Spherical aberration and the electromagnetic field in high-aperture systems

Taco D. Visser and Sjoerd H. Wiersma

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

Received February 5, 1990; revised manuscript received April 4, 1991; accepted April 8, 1991

We present a model for investigating the effect of spherical aberration on the electromagnetic field and the Poynting vector in the focal region of a high-aperture lens. The fields are obtained by integrating the vector equivalent of Kirchhoff's boundary integral over the aberrated wave front. We have studied both diffraction patterns and transfer functions. Our results differ significantly from those obtained by classical focusing theory. For example, the intensity peak is narrower. Also the intensiy distribution is no longer symmetric on the optical axis. A similar asymmetry has recently been measured.

1. INTRODUCTION

In two classical articles Wolf[1] and Richards and Wolf[2] used the Debye approximation in a vectorial diffraction theory to give a complete description of the electromagnetic field of an ideal high-aperture system. It is our aim in this paper to apply electromagnetic diffraction theory to high-aperture lenses with spherical aberration. In recent years the development of confocal microscopes has given rise to a number of papers dedicated to the study of high-aperture systems with aberrations.[3-6] All of these use either Fourier diffraction or classical focusing theory. A limitation of scalar theories is that they cannot accommodate polarization. Furthermore it has been pointed out by Wolf[1] that, especially in high-aperture systems, a scalar approach to electromagnetic phenomena is highly suspicious. The larger the aperture, the more the vector character of the electromagnetic field comes into play, since then the refracted rays make an appreciable angle with one another. For an aperture smaller than about 10°, the results of the electromagnetic and the classical focusing theory should more or less coincide.[2] Today's confocal microscopes, however, typically have an aperture of 120°. So, contrary to practice, an electromagnetic diffraction theory, as described by Born and Wolf[7] and Stamnes (Ref. 8, Sect. 15.4.1), should be used in this case when one wants to describe spherical aberration. In a different context, Ling and Lee[9] have studied the focusing of electromagnetic waves through a dielectric interface. (For an alternative account of their results, see Ref. 8.) When a converging spherical wave enters a medium with a different permittivity, its form will be altered, giving rise to spherical aberration. Ling and Lee, however, did not concentrate on the role of this aberration.

One possible approach to dealing with a composed optical system is to combine ray-tracing methods together with a diffraction theory (Ref. 8, Chap. 3).

Here we will use the electromagnetic Kirchhoff integral to investigate the effects of spherical aberration in a high-aperture lens. In a way this paper can be seen as an extension of the electromagnetic Kirchhoff theory by now including spherical aberration.

This paper is organized as follows: in Section 2 we present the vector-integral representations of the electric and the magnetic fields by using the vector Kirchhoff boundary integral. Section 3 is dedicated to the introduction of spherical aberration and its effect on the mathematical expression for the aberrated wave front. In Section 4 the electric and magnetic fields on the wave front in the exit pupil of the system are calculated for an incoming plane monochromatic wave with arbitrary polarization angle. In Section 5 we define the quantities of interest to us: the energy density and the Poynting vector that we shall both study in the vicinity of focus. Furthermore we make some remarks on computational aspects. In Section 6 a comparison is made between the vector theory and the classical focusing theory. Section 7 deals with diffraction patterns; our results are presented in figures and in a table. In Section 8 the influence of spherical aberration on the optical transfer function (OTF) is studied, and the results are compared with Fourier theory. Finally, our results are summarized in Section 9.

2. VECTOR-INTEGRAL REPRESENTATION OF THE FIELDS

The coordinate system that we use is depicted in Fig. 1. A polarized plane monochromatic wave with angular frequency ω comes in from the left, running parallel to the optical axis. The polarization angle α is defined as the angle between the incoming electric field $\mathbf{E}_{\text{inc}}$ and the positive x axis. Ideally the wave front after refraction, or rather that portion of it that approximately fills the exit pupil, coincides with a reference sphere. In the case of an aberration-free lens, the reference sphere is a sphere with the focal point as its center and with radius equal to R, the distance along the optical axis from the exit pupil to the focus. In the presence of aberrations, however, the actual wave front at the exit pupil, called S, will deviate from the reference sphere.

As we will now show, the electric and magnetic fields in the focal region can be expressed in terms of the wave

0740-3232/91/091404-07$05.00

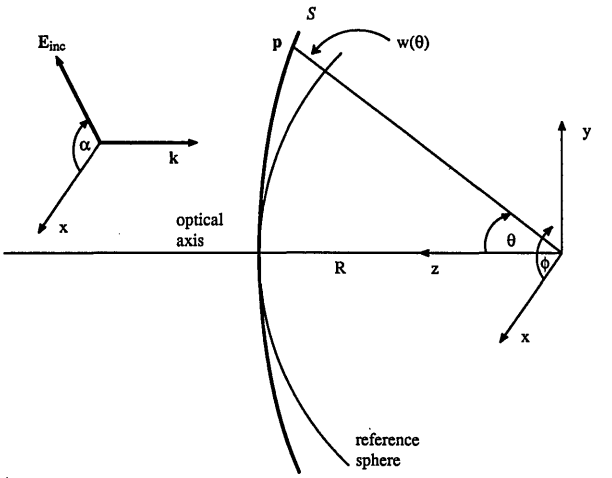

Fig. 1. Definition of the coordinate system. On the left are the wave vector $\mathbf{k}$ and the electric vector $\mathbf{E}_{\text{inc}}$, both before refraction. The angle between the x axis and $\mathbf{E}_{\text{inc}}$ is the polarization angle α. The origin is placed at focus, which is the center 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 is defined as usual as the angle between the positive x axis and the projection of $\mathbf{p}$ onto the xy plane. The z axis coincides with the optical axis.

