

Diffraction of converging electromagnetic waves

Taco D. Visser and Sjoerd H. Wiersma

Department of Molecular Cytology, University of Amsterdam, Plantage Muidergracht 14, 1018 TV Amsterdam, The Netherlands

Received November 8, 1991; revised manuscript received April 3, 1992; accepted April 3, 1992

Using the electromagnetic equivalent of the Kirchhoff diffraction integral, we investigate the effect of spherical aberration and defocus on the diffraction of Gaussian, uniform, and centrally obscured beams. We find, among other things, that in high-angular-aperture systems suffering from either spherical aberration or defocus the axial intensity distribution is no longer symmetric. Equations are derived for the axial intensity near focus for different beam profiles. Intensity contours in focal and meridional planes are depicted for both ideal and aberrated lenses. It is shown that, contrary to certain previous theories, our theory is valid for both high and low angular aperture systems.

1. INTRODUCTION

In a previous paper¹ we have given a detailed description of a new electromagnetic diffraction theory based on the vectorial equivalent of the Kirchhoff-Fresnel integral. The diffracted fields were obtained by integration over the (aberrated) wave front. We applied this theory to investigate the effect of spherical aberration on the electromagnetic field in the focal region of a high-aperture lens. Among other things, we found for an incoming plane wave that the intensity distribution on the axis was no longer symmetric around the peak. A similar feature has recently been measured.² Our aim in the present paper is twofold. First, we show that the vectorial theory of Richards and Wolf,^{3,4} which is valid for high-aperture values, and the paraxial scalar theory of Li and Wolf⁵ are both special cases of our approach. Second, we extend this electromagnetic model to include Gaussian beams, centrally obscured beams, and defocus. Equations are derived for the axial intensity distribution of such systems. It was found for high-aperture lenses with defocus that the displacement theorem (Ref. 6, Chap. 9), which predicts a mere shift of the diffraction pattern, no longer holds. The intensity distribution is now asymmetric and has a lower peak intensity.

It is seen in our study that one can clearly distinguish among three types of lens, namely, paraxial, low-aperture, and high-aperture systems, with Fresnel numbers of order 1, 10^2 , and 10^4 , respectively.

A study of different beam profiles, but for a scalar theory in the Fresnel approximation, has been carried out by Mahajan,⁷ who compared beams with the same total power. We also mention the research of Mansuripur, who describes a vectorial Fresnel diffraction theory.⁸ It should be noted that these Fresnel theories are based on four additional assumptions that are not present in our framework:

- (1) The scalar approximation is used (in the case of Mahajan).
- (2) The amplitude variation over the emerging wave front is neglected (see Section 6).
- (3) The obliquity factor is left out.⁹
- (4) Spherical wave fronts are modeled as quadratic surfaces.⁹

The present authors believe that this makes the Fresnel approach less suitable for the study of systems with high angular aperture. How much the results differ when the above approximations are made is a question that, at least here, can be answered only partially. Mahajan studied systems with Fresnel numbers between 1 and 100. We have explored a wider range of systems and found that high-Fresnel-number lenses (say, with $N = 10^4$) do not show the same behavior as lenses with $N = 100$. We compare our results with those of Mahajan in Section 10.

This paper is organized as follows. In Section 2 we briefly repeat the basic equations of our previous studies. In Section 3 the differences between our theory and that of Richards and Wolf, which is based on the Debye approximation, are discussed. In Section 4 we show that our theory, contrary to that of Richards and Wolf, is also valid for systems with low Fresnel numbers and accurately describes the focal shift phenomenon that occurs in such systems. Section 5 deals with the influence of spherical aberration and defocus on the form of the diffracted wave front. The electromagnetic fields on the deformed wave front, which are needed for the diffraction integrals, are calculated in Section 6. In Section 7 different beam profiles are described. A general integral expression is derived in Section 8 for the axial electromagnetic fields in the focal region of a high-aperture system. The symmetry properties of this equation are also discussed. In Section 9 it is shown that the equation can be solved in terms of Fresnel functions for the case of an ideal lens focusing a uniform wave. In Section 10 axial intensities for converging waves with defocus and spherical aberration are discussed. Different beam profiles are also compared. In Section 11 intensity contours in the focal plane and a meridional plane of different lenses

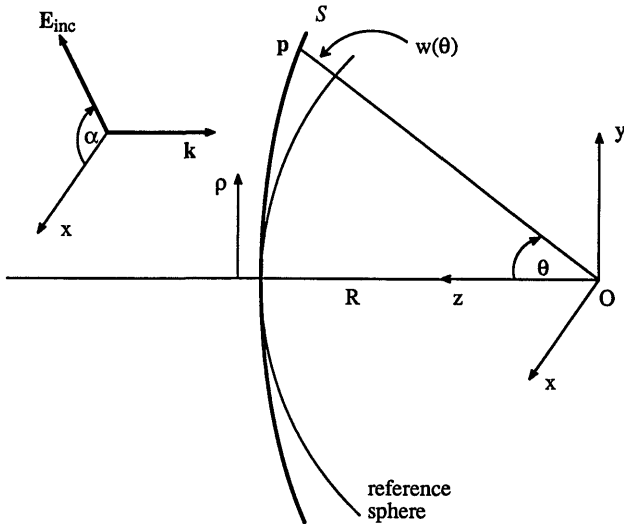


Fig. 1. Definition of the coordinate system. Shown at the left are the wave vector \mathbf{k} and the electric vector \mathbf{E}_{inc} , both before refraction by a lens with semiaperture angle Ω . The angle between the positive x axis and \mathbf{E}_{inc} is the polarization angle α . The origin is placed at focus, which is the center of the reference sphere with radius R . \mathbf{p} is a point on the aberrated wave front S where the Kirchhoff integral is evaluated. The aberration function $w(\theta)$ is defined as the difference between $|\mathbf{p}|$ and R . The azimuthal angle ϕ (not indicated) is defined as usual as the angle between the positive x axis and the projection of \mathbf{p} onto the xy plane. The lateral coordinate ρ equals $(x^2 + y^2)^{1/2}$. The incoming wave propagates along the negative z axis.

are presented. Finally, our results are summarized in Section 12.

2. DIFFRACTION INTEGRALS

We consider the diffraction of an incoming plane-polarized monochromatic electromagnetic wave by a high-aperture lens (Fig. 1). For the electric field $\tilde{\mathbf{E}}(\mathbf{x}, t)$ and the magnetic field $\tilde{\mathbf{B}}(\mathbf{x}, t)$, we have

$$\begin{aligned}\tilde{\mathbf{E}}(\mathbf{x}, t) &= \text{Re}[\mathbf{E}(\mathbf{x})\exp(-i\omega t)], \\ \tilde{\mathbf{B}}(\mathbf{x}, t) &= \text{Re}[\mathbf{B}(\mathbf{x})\exp(-i\omega t)],\end{aligned}\quad (1)$$

where $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ are the time-independent parts of the fields. In a previous paper¹ based on the vectorial Kirchhoff-Fresnel diffraction integral,^{10,11} two basic equations were derived for the electric and the magnetic field near focus in terms of \mathbf{E}_S and \mathbf{B}_S , the spatially dependent electric and magnetic fields on the wave front S , namely,

$$\mathbf{E}(\mathbf{x}) = \int_S [(-ikG + \nabla G \cdot \hat{\mathbf{n}})\mathbf{E}_S - \hat{\mathbf{n}}(\nabla G \cdot \mathbf{E}_S)]d\sigma, \quad (2)$$

$$\mathbf{B}(\mathbf{x}) = \int_S [(-ikG + \nabla G \cdot \hat{\mathbf{n}})\mathbf{B}_S - \hat{\mathbf{n}}(\nabla G \cdot \mathbf{B}_S)]d\sigma, \quad (3)$$

where $\hat{\mathbf{n}}$ denotes the inward normal of the wave front S and $k = 2\pi/\lambda$ is the wave number. The Green's function G and its derivative are given by

$$G(\mathbf{p}, \mathbf{x}) = \exp(ik|\mathbf{p} - \mathbf{x}|)/4\pi|\mathbf{p} - \mathbf{x}|, \quad (4)$$

$$\nabla G(\mathbf{p}, \mathbf{x}) = (1/|\mathbf{p} - \mathbf{x}| - ik)G\hat{\mathbf{e}}_G, \quad (5)$$

respectively. The unit vector $\hat{\mathbf{e}}_G$ is directed from a point \mathbf{p} on S , where the integrand is evaluated, to the point \mathbf{x} ,

where the field is calculated:

$$\hat{\mathbf{e}}_G = (\mathbf{x} - \mathbf{p})/|\mathbf{x} - \mathbf{p}|. \quad (6)$$

The integrations in Eqs. (2) and (3) extend over that portion of the wave front S that approximately fills the exit pupil; in other words, the θ integration extends from zero to the semiaperture angle Ω (see Fig. 1). In deriving these two expressions, we use the free-space form of the Green's function. That means that all further analysis describes imaging in vacuum or, to a high degree of accuracy, in air.

In the presence of defocus or spherical aberration the wave front S is no longer spherical. Its exact form and that of the surface element $d\sigma$ are studied in Section 5. Equations for the electromagnetic field on S are derived in Section 6.

The above expressions, which are valid for both low- and high-aperture systems, can be used to describe spherical aberration and defocus, even when the aberrations are large. Also, different beam profiles (such as centrally obscured, Gaussian, and uniform incoming beams) can be dealt with (see Section 10). In general, the equations must be solved numerically. For points on the axis, however, a further analysis is possible, as we show in Sections 4, 8, and 9. But first we show that both the theory of Richards and Wolf and that of Li and Wolf follow as special cases from these equations.

3. DEBYE APPROXIMATION

The electromagnetic theory presented in this paper not only can deal with several beam profiles and aberrations but also makes no use of the Debye approximation. The theory of Richards and Wolf,^{3,4} on the other hand, deals with the focusing of a uniform plane wave by a perfect high-aperture lens. They use the Kirchhoff diffraction integral together with the Debye approximation. It is on the consequences of this last point that we will elaborate in this section.

