrimination. The question then becomes whether it is possible to reduce the size of the focal spot indefinitely and thus increase resolution arbitrarily. We may think of a focused spot as the unit capable of extracting information from the system, much like the photon is the minimal unit of optical radiation in quantum mechanics. Accordingly, we may rephrase our problem by asking what is the minimal unit of classical information able to provide useful information by means of a focused beam. We show that there is a lower bound to the focal confinement beyond the diffraction limit. This result implies a fundamental limitation on the ability of an optical system to detect high-resolution volumetric information exceeding the diffraction-limited performance. It should be noted that there are alternative approaches to increase resolution, such as near-field microscopy [6]. But in this paper we consider only what we call optical super-resolution, or when a focused beam constitutes the probing stylus.

For simplicity, we assume diffraction within Fresnel approximation to describe the focusing by an aberration-free objective. In this case, the normalized complex field amplitude ψ due to a point source located at infinity can be written in adimensional coordinates as [1]

$$\psi(\eta, \delta\mu) = 2\int_0^1 T(r)J_0(\eta r)\exp(-i2\pi\delta\mu r^2)r\,dr\,, \tag{1}$$

where η is a normalized transverse coordinate defined as $\eta = 2\pi \mathrm{NA}\rho/\lambda$, with NA the numerical aperture of the system, λ the wavelength, and ρ the actual transverse coordinate. $\delta\mu$ corresponds to the axial coordinate given by $\delta\mu = \Delta z \mathrm{NA}^2/2\lambda$, where $\Delta z = z - f$ measures the

 0031-9007/98/81(18)/3844(4)$15.00

axial displacement along the optical z axis from the focal location f in actual units. The aperture coordinate r is normalized to one. The function $T(r)$ defines the complex transmission of the pupil, or pupil function, which assumes the value $T = 1$ in the diffraction limit.

Super-resolution is achieved by controlling the phase and/or amplitude of T. In practice, the pupil function is modified by means of filters [5] that can absorb part of the incident beam energy or alter the phase of the wave front. Most super-resolution techniques are based on either amplitude- or phase-only filters. Examples of amplitude-only filters include obscurations or sets of annuli that block light propagation in selected regions of the pupil. A Bessel beam [3], for instance, is obtained with an infinitely thin annulus. In this limit, the transverse spot size measured as the distance between the minima on either side of the focal spot assumes a value equal to 0.62 of the diffraction-limited spot size. Let G_T and G_A represent the spot size relative to the diffraction limit on the transverse focal plane and along the optical axis, respectively. Thus, one sees that the Bessel beam generates transverse super-resolution since it presents $G_T = 0.62 < 1$. However, along the axis the spot size increases considerably leading to an extended focal depth, as illustrated in Fig. 1. Of course, amplitude-only filters lead to an unavoidable loss of brightness. To ameliorate this situation and significantly increase the focal energy one can make use of phase-only filters [7] that absorb no light and offer comparable or better resolution with higher focal intensities. Nevertheless, the general behavior of both types of filters is qualitatively similar. Therefore, to preserve mathematical simplicity, we focus on amplitude-only pupil functions for most of this work by assuming that the pupil is modified by a strictly positive amplitude transmission T. However, our results are in accordance with those obtained with general forms of transmission, including complex functions.

While it has been observed that there is no limit to the resolution improvement [2] along the transverse focal direction, it is still to be answered by how much the axial resolution can be increased. To address this question, consider the axial field calculated at the first zero from the focal peak (in our adimensional units this value equals G_A). As a result, Eq. (1) allows us to write

$$\int_0^1 T(\sqrt{u})\cos(2\pi G_A u)\,du - i\int_0^1 T(\sqrt{u})\sin(2\pi G_A u)\,du = 0\,, \qquad (2)$$