number k, the inward normal $\mathbf{n}$ of the aberrated wave front S, and the fields on S. For the electric field $\tilde{\mathbf{E}}(\mathbf{x},t)$ and the magnetic field $\tilde{\mathbf{B}}(\mathbf{x},t)$, we have

$$\tilde{\mathbf{E}}(\mathbf{x},t) = \text{Re}[\mathbf{E}(\mathbf{x})\exp(-i\omega t)],$$
$$\tilde{\mathbf{B}}(\mathbf{x},t) = \text{Re}[\mathbf{B}(\mathbf{x})\exp(-i\omega t)], \tag{1}$$

where $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ are the time-independent parts of the fields. The idea is to obtain the fields near focus by evaluation of the vector equivalent of Kirchhoff's generalized boundary integral,[10] with the integral over the now distorted wave front S. For the electric field $\mathbf{E}(\mathbf{x})$ after refraction, we have

$$\mathbf{E}(\mathbf{x}) = \int_S [(\mathbf{E}_S \cdot \hat{\mathbf{n}})\nabla G + (\hat{\mathbf{n}} \times \mathbf{E}_S) \times \nabla G + ik(\hat{\mathbf{n}} \times \mathbf{B}_S)G]\mathrm{d}\sigma, \tag{2}$$

where the integral is over S, the aberrated wave front. $\mathbf{E}_S$ and $\mathbf{B}_S$ are the electric and magnetic fields on S. G is the Green's function associated with the Helmholtz equation. For G, we have

$$G(\mathbf{p},\mathbf{x}) = \frac{\exp(ik|\mathbf{p}-\mathbf{x}|)}{4\pi|\mathbf{p}-\mathbf{x}|}, \qquad \nabla G = G\left(\frac{1}{|\mathbf{p}-\mathbf{x}|} - ik\right)\hat{e}_G. \tag{3}$$

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

$$\hat{e}_G = \frac{\mathbf{x}-\mathbf{p}}{|\mathbf{x}-\mathbf{p}|}. \tag{4}$$

At the exit pupil, the wave vector, which we call $\hat{\mathbf{k}}_S$, is orthogonal to S and coincides with $\hat{\mathbf{n}}$. Therefore the first term of the integrand is zero, since $\mathbf{E}_S$ and $\hat{\mathbf{n}}$ are perpendicular on S. Using

$$\hat{\mathbf{n}} \times \mathbf{B}_S = -\sqrt{\mu\varepsilon}\mathbf{E}_S, \tag{5}$$

we get

$$\mathbf{E}(\mathbf{x}) = \int_S \mathrm{d}\sigma[(\hat{\mathbf{n}} \times \mathbf{E}_S) \times \nabla G - ikG\sqrt{\mu\varepsilon}\mathbf{E}_S]. \tag{6}$$

Expanding the triple cross product gives

$$\mathbf{E}(\mathbf{x}) = \int_S \mathrm{d}\sigma[(-ikG\sqrt{\mu\varepsilon} + \nabla G \cdot \hat{\mathbf{n}})\mathbf{E}_S - \hat{\mathbf{n}}(\nabla G \cdot \mathbf{E}_S)]. \tag{7}$$

The Kirchhoff vector integral for the magnetic field is[10]

$$\mathbf{B}(\mathbf{x}) = \int_S \mathrm{d}\sigma[-ik(\hat{\mathbf{n}} \times \mathbf{E}_S)G + (\hat{\mathbf{n}} \times \mathbf{B}_S) \times \nabla G + (\hat{\mathbf{n}} \cdot \mathbf{B}_S)\nabla G]. \tag{8}$$

Using the facts that, at S, we have

$$\hat{\mathbf{n}} \cdot \mathbf{B}_S = 0 \tag{9}$$

and that

$$\hat{\mathbf{n}} \times \mathbf{E}_S = (1/\sqrt{\mu\varepsilon})\mathbf{B}_S, \tag{10}$$

we eventually get

$$\mathbf{B}(\mathbf{x}) = \int_S \mathrm{d}\sigma\left[\left(\frac{-ik}{\sqrt{\mu\varepsilon}}G + \nabla G \cdot \hat{\mathbf{n}}\right)\mathbf{B}_S - (\nabla G \cdot \mathbf{B}_S)\hat{\mathbf{n}}\right]. \tag{11}$$

Equations (7) and (11) will serve as a basis for further calculations. In Sections 3 and 4 we will derive expressions for all quantities appearing in Eqs. (7) and (11) in polar coordinates for the case of spherical aberration.

