

Extended Nijboer-Zernike approach to aberration and birefringence retrieval in a high-numerical-aperture optical system

Joseph J.M. Braat

*Optics Research Group, Faculty of Applied Sciences, Delft University of Technology,
Lorentzweg 1, NL-2628 CJ Delft, The Netherlands*

Peter Dirksen

Philips Research Laboratories, WAG-02, NL-5656 AA Eindhoven, The Netherlands

Augustus J.E.M. Janssen

Philips Research Laboratories, WO-02, NL-5656 AA Eindhoven, The Netherlands

Arthur S. van de Nes

*Optics Research Group, Faculty of Applied Sciences, Delft University of Technology,
Lorentzweg 1, NL-2628 CJ Delft, The Netherlands*

Abstract

The judgment of the imaging quality of an optical system can be carried out by examining its through-focus intensity distribution. It has been shown in a previous paper that a scalar-wave analysis of the imaging process according to the extended Nijboer-Zernike theory allows the retrieval of the complex pupil function of the imaging system, including aberrations *and* transmission variations. However, the applicability of the scalar analysis is limited to systems with an NA-value of the order of 0,60 or less; beyond these values polarization effects become significant. In this scalar retrieval method, the complex pupil function is represented by means of the coefficients of its expansion in a series involving the Zernike polynomials. This representation is highly efficient, in terms of number and magnitude of the required coefficients, and lends itself quite well for matching procedures in the focal region. This distinguishes the method from the retrieval schemes in the literature, that are normally not based on Zernike-type expansions, and rather rely on point-by-point matching procedures.

In a previous paper, *J. Opt. Soc. Am.* **A20**, 2281-2292 (2003), we have incorporated the extended Nijboer-Zernike approach into the Ignatowsky - Richards/Wolf formalism for the vectorial treatment of optical systems with high NA. In the present paper we further develop this approach by defining an appropriate set of functions

that describe the energy density distribution in the focal region. Using this more refined analysis, we establish the set of equations that allow the retrieval of aberrations and birefringence from the intensity point-spread function in the focal volume for high-NA systems. It is shown that one needs four analyses of the intensity distribution in the image volume with different states of polarization in the entrance pupil. Only in this way it will be possible to retrieve the 'vectorial' pupil function that includes the effects of birefringence induced by the imaging system. A first numerical test example is presented that illustrates the importance of using the vectorial approach and the correct NA-value in the aberration retrieval scheme.

OCIS codes: 000.3860, 050.1960, 100.3190, 100.5070, 110.2990, 120.4820

1. Introduction

The characterization and the control of the quality of optical imaging systems with a high numerical aperture, like microscope objectives and projection imaging systems, is of great practical importance. This type of high-quality imaging systems is encountered in a manufacturing environment like that of the semiconductor industry where very precise projection lenses are used to define lines and spacings on Si-wafers well below 100 nm in lateral size and that are crucial for the functioning of advanced computer processors and memories. Once applied to the silicon wafer, the characterization of these very fine features also asks for high-quality optical inspection devices working well within the diffraction limit. A reliable description of the residual aberrations of these optical inspection systems is needed to successfully enable the precise reconstruction of the details of the features on the wafer. In practice, we also notice that the reliable production of advanced integrated circuits requires a constant quality monitoring of the optical projection lenses used in microlithography; this applies to the manufacturing process itself but it also has to be repeated during the lifetime of the apparatus in order to prevent any drift in the manufacturing conditions.

A wide-spread classical method for quality control of an optical system relies on interferometry to derive the wavefront function in the exit pupil of the optical system¹. A practical drawback of this method is the special requirement on the source: an at-wavelength coherent source should be available which is not always easily realized. An interferometric method also asks for the insertion of special optical components to realize the wavefront or amplitude splitting that is needed for interferometric measurements.

An alternative to interferometry is the measurement of the intensity impulse response of the imaging system. In the literature many papers have been published where the intensity distributions in the image plane are measured in the presence of a known object. In some cases, a simultaneous measurement of the intensity in the exit pupil is included. Using the a priori object information in analyzing the image, algorithms enable the reconstruction of the complex lens function, including the aberration phase²⁻¹² of the imaging system (inversion). The stability of these algorithms is not always guaranteed because of inherent non-uniqueness problems. Information obtained from a focal plane intensity pattern has to be transformed into a complex

lens pupil function. Noise in the image capturing process tends to make the inversion process unreliable and certain frequency bands can be irrevocably lost. Wiener filtering is often used to stabilize the inversion process but this method introduces an arbitrary parameter to the inversion process. Other approaches to retrieve the phase of the complex lens function use least-squares method or other optimization methods^{13–17} to compute the complex lens function that best fits the problem.

