Axial birefringence in high-numerical-aperture optical systems and the light distribution close to focus

Sjoerd Stallinga

Philips Research Laboratories, Professor Holstlaan 4, 5656 AA, Eindhoven, The Netherlands

Received October 19, 2000; revised manuscript received April 6, 2001; accepted April 10, 2001

The effects of birefringence on the light distribution in the focal region of a high-NA optical system are investigated with use of the Debye approach to vector diffraction theory. The attention is limited to uniaxially birefringent media with symmetry axis along the optical axis of the imaging system. The radially (p) and tangentially (s) polarized fields in the exit pupil produce spots in the focal region that are defocused with respect to each other. For small birefringence values the relative defocus causes a distortion and broadening of the spot; for larger values the two spots separate completely. As a corollary to the theory it is shown that there is a tangential tornadolike flow of energy in the focal region when the polarization in the entrance pupil is elliptical.

OCIS codes: 260.1440, 260.1960, 210.4770, 180.1790.

1. INTRODUCTION

The information on optical disks is read out by focusing light of wavelength λ and numerical aperture (NA) on the information layer through a protective cover layer of thickness d. To increase the information density the wavelength is decreased and the numerical aperture is increased, as the spot size scales with λ/NA. To that end the existing compact disk (CD) standard (λ = 780 nm, NA = 0.45, d = 1.2 mm) is supplemented with new standards such as those for digital versatile disk (DVD) (λ = 660 nm, NA = 0.60, d = 0.6 mm) and digital video recorder (DVR) (λ = 400 nm, NA = 0.85, d = 0.1 mm).[1] The high-NA values of these new standards introduce polarization effects that change the light distribution in the focal region compared with the distribution predicted by the scalar theory of light. The light distribution is also influenced by the effects of the cover layer of the disk. In particular, birefringence of the cover layer can seriously distort the light distribution around focus. Effects of polarization and birefringence can be described within the framework of vector-diffraction theory. The goal of this paper is to extend the vector diffraction theory in the literature to include the effects of a uniaxially birefringent cover layer with its symmetry axis perpendicular to the layer. Birefringence of this type is called axial birefringence. Next to the field of optical data storage the results of this paper may also have relevance for the field of confocal microscopy.

Vector diffraction in a single isotropic medium was originally considered by Wolf and co-workers,[2–5] and later by Mansuripur.[6,7] Focusing through an interface between two dielectric media was first investigated by Ling and Lee,[8] followed by a number of other authors.[9–15] A general description of focusing in anisotropic media was given by Stamnes and Jiang.[16] They applied their general theory to the particular case of focusing with a cylindrical lens into a birefringent medium with its uniaxial symmetry axis in the plane of incidence.[17,18] In this two-dimensional problem only p-polarized incident waves, which excite only the extraordinary wave in the uniaxial crystal, are considered. This stands in contrast to this paper, which deals with focusing with rotationally symmetric lenses, making the problem fully three-dimensional. Moreover, arbitrary states of polarization in the entrance pupil of the optical system are considered, implying that both the extraordinary and ordinary waves contribute to the field close to focus.

The content of this paper is as follows. The Debye approach for calculating the electromagnetic field close to focus is generalized to include axially birefringent media in Section 2. General formulas for the density and flow of electromagnetic energy are derived in Section 3. The effects of axial birefringence are discussed in Section 4. The paper is concluded with a summary of the main results in Section 5.

2. FOCUSING THROUGH AXIALLY BIREFRINGENT MEDIA

A. Angular Spectrum of Plane Waves for the Propagator Matrix

Consider media 1 and 2 separated by an interface in $z = -d$. The exit pupil of the optical system is in the plane $z = -R$, and the focal plane is $z = 0$ (see Fig. 1). Medium 1 is isotropic, and medium 2 is uniaxially birefringent with uniaxial symmetry axis $\hat{\mathbf{a}}$ along the optical axis (the z-axis). The electric and magnetic fields as a function of position $\mathbf{r}$ can be expressed as two-dimensional Fourier integrals, the so-called angular spectrum of plane waves:

$$\mathbf{E}(x,y,z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\mathrm{d}^2k}{(2\pi)^2}\mathbf{E}(k_x,k_y,z) \times \exp[(i(k_x x + k_y y)], \quad (1)$$

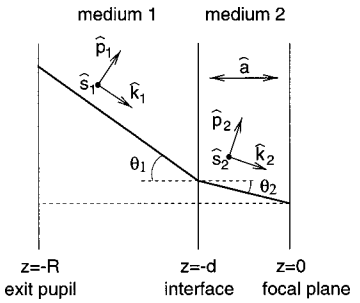

Fig. 1. Light path from the exit pupil in $z = -R$ to the focal plane in $z = 0$ with refraction at the interface in $z = -d$, the uniaxial symmetry axis $\hat{\mathbf{a}}$, the associated wave vectors $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$, and polarization vectors parallel to the plane of incidence $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$ and perpendicular to the plane of incidence $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$.

$$\mathbf{H}(x,y,z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\mathrm{d}^2k}{(2\pi)^2}\mathbf{H}(k_x,k_y,z) \times \exp[i(k_x x + k_y y)]. \tag{2}$$

The z component of the wave vector of the wave propagating forward in medium 1, with refractive index n_1, is defined by

$$k_{1z} = ({n_1}^2k^2 - {k_x}^2 - {k_y}^2)^{1/2}, \tag{3}$$

where $k = 2\pi/\lambda$, with λ being the wavelength in vacuum. For the wave propagating backward the sign of the z component just changes. The total wave vector is then $\mathbf{k}_1 = n_1 k\hat{\mathbf{k}}_1$, where the unit vector $\hat{\mathbf{k}}_1$ along $\mathbf{k}_1$ can be parametrized by the polar and azimuthal angles θ_1 and ϕ_1 as

$$\hat{\mathbf{k}}_1 = (\sin\theta_1\cos\phi_1, \sin\theta_1\sin\phi_1, \cos\theta_1). \tag{4}$$

The polarization vectors are the p and s vectors, defined by

$$\hat{\mathbf{p}}_1 = (\cos\theta_1\cos\phi_1, \cos\theta_1\sin\phi_1, -\sin\theta_1), \tag{5}$$

$$\hat{\mathbf{s}}_1 = (-\sin\phi_1, \cos\phi_1, 0). \tag{6}$$

The three unit vectors $\hat{\mathbf{p}}_1$, $\hat{\mathbf{s}}_1$, and $\hat{\mathbf{k}}_1$ form an orthonormal set, i.e., $\hat{\mathbf{p}}_1 \times \hat{\mathbf{s}}_1 = \hat{\mathbf{k}}_1$. Similar vectors $\hat{\mathbf{p}}_1^-$, $\hat{\mathbf{s}}_1^-$ and $\hat{\mathbf{k}}_1^-$ may be defined for the wave propagating backward by replacing θ_1 by $\pi - \theta_1$. The electric and magnetic fields in medium 1 can now be expressed as

$$\mathbf{E}(k_x,k_y,z) = (E_p^i\hat{\mathbf{p}}_1 + E_s^i\hat{\mathbf{s}}_1) \times \exp[ik_{1z}(z - z_1)] + (E_p^r\hat{\mathbf{p}}_1^- + E_s^r\hat{\mathbf{s}}_1^-) \times \exp[-ik_{1z}(z - z_1)], \tag{7}$$

$$\mu_0 c\mathbf{H}(k_x,k_y,z) = (-n_1E_s^i\hat{\mathbf{p}}_1 + n_1E_p^i\hat{\mathbf{s}}_1) \times \exp[ik_{1z}(z - z_1)] + (-n_1E_s^r\hat{\mathbf{p}}_1^- + n_1E_p^r\hat{\mathbf{s}}_1^-) \times \exp[-ik_{1z}(z - z_1)], \tag{8}$$

respectively, where μ_0 is the permeability of vacuum, c is the speed of light, and $-R \leqslant z_1 < -d$. The p and s amplitudes of the wave incident on the interface in $z = -d$ are denoted as E_p^i and E_s^i, whereas the p and s amplitudes of the wave reflected at the interface in $z = -d$ are denoted as E_p^r and E_s^r. All amplitudes are functions of k_x, k_y, and z_1.

In medium 2, the medium with axial birefringence, the ordinary and extraordinary refractive indices are n_o and n_e, respectively, and the wave vector z components of the ordinary and extraordinary modes are

$$k_{2z}^o = (n_o^2k^2 - k_x^2 - k_y^2)^{1/2}, \tag{9}$$

$$k_{2z}^e = \frac{n_o}{n_e}(n_e^2k^2 - k_x^2 - k_y^2)^{1/2}, \tag{10}$$

as follows from the dispersion relations of the two modes.[19] For the ordinary mode the total wave vector is $\mathbf{k}_2^o = n_o k\hat{\mathbf{k}}_2$, where the unit vector $\hat{\mathbf{k}}_2$ and the corresponding polarization vectors $\hat{\mathbf{p}}_2$ and $\hat{\mathbf{s}}_2$ can be expressed in the polar and azimuthal angles θ_2 and ϕ_2 as

$$\hat{\mathbf{k}}_2 = (\sin\theta_2\cos\phi_2, \sin\theta_2\sin\phi_2, \cos\theta_2), \tag{11}$$

$$\hat{\mathbf{p}}_2 = (\cos\theta_2\cos\phi_2, \cos\theta_2\sin\phi_2, -\sin\theta_2), \tag{12}$$

$$\hat{\mathbf{s}}_2 = (-\sin\phi_2, \cos\phi_2, 0). \tag{13}$$

The polarization of the ordinary mode is perpendicular to the plane of incidence, i.e., coincides with $\hat{\mathbf{s}}_2$, whereas the polarization of the extraordinary mode is in the plane of incidence, although not perpendicular to $\hat{\mathbf{k}}_2$. This results in expressions for the electric and magnetic fields, which can be obtained by rewriting textbook formulas,[19]

$$\mathbf{E}(k_x,k_y,z) = E_p^t(\hat{\mathbf{p}}_2 + \sigma\hat{\mathbf{k}}_2) \times \exp[ik_{2z}^e(z - z_2)] + E_s^t\hat{\mathbf{s}}_2 \times \exp[ik_{2z}^o(z - z_2)], \tag{14}$$

$$\mu_0 c\mathbf{H}(k_x,k_y,z) = -n_oE_s^t\hat{\mathbf{p}}_2\exp[ik_{2z}^o(z - z_2)] + \gamma n_oE_p^t\hat{\mathbf{s}}_2\exp[ik_{2z}^e(z - z_2)], \tag{15}$$

where $-d \leqslant z_2$ and where E_p^t and E_s^t are the p (extraordinary) and s (ordinary) amplitudes of the wave transmitted at the interface in $z = -d$, which are functions of k_x, k_y, and z_2. The numbers σ and γ are given by

$$\sigma = \frac{(n_e^2 - n_o^2)\sin\theta_2\cos\theta_2}{n_o^2\sin^2\theta_2 + n_e^2\cos^2\theta_2}, \tag{16}$$

$$\gamma = \frac{n_e}{(n_o^2\sin^2\theta_2 + n_e^2\cos^2\theta_2)}, \tag{17}$$

and deviate from 0 and 1, respectively, due to the axial birefringence.