3. SPHERICAL ABERRATION AND THE FORM OF THE WAVE FRONT

In a system with rotational symmetry with incident wave fronts orthogonal to the optical axis, the only aberration that can be present is a spherical aberration. (Of course in this case there can also be defocus, but this is usually not considered to be a true aberration.) In this section we will investigate how spherical aberration affects the form of the wave front. The aberration function w is defined as the deviation of the actual wave front from the Gaussian reference sphere. See Fig. 1. In the case of spherical aberration we have

$$w(x,y) = C(x^2 + y^2)^2, \tag{12}$$

where x and y are coordinates in the exit pupil and C is the aberration constant. Equivalently one can write

$$w(\theta) = CR^4 \sin^4\theta, \tag{13}$$

where R is the distance along the optical axis from the exit pupil to the focus. From Fig. 1 it is clear that r, the distance from the origin to the wave front, is given by

$$r = R + w. \tag{14}$$

Define now the wave front S as the collection of points

approximately filling the exit pupil. We want an explicit expression for S. Substituting Eq. (12) into Eq. (14) and using $r^2 = x^2 + y^2 + z^2$ give

$$S(x,y,z) = x^2 + y^2 + z^2 - R^2 - 2RC(x^2 + y^2)^2 - C^2(x^2 + y^2)^4 = 0. \quad (15)$$

A vector normal to S is obtained by taking the gradient of Eq. (15):

$$\mathbf{n}(\mathbf{x}) = -\nabla S, \quad (16)$$

or, explicitly,

$$\mathbf{n}(\mathbf{x}) = -\begin{bmatrix} 2x - 8RC(x^2 + y^2)x - 8C^2(x^2 + y^2)^3 x \\ 2y - 8RC(x^2 + y^2)y - 8C^2(x^2 + y^2)^3 y \\ 2z \end{bmatrix}. \quad (17)$$

One can re-express all points (x, y, z) in terms of θ and ϕ as follows:

$$x = r(\theta)\sin\theta\cos\phi,$$
$$y = r(\theta)\sin\theta\sin\phi,$$
$$z = r(\theta)\cos\theta. \quad (18)$$

We then get, for the unit inward normal vector on S,

$$\hat{\mathbf{n}} = -1/\{[2r(\theta)\sin\theta - \epsilon]^2 + 4r^2(\theta)\cos^2\theta\}^{1/2} \times \begin{bmatrix} [2r(\theta)\sin\theta - \epsilon]\cos\phi \\ [2r(\theta)\sin\theta - \epsilon]\sin\phi \\ 2r(\theta)\cos\theta \end{bmatrix}, \quad (19)$$

where we have defined

$$\epsilon = \epsilon(\theta) \equiv 8CRr^3(\theta)\sin^3\theta + 8C^2r^7(\theta)\sin^7\theta. \quad (20)$$

Note that, for zero aberration, i.e., $C = 0$ and $\epsilon(\theta) = 0$, we regain the normal to a perfect sphere. Now the normal appearing in Eqs. (7) and (11) is expressed in terms of θ and ϕ. From now on, we will omit the circumflex over unit vectors.

By using standard analysis, one can deduce an expression for the infinitesimal surface element $d\sigma$ of the aberrated wave front. It turns out that

$$d\sigma = \left\{1 + \frac{1}{r^2(\theta)}\left[\frac{dr(\theta)}{d\theta}\right]^2\right\}^{1/2} r^2(\theta)\sin\theta d\theta d\phi. \quad (21)$$

For an aberration-free system we have $dr(\theta)/d\theta = 0$, and thus the familiar spherical surface element $d\sigma = r^2 \sin\theta d\theta d\phi$ is recovered.

4. FIELDS ON THE WAVE FRONT

We now deduce an expression for $\mathbf{E}_S$, the (time-independent) electric field on the wave front, just after refraction. The effect of refraction on the polarization angle will be neglected. From the Fresnel equations it follows that this neglect is justified as long as the incoming wave vector does not make an appreciable angle with the normal of the refracting surface. This means that after refraction the electric vector makes the same angle with the meridional plane M as $\mathbf{E}_{\text{inc}}$ does. M is defined as the plane containing both the optical axis and the incoming wave vector $\mathbf{k}$. Furthermore, both the incident and the refracted electric vector lie on the same side of M. In Fig. 2 we have depicted the wave vector and the electric vector before and after refraction. At refraction the wave vector $\mathbf{k}$ is rotated into the direction of $\mathbf{n}$. 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 $\mathbf{E}_{\text{inc}}$ in the direction $\mathbf{n} \times \mathbf{k}$ and a rotated component 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 its length remains unchanged, but it is now pointing in the direction of $(\mathbf{n} \times \mathbf{k}) \times \mathbf{n}$. Summarizing all this, we have

$$\mathbf{E}_S = (\mathbf{n}\cdot\mathbf{k})^{1/2}\left(\frac{[\mathbf{E}_{\text{inc}} \cdot (\mathbf{n} \times \mathbf{k})](\mathbf{n} \times \mathbf{k})}{|\mathbf{n} \times \mathbf{k}|^2} + \frac{\{\mathbf{E}_{\text{inc}} \cdot [(\mathbf{n} \times \mathbf{k}) \times \mathbf{k}]\}(\mathbf{n} \times \mathbf{k}) \times \mathbf{n}}{|(\mathbf{n} \times \mathbf{k}) \times \mathbf{k}|\,|(\mathbf{n} \times \mathbf{k}) \times \mathbf{n}|}\right). \quad (22)$$

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 $(\mathbf{n}\cdot\mathbf{k})^{1/2}$. In Refs. 2 and 11 an analogous equation is derived for a perfect lens.