We propose an alternative method that is more practical and easily adapted to e.g. the on-line quality measurement of projection lenses: starting from a quasi point-source, we directly analyze the spatial intensity in the image *volume* of the imaging system. In some previous papers, the authors have described this method that enables the retrieval of the complex lens function from intensity data collected through the focal volume of the imaging system¹⁸. To this goal they developed and used a parametric semi-analytic description of the intensity in the focal region, the so-called *extended* Nijboer-Zernike theory^{19,20} of diffraction. While the analytic results from the classical Nijboer-Zernike theory were limited in practice to near best-focus image planes, the extended analysis enables analytic results that are valid and computationally reliable in an extended focal volume. By using this extended theory and the immediate relationship that is established in this theory between the complex field in the exit pupil and the intensity distribution in the focal volume, we can establish a relatively simple set of equations for the Zernike coefficients. These equations use the measured intensity distribution in the image volume as an input for a 'matching' operation with respect to the analytically calculated intensity distribution determined by the still unknown complex Zernike coefficients. The solution of this system of equations yields an effective representation of the complex pupil function of the imaging system even when a relatively small set of Zernike coefficients is used in the 'matching' process.

A. *From scalar-wave to vectorial-wave imaging in the presence of an aberrated optical system*

In the case of low-aperture scalar imaging the retrieval of the complex Zernike coefficients enabled a full description of the wavefront aberration of the optical system *and* of the possibly spatially varying transmission function of the lens system. The treatment of the imaging by high-NA optical systems first requires the extension of the forward calculated intensity pattern in the focal volume from the scalar case to the high numerical aperture vectorial case. With respect to the scalar imaging theory, several refinements are needed:

- The vectorial nature of the problem requires the calculation of the electromagnetic field vectors in the focal volume. The aberration-free case has been thoroughly examined in two well-known publications^{21,22}. The case of aberrated imaging systems^{23–27} has been analytically extended to the Zernike formalism in a recent paper by the present authors²⁸.
- Apart from the vectorial nature, the high-NA case also requires a more careful

treatment of the effect of defocusing where the originally chosen quadratic phase factor²² has to be refined

- The so-called radiometric effect²² has to be included. This can either be done in the diffraction integrals themselves or it can be included via a non-uniform transmission function of the imaging system.
- For a full description in the vectorial case of the optical wave exiting the imaging system, we have to specify the complex exit pupil function for two orthogonal polarization states. In practice, one will often use two orthogonal linear states of polarization, e.g. along the x - and y -direction.

To carry out our forward calculations, we now define the complex pupil function, using polar coordinates (ρ, θ) on the exit pupil sphere, according to

$$\begin{aligned} B^x(\rho, \theta) &= A^x(\rho, \theta) \exp [i2\pi W^x(\rho, \theta)] \\ B^y(\rho, \theta) &= A^y(\rho, \theta) \exp [i2\pi W^y(\rho, \theta)]. \end{aligned} \quad (1)$$

The A^x and A^y in this expression are real-valued functions and describe the field strengths in the x - and y -direction. W^x and W^y are also real-valued functions and they describe the wavefront aberration in units of λ , the wavelength of the polarized light. The wavefront aberration has been caused by geometrical and polarization-dependent optical path length variation and birefringence and the aberration applies to linear polarization states along the x - and y -direction in the exit pupil of the imaging system. Different values of W^x and W^y are caused not only by material or stress birefringence but they can also be originated by polarization-dependent phase jumps at discontinuities (e.g. air-glass transitions, optical surface coatings) or by diffraction at structures with dimensions of the order of the wavelength. Further on in this paper, we will treat the more general case of elliptically polarized light that is incident at the entrance pupil. Note that in Ref.(28) we included the radiometric effect for a large-field imaging system (equal to $(1 - s_0^2 \rho^2)^{-1/4}$ with $s_0 = \sin \alpha$ equal to the value of the geometrical NA) directly in the functions A^x and A^y ; this is because this radiometric effect can be considered as being an intrinsic property of the large-field imaging system as it has to obey the Abbe sine condition.²² However, in this paper we will not follow this convention because it leads to nonzero aberration coefficients β in the perfect imaging case and this is considered as a counterintuitive result.