The continuity of the x and y components of the electric and magnetic fields at the interface in $z = -d$ requires that $\phi_1 = \phi_2 = \phi$ and that $n_1\sin\theta_1 = n_o\sin\theta_2$ (Snel's law). In addition to this, the amplitudes E_p^t, E_s^t, E_p^r, and E_s^r are related to the amplitudes E_p^i and E_s^i by the p and s transmission and reflection coefficients. These coefficients follow from the continuity conditions at the interface as

$$t_p = \frac{2n_1 \cos\theta_1}{\gamma^2 n_1 \cos\theta_2 + \gamma n_o \cos\theta_1}, \tag{18}$$

$$t_s = \frac{2n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_o \cos\theta_2}, \tag{19}$$

$$r_p = \frac{n_1 \cos\theta_2 - \gamma n_o \cos\theta_1}{n_1 \cos\theta_2 + \gamma n_o \cos\theta_1}, \tag{20}$$

$$r_s = \frac{n_1 \cos\theta_1 - n_o \cos\theta_2}{n_1 \cos\theta_1 + n_o \cos\theta_2}, \tag{21}$$

where it is used that $1 + \sigma \tan\theta_2 = \gamma^2$. In the isotropic case $n_o = n_e$ ($\gamma = 1$, $\sigma = 0$) these expressions reduce to the well-known expressions for the Fresnel coefficients. The electric and magnetic fields in $z = z_2$ can now be written as

$$\begin{aligned}\mathbf{E}(k_x, k_y, z_2) = {} & t_p E_p^i (\hat{\mathbf{p}}_2 + \sigma\hat{\mathbf{k}}_2) \\ & \times \exp\{i[k_{2z}^e (z_2 + d) \\ & - k_{1z}(z_1 + d)]\} + t_s E_s^i \hat{\mathbf{s}}_2 \\ & \times \exp\{i[k_{2z}^o (z_2 + d) \\ & - k_{1z}(z_1 + d)]\},\end{aligned} \tag{22}$$

$$\begin{aligned}\mu_o c\mathbf{H}(k_x, k_y, z_2) = {} & -n_o t_s E_s^i \hat{\mathbf{p}}_2 \exp\{i[k_{2z}^o (z_2 + d) \\ & - k_{1z}(z_1 + d)]\} \\ & + \gamma n_o t_p E_p^i \hat{\mathbf{s}}_2 \exp\{i[k_{2z}^e (z_2 + d) \\ & - k_{1z}(z_1 + d)]\}.\end{aligned} \tag{23}$$

Two successive Fourier transforms are applied to these formulas. First, the incident Fourier amplitudes E_p^i and E_s^i at $z = z_1$ are calculated from the field $\mathbf{E}^i(x_1, y_1, z_1)$ in the plane $z = z_1$. Second, the electromagnetic fields in the plane $z = z_2$ are found by the inverse Fourier transforms of $\mathbf{E}(k_x, k_y, z_2)$ and $\mu_0 c\mathbf{H}(k_x, k_y, z_2)$. It then follows that ($\alpha, \beta = x, y, z$; summation over β is implied):

$$E_\alpha(\mathbf{r}_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}^2 r_1 G_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}_1) E_\beta(\mathbf{r}_1), \tag{24}$$

$$\mu_0 c H_\alpha(\mathbf{r}_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}^2 r_1 G'_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}_1) E_\beta(\mathbf{r}_1), \tag{25}$$

with the matrix elements ($\alpha, \beta = x, y, z$)

$$\begin{aligned}G_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}_1) = {} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathrm{d}^2 k}{(2\pi)^2} [t_p (\hat{p}_{2\alpha} + \sigma\hat{k}_{2\alpha}) \hat{p}_{1\beta} \\ & \times \exp(i\Phi^e) + t_s \hat{s}_{2\alpha}\hat{s}_{1\beta} \exp(i\Phi^o)],\end{aligned} \tag{26}$$

$$\begin{aligned}G'_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}_1) = {} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathrm{d}^2 k}{(2\pi)^2} [-n_o t_s \hat{p}_{2\alpha}\hat{s}_{1\beta} \\ & \times \exp(i\Phi^o) + \gamma n_o t_p \hat{s}_{2\alpha}\hat{p}_{1\beta} \exp(i\Phi^e)].\end{aligned} \tag{27}$$

The phases Φ^e and Φ^o are defined by

$$\begin{aligned}\Phi^e = {} & k_x(x_2 - x_1) + k_y(y_2 - y_1) \\ & + k_{2z}^e (z_2 + d) - k_{1z}(z_1 + d),\end{aligned} \tag{28}$$

$$\begin{aligned}\Phi^o = {} & k_x(x_2 - x_1) + k_y(y_2 - y_1) \\ & + k_{2z}^o (z_2 + d) - k_{1z}(z_1 + d).\end{aligned} \tag{29}$$

The propagator matrices are no longer functions of the difference position vector $\mathbf{r}_2 - \mathbf{r}_1$, because translational invariance is lost owing to the presence of the interface.

B. Approximate Expression for the Propagator Matrix

So far, the present results agree with those of Stamnes and Jiang.[16] In this paper three approximations are used that simplify the expressions derived for the propagator matrices considerably without compromising the quantitative validity of the expressions.

The first of these approximations is the small birefringence approximation, which is introduced in this paper for the first time. It turns out that in practice the birefringence of the cover layer $\Delta n = n_e - n_o$ is of the order 10^{-3} or less. Then it is sufficient to take the effects of birefringence into account by the phases alone. This means that n_e can be set equal to n_o everywhere except in the expressions for the phases. Then the numbers σ and γ are equal to 0 and 1, respectively, and the p and s transmission coefficients reduce to the classical Fresnel coefficients. The difference between the phases for the two modes is well approximated by

$$\begin{aligned}\Delta W = (k_{2z}^e - k_{2z}^o)(d + z_2) &= \frac{(k_{2z}^e)^2 - (k_{2z}^o)^2}{k_{2z}^e + k_{2z}^o}(d + z_2) \\ &\approx \frac{(k_{2z}^e)^2 - (k_{2z}^o)^2}{2k_{2z}^o}(d + z_2), \\ &= k(d + z_2)\Delta n \frac{\sin^2\theta_2}{\cos\theta_2}.\end{aligned} \tag{30}$$

This leads to propagator functions

$$\begin{aligned}G_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}_1) = {} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathrm{d}^2 k}{(2\pi)^2} [t_p \exp(i\Delta W) \hat{p}_{2\alpha}\hat{p}_{1\beta} \\ & + t_s \hat{s}_{2\alpha}\hat{s}_{1\beta}] \exp[i\Phi(\mathbf{k}_1, \mathbf{k}_2, \mathbf{r}_1, \mathbf{r}_2)],\end{aligned} \tag{31}$$

$$\begin{aligned}G'_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}_1) = {} & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathrm{d}^2 k}{(2\pi)^2} n_o [-t_s \hat{p}_{2\alpha}\hat{s}_{1\beta} \\ & + t_p \exp(i\Delta W) \hat{s}_{2\alpha}\hat{p}_{1\beta}] \\ & \times \exp[i\Phi(\mathbf{k}_1, \mathbf{k}_2, \mathbf{r}_1, \mathbf{r}_2)],\end{aligned} \tag{32}$$

with the phase

$$\begin{aligned}\Phi(\mathbf{k}_1, \mathbf{k}_2, \mathbf{r}_1, \mathbf{r}_2) = {} & k_x(x_2 - x_1) + k_y(y_2 - y_1) \\ & + k_{2z}(z_2 + d) - k_{1z}(z_1 + d),\end{aligned} \tag{33}$$

where $\mathbf{k}_2 = \mathbf{k}_2^o$.

The propagator matrices may be evaluated explicitly with use of a second approximation, the stationary phase approximation. Jiang and Stamnes mention the application of the stationary phase approximation to the calculation of the propagator function but give the result only for the two-dimensional case of focusing with a cylindrical lens. The stationary phase approximation is quite justi-

fied, as $|z_1 + d|$ and $|z_2 + d|$ can be taken to be much larger than the wavelength λ. The calculation is straightforward, following the methods described in the book of Stamnes[20] and leads to the following results: The stationary point for each of these integrals corresponds to the path of the geometrical optics light ray from $\mathbf{r}_1$ to $\mathbf{r}_2$, refracting at the interface according to Snell's law[8,14] (see Fig. 1). This implies that the wave vectors can be expressed in terms of angular variables θ_1, θ_2, and ϕ as follows:

$$k_x^s = kn_1 \sin\theta_1 \cos\phi = kn_2 \sin\theta_2 \cos\phi, \tag{34}$$

$$k_y^s = kn_1 \sin\theta_1 \sin\phi = kn_2 \sin\theta_2 \sin\phi, \tag{35}$$

$$k_{1z}^s = kn_1 \cos\theta_1, \tag{36}$$

$$k_{2z}^s = kn_2 \cos\theta_2, \tag{37}$$

where $n_2 \equiv n_o$. In turn, these angular variables are defined in terms of the coordinates by

$$x_2 - x_1 = [(R - d)\tan\theta_1 + d\tan\theta_2]\cos\phi, \tag{38}$$

$$y_2 - y_1 = [(R - d)\tan\theta_1 + d\tan\theta_2]\sin\phi. \tag{39}$$

The phase for the stationary point is k times the optical path length

$$\Phi(\mathbf{k}_1^s, \mathbf{k}_2^s, \mathbf{r}_1, \mathbf{r}_2) = k\,\mathrm{OPL}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{kn_1(z_1 + d)}{\cos\theta_1} + \frac{kn_2(z_2 + d)}{\cos\theta_2}, \tag{40}$$

and the Hessian (determinant of the matrix of second-order derivatives of the phase function with respect to k_x and k_y at the stationary point) is

$$Q = \left(\frac{z_2 + d}{kn_2\cos\theta_2} - \frac{z_1 + d}{kn_1\cos\theta_1}\right) \times \left(\frac{z_2 + d}{kn_2\cos^3\theta_2} - \frac{z_1 + d}{kn_1\cos^3\theta_1}\right). \tag{41}$$