From the definitions in Section 2 it follows that

$$\mathbf{E}_{\text{inc}} = (\cos\alpha, \sin\alpha, 0) \quad (23)$$

and also that

$$\mathbf{k} = (0, 0, -1). \quad (24)$$

The first factor of Eq. (22) then becomes

$$(\mathbf{n}\cdot\mathbf{k})^{1/2} = \left[\frac{2r(\theta)\cos\theta}{N}\right]^{1/2}, \quad (25)$$

where we have defined

$$N \equiv \{[2r(\theta)\sin\theta - \epsilon]^2 + 4r^2(\theta)\cos^2\theta\}^{1/2}. \quad (26)$$

By substituting Eqs. (23)–(26) and Eq. (19) into Eq. (22)

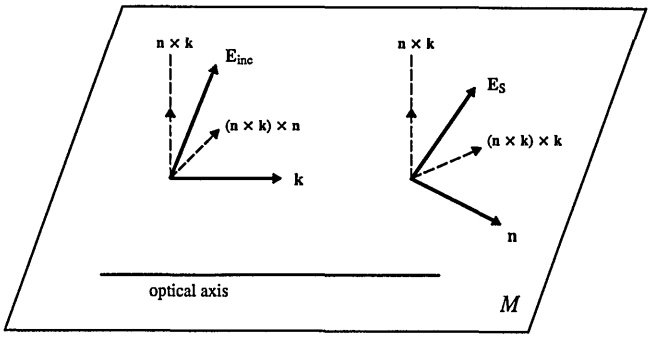

Fig. 2. Meridional plane M with the left-hand side being the situation before, and the right-hand side, immediately after refraction. Indicated are the wave vector, the electric vector, and its components in the plane and perpendicular to it. The vector $\mathbf{n}$ coincides with the refracted wave vector. $\mathbf{E}_S$ is the electric vector at the aberrated wave front. The vectors $(\mathbf{n} \times \mathbf{k}) \times \mathbf{k}$ and $(\mathbf{n} \times \mathbf{k}) \times \mathbf{n}$ both lie in the meridional plane.

and working out the cross products, we get, after a lengthy calculation for the electric field on the wave front S,

$$\mathbf{E}_S(\theta,\phi,\alpha) = [(2r\cos\theta)/N]^{1/2} \times \left\{ \cos\alpha \begin{bmatrix} \sin^2\phi + 2rN^{-1}\cos\theta\cos^2\phi \\ \cos\phi\sin\phi(2rN^{-1}\cos\theta - 1) \\ -N^{-1}(2r\sin\theta - \epsilon)\cos\phi \end{bmatrix} + \sin\alpha \begin{bmatrix} \cos\phi\sin\phi(2rN^{-1}\cos\theta - 1) \\ \cos^2\phi + 2rN^{-1}\cos\theta\sin^2\phi \\ -N^{-1}(2r\sin\theta - \epsilon)\sin\phi \end{bmatrix} \right\}, \tag{27}$$

in which we have dropped the explicit θ dependence of $r(\theta)$. Notice that for zero aberration (i.e., $\epsilon = 0$ and $N = 2r$) Eq. (27) reduces to the expression for the electric field given by Richards and Wolf[2] [Eq. (2.23)]. (In fact they take α to be zero.) In other words, for a perfect lens our energy projection over the wave front reduces to the so-called aplanatic projection as described in Ref. 8 (p. 462).

For the magnetic field we have

$$\mathbf{B}_S = \sqrt{\mu\varepsilon}\,\mathbf{n} \times \mathbf{E}_S. \tag{28}$$

This yields

$$\mathbf{B}_S(\theta,\phi,\alpha) = (\mu\varepsilon 2rN^{-1}\cos\theta)^{1/2} \times \left\{ \cos\alpha \begin{bmatrix} (1 - 2rN^{-1}\cos\theta)\sin\phi\cos\phi \\ -\cos^2\phi - 2rN^{-1}\cos\theta\sin^2\phi \\ (2r\sin\theta - \epsilon)N^{-1}\sin\phi \end{bmatrix} + \sin\alpha \begin{bmatrix} \sin^2\phi + 2rN^{-1}\cos\theta\cos^2\phi \\ (2rN^{-1}\cos\theta - 1)\sin\phi\cos\phi \\ (\epsilon - 2r\sin\theta)N^{-1}\cos\phi \end{bmatrix} \right\}. \tag{29}$$

Every quantity appearing in Eqs. (7) and (11) has now been expressed in terms of θ and ϕ. The electric and magnetic fields can now be calculated for any arbitrary point in the image space. These fields will be used in the definitions in Section 5.