Using the complex amplitude functions B^x and B^y above, we can evaluate the complex field in the exit pupil in the presence of a general superposition of two linear polarization states in the entrance pupil (using complex numbers a and b with $|a|^2 + |b|^2 = 1$ for normalization purposes) and this will lead to a distribution of, in general, elliptically polarized light in the exit pupil of the imaging system.

For the vectorial treatment of the imaging by an aberrated imaging system we have expanded the complex exit pupil functions from Eq.(1) with the aid of a set of complex Zernike coefficients $\beta_{n,x}^m$ or $\beta_{n,y}^m$ for the Zernike terms $\exp[im\theta]R_n^{|m|}(\rho)$. It was shown²⁸ that the three cartesian electric field components in the focal region,

corresponding to an initially linearly x -polarized incident wave are given by

$$\mathbf{E}^x(r, \phi, f) = -i\gamma s_0^2 \exp\left[\frac{-if}{u_0}\right] \sum_{n,m} i^m \beta_{n,x}^m \exp[im\phi] \times \begin{pmatrix} V_{n,0}^m + \frac{s_0^2}{2} V_{n,2}^m \exp[2i\phi] + \frac{s_0^2}{2} V_{n,-2}^m \exp[-2i\phi] \\ -\frac{is_0^2}{2} V_{n,2}^m \exp[2i\phi] + \frac{is_0^2}{2} V_{n,-2}^m \exp[-2i\phi] \\ -is_0 V_{n,1}^m \exp[i\phi] + is_0 V_{n,-1}^m \exp[-i\phi] \end{pmatrix}, \quad (2)$$

and for a y -polarized incident field a similar expression is found

$$\mathbf{E}^y(r, \phi, f) = -i\gamma s_0^2 \exp\left[\frac{-if}{u_0}\right] \sum_{n,m} i^m \beta_{n,y}^m \exp[im\phi] \times \begin{pmatrix} -i\frac{s_0^2}{2} V_{n,2}^m \exp[2i\phi] + i\frac{s_0^2}{2} V_{n,-2}^m \exp[-2i\phi] \\ V_{n,0}^m - \frac{s_0^2}{2} V_{n,2}^m \exp[2i\phi] - \frac{s_0^2}{2} V_{n,-2}^m \exp[-2i\phi] \\ -s_0 V_{n,1}^m \exp[i\phi] - s_0 V_{n,-1}^m \exp[-i\phi] \end{pmatrix}. \quad (3)$$

The functions $V_{n,j}^m(r, f)$ have been given in Ref.(28). A minor modification is introduced here concerning the radiometric effect (equivalent to a factor of $(1 - s_0^2 \rho^2)^{-1/4}$ in the integrand). The functions are now given by

$$V_{n,j}^m = \int_0^1 \frac{\rho^{|j|} \left(1 + \sqrt{1 - s_0^2 \rho^2}\right)^{-|j|+1}}{(1 - s_0^2 \rho^2)^{1/4}} \exp\left[\frac{if}{u_0} \left(1 - \sqrt{1 - s_0^2 \rho^2}\right)\right] \times R_n^{|m|}(\rho) J_{m+j}(2\pi r \rho) \rho d\rho. \quad (4)$$

In the formulae above (r, ϕ, f) are normalized cylindrical co-ordinates in the image space with the origin located at the geometrical focus, see Fig.1, with f being the normalized axial coordinate and (r, ϕ) the lateral polar coordinates. The normalization has been carried out with respect to the diffraction unit λ/NA in the radial r -direction and $f = -2\pi u_0 z/\lambda$ in the axial direction with $u_0 = 1 - \sqrt{1 - s_0^2}$. As mentioned before, the quantity $s_0 = \sin \alpha$ equals the (geometrical) numerical aperture of the imaging system ($u_0 = 1 - \cos \alpha$). The coefficients $\beta_{n,x}^m$ pertain to the Zernike circle polynomials $R_n^{|m|}(\rho)$ and are sufficient to describe the complex pupil function (both amplitude and phase) of the lens system under study in the case of linearly polarized incident light in the x -direction. The integers n, m satisfy $n - |m| \geq 0$ and even. An analytical approach to evaluate the $V_{n,j}^m$ -integral has been obtained²⁸ in the framework of the extended Nijboer-Zernike theory and the evaluation with the radiometric effect included is carried out along the same lines.