This gives rise to the propagator matrices:

$$G_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}_1) = \frac{1}{2\pi i\sqrt{Q}}[t_p \exp(i\Delta W)\hat{p}_{2\alpha}\hat{p}_{1\beta} + t_s\hat{s}_{2\alpha}\hat{s}_{1\beta}]\exp[ik\,\mathrm{OPL}(\mathbf{r}_1, \mathbf{r}_2)], \tag{42}$$

$$G'_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}_1) = \frac{n_2}{2\pi i\sqrt{Q}}[-t_s\hat{p}_{2\alpha}\hat{s}_{1\beta} + t_p \times \exp(i\Delta W)\hat{s}_{2\alpha}\hat{p}_{1\beta}] \times \exp[ik\,\mathrm{OPL}(\mathbf{r}_1, \mathbf{r}_2)]. \tag{43}$$

The appearance of the phase factor $\exp(i\Delta W)$ is the only sign of axial birefringence. The Debye approach is followed, according to which only the stationary points inside the exit pupil are taken into account, i.e. only those geometrical optics paths for which $({x_1}^2 + {y_1}^2)^{1/2} \leqslant a$, where a is the radius of the exit pupil. It follows that the integration domain in Eqs. (24) and (25) is then the exit pupil. It is mentioned that the quantitative validity of the Debye approach is limited to the cases where the Fresnel number (ratio of the aperture radius and the diffraction-limited spot size) is much larger than one.

A third approximation can be made for diffraction-limited imaging systems, which are the only relevant systems in the readout of optical disks. The dimensions of the focal region, the only region of interest, are of the order of the wavelength λ. For points this close to focus the phase of the propagator matrices may be approximated as

$$k\,\mathrm{OPL}(\mathbf{r}_1, \mathbf{r}_2) = k\,\mathrm{OPL}(\mathbf{r}_1, 0) + \mathbf{k}_2^s \cdot \mathbf{r}_2. \tag{44}$$

In addition to this the angular variables θ_1, θ_2, and ϕ are now associated with the geometrical optics light path through the pupil point $\mathbf{r}_1$ and the geometrical focus $\mathbf{r}_2 = 0$. It follows that the p and s polarization vectors and the wave vectors $\mathbf{k}_1^s$ and $\mathbf{k}_2^s$ depend on $\mathbf{r}_1$ alone. Finally, z_2 may be neglected in the expression for ΔW, as $|z_2| \ll d$. Clearly, $\mathbf{r}_2$ appears only in the propagator matrices in the phase factor $\exp(i\mathbf{k}_2^s \cdot \mathbf{r}_2)$.

C. Integral Expression for the Electromagnetic Field

The field in the exit pupil plane $z_1 = -R$ must be known in order to calculate the image field. It is assumed that the imaging system is aplanatic, invariant under rotations around the optical axis, and corrected for the defocus and spherical aberration (of all orders) arising from the cover layer. Furthermore, the system is taken to be illuminated with a collimated, aberration-free, and uniform beam of amplitude E_0, proportional to the (unit) polarization vector $(A_x, A_y, 0)$. According to Wolf[2] the p and s amplitudes in the exit pupil are then proportional to

$$B_p = E_0(\cos\phi A_x + \sin\phi A_y), \tag{45}$$

$$B_s = E_0(-\sin\phi A_x + \cos\phi A_y). \tag{46}$$

Energy conservation in the imaging step from entrance to exit pupil through the aplanatic system requires scaling of the amplitude with a factor

$$\sqrt{I} = \frac{R}{{n_1}^2 k(Q\cos\theta_1)^{1/2}}. \tag{47}$$

The factor $\cos\theta_1$ results from the fact that the ray intersects the exit pupil obliquely.[21] Finally, correction for focusing through the cover layer means that the phase in the pupil plane is given by $-k\,\mathrm{OPL}(\mathbf{r}_1, 0)$. In case the optical system is aberrated the field in the exit pupil acquires an additional phase $\exp(iW)$, where W is the classical aberration function. For example, the aberration function is $W = (k_{2z}^s - k_{1z}^s)d$ if the system is not corrected for focusing through the cover layer. The field in the exit pupil is then given by

$$\mathbf{E}^i(\mathbf{r}_1) = \sqrt{I}(B_p\hat{\mathbf{p}}_1 + B_s\hat{\mathbf{s}}_1)\exp[-ik\,\mathrm{OPL}(\mathbf{r}_1, 0) + iW]. \tag{48}$$

Combining this equation with Eqs. (24), (25), and (42)–(44) and changing the integration variables from x_1 and y_1 and θ_1 and ϕ leads to the following electromagnetic field in $\mathbf{r}_2$:

$$\mathbf{E}(\mathbf{r}_2) = \frac{E_0 R}{i\lambda}\int_0^{\beta_1}\int_0^{2\pi} \sin\theta_1\, \mathrm{d}\theta_1 \mathrm{d}\phi (\cos\theta_1)^{1/2} \times [t_p \exp(iW_p)(\cos\phi A_x + \sin\phi A_y)\hat{\mathbf{p}}_2 + t_s \exp(iW_s)(-\sin\phi A_x + \cos\phi A_y)\hat{\mathbf{s}}_2 \times \exp(i\mathbf{k}_2^s \cdot \mathbf{r}_2), \quad (49)$$

$$\mu_0 c\mathbf{H}(\mathbf{r}_2) = \frac{E_0 R}{i\lambda}\int_0^{\beta_1}\int_0^{2\pi} \sin\theta_1\, \mathrm{d}\theta_1\, \mathrm{d}\phi (\cos\theta_1)^{1/2} \times [-n_2 t_s \exp(iW_s)(-\sin\phi A_x + \cos\phi A_y)\hat{\mathbf{p}}_2 + n_2 t_p \exp(iW_p) \times (\cos\phi A_x + \sin\phi A_y)\hat{\mathbf{s}}_2]\exp(i\mathbf{k}_2^s \cdot \mathbf{r}_2). \quad (50)$$

The scalar aberration function W is combined with ΔW into the aberration functions of the p and s polarizations $W_p = W + \Delta W$ and $W_s = W$, respectively. The integration domain is the cone with polar angles θ_1 smaller than β_1, with $\beta_1 = \arctan(a/R)$. The maximum polar angle β_1 is related to the numerical aperture NA by $\mathrm{NA} = n_1 \sin\beta_1$.

Equations (49) and (50) show that the field in the proximity of focus is the coherent superposition of two spots, one coming from the field component that is radially polarized in the exit pupil (associated with the extraordinary mode), and the other coming from the field component that is tangentially polarized in the exit pupil (associated with the ordinary mode). The two spots have aberration functions W_p and W_s. Clearly, the effect of axial birefringence is the introduction of different aberrations for the two spots. The integral expressions (49) and (50) generalize Debye's integral to include effects of high NA, arbitrary polarization in the entrance pupil, the effects of focusing through a cover layer, and axial birefringence in that cover layer. The objective of Section 3 is to evaluate the diffraction integral from which the distribution and the flow of electromagnetic energy can then be calculated.

3. LIGHT DISTRIBUTION CLOSE TO FOCUS

A. Electric and Magnetic Fields

The integral (49) can be simplified and rewritten similar to known results in the literature, resulting in

$$E_x(u,v,\psi) = -i\pi E_0 N (T_0/n_1 n_2)^{1/2}\{[F_0(u,v) + F_2(u,v)\cos(2\psi)]A_x + F_2(u,v) \times \sin(2\psi)A_y\}, \quad (51)$$

$$E_y(u,v,\psi) = -i\pi E_0 N (T_0/n_1 n_2)^{1/2}\{F_2(u,v) \times \sin(2\psi)A_x + [F_0(u,v) - F_2(u,v) \times \cos(2\psi)]A_y\}, \quad (52)$$

$$E_z(u,v,\psi) = -2\pi E_0 N (T_0/n_1 n_2)^{1/2} F_1(u,v) \times [\cos(\psi)A_x + \sin(\psi)A_y]. \quad (53)$$

Here, the image point $\mathbf{r}_2$ is expressed in terms of the azimuthal coordinate ψ and the dimensionless axial and radial coordinates u and v, which are defined by

$$u = \frac{\mathrm{NA}^2}{n_2\lambda} z_2, \quad (54)$$

$$v = \frac{\mathrm{NA}}{\lambda}(x_2^2 + y_2^2)^{1/2}. \quad (55)$$

The three functions $F_k(u,v)$ for $k = 0,1,2$ are defined by

$$F_k(u,v) = \int_0^1 \mathrm{d}\rho \rho g_k(\rho) J_k(2\pi v\rho) \times \exp\left[-\frac{2\pi i u \rho^2}{1 + (1 - \rho^2 \sin^2\beta_2)^{1/2}}\right]. \quad (56)$$

The Bessel functions of the first kind $J_k(2\pi v\rho)$ arise from the integration over ϕ, which can be done analytically. The integration variable ρ is obtained from the substitution

$$\rho = n_1 \sin\theta_1/\mathrm{NA} = n_2 \sin\theta_2/\mathrm{NA} = \sin\theta_1/\sin\beta_1 = \sin\theta_2/\sin\beta_2, \quad (57)$$

where the angles β_k are related to the numerical aperture NA by $\sin\beta_k = \mathrm{NA}/n_k$, and the three functions $g_k(\rho)$ for $k = 0, 1, 2$ are defined by

$$g_0(\rho) = (1 - \rho^2 \sin^2\beta_1)^{-1/4} \frac{n_1 + n_2}{2n_1}[t_s \exp(iW_s) + t_p \exp(iW_p)(1 - \rho^2 \sin^2\beta_2)^{1/2}], \quad (58)$$

$$g_1(\rho) = (1 - \rho^2 \sin^2\beta_1)^{-1/4} \frac{n_1 + n_2}{2n_1} t_p \times \exp(iW_p)\rho \sin\beta_2, \quad (59)$$

$$g_2(\rho) = (1 - \rho^2 \sin^2\beta_1)^{-1/4} \frac{n_1 + n_2}{2n_1} \times [t_s \exp(iW_s) - t_p \times \exp(iW_p)(1 - \rho^2 \sin^2\beta_2)^{1/2}]. \quad (60)$$

The remaining undefined quantities in Eqs. (51)–(53) are the normal incidence transmission coefficient of the interface at $z = -d$,

$$T_0 = \frac{4n_1 n_2}{(n_1 + n_2)^2}, \quad (61)$$

and the Fresnel number,

$$N = \frac{n_1 R \sin^2\beta_1}{\lambda} = \frac{R \sin\beta_1}{(\lambda/\mathrm{NA})}. \quad (62)$$

Finally, it is noted that in Eqs. (51)–(53) an overall phase factor $\exp(2\pi i u/\sin^2\beta_2)$ is left out. Equations (51)–(53) for the electric field are valid for arbitrary states of polarization in the entrance pupil and describe the effect of axial birefringence. In this respect they generalize expressions previously obtained in the literature.