where the simple transformation $u = r^2$ was used. Both the real and imaginary parts must vanish in order to satisfy Eq. (2). It can be seen that this condition can be obeyed only with T nonzero if $G_A > 0.5$. In other words, the axial resolution as measured by the spot size can at most be doubled. For $G_A = 0.5$ the only solution is such that T is identically zero. Phase-only elements can achieve similar results with nonzero brightness, and estimates indicate [8] that the minimum spot size could go as low as $G_A = 4/\pi^2$ (about 0.41). However, such performance has never been observed, and it seems likely that $G_A = 0.5$ is, indeed, the best that can be done in terms of axial super-resolution with optical filters. Of course, we consider a well-defined focal spot with minima on either side assuming nearly zero value. If the focal spot is blurred as a result of the pupil function one can find the distance between minima to be less than 0.5. But then its effective size is much larger since the distance between minima no longer provides a meaningful measure of the spot size.

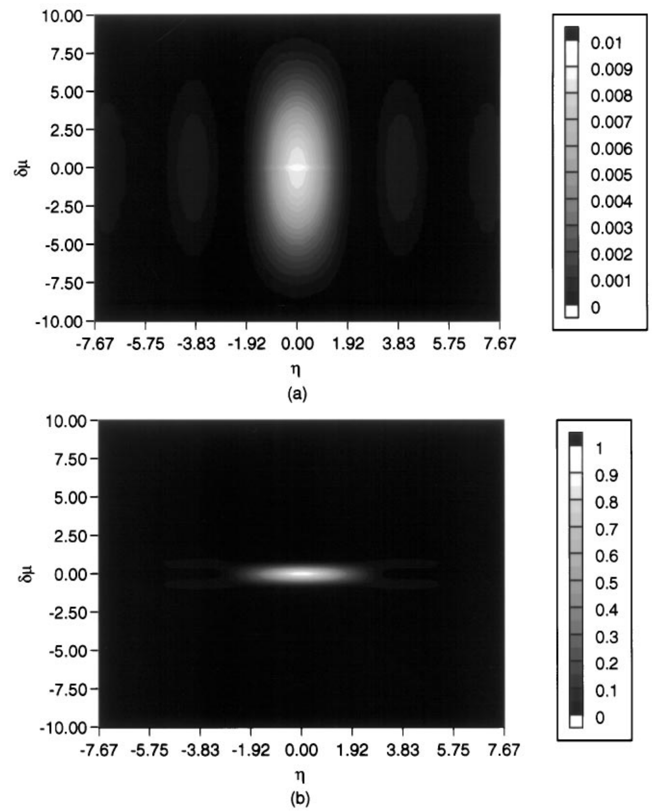

FIG. 1. (a) Three-dimensional diffraction pattern of a beam focused through a thin annulus compared to the (b) diffraction limit.

Considering that the lateral resolution can be continuously improved and that the axial resolution can at most be doubled, the question becomes what results can be obtained by increasing the axial and lateral resolution simultaneously. Although typical resolution improvement is limited to a single direction, there are also solutions [8] for three-dimensional super-resolution. An example, obtained with a simple binary phase-only filter, is shown in Fig. 2. One might be tempted to infer that the focal spot could be made as small as desired, with appropriate filter design. This reasoning, however, is not necessarily correct because the resolution improvement on the transverse