5. ENERGY DENSITY AND THE POYNTING VECTOR

The time-averaged electric and magnetic energy density distributions are defined[7] as

$$\langle W_E\rangle_\alpha = \beta_1 \mathbf{E}_\alpha \cdot \mathbf{E}_\alpha^*,$$
$$\langle W_M\rangle_\alpha = \beta_2 \mathbf{B}_\alpha \cdot \mathbf{B}_\alpha^*, \tag{30}$$

respectively. Here $*$ denotes complex conjugation and $\langle\ldots\rangle$ means averaging over a large number of periods. Furthermore the label α indicates that the electric and magnetic fields depend on the polarization angle α with $\mathbf{E}_S$ and $\mathbf{B}_S$, respectively. We do not bother to insert the exact constants of proportionality since we are interested only in the relative distribution around the focal position. The norm of the Poynting vector is the quantity to be compared with the intensity as calculated in a scalar theory. The time-averaged Poynting vector $\langle\mathbf{S}\rangle$ can be written as[7]

$$\langle\mathbf{S}\rangle_\alpha = \beta_3 \mathrm{Re}(\mathbf{E}_\alpha \times \mathbf{B}_\alpha^*). \tag{31}$$

Incidentally, one can show that, for an ideal high-aperture system, the norm of the time-averaged energy flow and the time-averaged total energy density are proportional on the optical axis[2]:

$$|\langle\mathbf{S}\rangle_\alpha| \propto \langle W\rangle_\alpha, \tag{32}$$

with $\langle W\rangle_\alpha$ being, of course, the sum of the two quantities appearing in Eqs. (30). The complex components of $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ yield a total of 12 real two-dimensional integrals. To obtain numerical stability, we split the integration over ϕ in each integral into two parts. So a total of 24 integrals have to be evaluated to calculate $\langle\mathbf{S}\rangle_\alpha$ in a single point. This was done with routine *D01FCF* of the NAG library.[12] This routine by van Dooren and de Ridder[13] approximates the integral with a seventh-degree rule. It uses an adaptive subdivision strategy and was later optimized by Genz and Malik.[14] The computation time depends strongly on the aberration constant.

It should be pointed out that several other numerical methods for dealing with diffraction integrals exist. We mention the research of Hopkins,[15] Ludwig,[16] Stamnes *et al.*,[17] and Stamnes (Ref. 8, Sec. 7.2).

6. COMPARISON WITH CLASSICAL FOCUSING THEORY

We will now proceed to compare our results with a third-order scalar diffraction theory. First we adopt the dimensionless axial and lateral optical coordinates u and v:

$$u = kz\sin^2\Omega,$$
$$v = k(x^2 + y^2)^{1/2}\sin\Omega, \tag{33}$$

with Ω the angular semiaperture.

Classical focusing theory is based on three assumptions: (1) a scalar field, (2) the paraxial approximation, and (3) the Debye approximation. For a detailed discussion on the validity of this theory we refer to Ref. 8. For the case of spherical aberration, classical focusing theory gives an analytical expression for the intensity distribution along the optical axis,[8] namely,[18]

$$I(u,0) = (1/4|\delta|)\{[C(t_1) + C(t_2)]^2 + [S(t_1) + S(t_2)]^2\}, \tag{34}$$

where

$$t_1 = (2\pi|\delta|)^{1/2}[1 - (u/8\pi\delta)],$$
$$t_2 = (2\pi|\delta|)^{1/2}(u/8\pi\delta), \tag{35}$$

and $C(t)$ and $S(t)$ are the Fresnel integrals defined as

$$C(t) = \left(\frac{2}{\pi}\right)^{1/2}\int_0^t \cos(x^2)\mathrm{d}x,$$
$$S(t) = \left(\frac{2}{\pi}\right)^{1/2}\int_0^t \sin(x^2)\mathrm{d}x. \tag{36}$$

Here δ is defined as the wave-front aberration at the edge of the exit pupil, measured in wavelengths. One can easily show[8] that classical focusing theory predicts an intensity distribution that is symmetric around $u = 4\pi\delta$. The Fresnel integrals can be calculated with NAG routines *S20ACF* and *S20ADF*.[12]

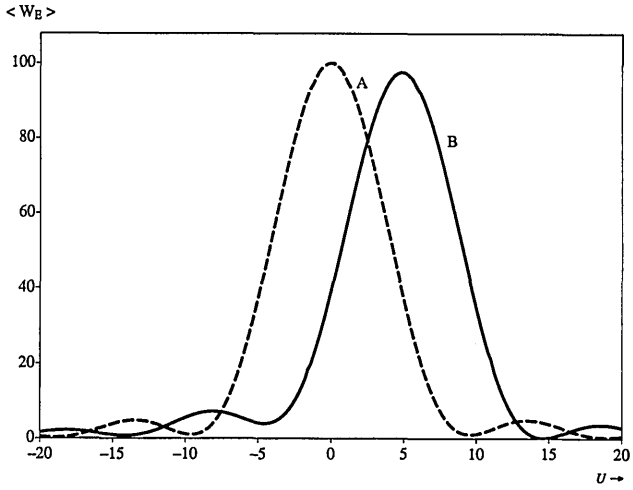

Fig. 3. Time-averaged electric density $\langle W_E \rangle$ (in arbitrary units) along the optical axis: A, perfect high-aperture lens; B, lens with spherical aberration equal to half a wavelength at the edge of the exit pupil. The semiaperture angle is 60°, and the polarization angle is 30°.