The scalar diffraction theory corresponds to the limit $\mathrm{NA}\to 0$. A series expansion shows that the functions $F_k(u,v)$ are proportional to NA^k for small NA. It follows that only $F_0(u,v)$ is relevant in the limit $\mathrm{NA}\to 0$. The electric field is then simply proportional to the product of $F_0(u,v)$ and the polarization in the entrance pupil $(A_x, A_y, 0)$. The function $g_0(\rho)$ may be approximated in this limit by

$$g_0(\rho) = \exp(iW_p) + \exp(iW_s). \tag{63}$$

In the birefringence free case this corresponds to the results already discussed in Richards and Wolf.[3]

Next, the magnetic field $\mathbf{H}$ is considered. It can be expressed in terms of integrals similar to the electric field. It turns out that the magnetic field expression can be obtained from the electric-field expression by substituting $A_x\to -A_y$ and $A_y\to A_x$, and also $t_p\exp(iW_p)\to t_s\exp(iW_s)$ and $t_s\exp(iW_s)\to t_p\exp(iW_p)$ and by an overall multiplication with n_2. In Subsection 3.B, the magnetic functions corresponding to $F_k(u,v)$ and $g_k(\rho)$ are indicated by $F_k'(u,v)$ and $g_k'(\rho)$.

B. Electromagnetic Energy Density

The intensity of the α component of the electric field is $I_\alpha = \frac{1}{2}\epsilon_0 n_2^2|E_\alpha|^2$. The contribution from the birefringence may be neglected, as Δn is assumed to be much smaller than n_2. Introducing a scaling factor

$$I_0 = \frac{1}{2}\epsilon_0\frac{n_2}{n_1}T_0\pi^2N^2, \tag{64}$$

it is found that

$$\begin{aligned}\frac{I_x}{I_0} &= |F_0 + F_2\cos(2\psi)|^2|A_x|^2 \\ &\quad + 2\,\mathrm{Re}\{[F_0 + F_2\cos(2\psi)]F_2^*\sin(2\psi)A_xA_y^*\} \\ &\quad + |F_2|^2\sin^2(2\psi)|A_y|^2 \\ &= \frac{1}{2}[|F_0|^2 + |F_2|^2 + 2\,\mathrm{Re}(F_0F_2^*)\cos(2\psi)] \\ &\quad\times(|A_x|^2 + |A_y|^2) + \frac{1}{2}[|F_0|^2 + |F_2|^2\cos(4\psi) \\ &\quad + 2\,\mathrm{Re}(F_0F_2^*)\cos(2\psi)](|A_x|^2 - |A_y|^2) \\ &\quad + \left[\mathrm{Re}(F_0F_2^*)\sin(2\psi) + \frac{1}{2}|F_2|^2\sin(4\psi)\right] \\ &\quad\times\mathrm{Re}(A_x^*A_y) - \mathrm{Im}(F_0F_2^*)\sin(2\psi)\mathrm{Im}(A_x^*A_y),\end{aligned} \tag{65}$$

$$\begin{aligned}\frac{I_y}{I_0} &= |F_2|^2\sin^2(2\psi)|A_x|^2 + 2\,\mathrm{Re}\{[F_0 - F_2 \\ &\quad\times\cos(2\psi)]F_2^*\sin(2\psi)A_x^*A_y\} \\ &\quad + |F_0 - F_2\cos(2\psi)|^2|A_y|^2 \\ &= \frac{1}{2}[|F_0|^2 + |F_2|^2 - 2\,\mathrm{Re}(F_0F_2^*)\cos(2\psi)] \\ &\quad\times(|A_x|^2 + |A_y|^2) - \frac{1}{2}[|F_0|^2 \\ &\quad + |F_2|^2\cos(4\psi) - 2\,\mathrm{Re}(F_0F_2^*)\cos(2\psi)] \\ &\quad\times(|A_x|^2 - |A_y|^2) + \left[\mathrm{Re}(F_0F_2^*)\sin(2\psi)\right. \\ &\quad \left. - \frac{1}{2}|F_2|^2\sin(4\psi)\right]\mathrm{Re}(A_x^*A_y) \\ &\quad + \mathrm{Im}(F_0F_2^*)\sin(2\psi)\mathrm{Im}(A_x^*A_y),\end{aligned} \tag{66}$$

$$\begin{aligned}\frac{I_z}{I_0} &= 4|F_1|^2|\cos(\psi)A_x + \sin(\psi)A_y|^2 \\ &= 2|F_1|^2(|A_x|^2 + |A_y|^2) + 2|F_1|^2 \\ &\quad\times\cos(2\psi)(|A_x|^2 - |A_y|^2) \\ &\quad + 4|F_1|^2\sin(2\psi)\mathrm{Re}(A_x^*A_y).\end{aligned} \tag{67}$$

These expressions can be rewritten by using the Stokes parameters, which describe the state of polarization in the entrance pupil:

$$S_0 = |A_x|^2 + |A_y|^2 = 1, \tag{68}$$

$$S_1 = |A_x|^2 - |A_y|^2 = \cos(2\epsilon)\cos(2\theta), \tag{69}$$

$$S_2 = 2\,\mathrm{Re}(A_x^*A_y) = \cos(2\epsilon)\sin(2\theta), \tag{70}$$

$$S_3 = 2\,\mathrm{Im}(A_x^*A_y) = \sin(2\epsilon). \tag{71}$$

Here ϵ is the ellipticity angle and θ is the angle between the long axis of the polarization ellipse and the x axis. The resulting expressions are

$$\begin{aligned}\frac{I_x}{I_0} &= \frac{1}{2}|F_0|^2[1 + \cos(2\epsilon)\cos(2\theta)] + \mathrm{Re}(F_0F_2^*) \\ &\quad\times[\cos(2\epsilon)\cos(2\psi - 2\theta) + \cos(2\psi)] \\ &\quad - \mathrm{Im}(F_0F_2^*)\sin(2\epsilon)\sin(2\psi) + \frac{1}{2}|F_2|^2 \\ &\quad\times[1 + \cos(2\epsilon)\cos(4\psi - 4\theta)],\end{aligned} \tag{72}$$

$$\begin{aligned}\frac{I_y}{I_0} &= \frac{1}{2}|F_0|^2[1 - \cos(2\epsilon)\cos(2\theta)] + \mathrm{Re}(F_0F_2^*) \\ &\quad\times[\cos(2\epsilon)\cos(2\psi - 2\theta) - \cos(2\psi)] \\ &\quad + \mathrm{Im}(F_0F_2^*)\sin(2\epsilon)\sin(2\psi) + \frac{1}{2}|F_2|^2 \\ &\quad\times[1 - \cos(2\epsilon)\cos(4\psi - 4\theta)],\end{aligned} \tag{73}$$

$$\frac{I_z}{I_0} = 2|F_1|^2 + 2|F_1|^2\cos(2\epsilon)\cos(2\psi - 2\theta). \quad (74)$$

The total intensity is then

$$\frac{I}{I_0} = |F_0|^2 + 2|F_1|^2 + |F_2|^2 + 2[|F_1|^2 + \mathrm{Re}(F_0F_2^*)]\cos(2\epsilon)\cos(2\psi - 2\theta). \quad (75)$$

Equations (72)–(75) generalize the results of Richards and Wolf and subsequent authors to arbitrary states of polarization in the entrance pupil. So far, only the special case of a linear polarization ($\epsilon = 0$) oriented along the x axis ($\theta = 0$) has been considered. In that particular case I_y has a fourfold rotation symmetry, with maxima at $\psi = \pi/4$, $3\pi/4$, $5\pi/4$, and $7\pi/4$ and zeros at $\psi = 0$, $\pi/2$, π, and $3\pi/2$, whereas I_z has a twofold rotation symmetry, with maxima at $\psi = 0$ and π and zeros at $\psi = \pi/2$ and $3\pi/2$. It follows from the behavior of the functions F_k for small NA that I_x is of the order unity, I_y is of the order NA^4, and I_z is of the order NA^2. The total intensity as a function of ψ has a maximum in the plane of the polarization and a minimum in the plane perpendicular to this plane, meaning that the spot is elongated in the direction of the polarization. This behavior is also found with the computer program Diffract.[6,7] For a circular polarization ($\epsilon = \pi/4$) it is found that

$$\frac{I_x}{I_0} = \frac{1}{2}(|F_0|^2 + |F_2|^2) + \mathrm{Re}(F_0F_2^*)\cos(2\psi) - \mathrm{Im}(F_0F_2^*)\sin(2\psi), \quad (76)$$

$$\frac{I_y}{I_0} = \frac{1}{2}(|F_0|^2 + |F_2|^2) - \mathrm{Re}(F_0F_2^*)\cos(2\psi) + \mathrm{Im}(F_0F_2^*)\sin(2\psi), \quad (77)$$

$$\frac{I_z}{I_0} = 2|F_1|^2. \quad (78)$$

The dependence on ψ signifies a broken rotational symmetry, which is related to the fact that the point of observation is not necessarily on the optical axis. The total intensity, however, is equal to

$$\frac{I}{I_0} = |F_0|^2 + 2|F_1|^2 + |F_2|^2, \quad (79)$$

for all directions ψ, i.e., the elongation of the spot has disappeared. It follows that this elongation is purely a polarization effect. Numerical calculations show that the NA hardly influences the intensity distribution [Eq. (79)], apart from the scaling contained in the dimensionless coordinates u and v. The most significant change is in the local minima and maxima, which are less sharp for larger NA; i.e., the oscillations in the intensity profile are smoothened. Finally, it is mentioned that the intensities for unpolarized light can be found by setting the Stokes parameters $S_1 = S_2 = S_3 = 0$.