B. Procedure for the retrieval of the complex pupil function of a high-NA imaging system

In the present paper, we use the results of the forward calculation according to Eqs.(2)-(4) to obtain semi-analytic expressions for the intensity distribution through the focal

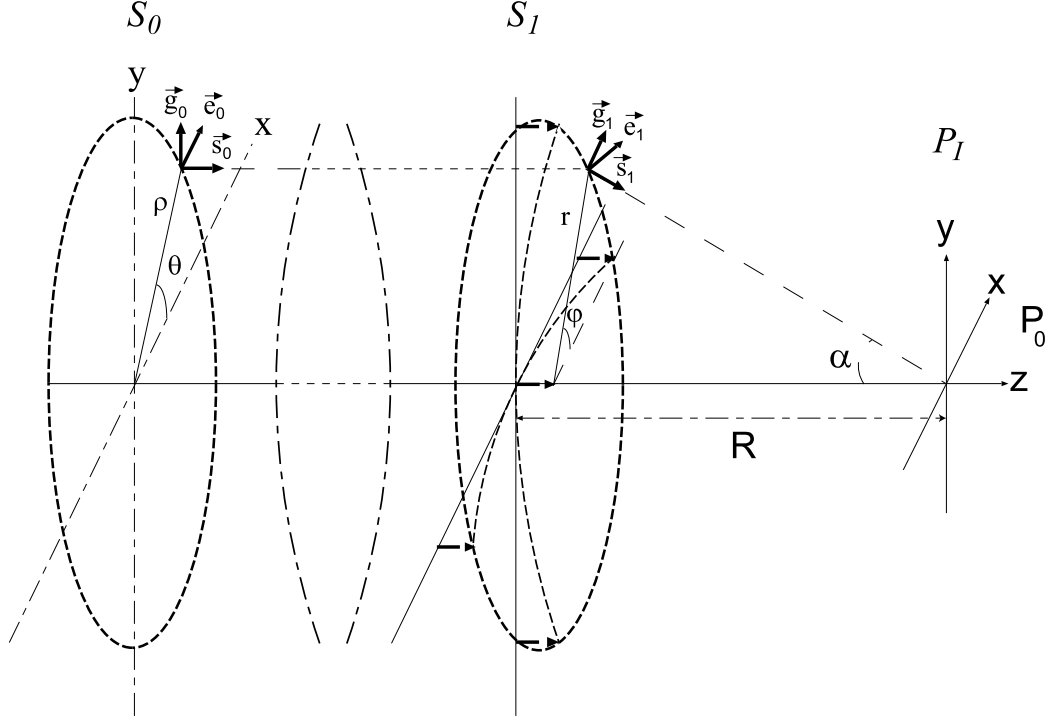


Fig. 1. The propagation of light in a high-NA optical system. The incident field is specified in the entrance pupil S_0 of the system (polar co-ordinates (ρ, θ)) with the aid of the amplitude and phase of e.g. the tangential and radial field vectors (\vec{g}_0 and \vec{e}_0) and the unit propagation vector \vec{s}_0 . After traversal of the optical system, the field vectors and the propagation direction on the exit pupil sphere S_1 (cylindrical co-ordinates (r, ϕ, z)) are specified by, respectively, the vectors \vec{g}_1 , \vec{e}_1 and \vec{s}_1 . The numerical aperture of the imaging system is given by $NA = n \sin \alpha$, with n equal to the refractive index of the image space. The nominal image plane position is given by P_I . The description of the field vectors according to the scheme in the paper requires a distance R that is rather large so that the aberrations of the system do not significantly influence the directions of the electric field vectors \vec{g}_1 and \vec{e}_1 in image space.

volume as a function of the cylindrical co-ordinates (r, ϕ, f) . Although the squaring operation of the complex amplitudes of the electric field components leads to rather complicated expressions, a systematic notation has turned out to be possible. From this notation it becomes clear that each separate aberration term with an azimuthal dependence of order m gives rise to azimuthal components in the resulting intensity distribution of orders $m - 2$, m and $m + 2$. If more than one aberration type is present in the pupil function, cross-terms are present with sum and difference orders $m_1 \pm m_2 \pm 0, 2, 4$ of the azimuthal dependence. The complete path towards the reconstruction of the complex pupil function uses the basic ideas developed in a for-