In a classic paper on the Huygens-Fresnel principle, Debye¹² made a certain approximation for the Green's function and its derivative defined in our Eqs. (4) and (5) above. We now show that applying these approximations in Eqs. (2) and (3) leads, for an aberration-free system, to the theory of Richards and Wolf. Incidentally, a connection between the vectorial Kirchhoff diffraction integral and the results of Richards and Wolf has also been pointed out by Sheppard *et al.*¹³ Let us then follow Debye and approximate Eqs. (4) and (5) for a point \mathbf{x} close to the focus by

$$G(\mathbf{p}, \mathbf{x}) \approx (1/4\pi R)\exp[ik(R - \hat{\mathbf{q}} \cdot \mathbf{x})], \quad (7)$$

$$\nabla G(\mathbf{p}, \mathbf{x}) \approx (-ik/4\pi R)\exp[ik(R - \hat{\mathbf{q}} \cdot \mathbf{x})]\hat{\mathbf{n}}, \quad (8)$$

respectively. Here R is the distance, measured along the axis, from the exit pupil to the Gaussian focus and $\hat{\mathbf{q}}$ is a unit vector pointing from the focus in the direction of the point \mathbf{p} on the wave front. The electric field \mathbf{E}_S is tangential to the wave front:

$$\hat{\mathbf{n}} \cdot \mathbf{E}_S = 0. \quad (9)$$

Substituting Eq. (9) together with approximations (7) and

(8) into Eq. (2) yields

$$\mathbf{E}(\mathbf{x}) = \frac{-ik}{2\pi R} \exp(ikR) \int_S \mathbf{E}_S \exp[-ik(\hat{\mathbf{q}} \cdot \mathbf{x})] d\sigma. \quad (10)$$

When no aberrations are present, the surface element $d\sigma$ is spherical:

$$d\sigma = R^2 d\Gamma, \quad (11)$$

where Γ is the solid angle subtended by the aperture at the focal point. Inserting Eq. (11) into Eq. (10) gives

$$\mathbf{E}(\mathbf{x}) = \frac{-ikR}{2\pi} \exp(ikR) \int_{\Gamma} \mathbf{E}_S \exp[-ik(\hat{\mathbf{q}} \cdot \mathbf{x})] d\Gamma. \quad (12)$$

In the same manner, applying approximations (7) and (8) in Eq. (3) for the magnetic field, while using that

$$\hat{\mathbf{n}} \cdot \mathbf{B}_S = 0, \quad (13)$$

we obtain

$$\mathbf{B}(\mathbf{x}) = \frac{-ikR}{2\pi} \exp(ikR) \int_{\Gamma} \mathbf{B}_S \exp[-ik(\hat{\mathbf{q}} \cdot \mathbf{x})] d\Gamma. \quad (14)$$

Equations (12) and (14) are vector generalizations of the Debye integral.¹² They were first derived in a different manner by Wolf [Ref. 3, Eqs. (3.3) and (3.5)] (note the use of a differently oriented system of coordinates). From Eqs. (12) and (14) one sees that approximations (7) and (8) entail the assumption that the diffracted fields are a superposition of plane waves with amplitudes \mathbf{E}_S and \mathbf{B}_S whose propagation vectors all lie within the geometrical light cone. In this approximation the contribution to the angular spectrum that is due to diffraction at the aperture's edge is neglected. In our approach—even for an aberration-free system—the presence of wave vectors outside Γ can be readily established if we consider Eq. (4) for a point \mathbf{x} on the axis that lies closer to the lens than the focus. However, we find numerically for uniform beams diffracted by an aberration-free lens with a Fresnel number greater than 100 that our theory agrees well with that of Richards and Wolf. The Fresnel number N is defined as

$$N = a^2/\lambda R, \quad (15)$$

where a is the radius of the lens and λ is the wavelength in vacuum.

The Debye integral predicts an intensity distribution that is symmetric around the focal plane,^{6,14} whereas in the Kirchhoff approximation this is not always the case.¹⁵ Furthermore, these two Debye diffraction integrals describe what is called the phase anomaly at focus. Let the incident field be linearly polarized along the x axis, and let Φ be the phase of the diffracted field. One then has⁴

$$\Phi(x, y, -z) = -\Phi(x, y, z) + \pi \pmod{2\pi}. \quad (16)$$

In words this means that the phase of the field makes a sudden jump at the focus. We will derive a more general result in Section 8.

Wolf and Li¹⁶ showed that a representation as in Eqs. (12) and (14) is valid only if the Fresnel number of the system is much greater than unity. For systems with a low Fresnel number the intensity symmetry around the focal plane vanishes. As is well known, for such systems the maximum intensity occurs no longer at the geometrical focus but at a point that lies closer to the diffracting

lens. This so-called focal shift was explained by, among others, Li and Wolf,^{5,17-19} using a paraxial scalar theory based on the Huygens-Fresnel principle. The focal shift is the subject of Section 4.

4. AXIAL FIELDS FOR LOW FRESNEL NUMBERS

We shall now derive expressions for the diffracted electromagnetic field for points lying on the axis of an ideal system. In this section we study low-Fresnel-number systems. It is assumed that the radius of the exit pupil is small compared with the focal length R . So for ρ , the lateral distance from a point on the wave front to the axis, we have

$$\rho^2/R^2 \ll 1. \quad (17)$$

Let p_z be the z coordinate of a point \mathbf{p} on the spherical wave front that lies at a distance ρ from the axis (Fig. 1 above). We have

$$p_z = (R^2 - \rho^2)^{1/2}. \quad (18)$$

The above assumption then yields

$$p_z \approx R - \rho^2/2R. \quad (19)$$

We can now write the distance $|\mathbf{p} - \mathbf{x}|$ for an axial point \mathbf{x} with coordinates $(0, 0, z)$ in the paraxial approximation as

$$|\mathbf{p} - \mathbf{x}|^2 = \rho^2 + (p_z - z)^2 \approx (R - z)^2 + z\rho^2/R. \quad (20)$$

Neglecting a term $\rho^4/4R^2$ that we assume to be small compared with $z\rho^2/R$ gives

$$s(\theta) \equiv |\mathbf{p} - \mathbf{x}| \approx [(R - z)^2 + zR \sin^2 \theta]^{1/2}. \quad (21)$$

In order to solve Eq. (2) for the diffracted \mathbf{E} field, we first focus on its two scalar products. Consider Eq. (6). Let $(R \sin \theta \cos \phi, R \sin \theta \sin \phi, R \cos \theta)$ be the Cartesian coordinates of the point \mathbf{p} . And let \mathbf{x} , a point on the axis, again be given by $(0, 0, z)$. Then we have

$$\hat{\mathbf{e}}_G(\theta, \phi, z) = \frac{-R}{s(\theta)} \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta - z/R \end{pmatrix}. \quad (22)$$

For the inward unit normal we have of course

$$\hat{\mathbf{n}}(\theta, \phi) = \begin{pmatrix} -\sin \theta \cos \phi \\ -\sin \theta \sin \phi \\ -\cos \theta \end{pmatrix}. \quad (23)$$

This yields the first scalar product:

$$\hat{\mathbf{e}}_G \cdot \hat{\mathbf{n}} = \frac{R}{s(\theta)} \left(1 - \frac{z}{R} \cos \theta \right). \quad (24)$$

In a previous paper¹ we showed that $\mathbf{E}_{S,a}(\theta, \phi)$, the electric field on the wave front, can be written as

$$\mathbf{E}_{S,a}(\theta, \phi) = (\cos \theta)^{1/2} \left[(\cos \alpha) \begin{pmatrix} \sin^2 \phi + \cos \theta \cos^2 \phi \\ (\cos \phi \sin \phi)(\cos \theta - 1) \\ -\sin \theta \cos \phi \end{pmatrix} + (\sin \alpha) \begin{pmatrix} (\cos \phi \sin \phi)(\cos \theta - 1) \\ \cos^2 \phi + \cos \theta \sin^2 \phi \\ -\sin \theta \sin \phi \end{pmatrix} \right], \quad (25)$$

where α is the polarization angle (see Fig. 1). A little algebra then yields the second dot product:

$$\mathbf{E}_s \cdot \hat{\mathbf{e}}_G = \frac{(-z \sin \theta)(\cos \theta)^{1/2}}{s(\theta)} (\cos \alpha \cos \phi + \sin \alpha \sin \phi). \quad (26)$$

Because of rotational symmetry we may, without loss of generality, assume the incident electric field to be linearly polarized along the x axis. (A method for dealing with unpolarized light is described in our previous paper.¹) Our result will also apply to unpolarized light as long as we confine ourselves to the axis of symmetry. So in Eq. (25) we have $\cos \alpha = 1$ and $\sin \alpha = 0$. Next substitute Eqs. (23)–(26) into Eq. (2) for the electric field. Carrying out the integration over ϕ yields that both the y and z components of the diffracted axial electric field vanish:

$$\mathbf{E}_y(0, 0, z) = 0, \quad (27)$$

$$\mathbf{E}_z(0, 0, z) = 0. \quad (28)$$

So only the x component of the diffracted electric vector, which we will shortly calculate, is nonzero. Hence the electric-field vector at each point on the axis of revolution in the image space is linearly polarized, and its direction is the same as the direction of the electric vector in the object space.