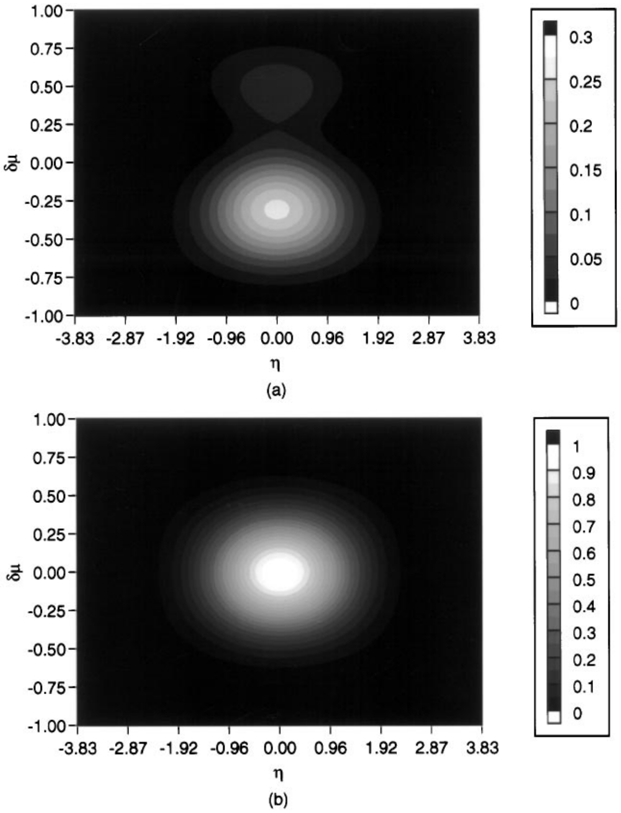

FIG. 2. (a) Three-dimensional confinement of a focal spot compared to the (b) diffraction limit, in the confocal imaging mode (Ref. [5]).

plane is dependent on the improvement found along the axial plane, and vice versa. In fact, there is a minimum possible spot size, which we shall calculate next.

Consider initially the expression for the transverse field, $\delta\mu = 0$ in Eq. (1), calculated at the first zero from the focal peak. If the Bessel function of order zero is expanded in a power series around the origin we can write

$$S^{1/2} = -2\sum_{n=1}^{\infty}(-1)^n c_n^T G_T^{2n}, \tag{3}$$

where S is the focal peak intensity with respect to the diffraction limit (Strehl ratio) and the coefficient c_n^T for the transverse expansion is given by

$$c_n^T = \frac{\eta_{1A}^{2n}}{2^{2n}(n!)^2}\int_0^1 T(r)r^{2n+1}\,dr = \frac{\eta_{1A}^{2n}}{2^{2n}(n!)^2} I_n^T, \tag{4}$$

where η_{1A} denotes the first zero of the diffraction-limited transverse pattern in adimensional units.

The axial field, $\eta = 0$, can similarly be expanded, and the result is

$$S^{1/2} = -i2\sum_{n=1}^{\infty}(-1)^n c_{2n-1}^A G_A^{2n-1} - 2\sum_{n=1}^{\infty}(-1)^n c_{2n}^A G_A^{2n}, \tag{5}$$

where the coefficient for the axial expansion is given by

$$c_n^A = \frac{(2\pi)^n}{n!} I_n^T. \tag{6}$$

Combining Eqs. (3) and (5) one obtains a general relation between the transverse and axial spot sizes, satisfied for a general functional dependence of the pupil function T. There are, in fact, two conditions to be satisfied:

$$\sum_{n=1}^{\infty}(-1)^n c_{2n-1}^A G_A^{2n-1} = 0, \tag{7a}$$

$$\sum_{n=1}^{\infty}(-1)^n c_n^T G_T^{2n} = \sum_{n=1}^{\infty}(-1)^n c_{2n}^A G_A^{2n}. \tag{7b}$$

As the axial spot size decreases towards 0.5 the pupil function T tends to zero and consequently the integral I_k^T also tends to zero. Thus each term in the above series expansions becomes increasingly small. In the limit when G_A tends to 0.5, the summation in Eq. (7a) is arbitrarily close to zero, and we can retain only the first terms in each summation of Eq. (7b). Given a real $\delta > 0$ such that $|G_A - 0.5| < \delta$, there exists an arbitrarily small real $\varepsilon > 0$ such that $I_n^T < \varepsilon$. Since T is real valued for $0 \le r \le 1$, we can define an effective transmission $T_0(\delta) \neq 0$ such that $I_1^{T_0} = \varepsilon_1 < \varepsilon$, by taking $T_0 = 4\varepsilon_1$. The effective transmission T_0 satisfies the limiting condition on G_A for all n. In general, this approximation cannot be expected to hold equally well for any two arbitrary values of n, but this is not required here since we need only to retain the first two terms of Eq. (7b). With this result we calculate I_2^T/I_1^T to be $2/3$, which is actually required only to be a lower bound for three-dimensional super-resolution [9]. After some more calculations we find