As stated above, a scalar theory cannot accommodate the polarization of light. Therefore this comparison is limited to the case of unpolarized light. In our approach the electromagnetic properties of unpolarized light can be studied by integrating over the polarization angle α in Eqs. (30). We can write the time-averaged Poynting vector for unpolarized light as

$$\langle \mathbf{S} \rangle = \frac{1}{2\pi} \int_0^{2\pi} \langle \mathbf{S} \rangle_\alpha \mathrm{d}\alpha . \tag{37}$$

As we will now show, this integral can be performed analytically. The Poynting vector $\langle \mathbf{S} \rangle_\alpha$ depends only on α from $\mathbf{E}_S$ and $\mathbf{B}_S$. From Eqs. (7) and (11) it is clear that $\mathbf{E}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ depend linearly on $\mathbf{E}_S$ and $\mathbf{B}_S$, so, according to Eqs. (27) and (29), we can write symbolically

$$\begin{aligned} \mathbf{E}_\alpha(\mathbf{x}) &= \mathbf{C} \cos \alpha + \mathbf{D} \sin \alpha , \\ \mathbf{B}_\alpha(\mathbf{x}) &= \mathbf{F} \cos \alpha + \mathbf{H} \sin \alpha . \end{aligned} \tag{38}$$

Thus we have

$$\langle \mathbf{S} \rangle = \frac{1}{2\pi} \mathrm{Re} \left\{ \int_0^{2\pi} [(\mathbf{C} \times \mathbf{F}^*) \cos^2 \alpha + (\mathbf{D} \times \mathbf{H}^*) \sin^2 \alpha + (\mathbf{C} \times \mathbf{H}^* + \mathbf{D} \times \mathbf{F}^*) \sin \alpha \cos \alpha] \mathrm{d}\alpha \right\} . \tag{39}$$

This integration can be done in a trivial manner and yields

$$\langle \mathbf{S} \rangle = \mathrm{Re}(\mathbf{C} \times \mathbf{F}^* + \mathbf{D} \times \mathbf{H}^*), \tag{40}$$

again up to a constant of proportionality. This concludes our calculation of $\langle \mathbf{S} \rangle$.

7. DIFFRACTION PATTERN

As above, we take δ to be the wave-front aberration at the edge of the exit pupil, measured in wavelengths. All our results are for δ in the range of 0–2, which is more than would be permitted in practice. We have assumed imaging in air. First we calculated $\langle \mathbf{S} \rangle_\alpha$ and $\langle W_E \rangle_\alpha$ for an aberration-free system (i.e., $C = 0$) and found for a semiangular aperture Ω of 60°, good agreement with all the results of Richards and Wolf,[2] although, as stated above, contrary to our theory their theory is based on the Debye approximation.

In Fig. 3 the time-averaged electric energy distribution (which is presumably what a photographic plate records) is depicted for both an ideal high-aperture lens and one with spherical aberration with $\delta = 0.5$. In the aberrated case the peak has shifted from the focal point, as could be expected. But also the symmetry of the distribution on the u axis is now broken.

Comparing the intensities with those as given by Eq. (34), we find, just as in the ideal case,[2] that for semiapertures smaller than 5° our results agree well with classical focusing theory (for all values of δ that we examined). For high apertures, however, we find that there are significant differences between the predictions of the two theories; for a typical example, see Fig. 4. According to the electromagnetic theory, spherical aberration causes the axial distribution to be no longer symmetric around the maximum: The minima before the peak are much smaller than those behind it and can even become zero. (Notice that the axial distribution for a perfect high-aperture lens does not have any zeros.) Recent measurements of the axial response function of a confocal microscope with various amounts of spherical aberration by Wilson and Carlini (Ref. 4, Fig. 8) seem to confirm this feature. Also the peak shift away from the Gaussian image plane is now less than $4\pi\delta$. Finally, the peak width as predicted by the electromagnetic theory is less than that of the classical theory. These are all general features, which are also found for other values of δ and Ω. More results can be found in Table 1, as follows:

(1) With the same δ, the peak width [or rather, the full width at half-maximum (FWHM)] gets less as the aperture increases.

(2) With Ω constant, the FWHM gets larger when δ increases. Both were to be expected.