For the x component the result of the integration over ϕ is

$$\begin{aligned} \mathbf{E}_x(0, 0, z) = & \int_0^\Omega \left(G(\theta, z)(\cos \theta)^{1/2}(\pi + r \cos \theta) \right. \\ & \times \left\{ -ik + \left[\frac{1}{s(\theta)} - ik \right] \frac{R}{s(\theta)} \left(1 - \frac{z}{R} \cos \theta \right) \right\} \\ & + \pi(\sin^2 \theta) G(\theta, z)(\cos \theta)^{1/2} \left[\frac{1}{s(\theta)} - ik \right] \frac{-z}{s(\theta)} \\ & \times R^2 \sin \theta d\theta, \end{aligned} \quad (29)$$

where Ω denotes the semiaperture angle. Rearranging terms gives

$$\begin{aligned} \mathbf{E}_x(0, 0, z) = & \int_0^\Omega \frac{\exp[iks(\theta)]}{4s(\theta)} (\cos \theta)^{1/2} R^2 (\sin \theta) (1 + \cos \theta) \\ & \times \left\{ -ik + \left[\frac{1}{s(\theta)} - ik \right] \left[\frac{R - z}{s(\theta)} \right] \right\} d\theta. \end{aligned} \quad (30)$$

Next make s the new variable of integration. From Eq. (21) we have

$$d\theta = \frac{s ds}{Rz \sin \theta \cos \theta}. \quad (31)$$

This leads to

$$\begin{aligned} \mathbf{E}_x(0, 0, z) = & \frac{R}{4z} \int_{s(0)}^{s(\Omega)} (\cos^{-1/2} \theta + \cos^{1/2} \theta) \exp(iks) \\ & \times \left(-ik + \frac{R - z}{s^2} - ik \frac{R - z}{s} \right) ds. \end{aligned} \quad (32)$$

In this section we are dealing with small-aperture angles [see relation (17)]. For $\Omega < 5^\circ$ the cosine factor can be

approximated to high accuracy by

$$\cos^{-1/2} \theta + \cos^{1/2} \theta \approx 2. \quad (33)$$

Integration of the first term inside the large parentheses in Eq. (32) is now trivial and yields

$$(R/2z) \{ \exp[iks(0)] - \exp[iks(\Omega)] \}. \quad (34)$$

The second term inside the large parentheses can be integrated by parts:

$$\begin{aligned} \int_{s(0)}^{s(\Omega)} \frac{1}{s^2} \exp(iks) ds = & \left[\frac{-1}{s} \exp(iks) \right]_{s(0)}^{s(\Omega)} \\ & + \int_{s(0)}^{s(\Omega)} \frac{ik}{s} \exp(iks) ds. \end{aligned} \quad (35)$$

The last term of Eq. (35) cancels with the third term inside the large parentheses in Eq. (32). Inserting these last two results gives

$$\begin{aligned} \mathbf{E}_x(0, 0, z) = & \frac{R}{2z} \left\{ \exp[iks(0)] \left[\frac{R - z}{s(0)} + 1 \right] \right. \\ & \left. - \exp[iks(\Omega)] \left[\frac{R - z}{s(\Omega)} + 1 \right] \right\}. \end{aligned} \quad (36)$$

This expression can be somewhat simplified with one further approximation. According to Eq. (21) we can, except in the exponent, substitute for low Fresnel numbers:

$$s(\Omega) \approx s(0) = R - z. \quad (37)$$

Inserting this approximation into Eq. (36) finally yields

$$\mathbf{E}_x(0, 0, z) = (R/z) \{ \exp[ik(R - z)] - \exp[iks(\Omega)] \}. \quad (38)$$

Any physical intensity measurement ultimately depends on an atomic transition. The probability that such a transition occurs is, according to quantum mechanics, proportional to $|\mathbf{E}(\mathbf{x}, t)|^2$ (see, e.g., Ref. 20, p. 137). Another way of putting it is that the intensity is proportional to the time-averaged electric energy density $\langle W_E \rangle$, which is defined as (Ref. 6, p. 33)

$$\langle W_E \rangle \propto \mathbf{E} \cdot \mathbf{E}^*. \quad (39)$$

This leads to

$$\langle W_E \rangle_\Omega(0, 0, z) \propto (R/z)^2 (2 - 2 \cos[k(R - z - s(\Omega))]). \quad (40)$$

In order to link this result to those of Li and Wolf,⁵ we return to Eq. (38) and expand the term $s(\Omega)$ with the help of Eq. (21):

$$s(\Omega) = (R - z) \left[1 + \frac{zR \sin^2 \Omega}{2(R - z)^2} - \frac{z^2 R^2 \sin^4 \Omega}{8(R - z)^4} + \dots \right]. \quad (41)$$

The last and higher-order terms of this expansion may be discarded in the exponent if

$$|kz^2 R^2 (\sin^4 \Omega) / 8(R - z)^3| \ll 2\pi. \quad (42)$$

Substituting definition (15) for the Fresnel number N while using the fact that for the lens radius a we have $a =$

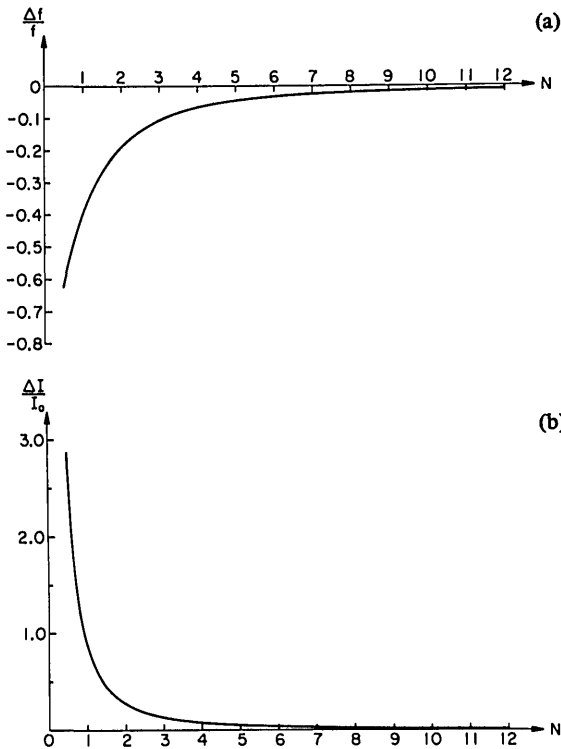


Fig. 2. (a) Relative focal shift $\Delta f/f$ and (b) the corresponding relative excess $\Delta I/I_0 = (I_{\max} - I_0)/I_0$ of the maximum intensity I_{\max} along the axis over the intensity I_0 at the geometrical focus, both for systems of different Fresnel numbers N . These plots are reproduced by permission from Y. Li and E. Wolf.⁵

$R \sin \Omega$ yields

$$z^2 \lambda N^2 / 8(R - z)^3 \ll 1. \quad (43)$$

For Fresnel numbers of order 1 and $|z| < R/2$, this demand is indeed satisfied. In that case we may write Eq. (38) as

$$E_x(0, 0, z) = Q \frac{\exp(-ikz)}{z} \left\{ 1 - \exp \left[\frac{ik}{2} \left(\frac{z}{R} \right) \frac{a^2}{R - z} \right] \right\}, \quad (44)$$

where all the constants are absorbed into the factor Q . Apart from a minus in z , which is due to another definition of coordinates, this is equivalent to Eq. (2.15) of Li and Wolf.⁵ They use the Huygens-Fresnel principle as a starting point for a paraxial scalar theory. In their paper they show that the above expression implies that the maximum intensity occurs no longer at the geometrical focus but closer to the lens. This focal shift can be up to 50% of the geometrical focal length, as we can see from Fig. 2. The values that Eq. (44) predicts for the focal shift were later shown by Li and Platzner²¹ to agree well with experimental values. (Incidentally, in their Table 1 the aperture radii are given in centimeters; this should be millimeters.) For completeness' sake we mention that numerical integration of Eqs. (2) and (44) yields the same focal shift within 1%.

Summarizing, we find that our electromagnetic theory predicts (with the same accuracy) the focal shift phenomenon that occurs in low Fresnel number focusing. So, contrary to the theory of Richards and Wolf, our equations

can also be applied to the study of polarization effects and energy flows in paraxial systems having Fresnel numbers of order 1. In other words, Eqs. (2)–(6) above unify the electromagnetic high-aperture theory of Richards and Wolf and the paraxial scalar theory of Li and Wolf. In the remainder of this paper we will show how our theory can be applied to the study of the effect of aberrations and beam profiles in systems with higher Fresnel numbers.

5. SPHERICAL ABERRATION AND DEFOCUS

In a system with rotational symmetry the form of the third-order aberration function $w(\rho)$ is

$$w(\rho) = C\rho^4 + D\rho^2, \quad (45)$$

where C and D are the coefficients of spherical aberration and defocus, respectively, and ρ denotes a lateral distance in the exit pupil (see Fig. 1) for which we have

$$\rho = R \sin \theta. \quad (46)$$

Let $r(\rho)$ be the distance from the Gaussian focus to the wave front at the exit pupil. Then we have

$$r(\rho) = R + w(\rho) \quad (47)$$

$$= R + D\rho^2 + C\rho^4. \quad (48)$$

Next we derive an explicit expression for S , the deformed wave front at the exit pupil. Neglecting second-order terms in $w(\rho)$ (i.e., terms of order λ^2) gives the following for a point (x, y, z) on the wave front:

$$x^2 + y^2 + z^2 = R^2 + 2RC\rho^4 + 2RD\rho^2. \quad (49)$$

[One second-order term in $w(\rho)$ is $D^2\rho^4$, which, since $w(\rho)$ is of the order of a wavelength, may be neglected compared with $2RC\rho^4$.] Substitution of $\rho^2 = x^2 + y^2$ gives

$$S(x, y, z) = (1 - 2RD)(x^2 + y^2) + z^2 - 2RC(x^4 + 2x^2y^2 + y^4) - R^2 = 0. \quad (50)$$

The inward normal \mathbf{n} to this surface is found by calculation of the gradient:

$$\mathbf{n}(\mathbf{x}) = -\nabla S(x, y, z). \quad (51)$$

So we have

$$\mathbf{n}(\mathbf{x}) = \begin{pmatrix} (8RCy^2 + 4RD - 2)x + 8RCx^3 \\ (8RCx^2 + 4RD - 2)y + 8RCy^3 \\ -2z \end{pmatrix}. \quad (52)$$

Substituting

$$\begin{aligned} x &= r(\theta)(\sin \theta \cos \phi), & y &= r(\theta)(\sin \theta \sin \phi), \\ z &= r(\theta)(\cos \theta) \end{aligned} \quad (53)$$

and normalizing the result gives

$$\hat{\mathbf{n}}(\theta, \phi) = \frac{1}{K} \begin{pmatrix} \beta \sin \theta \cos \phi \\ \beta \sin \theta \sin \phi \\ -2 \cos \theta \end{pmatrix}, \quad (54)$$

where we have defined β as

$$\beta = -2 + 4RD + 8RCr^2(\theta)(\sin^2 \theta) \quad (55)$$

and K as

$$K = (\beta^2 \sin^2 \theta + 4 \cos^2 \theta)^{1/2}. \quad (56)$$

Henceforth we shall omit the circumflex over normalized vectors.