$$G_T = \sqrt{\frac{16\pi^2}{3\eta_{1A}^2}}\,G_A. \tag{8}$$

For a circular aperture, $G_T = 1.89 G_A$ since in this case $\eta_{1A} \cong 1.22\pi$. Consequently, if $G_A = 0.5$ then $G_T = 0.94$. Although Eq. (8) assumes a strictly positive pupil function, it is in excellent agreement [8] with results obtained for phase-only filters. Now, taking as a simple measure of the spot size the rectangle determined by the axial and lateral spot dimensions, a focal spot has to satisfy the more general relation

$$G_A G_T \ge g, \tag{9}$$

where g is a geometric coefficient that depends on the shape of the aperture and is defined by $g^2 = \pi^2/3\eta_{1A}^2$, or $g \cong 0.47$ in the circular case. A quick estimate for a square aperture assuming changes only in the transverse pattern yields $g \cong 0.58$. We can now use our definition of adimensional coordinates to express the size of the smallest focal spot σ_0 in actual coordinates as

$$\sigma_0 = \sqrt{\frac{1}{3}\,\frac{\lambda^2}{\mathrm{NA}^3}}\,, \tag{10}$$

which interestingly is dependent only on the numerical aperture and wavelength while the diffraction-limited spot size (and in general a spot modified by a pupil function) also depends on the aperture geometry.

For a more direct comparison, if the diffraction-limited size is written as σ_{DL}, we can express the minimum possible spot size in the following simple form:

$$\sigma_0 = g\sigma_{\mathrm{DL}}\,. \tag{11}$$

Consequently, the minimum possible focal spot is about half of the diffraction-limited spot size, independent of the pupil function used to increase resolution. Equation (11) defines a lower bound for the reduction of the three-dimensional spot size beyond the diffraction limit. How close one can approach the lower bound can be determined only by actual filter optimization. Nevertheless, the spot σ_0 can be regarded as the minimal focal spot able to probe volumetric information. In this respect it is interesting to again evoke the analogy with the photon, which can be considered the optical unit of quantum information. For the photon the energy is well defined but one cannot ascertain a definite spatial localization. For the focused beam, however, the spatial distribution is well defined, but its energy content becomes indefinite. Therefore, the photon and the minimal focal spot play similar roles in the appropriate quantum and classical limits. It should be noted that our results are also valid in the confocal imaging mode [5] since it preserves the zeros of the conventional imaging, as defined by Eq. (1). It is for confocal microscopy applications that spatial confinement of the focal beam presents more interest. It is also interesting to notice that the general result, Eq. (9), represents an analogous relation to the Heisenberg uncertainty principle but now applied to classical fields. That is, it expresses how the access to more information (higher resolution) in a given direction leads to a loss of information in another.

The author is grateful to Professor Colin Sheppard for valuable discussions and a critical reading of the manuscript.

[1] M. Born and E. Wolf, *Principles of Optics* (Pergamon, New York, 1975).
[2] G. Toraldo di Francia, Nuovo Cimento Suppl. **9**, 426–438 (1952).
[3] J. Durnin, J.J. Miceli, and J.H. Eberly, Phys. Rev. Lett. **58**, 1499–1501 (1987).
[4] C.J.R. Sheppard and Z.S. Hegedus, J. Opt. Soc. Am. A **5**, 643–647 (1988).
[5] T. Wilson, *Confocal Microscopy* (Academic, London, 1990).
[6] E. Betzig and J.K. Trautman, Science **257**, 189–195 (1992).
[7] T.R.M. Sales and G.M. Morris, Opt. Lett. **22**, 582–584 (1997).
[8] T.R.M. Sales, Ph.D. Dissertation, University of Rochester, Rochester, NY, 1997.
[9] This result can be verified explicitly for specific forms of transmission function found in the literature. For an interesting example that satisfies $G_A \to 0.5$ and where the ratio I_2^T/I_1^T can be very simply calculated, see M. Martínez-Corral, P. Andrés, J. Ojeda-Castañeda, and G. Saavedra, Opt. Commun. **119**, 491–498 (1995).