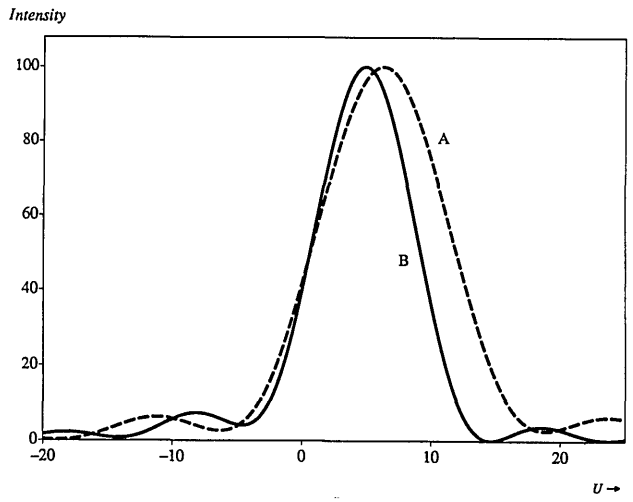

Fig. 4. Comparison of scalar and electromagnetic diffraction theories (in arbitrary units) for $\delta = 0.5$ along the u axis: A, the intensity for the scalar theory; B, the norm of the time-averaged Poynting vector. In this example the semiaperture angle was 60°. In this case the vector theory yielded a Strehl ratio of 0.97. Notice that, unlike in the case of a perfect high-aperture lens, we find that one of the first minima is zero.

Table 1. Vector Theory, Position, and Full Width at Half-Maximum for the Electromagnetic Diffraction Theory[a]

	10°		40°		60°	
δ	Peak	FWHM	Peak	FWHM	Peak	FWHM
0.5	6.20	11.15	5.61	9.96	4.88	8.51
1.0	12.50	11.70	11.24	10.26	9.80	8.67
1.5	18.75	13.16	16.87	11.07	14.74	9.01

[a]The position of the peak and the FWHM for the intensity distribution on the optical axis for various values of spherical aberration and for different semiapertures. The amount of aberration is measured in wavelengths at the edge of the exit pupil. All FWHM's and positions are in optical units along the u axis.

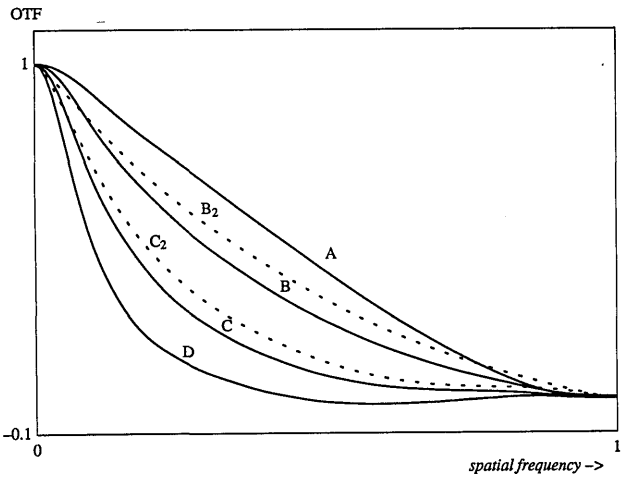

Fig. 5. Incoherent OTF for various amounts of spherical aberration as predicted by electromagnetic Kirchhoff theory. The results are for a lens with a semiaperture of 60° that images in air. Indicated is the amount of aberration at the edge of the pupil in wavelengths: A, ideal lens; B, 0.25; C, 0.50; D, 0.75. The dashed curves are the transfer functions according to Fourier theory: B_2, 0.25; C_2, 0.50.

(3) For the same δ, the peak shift away from the Gaussian image plane decreases as Ω increases.

It should be stressed that a nonparaxial scalar theory[3] also predicts an asymmetric axial intensity distribution. That theory, however, differs in two respects from ours. First, the FWHM is greater than that of the vectorial theory (for instance, for $\Omega = 60°$ and $\delta = 0.5$, the difference is 10%). Second, in the same example the scalar theory does not yield a zero for the first minimum, whereas ours does (see Fig. 4). More precise measurements are needed to determine the limits of validity of either of the two approaches.

8. OPTICAL TRANSFER FUNCTION

Our knowledge of the influence of spherical aberration on the diffraction pattern enables us to calculate the OTF. The incoherent OTF, which we call $C(m, n)$, can be written as an inverse Fourier transform (Ref. 7, p. 485):

$$C(m,n) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |h(x,y)|^2 \exp[2\pi i(mx + ny)]\mathrm{d}x\mathrm{d}y, \quad (41)$$

where m and n are spatial frequencies. The amplitude point-spread function is denoted $h(x, y)$. The integrals are over the Gaussian image plane. Note that, with spherical aberration, this is no longer the plane in which the maximum intensity occurs. Now clearly $|h(x, y)|^2$ is the intensity distribution in the Gaussian image plane of the lens if we have an incoming plane wave. So, in our approach, we can substitute

$$|h(x,y)|^2 = |\langle \mathbf{S}(x,y)\rangle|^2_{\text{unpolarized}}. \quad (42)$$

Because of rotational symmetry, $\langle \mathbf{S}\rangle$ depends only on v, the optical coordinate of Eqs. (33), which implies that the two-dimensional Fourier transform in Eq. (41) can be written as a one-dimensional Fourier–Bessel (or zeroth-order Hankel) transform:

$$C(p) = 2\pi \int_0^{\infty} |\langle \mathbf{S}(v)\rangle|^2 v J_0(2\pi p v)\mathrm{d}v, \quad (43)$$

with J_0 the zeroth-order Bessel function. We have studied cases for which $\delta < 1.0$. In Fig. 5 we have depicted the OTF for several degrees of spherical aberration. When the aberration at the pupil's edge is 0.75λ (curve D) the OTF becomes negative for certain spatial frequencies.