The surface element $d\sigma$ on the now deformed wave front S is easily shown to be

$$d\sigma = \left\{ 1 + \frac{1}{r^2(\theta)} \left[\frac{dr(\theta)}{d\theta} \right]^2 \right\}^{1/2} r^2(\theta)(\sin \theta) d\theta d\phi. \quad (57)$$

This can be calculated with the use of

$$\frac{dr(\theta)}{d\theta} = \frac{dw(\theta)}{d\theta} = \frac{\partial w}{\partial \rho} \frac{\partial \rho}{\partial \theta}, \quad (58)$$

where

$$\frac{\partial \rho}{\partial \theta} = R \cos \theta. \quad (59)$$

6. FIELDS ON THE WAVE FRONT

We now deduce an expression for \mathbf{E}_S , the (time-independent) electric field on the diffracted wave front, in terms of the incoming electric wave amplitude \mathbf{E}_{inc} . We shall assume that the diffracted rays emerge from the exit pupil at the same lateral distance from the axis as that at which they entered the system's entrance pupil. The effect of refraction on the polarization angle will be neglected. From the Fresnel equations it follows that this is justified as long as the incoming wave vector does not make an appreciable angle with the normal of the refracting surfaces. This means that after refraction the electric vector makes the same angle with the meridional plane M as \mathbf{E}_{inc} does. M is defined as the plane containing both the central axis of the system and the incoming wave vector \mathbf{k} . Also, \mathbf{n} , the inward normal to the aberrated wave front, lies in the meridional plane. Furthermore, both the incident and the refracted electric vectors lie upon the same side of M . In Fig. 3 we have depicted the wave vector and the electric vector before and after refraction. Obviously, at refraction the wave vector \mathbf{k} is rotated toward \mathbf{n} . The normal to M is easily seen to be $\mathbf{n} \times \mathbf{k}$. Because of these considerations, the refracted field \mathbf{E}_S can be written as the sum of an unchanged component of \mathbf{E}_{inc} [in the direction $(\mathbf{n} \times \mathbf{k})$] and a rotated component that lies in the plane M . The latter component before refraction has a magnitude equal to $\mathbf{E}_{\text{inc}} \cdot [(\mathbf{n} \times \mathbf{k}) \times \mathbf{k}]$. After refraction this component still has the same length but now points in the direction $(\mathbf{n} \times \mathbf{k}) \times \mathbf{n}$. Summarizing, we have

$$\mathbf{E}_S = (\mathbf{n} \cdot \mathbf{k})^{1/2} \left\{ \frac{\mathbf{E}_{\text{inc}} \cdot (\mathbf{n} \times \mathbf{k})}{|\mathbf{n} \times \mathbf{k}|^2} (\mathbf{n} \times \mathbf{k}) + \frac{\mathbf{E}_{\text{inc}} \cdot [(\mathbf{n} \times \mathbf{k}) \times \mathbf{k}]}{[(\mathbf{n} \times \mathbf{k}) \times \mathbf{k}] \cdot [(\mathbf{n} \times \mathbf{k}) \times \mathbf{n}]} [(\mathbf{n} \times \mathbf{k}) \times \mathbf{n}] \right\}. \quad (60)$$

The first term on the right-hand side is the unchanged component, and the second one is the rotated component

of the incoming field. Because the incoming plane wave is changed by the lens into an (aberrated) spherical wave, the energy flux is smeared out. Conservation of energy then leads to an angular-dependent prefactor $(\mathbf{n} \cdot \mathbf{k})^{1/2}$ (cf. Ref. 22). In Ref. 4 an analogous equation is derived for a perfect lens.

Consider an incoming plane wave traveling along the axis of the optical system. The incoming electric field is assumed to be linearly polarized with polarization angle α (see Fig. 1). (For the case of unpolarized light we again refer to Ref. 1.) So we have

$$\mathbf{E}_{\text{inc}; \alpha} = (\cos \alpha, \sin \alpha, 0), \quad (61)$$

and furthermore

$$\mathbf{k} = (0, 0, -1). \quad (62)$$

The first factor on the right-hand side of Eq. (60) then becomes

$$(\mathbf{n} \cdot \mathbf{k})^{1/2} = [(2 \cos \theta)/K]^{1/2}, \quad (63)$$

where K is defined as in Eq. (56). Working out all the vector products and substituting the above expressions together with Eq. (54) for $\hat{\mathbf{n}}$ into Eq. (60), we finally obtain the following for the electric field on the deformed wave front S :

$$\begin{aligned} \mathbf{E}_{S; \alpha}(\theta, \phi) = & \left(\frac{2 \cos \theta}{K} \right)^{1/2} \\ & \times \left[(\cos \alpha) \begin{pmatrix} \sin^2 \phi + 2K^{-1} \cos \theta \cos^2 \phi \\ (\cos \phi \sin \phi)(2K^{-1} \cos \theta - 1) \\ \beta K^{-1} \sin \theta \cos \phi \end{pmatrix} \right. \\ & \left. + (\sin \alpha) \begin{pmatrix} (\cos \phi \sin \phi)(2K^{-1} \cos \theta - 1) \\ \cos^2 \phi + 2K^{-1} \cos \theta \sin^2 \phi \\ \beta K^{-1} \sin \theta \sin \phi \end{pmatrix} \right]. \end{aligned} \quad (64)$$

Note that for a uniform wave and a lens without spherical aberration or defocus (i.e., $\beta = -2$ and $K = 2$), Eq. (64) reduces to expression (2.23) for the electric field given by Richards and Wolf.⁴ (They take α to be identical to zero.) In other words, for a perfect lens our energy projection over the wave front reduces to the so-called aplanatic projection as described on p. 462 of Ref. 23.

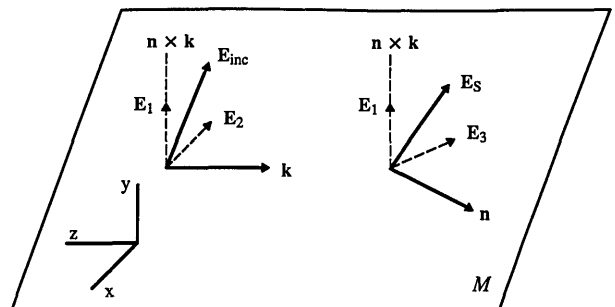


Fig. 3. Meridional plane M (left) before refraction and (right) immediately after refraction. Indicated are the wave vector \mathbf{k} and the electric vector and its components in the plane (\mathbf{E}_2 and \mathbf{E}_3) and perpendicular to it (\mathbf{E}_1). The vector \mathbf{n} coincides with the refracted wave vector. The incoming wave propagates in the minus z direction.

For the magnetic field on the wave front we have

$$\mathbf{B}_S = \mathbf{n} \times \mathbf{E}_S. \quad (65)$$

Using Eqs. (54) and (56) yields

$$\begin{aligned} \mathbf{B}_{S;\alpha}(\theta, \phi) = & \left(\frac{2 \cos \theta}{K} \right)^{1/2} \\ & \times \left[(\cos \alpha) \begin{pmatrix} (1 - 2K^{-1} \cos \theta)(\sin \phi \cos \phi) \\ -\cos^2 \phi - 2K^{-1} \cos \theta \sin^2 \phi \\ -\beta K^{-1} \sin \theta \sin \phi \end{pmatrix} \right. \\ & \left. + (\sin \alpha) \begin{pmatrix} \sin^2 \phi + 2K^{-1} \cos \theta \cos^2 \phi \\ (2K^{-1} \cos \theta - 1)(\sin \phi \cos \phi) \\ \beta K^{-1} \sin \theta \cos \phi \end{pmatrix} \right]. \end{aligned} \quad (66)$$

Now that both fields on the wave front and the form of the aberrated wave front are found, the electric and magnetic fields can be calculated for any arbitrary point in the image space by the use of Eqs. (2)–(6) above. For Gaussian beams and beams with a central obscuration the incoming fields, and hence \mathbf{B}_S and \mathbf{E}_S , depend on the lateral coordinate ρ (see Fig. 1 above). This is the subject of Section 7.

7. BEAM PROFILES AND OBSCURATIONS

Consider an incoming plane wave that is linearly polarized and has a Gaussian profile. We then have the following for \mathbf{E}_{inc} and \mathbf{B}_{inc} , the amplitudes of the incoming electromagnetic field:

$$\begin{aligned} \mathbf{E}_{\text{inc}}(\rho) &= \mathbf{E}_{\text{gau};\alpha}^w \exp(-\rho^2/2w^2), \\ \mathbf{B}_{\text{inc}}(\rho) &= \mathbf{B}_{\text{gau};\alpha}^w \exp(-\rho^2/2w^2), \end{aligned} \quad (67)$$

where w is the $1/e$ intensity radius of the beam and ρ is the lateral distance from the axis. The polarization angle α is the angle between the positive x axis and the incoming electric field (see Fig. 1 above). The subscript gau indicates the Gaussian beam profile. The time-averaged Poynting vector at the entrance pupil, $\langle \mathbf{S} \rangle_{\text{inc}}$, is, up to a constant, defined as (Ref. 6, Section 1.4.3)

$$\langle \mathbf{S} \rangle_{\text{inc}} = \text{Re}(\mathbf{E}_{\text{inc}} \times \mathbf{B}_{\text{inc}}^*). \quad (68)$$

When the incoming field is linearly polarized, the vectors on the right-hand side of Eq. (68) are real, which yields

$$\langle \mathbf{S} \rangle_{\text{inc}}(\rho) = (\mathbf{E}_{\text{gau};\alpha}^w \times \mathbf{B}_{\text{gau};\alpha}^w) \exp(-\rho^2/w^2). \quad (69)$$

In order to make a quantitative comparison between beams with different profiles, we will consider beams with the same total power. First, we briefly calculate the power of a Gaussian beam running parallel to the axis. Let P_{gau} denote this quantity, which equals the integral of the Poynting vector over the lens pupil. Then, for a lens with radius a , we have

$$\begin{aligned} P_{\text{gau}} &= \int_0^a |\mathbf{E}_{\text{inc}} \times \mathbf{B}_{\text{inc}}| 2\pi \rho d\rho \\ &= \pi w^2 \left[1 - \exp\left(-\frac{a^2}{w^2}\right) \right] |\mathbf{E}_{\text{gau};\alpha}^w \times \mathbf{B}_{\text{gau};\alpha}^w|. \end{aligned} \quad (70)$$

For a uniform beam the total power transmitted by the lens, which we call P_{uni} , equals

$$P_{\text{uni}} = \pi a^2 |\mathbf{E}_{\text{uni};\alpha} \times \mathbf{B}_{\text{uni};\alpha}|, \quad (71)$$

where $\mathbf{E}_{\text{uni};\alpha}$ and $\mathbf{B}_{\text{uni};\alpha}$ are both assumed to be normalized to unity. Demanding that the total power of the two beams be equal yields

$$\mathbf{E}_{\text{gau};\alpha}^w = (a/w) [1 - \exp(-a^2/w^2)]^{-1/2} \mathbf{E}_{\text{uni};\alpha}. \quad (72)$$

So, because of linearity, Eq. (64) for the electric field on the wave front must be multiplied by two factors when the incoming beam has a Gaussian profile and the same power as a uniform beam:

$$\begin{aligned} \mathbf{E}_{S;\text{gau};\alpha}(\theta, \phi) &= \mathbf{E}_{S;\text{uni};\alpha}(\theta, \phi) \frac{a}{w} \left[1 - \exp\left(-\frac{a^2}{w^2}\right) \right]^{-1/2} \exp\left(-\frac{\rho^2}{2w^2}\right). \end{aligned} \quad (73)$$

As we mentioned in Section 2, our equations are valid for free space. Hence the magnetic field has the same norm as that of the electric field. So the magnetic field \mathbf{B}_S must then be multiplied by the same two factors.