Expressions for the magnetic energy density $\frac{1}{2}\mu_0\mathbf{H}^2 = \frac{1}{2}\epsilon_0(\mu_0 c\mathbf{H})^2$ can be found by making the substitutions $A_x \to -A_y$ and $A_y \to A_x$ and $F_k \to F_k'$. In terms of the polarization parameters this means $\theta \to \theta + \pi/2$, i.e. a rotation of the polarization ellipse over $\pi/2$. This implies that the distribution of the magnetic energy density depends on the F_k' in the same way that the distribution of the electric energy density depends on the F_k, up to an additional rotation over $\pi/2$.

C. Flow of Electromagnetic Energy

The flow of electromagnetic energy is given by the Poynting vector:

$$\mathbf{S} = \mathrm{Re}(\mathbf{E} \times \mathbf{H}^*). \quad (80)$$

It follows from substitution of the known expressions for the electric and magnetic fields that

$$\begin{aligned}\frac{\mu_0 cS_x}{2I_0} &= \mathrm{Re}(-2i\{F_2\sin(2\psi)A_x + [F_0 - F_2\cos(2\psi)]A_y\}F_1'^*[-\sin(\psi)A_x^* + \cos(\psi)A_y^*] \\ &\quad + 2iF_1[\cos(\psi)A_x + \sin(\psi)A_y]\{[F_0'^* - F_2'^*\cos(2\psi)]A_x^* - F_2'^*\sin(2\psi)A_y^*\}) \\ &= \mathrm{Im}[(F_0 - F_2)F_1'^* - F_1(F_0'^* - F_2'^*)]\cos(\psi) - \mathrm{Re}[(F_0 + F_2)F_1'^* + F_1(F_0'^* + F_2'^*)]\sin(2\epsilon)\sin(\psi) - \mathrm{Im}(F_0F_1'^* \\ &\quad + F_1F_0'^*)\cos(2\epsilon)\cos(\psi - 2\theta) + \mathrm{Im}(F_0F_2'^* + F_2F_0'^*)\cos(2\epsilon)\cos(3\psi - 2\theta), \end{aligned} \quad (81)$$

$$\begin{aligned}\frac{\mu_0 cS_y}{2I_0} &= \mathrm{Re}(-2iF_1[\cos(\psi)A_x + \sin(\psi)A_y] \\ &\quad \times\{F_2'^*\sin(2\psi)A_x^* - [F_0'^* + F_2'^*\cos(2\psi)]A_y^*\} \\ &\quad + 2i\{[F_0 + F_2\cos(2\psi)]A_x \\ &\quad + F_2\sin(2\psi)A_y\}F_1'^*[-\sin(\psi)A_x^* + \cos(\psi)A_y^*]) \\ &= \mathrm{Im}[(F_0 - F_2)F_1'^* - F_1(F_0'^* - F_2'^*)]\sin(\psi) \\ &\quad + \mathrm{Re}[(F_0 + F_2)F_1'^* + F_1(F_0'^* \\ &\quad + F_2'^*)]\sin(2\epsilon)\cos(\psi) + \mathrm{Im}[F_0F_1'^* \\ &\quad + F_1F_0'^*]\cos(2\epsilon)\sin(\psi - 2\theta) + \mathrm{Im}[F_0F_2'^* \\ &\quad + F_2F_0'^*]\cos(2\epsilon)\sin(3\psi - 2\theta), \end{aligned} \quad (82)$$

$$\begin{aligned}\frac{\mu_0 cS_z}{2I_0} &= \mathrm{Re}(\{[F_0 + F_2\cos(2\psi)]A_x + F_2\sin(2\psi)A_y\} \\ &\quad \times \{[F_0'^* - F_2'^*\cos(2\psi)]A_x^* - F_2'^*\sin(2\psi)A_y^*\} \\ &\quad - \{F_2\sin(2\psi)A_x + [F_0 - F_2\cos(2\psi)]A_y\} \\ &\quad \times \{F_2'^*\sin(2\psi)A_x^* - [F_0'^* + F_2'^*\cos(2\psi)]A_y^*\}) \\ &= \mathrm{Re}(F_0F_0'^* - F_2F_2'^*) - \mathrm{Re}(F_0F_2'^* \\ &\quad - F_2F_0'^*)\cos(2\epsilon)\cos(2\psi - \theta). \end{aligned} \quad (83)$$

These equations generalize those of Richards and Wolf[3,5] to describe the effects of arbitrary states of polarization in

the entrance pupil, axial birefringence, and of focusing through a cover layer. The x and y components of the Poynting vector can be rewritten in terms of a radial and a tangential component, giving:

$$\frac{\mu_0 c S_r}{2I_0} = \cos(\psi)\frac{\mu_0 c S_x}{2I_0} + \sin(\psi)\frac{\mu_0 c S_y}{2I_0}$$
$$= \mathrm{Im}[(F_0 - F_2)F_1'^* - F_1(F_0'^* - F_2'^*)]$$
$$- \mathrm{Im}[(F_0 - F_2)F_1'^* + F_1(F_0'^* - F_2'^*)]$$
$$\times \cos(2\epsilon)\cos(2\psi - 2\theta), \quad (84)$$

$$\frac{\mu_0 c S_\psi}{2I_0} = -\sin(\psi)\frac{\mu_0 c S_x}{2I_0} + \cos(\psi)\frac{\mu_0 c S_y}{2I_0}$$
$$= \mathrm{Re}[(F_0 + F_2)F_1'^* + F_1(F_0'^* + F_2'^*)]\sin(2\epsilon)$$
$$+ \mathrm{Im}[(F_0 + F_2)F_1'^* + F_1(F_0'^* + F_2'^*)]$$
$$\times \cos(2\epsilon)\sin(2\psi - 2\theta), \quad (85)$$

$$\frac{\mu_0 c S_z}{2I_0} = \mathrm{Re}(F_0 F_0'^* - F_2 F_2'^*) - \mathrm{Re}(F_0 F_2'^* - F_2 F_0'^*)$$
$$\times \cos(2\epsilon)\cos(2\psi - 2\theta). \quad (86)$$

Several new terms are introduced here compared with the equations of Richards and Wolf.[3,5] The terms involving the azimuthal angle ψ appear owing to the difference between the electric and magnetic functions F_k and F_k', which in turn is rooted in the difference between the p and s Fresnel coefficients. Clearly, these terms represent effects of the interface in $z = -d$. The effect is zero for circularly polarized light ($\epsilon = \pm\pi/4$) and maximum for linearly polarized light. Calculations for $n_1 = 1$, $n_2 = 1.5$, NA $= 0.85$, $\Delta n = 0$ (no birefringence effects), and $\epsilon = 0$ (linear polarization) show that the difference between the functions F_k and F_k' is of the order 10^{-3} or less. The additional flow terms due to the difference between the F_k and F_k' are of the order 10^{-3} or less as well. Consequently, these terms may be neglected, leading to the approximation

$$\frac{\mu_0 c S_r}{2I_0} = 2\,\mathrm{Im}[(F_0 - F_2)F_1^*], \quad (87)$$

$$\frac{\mu_0 c S_\psi}{2I_0} = 2\,\mathrm{Re}[(F_0 + F_2)F_1^*]\sin(2\epsilon), \quad (88)$$

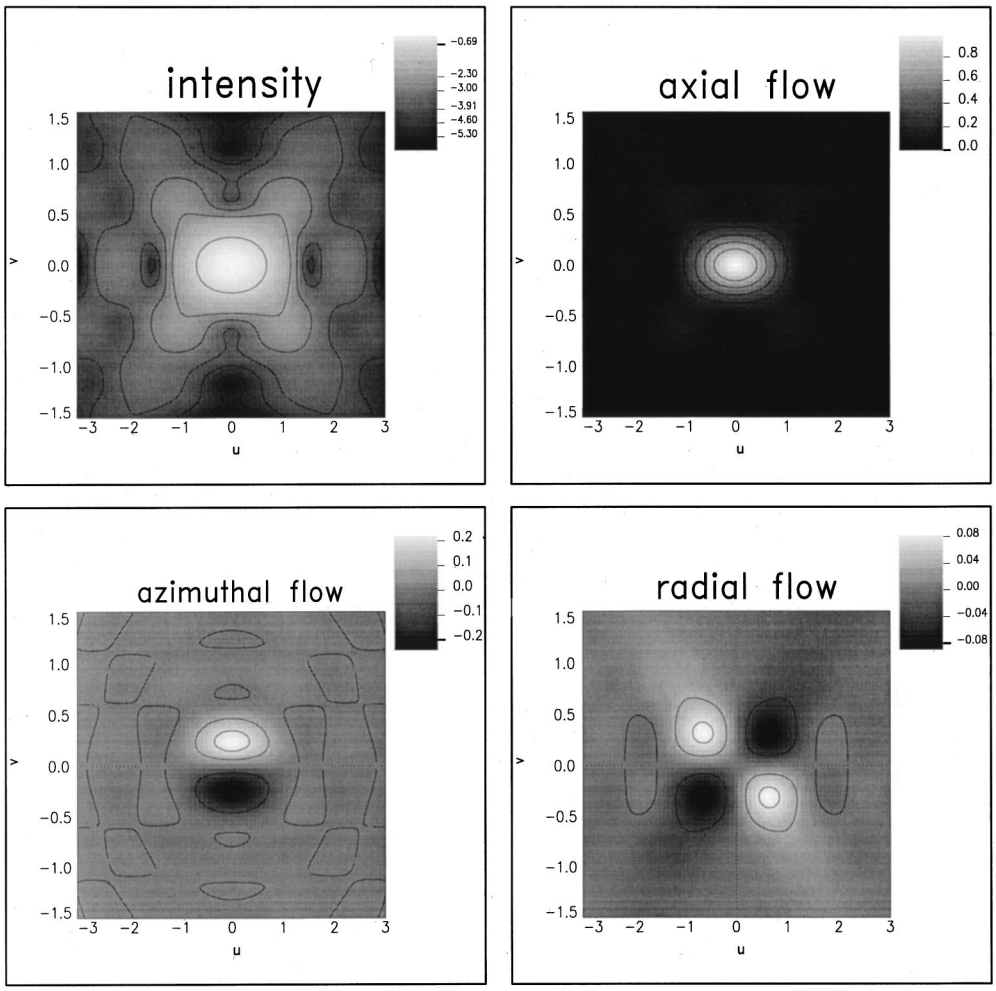

Fig. 2. Intensity on a logarithmic scale (upper left), and the axial (upper right), radial (lower left), and azimuthal (lower right) flow components in the focal region for NA $= 0.85$, $n_1 = n_2 = 1$, $\Delta n = 0$, and for circular polarization in the entrance pupil.