The usual approach to the OTF is to calculate the autocorrelation function of aperture aberration functions. For spherical aberration, this calculation was first done by Black and Linfoot.[19] Using Fourier theory, they derived that[20]

$$C(p) = \frac{4}{\pi}\int_0^{(1-p^2/4)^{1/2}} \mathrm{d}t \int_0^{(1-t^2)^{1/2}-p/2} \cos\left[8\pi\delta ps\left(s^2 + t^2 + \frac{p^2}{4}\right)\right]\mathrm{d}s. \quad (44)$$

When comparing this with our approach, one may indeed ask whether Fourier optics is valid in high-aperture systems. Black and Linfoot and the present authors do not think so, but, as stated in Section 1, it is commonly used in the literature. So if we then compare this approach with the electromagnetic theory, we find (for $\delta < 1.0$) that the latter typically predicts transfer functions that are worse. See Fig. 5.

9. CONCLUSION

We have studied the effect of spherical aberration on the electromagnetic field in high-aperture systems using a Kirchhoff theory. We find that the distribution of the time-averaged electric energy distribution is not symmetric along the u axis. The optical transfer function according to this new theory, typically is worse than that predicted by Fourier optics. We have shown that the electromagnetic Kirchhoff theory yields results that differ from the scalar classical focusing theory (based on the paraxial and the Debye approximations) when applied to high-aperture systems with spherical aberration. The most prominent difference is that the intensity distribution is no longer symmetric on the optical axis. Recent experimental evidence seems to confirm this feature. A similar asymmetry is also predicted by a nonparaxial scalar theory. This theory, however, differs in other respects from ours. Further measurements are needed to decide between these two approaches.

ACKNOWLEDGMENTS

We thank F. C. A. Groen for stimulating discussion and K. J. F. Gaemers and W. A. van Leeuwen for reading the manuscript.

REFERENCES AND NOTES

1. E. Wolf, "Electromagnetic diffraction in optical systems, I. An integral representation of the image field," Proc. R. Soc. London Ser. A **253,** 349–357 (1959).
2. B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems, II. Structure of the image field in an aplanatic system," Proc. R. Soc. London Ser. A **253,** 358–379 (1959).
3. C. J. R. Sheppard, "Aberration in high aperture conventional and confocal imaging systems," Appl. Opt. **27,** 4782–4786 (1988).
4. T. Wilson and A. R. Carlini, "The effect of aberrations on the axial response of confocal imaging systems," J. Microsc. **154,** 243–256 (1989).
5. A. M. Hamed, "Excentration errors combined with wave front aberrations in a coherent scanning microscope," Optik **82,** 1–4 (1989).
6. T. Wilson and C. J. R. Sheppard, *Theory and Practice of Scanning Optical Microscopy* (Academic, London, 1984).
7. M. Born and E. Wolf, *Principles of Optics,* 6th ed. (Pergamon, Oxford, 1980).
8. J. J. Stamnes, *Waves in Focal Regions* (Hilger, Bristol, UK, 1986).
9. H. Ling and S.-W. Lee, "Focusing of electromagnetic waves through a dielectric interface," J. Opt. Soc. Am. A **1,** 965–973 (1984).
10. J. D. Jackson, *Classical Electrodynamics,* 2nd ed. (Wiley, New York, 1975).
11. H. T. M. van der Voort and G. J. Brakenhoff, "3-D image formation in high aperture fluorescence confocal microscopy: a numerical analysis," J. Microsc. **158,** 43–54 (1990).
12. Numerical Algorithms Group, Inc., FORTRAN Library Manual, Mark 13 (Numerical Algorithms Group, Oxford, 1988).
13. P. van Dooren and L. de Ridder, "An adaptive algorithm for numerical integration over an n-dimensional cube," J. Comput. Appl. Math. **2,** 207–217 (1976).
14. A. C. Genz and A. A. Malik, "An adaptive algorithm for numerical integration over an N-dimensional rectangular region," J. Comput. Appl. Math. **6,** 295–299 (1980).
15. H. H. Hopkins, "The numerical evaluation of the frequency response of optical systems," Proc. Phys. Soc. B **70,** 1002–1005 (1957).
16. A. C. Ludwig, "Computation of radiation patterns involving numerical double integration," IEEE Trans. Antennas Propag. **AP-16,** 767–769 (1968).
17. J. J. Stamnes, B. Spelkavik, and H. M. Pedersen, "Evaluation of diffraction integrals using local phase and amplitude approximations," Opt. Acta **30,** 207–222 (1983).
18. We are of the humble opinion that the minus signs in Eq. (12.93g) of Stamnes[8] are misprints; cf., Ref. 4.
19. G. Black and E. H. Linfoot, "Spherical aberration and the information content of optical images," Proc. R. Soc. London Ser. A **239,** 522–540 (1957).
20. In making the transition from their Eqs. (21) to (22), Black and Linfoot make a small error.