Next, suppose that we have a uniform beam and a central obscuration in the exit pupil. Let a fraction ϵ of the total lens radius a (which equals $R \sin \Omega$) be obscured. This can be taken into account if we permit the integration in Eqs. (2) and (3) above to extend from $\sin^{-1}(\epsilon \sin \Omega)$ to Ω instead of from zero to Ω . It is straightforward to show that, in order to have the same total energy input as that of an unobscured uniform wave, we must require that the incoming electric field, which we call $\mathbf{E}_{\text{obs};\alpha}^\epsilon$, satisfy

$$\mathbf{E}_{\text{obs};\alpha}^\epsilon = (1 - \epsilon^2)^{-1/2} \mathbf{E}_{\text{uni};\alpha}. \quad (74)$$

So, for the description of centrally obscured beams, two adaptations are necessary in the formalism. First, the lower limit of integration in Eqs. (2) and (3) must be altered. Second, both fields \mathbf{E}_S and \mathbf{B}_S [Eqs. (64) and (66), respectively] must be multiplied by a factor $(1 - \epsilon^2)^{-1/2}$. The above expressions for different beam profiles will be used in Section 10 below.

8. AXIAL FIELDS FOR HIGH ANGULAR APERTURE LENSES

In the same manner as in Section 4 we shall deduce an expression for the axial electric field, but now for a high angular aperture lens with spherical aberration and defocus. This is done starting from Eq. (2) above. We assume the incoming electric field to be linearly polarized along the x axis. First, consider the form that the unit vector $\hat{\mathbf{e}}_G$ of Eq. (6) has in general. A point \mathbf{p} on the now deformed wave front S lies at a distance $r(\theta) = R + w(\theta)$ from the origin (see Fig. 1 above). So, for an axial point \mathbf{x} with coordinates $(0, 0, z)$, we have

$$\begin{aligned} \hat{\mathbf{e}}_G(\theta, \phi, z) &= \frac{-r(\theta)}{[r^2(\theta) + z^2 - 2zr(\theta)(\cos \theta)]^{1/2}} \begin{bmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta - z/r(\theta) \end{bmatrix}. \end{aligned} \quad (75)$$

Next, calculate the term $\hat{\mathbf{e}}_G \cdot \hat{\mathbf{n}}$ of Eq. (2) for the general

aberrated case with the help of Eq. (54):

$$\hat{\mathbf{e}}_G \cdot \hat{\mathbf{n}} = \frac{-r(\theta)}{Kt(\theta)} [\beta \sin^2 \theta - 2 \cos^2 \theta + 2zr^{-1}(\theta)(\cos \theta)], \quad (76)$$

where we have abbreviated the distance $|\mathbf{p} - \mathbf{x}|$ from the denominator of Eq. (75) as $t(\theta)$; that is,

$$t(\theta) = [r^2(\theta) + z^2 - 2zr(\theta)(\cos \theta)]^{1/2}. \quad (77)$$

Since there is no ϕ dependence either in this product or in the Green's function G , multiplication by \mathbf{E}_S [see Eq. (64) above] and integration over ϕ yields zero for the y and z components. Subsequently we calculate the term $\mathbf{E}_S \cdot \hat{\mathbf{e}}_G$ and find that

$$\begin{aligned} \mathbf{E}_S \cdot \hat{\mathbf{e}}_G &= \frac{-r(\theta)(2 \cos \theta)^{1/2}}{K\{K[r^2(\theta) + z^2 - 2zr(\theta)(\cos \theta)]\}^{1/2}} \\ &\times \left[\sin \theta \cos \theta(2 + \beta) - \frac{z}{r(\theta)} \beta \sin \theta \right] (\cos \phi). \end{aligned} \quad (78)$$

So the ϕ dependence of this product is simply through a cosine. Returning to Eq. (2) above, we see that this product must be multiplied by $\hat{\mathbf{n}}$ [see Eq. (54)] and integrated with respect to ϕ from 0 to π . The results for the y and z components clearly vanish. This means that the total expressions for the y and z components of the axial electric field are both zero:

$$\mathbf{E}_y(0, 0, z) = 0, \quad (79)$$

$$\mathbf{E}_z(0, 0, z) = 0. \quad (80)$$

So, just as for the ideal paraxial system, we find for a high-angular-aperture system suffering from spherical aberration and defocus that the y and z components of the diffracted axial electric field vanish. Since we had just assumed that the incoming beam was linearly polarized along the x axis, this means the following for the state of polarization of such systems: the electric-field vector at each point on the z axis in the image space is linearly polarized, and its direction is the same as the direction of the electric vector in the object space.

In order to calculate $\mathbf{E}_x(0, 0, z)$, we first make some approximations. If the wave aberration does not exceed a few wavelengths, the surface element of Eq. (57) above can be taken equal to the spherical surface element:

$$d\sigma \approx R^2 \sin \theta d\theta d\phi. \quad (81)$$

Under the same circumstances it can be seen from Eq. (56) that K differs from 2 only by a term of order $\mathcal{O}(\lambda/R)$. So we can make the approximation

$$K \approx 2. \quad (82)$$

Also, from Eq. (55) it follows that β is then almost equal to -2 . Thus

$$\beta \approx -2. \quad (83)$$

Furthermore, if we restrict the analysis to the focal region, i.e., to points z for which $z \ll R$, as we will do henceforth, we can write

$$r(\theta) \approx t(\theta) \approx R. \quad (84)$$

Applying these approximations, we obtain

$$\hat{\mathbf{e}}_G \cdot \hat{\mathbf{n}} = 1 - (z/R)(\cos \theta), \quad (85)$$

$$\mathbf{E}_S \cdot \hat{\mathbf{e}}_G = (-z/R) \cos^{1/2} \theta \sin \theta \cos \phi. \quad (86)$$

Inserting all this into Eq. (2) for the x component (except in the exponent) and integrating over ϕ lead for a uniform wave to

$$\begin{aligned} \mathbf{E}_x(0, 0, z) &= \Delta(R, k) \int_0^\pi \exp[ikt(\theta)] (\sin \theta) \\ &\times (\cos^{1/2} \theta + \cos^{3/2} \theta) d\theta, \end{aligned} \quad (87)$$

where we use the abbreviation

$$\Delta(R, k) \equiv \frac{R}{4} \left[-ik + \left(\frac{1}{R} - ik \right) \left(1 - \frac{z}{R} \right) \right] \quad (88)$$

$$\approx \frac{1}{4} - \frac{1}{2} ikR. \quad (89)$$

The last step is justified because in the focal region, to which we confined ourselves [see above Eq. (84)], we can assume that $z/R \ll 1$. In that case the factor Δ of Eq. (88) loses its z dependence. The variable $t(\theta)$ appearing in the integral in Eq. (87) is given in terms of the aberration coefficients C and D by the use of Eqs. (48) and (77) above. We emphasize that this expression for the electric field is generally valid (for axial points with coordinate $z \ll R$), both for low- and high-aperture lenses, with or without spherical aberration and defocus. By using relation (39), which says that the intensity is proportional to the square of the norm of the electric field, we now also have an expression for the axial intensity. It should be noted that the axial intensity is the same for both polarized and unpolarized incoming light because of rotational symmetry.

In a completely analogous way one can deduce that

$$\begin{aligned} \mathbf{B}_x(0, 0, z) &= 0, \quad \mathbf{B}_y(0, 0, z) = -\mathbf{E}_x(0, 0, z), \\ \mathbf{B}_z(0, 0, z) &= 0. \end{aligned} \quad (90)$$

Of course, one can also obtain this result by applying Eq. (65) above to the axial fields. Equations (90) have an important implication: for a high-aperture lens affected by spherical aberration and/or defocus, the time-averaged electric energy density [relation (39)] and the norm of the Poynting vector [Eq. (68)] have the same form on the axis. This was already shown to be true for a perfect lens by Richards and Wolf.⁴

It should be remarked that Eq. (87) has an interesting symmetry property. That is, by reversing the sign of the aberration function $w(\theta)$ one obtains an axial intensity distribution that is the mirror image in the $z = 0$ plane of the original distribution. This can easily be seen as follows. The variables z and $w(\theta)$ appear only in the exponent of the integrand in Eq. (87). Assume that both x and $w(\theta)$ are of the order of the wavelength λ . We can then expand the exponent as

$$\begin{aligned} &\exp[ikt(\theta)] \\ &\approx \exp \left\{ i \frac{2\pi}{\lambda} (R^2 + z^2)^{1/2} \left[1 + \frac{Rw(\theta) - zR \cos \theta}{R^2 + z^2} \right] \right\}, \end{aligned} \quad (91)$$

where terms of order $\exp[i\mathcal{O}(\lambda/R)]$ and higher are approximated as unity. The first term in the square brackets times the prefactor has no θ dependence and can be set in front of the integration sign. Under the transformations

$$z \rightarrow -z, \quad w(\theta) \rightarrow -w(\theta), \quad (92)$$

the first term is invariant and the second term changes sign. Referring back to Eq. (87), one sees that under this transformation the electric-field amplitude undergoes a mere phase shift. Hence

$$\langle W_E \rangle_{w(\theta)}(0, 0, z) = \langle W_E \rangle_{-w(\theta)}(0, 0, -z), \quad (93)$$

which concludes the proof of our assertion. In Sections 9 and 10 we study the integral of Eq. (87) for different cases.