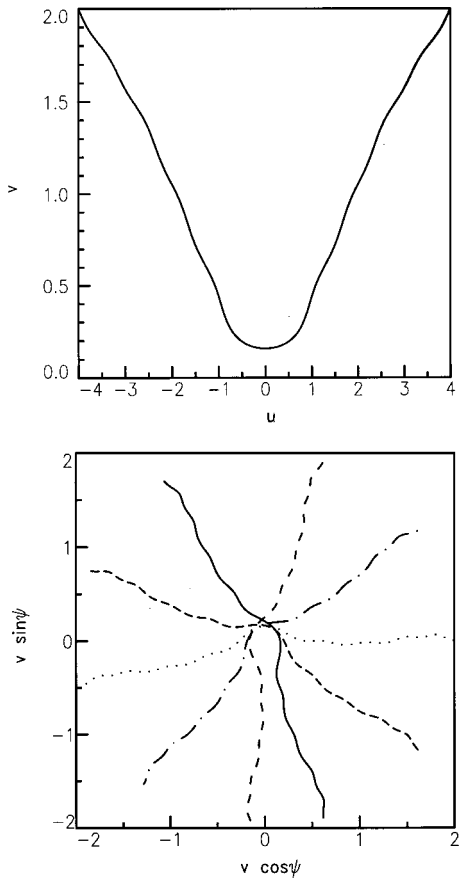

Fig. 3. Projection of flow lines on the meridional plane (left) and on the focal plane (right) for NA = 0.85, $n_1 = n_2 = 1$, $\Delta n = 0$, and $\epsilon = \pi/4$ (circular polarization). Starting points of the flow lines are on the circle $v = 2$ in the plane $u = -4$. The flow lines make a rotation around the optical axis of approximately π in the focal region.

$$\frac{\mu_0 c S_z}{2I_0} = |F_0|^2 - |F_2|^2. \tag{89}$$

The absence of the azimuthal angle ψ in these expressions implies that the flow pattern is invariant under rotations around the optical axis.

The remaining contribution to the tangential component S_ψ is absent for linearly polarized light and maximum for circularly polarized light. It therefore describes the effects of ellipticity. This term has remained unnoticed by previous authors, as they have only considered linearly polarized light. Ellipticity has the consequence that the Poynting vector is no longer confined to the meridional plane. It follows that the electromagnetic energy flows through the focal region similar to the flow of air in a tornado. This effect can be understood physically from the fact that elliptically polarized light carries angular momentum. Figure 2 shows the distribution of intensity and of the axial, radial, and azimuthal components of the Poynting vector in the focal region for NA = 0.85, $n_1 = n_2 = 1$, $\Delta n = 0$, and $\epsilon = \pi/4$ (circular polarization). Clearly, there is a toroidal region around the focal plane with relatively large azimuthal flow. The maximum is close to the circle $v = 0.24$, which is roughly halfway between the optical axis and the first minimum of the scalar Airy distribution at $v = 0.61$. Figure 3 shows projections of flow lines on the meridional and focal planes for the same parameters as used for Fig. 2. Flow lines $\mathbf{r}(t) = [u(t), v(t), \psi(t)]$ satisfy the set of first-order differential equations:

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = l\,\frac{\mu_0 c \mathbf{S}(\mathbf{r})}{2I_0}, \tag{90}$$

where t is a dimensionless number parametrizing the flow line and l is a unit of length. This differential equation can easily be integrated with the Runge–Kutta method. The projection on the meridional plane is the line $[u(t),v(t)]$; the projection on the focal plane is the line $[v(t)\cos\psi(t),\ v(t)\sin\psi(t)]$. Starting points of the flow lines in Fig. 3 are on the circle with radius $v = 2$ in the plane $u = -4$. The flow lines converge on the optical axis for $u < 0$ and diverge from the optical axis for $u > 0$. In the focal region close to $u = 0$ the flow lines spiral around the optical axis, making a nearly π rotation. The rotation takes place largely in the toroidal region with the large azimuthal Poynting vector component indicated in Fig. 2. This spiraling of flow lines around the optical axis stands in contrast to the linearly polarized case where the azimuthal angle ψ does not change at all.

It was originally noted by Wolf and co-workers[3,5] that there are regions of anomalous flow, i.e., regions where the electromagnetic energy flows backwards toward the focusing lens. These regions are defined by a negative Poynting vector z component. The flow pattern in the neighborhood of this region is quite extraordinary when the polarization in the entrance pupil is circular and can be analyzed as follows. First it is noted that in the focal plane $u = 0$ the functions F_k are real. It then follows that the radial flow is zero ($S_r = 0$). The axial flow S_z is zero if $F_0 = F_2$ or $F_0 = -F_2$. The points closest to the optical axis for which this is the case are $v = 0.57$ and $v = 0.65$, respectively (for $n_1 = n_2 = 1$ and NA = 0.85). These points are close to the first minimum of the Airy distribution of scalar diffraction theory at $v = 0.61$. For

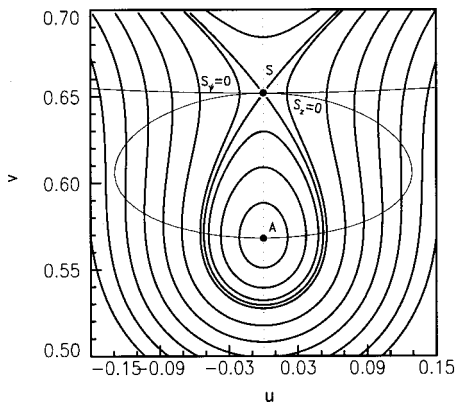

Fig. 4. Anomalous flow projected on the meridional plane close to the first minimum of the scalar diffraction Airy pattern. The singular point with zero Poynting vector is indicated by "S"; the point where the flow is purely azimuthal is indicated by "A." The thin curves have zero axial flow ($S_z = 0$) and zero azimuthal flow ($S_\psi = 0$). The line with zero radial flow ($S_r = 0$) is dotted and coincides with the focal plane $u = 0$. Flow lines within the toroidal region defined by the flow lines through S spiral around the circle through A.

Table 1. Sensitivity for Vertical Birefringence for the Three Optical Recording Standards[a]

Standard	λ (nm)	NA	d (mm)	n_o	B	Δn_{dl}	$\Delta n_{\mathrm{dl/10}}$
CD	780	0.45	1.2	1.5	6.31×10^4	1.8×10^{-3}	5.6×10^{-4}
DVD	660	0.60	0.6	1.5	6.96×10^4	1.7×10^{-3}	5.4×10^{-4}
DVR	400	0.85	0.1	1.5	2.12×10^4	3.1×10^{-3}	9.7×10^{-4}

[a] The coefficient B is the coefficient in the expression for the Strehl ratio $S = 1 - B\Delta n^2$, Δn_{dl} is the birefringence for which the spot is diffraction limited, and $\Delta n_{\mathrm{dl/10}}$ is the birefringence for which $S = 0.98$.

$v = 0.65$ the azimuthal flow S_ψ is zero as well. Clearly, the circle in the focal plane with radius 0.65 is a line with singular flow, i.e. a line where $\mathbf{S} = 0$. For $v = 0.57$ the flow is purely azimuthal. As a consequence, the circle in the focal plane with radius 0.57 is a flow line. The singular point S, the point with purely azimuthal flow A, and the lines $S_z = 0$ and $S_\psi = 0$ are shown in Fig. 4. Projection of the flow lines on the meridional plane are also drawn in Fig. 4. The flow lines through S enclose a region with toroidal topology in which the flow lines spiral around the circle through A. These flow lines are confined to the region with toroidal topology and form closed loops when the number of turns around A per revolution around the optical axis is a rational fraction. This is only the case for some particular values of the NA. Finally, it is mentioned that similar regions with anomalous flow can be found further away from the optical axis.

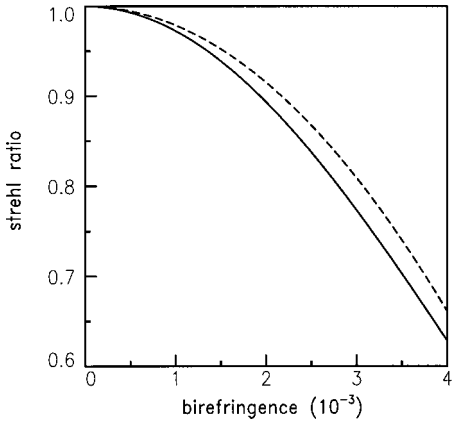

Fig. 5. Strehl ratio for NA = 0.85, λ = 400 nm, $n_1 = 1$, $n_2 = 1.5$, and $d = 100\ \mu$m, according to the analytical approximation [Eq. (98)] (solid curve) and according to exact numerical calculations (dashed curve).

4. EFFECTS OF BIREFRINGENCE

A. Focus Shift, Strehl Ratio, and Spot Width

We now turn to the case with nonzero birefringence. It is assumed that the focusing lens is corrected for aberrations introduced by an isotropic cover layer of thickness d and refractive index $n_2 = n_o$. Then

$$W_s = 0, \tag{91}$$

$$W_p = \Delta W = \frac{kd\Delta n \rho^2 \mathrm{NA}^2/{n_2}^2}{(1 - \rho^2 \mathrm{NA}^2/{n_2}^2)^{1/2}}. \tag{92}$$

Expanding ΔW in the radial pupil coordinate ρ gives to lowest order a defocus term

$$\Delta W = kd\Delta n(\mathrm{NA}/n_2)^2\rho^2 + \ldots . \tag{93}$$

Clearly, the spot arising from the radial polarization field in the exit pupil is defocused compared with the spot arising from the tangential polarization field. It follows that for large values of the birefringence the spot is split into two parts, a radial spot and a tangential polarization spot. For small values of the birefringence the two spots overlap sufficiently to regard them as a single, broadened spot. The separation of the two spots is equal to $2d\Delta n/n_2$, as the relative defocus is proportional to $kd\Delta n(\mathrm{NA}/n_2)^2$. This means that the maximum of the combined spot is shifted over a distance $d\Delta n/n_2$. The defocus is then evenly distributed over both spots. A measure of spot broadening is the deviation of the Strehl ratio from one. Because both spots are defocused relative to the point with maximum intensity this will be proportional to $(kd\Delta n)^2(\mathrm{NA}/n_2)^4$.