9. IDEAL HIGH ANGULAR APERTURE LENS

Consider Eq. (87) above for a lens without aberration (i.e., $C = 0$ and $D = 0$) that focuses a uniform beam. In that case the approximations listed in relations (81)–(83) are all exact, and according to Eq. (77) we have the following for $t(\theta)$:

$$t(\theta) = (R^2 + z^2 - 2zR \cos \theta)^{1/2}. \quad (94)$$

This can be expanded as

$$t(\theta) \approx (R^2 + z^2)^{1/2} \left(1 - \frac{zR}{R^2 + z^2} \cos \theta \right). \quad (95)$$

This is legitimate if for the next order we have

$$k(R^2 + z^2)^{1/2} \frac{(zR)^2}{2(R^2 + z^2)^2} \cos^2 \theta \ll 2\pi. \quad (96)$$

For the left-hand side we have

$$k \frac{(zR)^2}{2(R^2 + z^2)^{3/2}} \cos^2 \theta \leq \frac{kz^2}{2R}. \quad (97)$$

In order to see for which values of z the expansion is valid, we put

$$|z| \leq y\lambda, \quad (98)$$

where λ is the wavelength. This means that for $\lambda = 5 \times 10^{-7}$ m and $R = 10^{-2}$ m that we must require that

$$y \ll 200. \quad (99)$$

For axial points with coordinate z such that $|z| < 20\lambda$, the error introduced is within 1%. This is a further quantification of the assumption leading to relation (84) above. Now inserting the expansion of $t(\theta)$ into Eq. (87) yields the following integral for the axial electric field:

$$\begin{aligned} \mathbf{E}_x(0, 0, z) \\ = \Gamma(R, z, k) \int_0^\pi \exp(i\gamma \cos \theta) (\sin \theta) (\cos^{1/2} \theta + \cos^{3/2} \theta) d\theta, \end{aligned} \quad (100)$$

where we have defined

$$\gamma = -kzR/(R^2 + z^2)^{1/2}, \quad (101)$$

$$\Gamma(R, z, k) = \Delta(R, k) \exp[ik(R^2 + z^2)^{1/2}]. \quad (102)$$

Changing the variable of integration to $\xi = \cos^{1/2} \theta$ yields

$$\mathbf{E}_x(0, 0, z) = -2\Gamma(R, z, k) \int_1^{\cos^{1/2} \Omega} \exp(i\gamma \xi^2) (\xi^2 + \xi^4) d\xi. \quad (103)$$

In calculating the integral on the right-hand side of this equation, which we call $A(z)$, we assume that γ is positive. For negative values of γ the calculation yields the complex conjugate of the result. (For γ equal to zero the integration is trivial.) Integration by parts of the first term gives

$$\begin{aligned} \int_1^{\cos^{1/2} \Omega} \exp(i\gamma \xi^2) \xi^2 d\xi &= \frac{1}{2i\gamma} [\xi \exp(i\gamma \xi^2)]_1^{\cos^{1/2} \Omega} \\ &\quad - \frac{1}{2i\gamma} \int_1^{\cos^{1/2} \Omega} \exp(i\gamma \xi^2) d\xi. \end{aligned} \quad (104)$$

Since γ is positive, the last integral can be written as

$$\begin{aligned} \frac{i}{2\gamma} \left(\frac{\pi}{2\gamma} \right)^{1/2} (C[(\gamma \cos \Omega)^{1/2}] - C[(\gamma)^{1/2}] \\ + i\{S[(\gamma \cos \Omega)^{1/2}] - S[(\gamma)^{1/2}]\}), \end{aligned} \quad (105)$$

where S and C are the Fresnel integrals, which are defined as²⁴ (these functions are called S_1 and C_1 , respectively, in Ref. 24)

$$\begin{aligned} S(x) &= \left(\frac{2}{\pi} \right)^{1/2} \int_0^x \sin(t^2) dt, \\ C(x) &= \left(\frac{2}{\pi} \right)^{1/2} \int_0^x \cos(t^2) dt. \end{aligned} \quad (106)$$

Partial integration of the second term yields

$$\begin{aligned} \int_1^{\cos^{1/2} \Omega} \exp(i\gamma \xi^2) \xi^4 d\xi &= -\frac{i}{2\gamma} [\xi^3 \exp(i\gamma \xi^2)]_1^{\cos^{1/2} \Omega} \\ &\quad + \frac{3i}{2\gamma} \int_1^{\cos^{1/2} \Omega} \exp(i\gamma \xi^2) \xi^2 d\xi. \end{aligned} \quad (107)$$

The last integral is again of the form of Eq. (104). All this gives (for $z < 0$)

$$\begin{aligned} A(z) &= \frac{-i}{2\gamma} \left(\exp(i\gamma \cos \Omega) (\cos^{1/2} \Omega) \left(1 + \frac{3i}{2\gamma} + \cos \Omega \right) \right. \\ &\quad - [\exp(i\gamma)] \left(2 + \frac{3i}{2\gamma} \right) - \left(\frac{\pi}{2\gamma} \right)^{1/2} \left(1 + \frac{3i}{2\gamma} \right) \\ &\quad \times (C[(\gamma \cos \Omega)^{1/2}] - C[(\gamma)^{1/2}] \\ &\quad \left. + i\{S[(\gamma \cos \Omega)^{1/2}] - S[(\gamma)^{1/2}]\} \right). \end{aligned} \quad (108)$$

For the mathematically inclined reader we remark that this result can also be expressed in terms of the error function. So, for the axial electric field of a perfect lens, we finally find that

$$\mathbf{E}_x(0, 0, z) = \begin{cases} -2\Gamma(R, z, k)A(z) & \text{if } z < 0 \\ -2\Gamma(R, z, k)A^*(-z) & \text{if } z > 0 \\ -2\Gamma(R, z, k)[\frac{1}{3}\cos^{3/2}\Omega + \frac{1}{5}\cos^{5/2}\Omega - \frac{8}{15}] & \text{if } z = 0 \end{cases} \quad (109)$$

The expression for $z = 0$ is obtained by direct integration of Eq. (103). This concludes our evaluation of Eq. (87) for an ideal lens focusing a uniform wave.

10. ABERRATIONS AND PROFILES

We do not know an analytical solution of the integral (87) above for the general aberrated case, but it can be easily integrated numerically, since the integrand is not rapidly fluctuating. First we concentrate on the role of defocus [see Eq. (45) above]. Defocus is usually not considered to be a true aberration on the basis of the so-called displacement theorem (Ref. 6, Sec. 9.1.2), which we here repeat:

Displacement Theorem. The addition to an aberration function of a term $D\rho^2$, where D is a constant of order λ , results in no change in the three-dimensional intensity distribution near focus apart from an axial displacement of the distribution as a whole of $2(R/a)^2D$.

This theorem is deduced with the use of a paraxial scalar theory (and the Debye approximation). The assumption of paraxiality implies that its validity for a high-aperture system is questionable. When integrating Eq. (87) for semiapertures up to 10° , we find good agreement with the theorem. The form of the axial intensity distribution for a lens suffering from defocus remains the same as that for a perfect lens. Also, the axial displacement agrees with the predicted value. For a high-aperture lens, however, we find that the theorem no longer holds. Not only is there a shift in the intensity pattern but the peak intensity drops. Moreover, the axial intensity distribution is then no longer symmetric, even for small aberrations (see Fig. 4). For somewhat larger aberrations, the asymmetry gets even worse. In Fig. 4 we have used the dimensionless axial optical coordinate u :

$$u = kz \sin^2 \Omega, \quad (110)$$

where Ω is the semiaperture angle. Below we shall also use the lateral coordinate v , defined as

$$v = k(x^2 + y^2)^{1/2} \sin \Omega. \quad (111)$$

For a lens affected by spherical aberration we find that, as reported by us previously,¹ for a low-aperture system there is good agreement with the scalar classical focusing theory.²³ That is, the axial intensity distribution is symmetric around $u = 4\pi\delta$, where δ is the wave aberration at the pupil's edge measured in wavelengths. For high-aperture lenses, however, we find that the distribution is no longer symmetric (see Fig. 5). A similar asymmetry

has been reported in an experimental paper by Wilson and Carlini.²

Next we will focus on different beam profiles. According to Eq. (87), we have the following for the axial

time-averaged electric energy density of an aberrated uniform wave:

$$\langle W_E \rangle_{\text{uni}}(0, 0, z) = |\Delta(R, k)|^2 \left| \int_0^\Omega \exp[ikt(\theta)] (\sin \theta) (\cos^{1/2} \theta + \cos^{3/2} \theta) d\theta \right|^2, \quad (112)$$

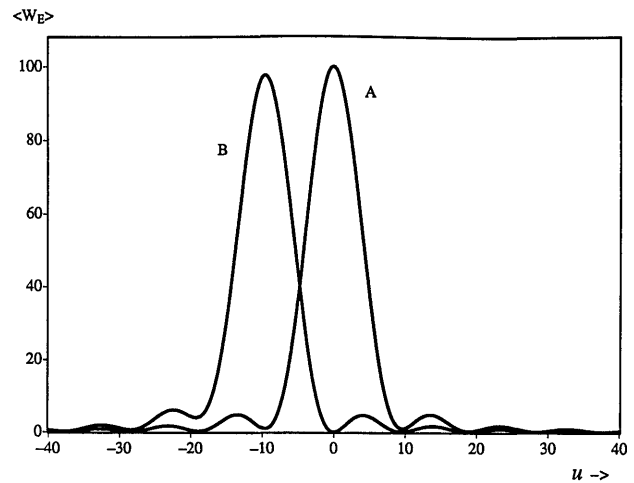


Fig. 4. Axial intensity (i.e., the time-averaged electric energy density) for A, a perfect lens and B, one with 1-wavelength defocus at the aperture's edge. The semiaperture angle is 60° , and the Fresnel number is $N = 1.5 \times 10^4$. Notice that the displacement theorem no longer holds: the peak intensity is now lower, and the distribution is asymmetric.

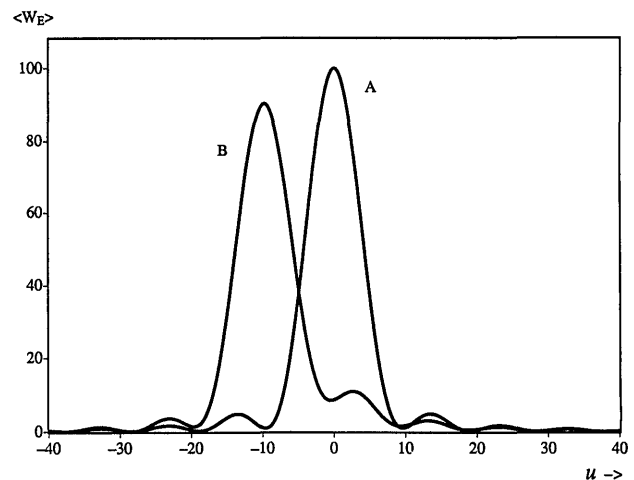


Fig. 5. Axial intensity (i.e., the time-averaged electric energy density) for A, a perfect lens and B, one with 1-wavelength spherical aberration at the aperture's edge. The semiaperture angle is 60° , and the Fresnel number is $N = 1.5 \times 10^4$. Note that the peak has shifted and the distribution is now asymmetric.

where

$$\Delta(R, k) \equiv \frac{1}{4} - \frac{1}{2}ikR. \quad (113)$$

For a Gaussian beam Eq. (112) for the axial intensity must, according to Section 7, be altered into

$$\begin{aligned} \langle W_E \rangle_{\text{gau}}(0, 0, z) &= |\Lambda(R, k, w)|^2 \\ &\times \left| \int_0^\Omega \exp \left[ikt(\theta) - \frac{R^2 \sin^2 \theta}{2w^2} \right] \right. \\ &\times (\sin \theta)(\cos^{1/2} \theta + \cos^{3/2} \theta) d\theta \Big|^2, \end{aligned} \quad (114)$$

where we have used the abbreviation

$$\Lambda(R, k, w) \equiv \Delta(R, k)(a/w)[1 - \exp(-a^2/w^2)]^{-1/2}. \quad (115)$$

For a centrally obscured beam we have [see Eq. (74) above]

$$\begin{aligned} \langle W_E \rangle_{\text{obs}}(0, 0, z) &= |\Psi(R, k, \epsilon)|^2 \\ &\times \left| \int_{\sin^{-1}(\epsilon \sin \Omega)}^\Omega \exp[ikt(\theta)](\sin \theta)(\cos^{1/2} \theta + \cos^{3/2} \theta) d\theta \right|^2, \end{aligned} \quad (116)$$

where

$$\Psi(R, k, \epsilon) \equiv \Delta(R, k)(1 - \epsilon^2)^{-1/2}. \quad (117)$$

The time-averaged electric energy density [relation (39) above], which is the quantity that we take as the intensity, is symmetric around $z = 0$ in all three cases (uniform, Gaussian, and centrally obscured beams). For a uniform beam this can be seen as follows. The intensity is proportional to the square of the norm of the electric field. The symmetry is then evident from Eqs. (109) above:

$$\langle W_E \rangle_{\text{uni}}(0, 0, z) = \langle W_E \rangle_{\text{uni}}(0, 0, -z). \quad (118)$$

Since generalizing to Gaussian and centrally obscured beams introduces no additional z dependence, this result also holds for such beam profiles. Furthermore, we remark that Eqs. (112), (114), and (116) are, because the system is symmetric with respect to rotation, also valid for unpolarized light. Next we compare the axial diffraction patterns for these three types of beams.

In Fig. 6 we have plotted the axial intensity for different beam profiles with the same total power for an ideal low-angular-aperture lens ($\Omega = 10^\circ$, $N = 603$), using Eqs. (112), (114), and (116). The minima of both the uniform and centrally obscured beams are zero, whereas the minima of the Gaussian beam are nonzero. The intensity at the focus of a uniform wave is higher than that of a centrally obscured or Gaussian beam. This was found to be the case for high- as well as low-angular-aperture systems and is in agreement with the paper of Mahajan,⁷ who, using a scalar theory in the Fresnel approximation, studied systems with Fresnel numbers up to 100. (As we remarked in Section 1, Mahajan uses four additional approximations.)

The Gaussian peak decreases with increasing a/w . For the centrally obscured beam it was found that with increasing linear obscuration ratio ϵ the central peak

and the secondary maxima both decreased. The axial full width at half-maximum, however, increases with increasing ϵ .

In Fig. 7 the same situation is depicted for a high-angular-aperture lens ($\Omega = 60^\circ$, $N = 1.5 \times 10^4$). We now find a different behavior: the first minima of all three beam profiles are now nonzero. This is in contrast to the low-aperture situation, in which we, like Mahajan, calculated zeros for both uniform and centrally obscured beams. Note that the first minimum of the Gaussian beam (curve B) is higher in a high-aperture system than in a system with a low aperture. This minimum also increases with increasing a/w and even tends to disappear. For weakly truncated beams, i.e., for $a/w > 2$, the diffraction pattern degenerates into a single peak with points of inflection rather than minima or maxima.

We have also studied the influence of spherical aberration on beams with different profiles. The Strehl ratio,

$\langle W_E \rangle$

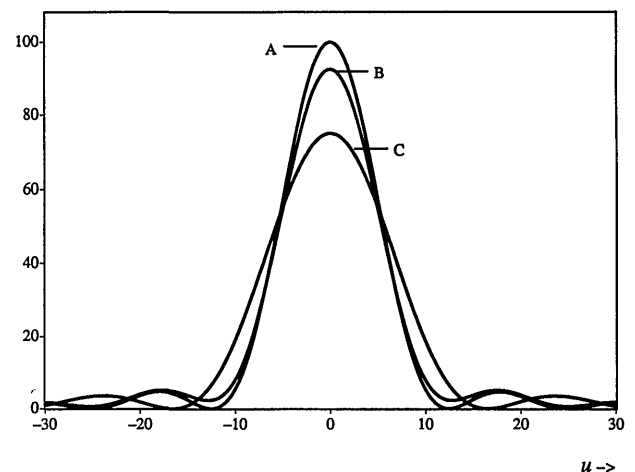


Fig. 6. Axial intensity (i.e., the time-averaged electric energy density) for a perfect lens, with semiaperture angle $\Omega = 10^\circ$ and Fresnel number $N = 603$, focusing beams with different profiles. All the beams have the same total power. Curve A is for a uniform beam, curve B is for a Gaussian beam with $(a/w) = 2.0$, and curve C is for a beam with a central obscuration such that the linear obscuration ratio is $\epsilon = 0.5$.

$\langle W_E \rangle$

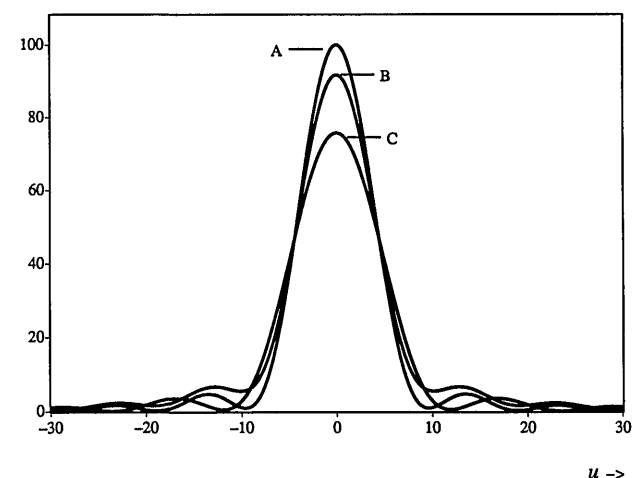


Fig. 7. Same as Fig. 6 but for semiaperture angle $\Omega = 60^\circ$ and Fresnel number $N = 1.5 \times 10^4$.

Table 1. Strehl Ratio for Different Beam Profiles with Spherical Aberration^a

Profile	Spherical Aberration			
	0.25 λ	0.50 λ	1.00 λ	2.00 λ
Uniform	0.994	0.975	0.905	0.664
Obscured, $\epsilon = 0.50$	0.999	0.995	0.980	0.924
Obscured, $\epsilon = 0.95$	0.999	0.999	0.999	0.999
Gaussian, $a/w = 0.03$	0.994	0.975	0.904	0.663
Gaussian, $a/w = 1.41$	0.993	0.974	0.901	0.651

^aStrehl ratio according to electromagnetic diffraction theory for a lens with semiaperture angle $\Omega = 60^\circ$ and Fresnel number $N = 1.5 \times 10^4$ with spherical aberration. The amount of spherical aberration is expressed in wavelengths at the edge of the exit pupil. Note that the centrally obscured beam is significantly less sensitive to spherical aberration than the uniform and Gaussian profiles.

which is the ratio of the peak intensity of the unaberrated beam to the peak intensity of the beam with aberration, is given in Table 1. From the table it can be seen that the uniform beam, the weakly truncated Gaussian beam, and the strongly truncated Gaussian beam show remarkably similar behaviors. Furthermore, it should be noted that the two centrally obscured beams are, in terms of the Strehl ratio, much less sensitive to spherical aberration than are the other two profiles. A possible explanation of this effect is that the phase difference between the rays meeting in the focal region is less than that of the other two profiles.

11. INTENSITY DISTRIBUTION IN THE FOCAL AND MERIDIONAL PLANES

By numerical integration of Eqs. (2) and (64) above, the intensity (by which we mean the time-averaged electric energy density $\langle W_E \rangle$; see Section 4) can be calculated for nonaxial points. This was done with the routine D01DAF of the Numerical Algorithm Group library.²⁵ The basis of this routine is described in two papers by Patterson.²⁶ The contours were obtained with the help of the Spyglass Transform software package.²⁷

Intensity contours in planes perpendicular to the z axis are shown in Figs. 8 and 9. In all our examples the lens has a semiaperture angle $\Omega = 60^\circ$ and a Fresnel number $N = 1.5 \times 10^4$. Also, in all the cases, the incident beam is linearly polarized along the x axis. All the beams have the same total power. In Fig. 8 the intensity contours in the focal plane of an ideal lens are drawn. Such contours were first studied by Boivin and Wolf,²⁸ who applied the aforementioned theory of Richards and Wolf^{3,4} to study an ideal lens with a semiaperture angle $\Omega = 45^\circ$. From Fig. 8 it can be seen that the intensity distribution is mirror symmetric with respect to both the planes $\phi = 0^\circ$ and $\phi = 90^\circ$. Near the focus the contours are approximately elliptical, with their major axis along the polarization direction. These ellipses become more elongated when the semiaperture angle increases (c.f. the figures in Boivin and Wolf²⁸). The minima along the x axis ($\phi = 0^\circ$) are greater than zero, whereas those along the y axis ($\phi = 90^\circ$) are zeros. This is in contrast to the axial intensity distribution, for which no zeros are found for an ideal high-aperture lens. The depicted plane has a width of 3.7λ .