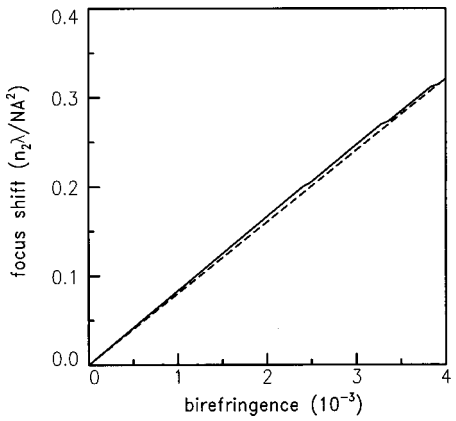

Fig. 6. Focus shift for NA = 0.85, λ = 400 nm, $n_1 = 1$, $n_2 = 1.5$, and $d = 100\ \mu$m, according to the analytical approximation [Eq. (97)] (solid curve) and according to exact numerical calculations (dashed curve).

An explicit calculation in the limit of small Δn and small NA confirms this qualitative argument. To calculate the focus shift and Strehl ratio the intensity distribution must be calculated. The point of maximum intensity (focus) must be on the optical axis because of rotational symmetry. The relevant function is then the integral $F_0(u, 0)$. This integral can be evaluated numerically,[22] but an analytical treatment appears to be quite possible for small values of the numerical aperture and birefringence. In that limiting case,

$$F_0(u,0) = \int_0^1 d\rho\rho\left[1 + \pi i(2p - u)\rho^2 - \frac{1}{2}\pi^2 \times (2p - u)^2\rho^4 + 1 - \pi i u\rho^2 - \frac{1}{2}\pi^2 u^2\rho^4\right] = 1 + i\frac{\pi}{2}(p - u) - \frac{\pi^2}{12}[(2p - u)^2 + u^2], \tag{94}$$

where the quantity p is defined by

$$p = \frac{d\Delta n}{\lambda}\frac{\mathrm{NA}^2}{{n_2}^2}. \tag{95}$$

Terms of order NA^2 and NA^4 that do not depend on u or p are left out, as they divide out in the expression for the Strehl ratio (the ratio of the intensity and the maximum intensity in the zero birefringence case). This Strehl ratio follows as

$$S = 1 - \left[\frac{\pi^2}{6}(2p - u)^2 + \frac{\pi^2}{6}u^2 - \frac{\pi^2}{4}(p - u)^2\right] = 1 - \frac{\pi^2}{12}(5p^2 - 2pu + u^2). \tag{96}$$

Maximum intensity is obtained when $u = p$, meaning that the focus is no longer in the plane $z = 0$, but instead is shifted over a distance

$$\Delta z = \frac{d\Delta n}{n_2}. \tag{97}$$

At this shifted focus point the Strehl ratio is then

$$S = 1 - \frac{\pi^2}{3}p^2 = 1 - \frac{1}{12}\left(\frac{2\pi d\Delta n}{\lambda}\right)^2\left(\frac{\mathrm{NA}}{n_2}\right)^4. \tag{98}$$

The results for the focus shift and the Strehl ratio agree with the qualitative argument. The spot is diffraction limited (Strehl ratio 0.8) if the retardance satisfies

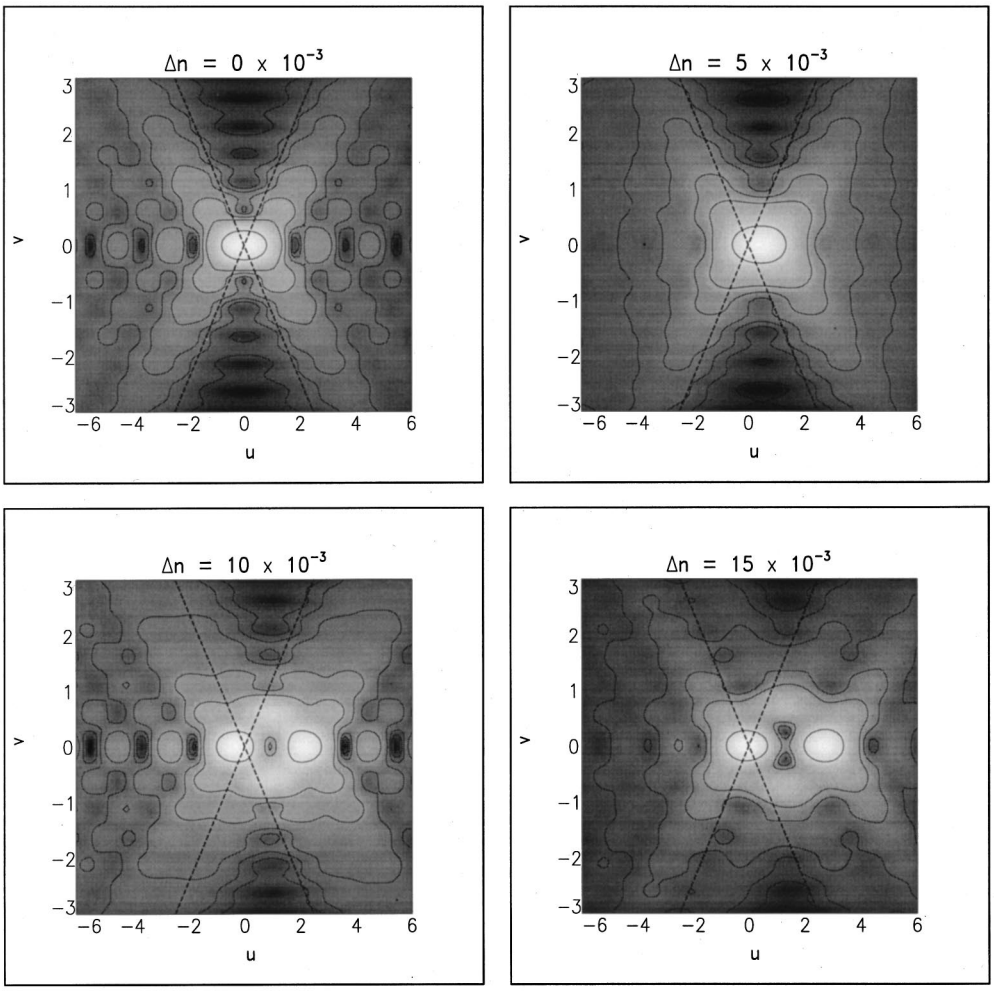

Fig. 7. Intensity distribution in the meridional plane for $\Delta n = 0$ (upper left), $\Delta n = 5 \times 10^{-3}$ (upper right), $\Delta n = 10 \times 10^{-3}$ (lower left), and $\Delta n = 15 \times 10^{-3}$ (lower right). The intensity increases from black to white. The full lines are isophotes for intensities 0.50, 0.10, 0.05, 0.02, 0.01, 0.005, and 0.001; the dashed lines indicate the geometrical light cone $v = un_2/(n_2^2 - \mathrm{NA}^2)^{1/2}$. The parameters used in the calculation are NA = 0.85, $\lambda = 400$ nm, $n_1 = 1$, $n_2 = 1.5$, and $d = 100\ \mu$m.

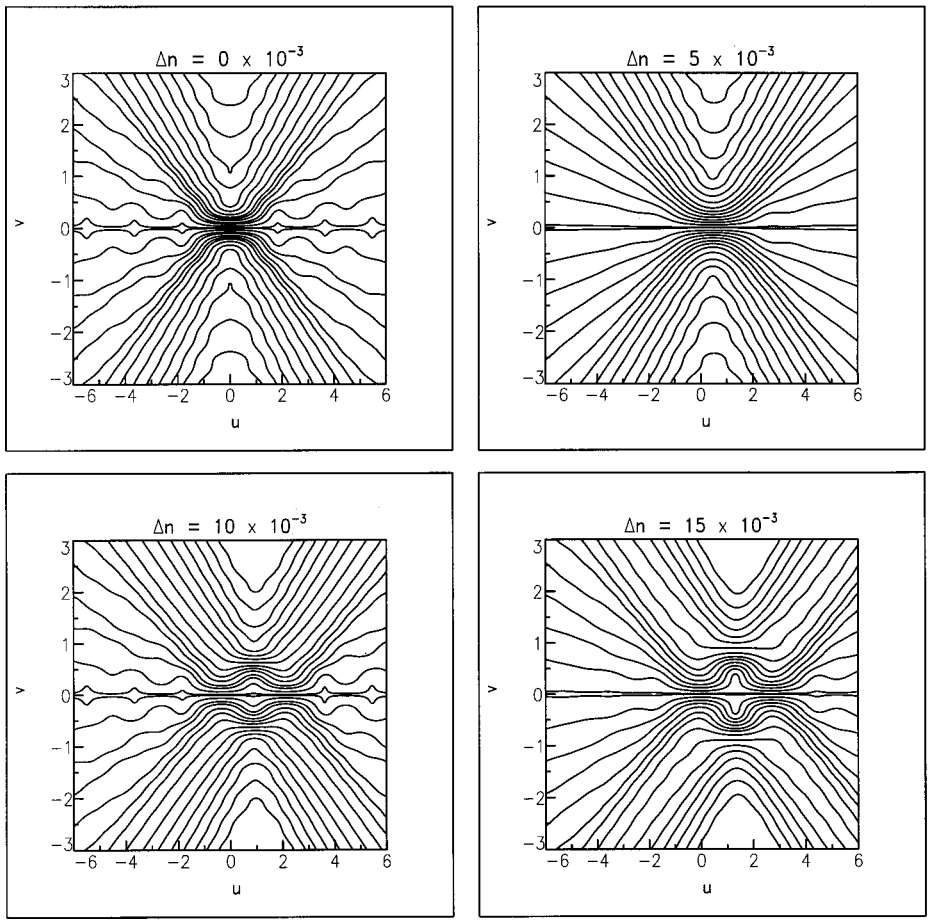

Fig. 8. Projection of flow lines on the meridional plane for circular polarization in the entrance pupil with $\Delta n = 0$ (upper left), $\Delta n = 5 \times 10^{-3}$ (upper right), $\Delta n = 10 \times 10^{-3}$ (lower left), and $\Delta n = 15 \times 10^{-3}$ (lower right). The parameters used in the calculation are NA = 0.85, $\lambda = 400$ nm, $n_1 = 1$, $n_2 = 1.5$, and $d = 100\ \mu$m. These are the same parameters used in the calculations of the intensity distributions shown in Fig. 7.

$$d\Delta n \leqslant d\Delta n_{\mathrm{dl}} = \sqrt{\frac{12}{5}}\,\frac{\lambda}{2\pi}\left(\frac{n_2}{\mathrm{NA}}\right)^2. \tag{99}$$

The birefringence for which 10% of the diffraction limited margin is taken (Strehl ratio 0.98) seems a suitable candidate for the maximum tolerable birefringence in a practical system. This birefringence value is $\Delta n_{\mathrm{dl/10}} = \Delta n_{\mathrm{dl}}/\sqrt{10}$.