In Fig. 9 the results are shown for the same beam and the same lens, only now with spherical aberration equal to

one wavelength at the edge of the exit pupil. The contours are drawn in the receiving plane perpendicular to the axis at $u = -9.82$. This is the plane in which the maximum intensity occurs. The minima along the x axis are now somewhat higher than in the ideal case, and those along the y axis are slightly greater than zero.

In Fig. 10 contours of the time-averaged electric energy density are shown in a section of the meridional plane with $\phi = 0^\circ$, which is the plane containing the polarization vector. The contours are for an aberration-free lens

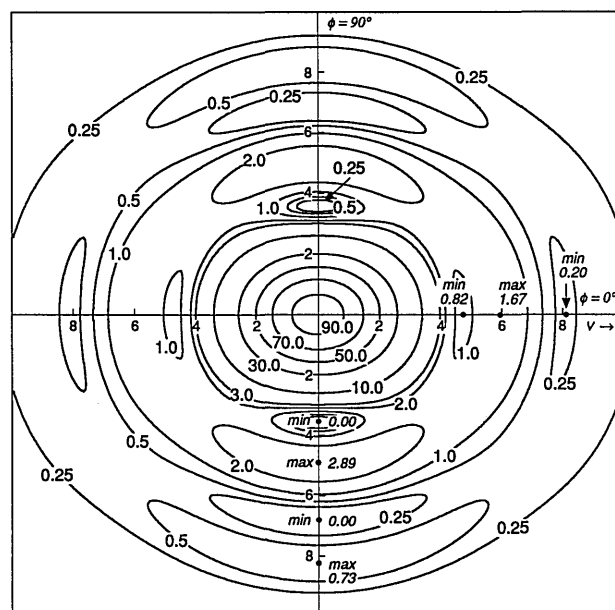


Fig. 8. Contours of the time-averaged electric energy density in the focal plane of an ideal lens with $\Omega = 60^\circ$ and $N = 1.5 \times 10^4$. The incident uniform beam is linearly polarized along the x axis. The intensities are normalized to 100 at focus.

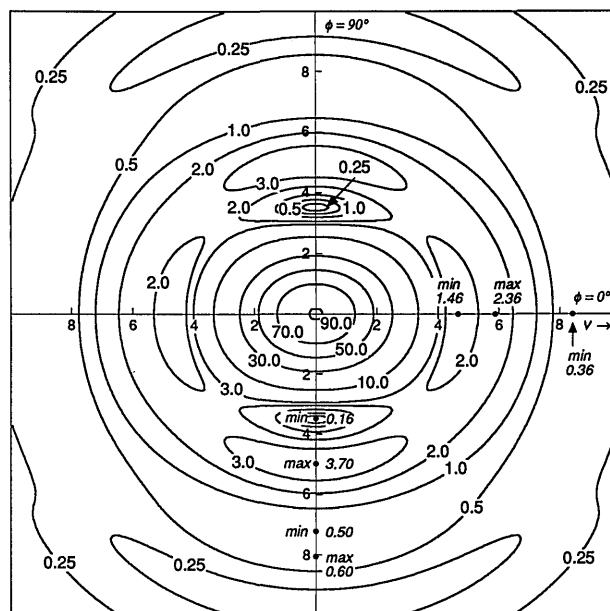


Fig. 9. Contours of the time-averaged electric energy density in the receiving plane $u = -9.82$ perpendicular to the axis for a lens with spherical aberration equal to 1 wavelength at the pupil's edge and $\Omega = 60^\circ$ and $N = 1.5 \times 10^4$. This is the plane in which the maximum intensity is found. The Strehl ratio is 0.905.

for a low-aperture lens, but the peak intensity is lower and not even symmetric.

- For a high-aperture lens with spherical aberration, we also found an asymmetric intensity profile on the axis.

- The peak intensity of a uniform beam is always higher than that of a Gaussian or centrally obscured beam when the three have the same total power.

- The peak intensity of a Gaussian beam decreases with decreasing beam width when the total beam power is kept constant.

- The peak intensity of a centrally obscured beam decreases with increasing obscuration ratio when the total beam power is kept constant.

- The first minima on the axis of a Gaussian beam are, contrary to the case of a uniform beam, not zero for low-aperture angles.

- In the high-angular-aperture case the first Gaussian minima on the axis are higher than those of a uniform beam.

- For weakly truncated Gaussian beams the axial intensity degenerates into a single peak without maxima or minima.

- For high-angular-aperture lenses, Gaussian and uniform beams are equally sensitive to spherical aberration in terms of the Strehl ratio. A centrally obscured beam is much less affected.

Finally, we have seen that the dependency of intensity contours in the focal and meridional planes on beam profiles and lens aberrations is rather large.

ACKNOWLEDGMENTS

We thank H. A. Ferwerda, H. J. Frankena, and F. C. A. Groen for suggesting several improvements in the manuscript. Oof Oud and Norbert Vischer greatly helped us in making the contour plots. We are grateful to Y. Li and E. Wolf for granting us permission to reproduce Fig. 2 from their 1981 paper.

REFERENCES AND NOTES

1. T. D. Visser and S. H. Wiersma, "Spherical aberration and the electromagnetic field in high aperture systems," *J. Opt. Soc. Am. A* **8**, 1404–1410 (1991).
2. T. Wilson and A. R. Carlini, "The effect of aberrations on the axial response of confocal imaging systems," *J. Microsc. (Oxford)* **154**, 243–256 (1989).
3. E. Wolf, "Electromagnetic diffraction in optical systems I," *Proc. R. Soc. London Ser. A* **253**, 349–357 (1959).
4. B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems II," *Proc. R. Soc. London Ser. A* **253**, 358–379 (1959).
5. Y. Li and E. Wolf, "Focal shifts in diffracted converging spherical waves," *Opt. Commun.* **39**, 211–215 (1981).
6. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, Oxford, 1980).
7. V. N. Mahajan, "Uniform versus Gaussian beams: a comparison of the effects of diffraction, obscuration and aberrations," *J. Opt. Soc. Am. A* **3**, 470–485 (1986).
8. M. Mansuripur, "Certain computational effects of vector diffraction problems," *J. Opt. Soc. Am. A* **6**, 768–805 (1989).
9. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, San Francisco, Calif., 1968).
10. C. J. Bouwkamp, "Diffraction theory," *Rep. Prog. Phys.* **17**, 35–100 (1954). The last term of Eq. (5.9) may safely be neglected, as we shall henceforth do.
11. J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1975).
12. P. J. W. Debye, "Das Verhalten von Lichtwellen in der Nähe eines Brennpunktes oder einer Brennnlinie," *Ann. Phys. (Leipzig)* **30**, 755–776 (1909).
13. C. J. R. Sheppard, A. Choudhury, and J. Gannaway, "Electromagnetic field near the focus of wide-angular lens and mirror systems," *Microwaves Opt. Acoust.* **1**, 129–132 (1977).
14. E. Collett and E. Wolf, "Symmetry properties of focused fields," *Opt. Lett.* **5**, 264–266 (1980).
15. J. J. Stamnes, "Focusing at small angular apertures in the Debye and Kirchhoff approximations," *Opt. Commun.* **40**, 81–85 (1981).
16. E. Wolf and Y. Li, "Conditions for the validity of the Debye integral representation of focused fields," *Opt. Commun.* **39**, 205–210 (1981).
17. Y. Li and E. Wolf, "Focal shift in focused truncated Gaussian beams," *Opt. Commun.* **42**, 151–156 (1982).
18. Y. Li, "Encircled energy for systems of different Fresnel numbers," *Optik (Stuttgart)* **64**, 207–218 (1983).
19. Y. Li, "Three-dimensional intensity distribution in low-Fresnel-number focusing systems," *J. Opt. Soc. Am. A* **4**, 1349–1353 (1987).
20. A. Yariv, *Theory and Applications of Quantum Mechanics* (Wiley, New York, 1982).
21. Y. Li and H. Platzer, "An experimental investigation of diffraction patterns in low-Fresnel-number focusing systems," *Opt. Acta* **30**, 1621–1643 (1983).
22. H. H. Hopkins, "The Airy disc formula for systems of high relative aperture," *Proc. Phys. Soc. London* **55**, 116–128 (1943).
23. J. J. Stamnes, *Waves in Focal Regions* (Hilger, Bristol, UK, 1986).
24. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (U.S. Government Printing Office, Washington, D.C., 1965).
25. Numerical Algorithm Group FORTRAN Library Manual, Mark 14 (NAG Ltd., Oxford, UK, 1991).
26. T. N. L. Patterson, "The optimum addition of points to quadrature formulae," *Math. Comput.* **22**, 847–856 (1968); "On some Gauss and Lobatto based integration formulae," *Math. Comput.* **22**, 877–881 (1968).
27. Spyglass Transform, version 2.0, Spyglass Inc., Champaign, Ill.
28. A. Boivin and E. Wolf, "Electromagnetic field in the neighborhood of the focus of a coherent beam," *Phys. Rev.* **138**, B1561–B1565 (1965).

Diffraction of converging electromagnetic waves: erratum

Taco D. Visser and Sjoerd H. Wiersma

*Department of Molecular Cytology, University of Amsterdam, Plantage Muidergrecht 14,
1018 TV Amsterdam, The Netherlands*

In our recent paper¹ Eq. (29) should read as follows:

$$\begin{aligned} \mathbf{E}_x(0, 0, z) = & \int_0^\Omega \left(G(\theta, z)(\cos \theta)^{1/2}(\pi + \pi \cos \theta) \right. \\ & \times \left\{ -ik + \left[\frac{1}{s(\theta)} - ik \right] \frac{R}{s(\theta)} \left(1 - \frac{z}{R} \cos \theta \right) \right\} \\ & + \pi(\sin^2 \theta)G(\theta, z)(\cos \theta)^{1/2} \left[\frac{1}{s(\theta)} - ik \right] \frac{-z}{s(\theta)} \Bigg) \\ & \times R^2 \sin \theta d\theta. \end{aligned} \quad (29)$$

REFERENCE

1. T. D. Visser and S. H. Wiersma, "Diffraction of converging electromagnetic waves," *J. Opt. Soc. Am. A* **9**, 2034–2047 (1992).