The optical recording standards CD, DVD, and DVR may be compared to each other in terms of the sensitivity for vertical birefringence in the cover layer. With use of Eq. (98) this sensitivity can be easily quantified. Table 1 shows the coefficient of Δn^2 in the expression for the Strehl ratio, the birefringence for which the spot is diffraction limited and the birefringence for which 10% of the diffraction-limited margin is taken. It turns out that the CD and DVD standards are comparable in birefringence sensitivity, whereas the newly proposed DVR standard has a distinctly lower sensitivity. The values for the maximum tolerable birefringence are in the range 5×10^{-4}–10×10^{-4}.

The analytical approximations leading to Eqs. (97) and (98) may be compared with exact numerical calculations of the focus shift and Strehl ratio as a function of birefringence. Figs. 5 and 6 show the exact and approximated Strehl ratio and focus shift, respectively, for the parameters pertaining to the DVR optical recording standard (see Table 1). Quantitative agreement between the exact and approximate focus shift and Strehl ratio cannot be expected *a priori* because of the high NA value of 0.85. It turns out that the results defy this expectation as the quantitative agreement is quite satisfactory, especially for the focus shift.

B. Isophotes and Flow Lines

Figure 7 shows the intensity distribution in the meridional plane for different values of the axial birefringence. The full lines are different isophotes (lines of equal intensity). The input parameters of the calculations are the parameters pertaining to the DVR optical recording standard. The polarization in the entrance pupil was taken to be circular. For $\Delta n = 5 \times 10^{-3}$ (just above the diffraction limit 3×10^{-3}) the spot is broadened and deformed compared with the zero birefringence case. For $\Delta n = 10 \times 10^{-3}$ the tangentially and radially polarized spots are separated. The tangentially polarized spot (ordinary mode) is centered at the geometrical focus $u = v = 0$, whereas the radially polarized spot (extraordinary mode) is shifted. The separation of the two spots increases with Δn, as follows from comparison of the fig-

ures for $\Delta n = 10 \times 10^{-3}$ and $\Delta n = 15 \times 10^{-3}$. The results of these numerical calculations are in agreement with the discussion in the previous subsection. In case of a linear polarization in the entrance pupil the elongation of the spot is different for the tangentially and radially polarized spots, leading to an almost circular spot halfway between the two spots.[23] This resembles the light distribution of an astigmatic spot in the scalar diffraction theory.

The axial birefringence gives rise to a distorted flow profile in the focal region. Figure 8 shows the projection of the three-dimensional flow lines $[u(t), v(t), \psi(t)]$ on the meridional plane, i.e. the lines $[u(t),v(t)]$, for different values of the birefringence, starting from equidistant points on the line $u = -6$, $\psi = 0$. The flow lines do not give information about the nonzero azimuthal component of the Poynting vector related to the circular polarization of the light. However, Fig. 8 clearly shows the distortion of the flow profile by the birefringence. The broadening and reduction of intensity for $\Delta n = 5 \times 10^{-3}$ compared with $\Delta n = 0 \times 10^{-3}$ (upper right picture compared with upper left picture) is clearly visible from the decreased concentration of flow lines in the central region. The subsequent splitting of the focus as the birefringence is increased further to $\Delta n = 10 \times 10^{-3}$, and $\Delta n = 15 \times 10^{-3}$ is evident from the lower two pictures.

5. SUMMARY AND CONCLUSION

The influence of axial birefringence on the distribution and flow of light close to focus is investigated with use of vector diffraction theory. Integral expressions for the electric and magnetic fields in the focal region are derived. The Debye approach is followed and the so-called small birefringence approximation is introduced. The latter approximation is quite justified in many practical cases, such as the readout of optical disks through a birefringent cover layer. Expressions for the intensity and the Poynting vector are derived on the basis of the diffraction integrals; these expressions are valid for arbitrary states of polarization in the entrance pupil of the optical system. For the case of circular polarization the light distribution is rotationally symmetric, and the Poynting vector has a tangential component, leading to a tornado-like flow of energy in the focal region.

The focal spot is the sum of contributions of the field that is radially polarized and the field that is tangentially polarized in the exit pupil. The radial polarization field corresponds to the p or extraordinary wave, and the tangential polarization field corresponds to the s or ordinary wave. Axial birefringence introduces a relative defocus of the tangential and radial polarization spots. For small values of the birefringence this causes a deformation, broadening, and displacement of the spot; for large values the two spots separate completely. Analytical expressions for the Strehl ratio and the focal shift are derived that are strictly valid for small NA and birefringence only. Numerical calculations show that they are quantitatively valid, even for large NA. The expression for the Strehl ratio allows for an estimate of the maximum tolerable birefringence. The spot is diffraction limited if the Strehl ratio equals 0.8. Taking 10% of the diffraction-limited margin the maximum tolerable birefringence is the value for which the Strehl ratio equals 0.98. It turns out that for the optical recording standards CD, DVD, and DVR, the maximum allowable birefringence is in the range $0.5–1.0 \times 10^{-3}$.

Further work along the lines presented in this paper is quite possible. Modifications of the p and s Fresnel coefficients due to focusing through an interference stack or through a gap with subwavelength thickness (allowing for evanescent wave coupling) are relatively simple. Analytical treatment of the azimuthal part of the diffraction integral is no longer possible when the rotational symmetry of the system is broken by aberrations such as coma or astigmatism, or by biaxial birefringence. These nonrotationally symmetric cases require a fully numerical treatment of the diffraction integral, which may complicate matters considerably.

ACKNOWLEDGMENTS

Martijn Dekker is thanked for discussions on the subject of birefringence in optical disk readout.

S. Stallinga can be reached at the address on the title page or by e-mail at sjoerd.stallinga@philips.com.

REFERENCES

1. T. Narahara, S. Kobayashi, M. Hattori, Y. Shimpuku, G. J. van den Enden, J. A. H. M. Kahlman, M. van Dijk, and R. van Woudenberg, "Optical disc system for digital video recording," Jpn. J. Appl. Phys. **39**, 912–919 (1959).
2. 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).
3. 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).
4. A. Boivin and E. Wolf, "Electromagnetic field in the neighborhood of the focus of a coherent beam," Phys. Rev. B **138**, 1561–1565 (1965).
5. A. Boivin, J. Dow, and E. Wolf, "Energy flow in the neighborhood of the focus of a coherent beam," J. Opt. Soc. Am. **57**, 1171–1175 (1967).
6. M. Mansuripur, "Distribution of light at and near the focus of high-numerical-aperture objectives," J. Opt. Soc. Am. A **3**, 2086–2093 (1986).
7. M. Mansuripur, "Certain computational aspects of vector diffraction problems," J. Opt. Soc. Am. A **6**, 786–805 (1989).
8. H. Ling and S.-W. Lee, "Focusing of electromagnetic waves through a dielectric interface," J. Opt. Soc. Am. A **1**, 965–973 (1984).
9. P. Török, P. Varga, Z. Laczik, and G. R. Booker, "Electromagnetic diffraction of light focused through a planar interface between materials of mismatched refractive indices: an integral representation," J. Opt. Soc. Am. A **12**, 325–332 (1995).
10. P. Török, P. Varga, and G. R. Booker, "Electromagnetic diffraction of light focused through a planar interface between materials of mismatched refractive indices: structure of the electromagnetic field. I," J. Opt. Soc. Am. A **12**, 2136–2144 (1995).
11. P. Török, P. Varga, and G. Németh, "Analytical solution of the diffraction integrals and interpretation of wave-front distortion when light is focused through a planar interface between materials of mismatched refractive indices," J. Opt. Soc. Am. A **12**, 2660–2671 (1995).
12. S. H. Wiersma and T. D. Visser, "Defocusing of a converging

electromagnetic wave by a plane dielectric interface," J. Opt. Soc. Am. A **13**, 320–325 (1996).
13. S. H. Wiersma, P. Török, T. D. Visser, and P. Varga, "Comparison of different theories for focusing through a plane interface," J. Opt. Soc. Am. A **14**, 1482–1490 (1997).
14. V. Dhayalan and J. J. Stamnes, "Focusing of electromagnetic waves into a dielectric slab: I. exact and asymptotic results," Pure Appl. Opt. **6**, 33–52 (1997).
15. D. G. Flagello, T. Milster, and A. E. Rosenbluth, "Theory of high-NA imaging in homogeneous thin films," J. Opt. Soc. Am. A **13**, 53–64 (1996).
16. J. J. Stamnes and D. Jiang, "Focusing of electromagnetic waves into a uniaxial crystal," Opt. Commun. **150**, 251–262 (1998).
17. D. Jiang and J. J. Stamnes, "Numerical and asymptotic results for focusing of two-dimensional waves in uniaxial crystals," Opt. Commun. **163**, 55–71 (1999).
18. D. Jiang and J. J. Stamnes, "Numerical and experimental results for focusing of two-dimensional electromagnetic waves into uniaxial crystals," Opt. Commun. **174**, 321–334 (2000).
19. P. Yeh, *Optical Waves in Layered Media* (Wiley, New York, 1988).
20. J. J. Stamnes, *Waves in Focal Regions* (Hilger, Bristol, UK 1986).
21. M. Mansuripur, "Distribution of light at and near the focus of high-numerical-aperture objectives: erratum; Certain computational aspects of vector diffraction problems: erratum," J. Opt. Soc. Am. A **10**, 382–383 (1993).
22. A. B. Marchant, "Cover sheet aberrations in optical recording," in *Optical Disk Systems and Applications*, E. V. LaBudde, ed., Proc. SPIE **421**, 43–49 (1983).
23. M. Mansuripur, *The Physical Principles of Magneto-Optical Recording* (Cambridge U. Press, Cambridge, UK, 1995).