mer publication¹⁸ on scalar retrieval, but now extended to the high-NA case. First, in Section 2, we derive the expression for the electric energy density in the focal region in the presence of aberrations and nonuniform pupil transmission, using the Zernike coefficients that correspond to the specific exit pupil function. The expression for the Poynting vector can be obtained also but will not be used here. This paper focuses on the aberration retrieval process. In Section 3 we introduce the energy density formulae that arise when the lens defects are relatively small and not too far beyond the so-called diffraction limit, both regarding the lens aberration and lens transmission defects. These formulae are derived in the presence of an arbitrary state of polarization of the incident light and we present some special cases and combinations of intensity patterns that enable a stable retrieval scheme for the complex lens pupil function. In Section 4 we extend our analysis to the practical case where the optical system under study exhibits residual (linear) birefringence. The explicit dependence of the detected intensity patterns is derived in the presence of both geometrical aberrations, transmission defects *and* polarization-sensitive 'aberrations' due to linear birefringence that is spatially varying over the exit pupil. In Section 5 we present the basic steps that have to be taken to retrieve the complex lens function from a three-dimensional aggregate of intensity data in the focal volume. In Section 6 we present a detailed analysis of the functions that are used to describe the complex amplitude and intensity distribution in the focal region. Apart from the aberration-free case we also present graphs of the focal intensity distribution in the presence of typical aberrations that illustrate the interaction between vectorial image formation effects and aberrational image degradation. Finally, we present a numerical retrieval example that is relevant in practice, viz. the high-NA retrieval of the lens function in the presence of illumination with natural (unpolarized) light. This case is frequently encountered in projection systems for lithography and it leads to a simplified version of our analysis, closely resembling the one we get in the scalar case. Finally, Section 7 is devoted to some conclusions on this theoretical and numerical study about the retrieval of lens data in high numerical aperture imaging systems.

2. Explicit expression for the electric energy density

For the retrieval of lens data we need the expression for the light intensity in the focal volume of the imaging system. To this end we consider the time averaged value of the electric field energy density $\langle w_e \rangle$ and, for a harmonic field in a homogeneous medium with a dielectric constant $\epsilon = n_r^2$, we obtain (see Eq.(2) above)

$$\langle w_e \rangle = \frac{\epsilon_0}{4} n_r^2 |\mathbf{E}|^2 . \quad (5)$$

The electric field components in the presence of aberrations in a high-NA system are used to compute the scalar product $\mathbf{E}^* \cdot \mathbf{E}$.

To examine the energy flow through the focal region, we should evaluate the time-averaged values of the Cartesian components of the Poynting vector \mathbf{S} and this leads to the expression

$$\langle \mathbf{S} \rangle = \frac{\epsilon_0 c^2}{2} \text{Re}[\mathbf{E} \times \mathbf{B}^*] , \quad (6)$$

with \mathbf{B} the magnetic induction for which a similar expression holds as for the electric field strength.²⁸ Although the divergence of the Poynting vector would also allow us to solve the retrieval problem, we prefer to use the expression for the electric energy density because $\langle w_e \rangle$ is the electromagnetic quantity directly relevant for the calculation of the locally absorbed electromagnetic energy (exposure)²⁹.

A straightforward notation of $\langle w_e \rangle$ leads to a rather lengthy expression involving a quadruple sum over the indices m , n , m' and n' that occur in the expressions for the products of electric field components as derived from Eq.(2). In general, we encounter expressions that can be written as

$$\begin{aligned}
G_{kl}(\alpha, \beta) &= \sum_{n,m} i^m \exp [im\phi] \alpha_n^m V_{n,k}^m(r, f) \exp [ik\phi] \\
&\quad \times \sum_{n',m'} i^{-m'} \exp [-im'\phi] \beta_{n'}^{m'} V_{n',l}^{m'*}(r, f) \exp [-il\phi] \\
&= \sum_{n,m,n',m'} \exp [i(m-m')\pi/2] \exp [i(m-m'+k-l)\phi] \\
&\quad \times \alpha_n^m \beta_{n'}^{m'*} V_{n,k}^m(r, f) V_{n',l}^{m'*}(r, f) , \quad (7)
\end{aligned}$$

where the function G_{kl} has the sets of Zernike coefficients α_n^m and β_n^m as variables (in shorthand notation written as α and β in the argument of G_{kl}).

For the retrieval of Zernike coefficients, it is important to make explicit the azimuthal dependence of the intensity distribution in the focal region. To this end, we write a quadruple series $\sum_{n,m,n',m'} a_{n,m,n',m'}$ as in Eq.(7) according to the following diagonal summation scheme

$$\begin{aligned}
\sum_{m,m',n,n'} a_{m,m';n,n'} &= \sum_{n=n_1}^{n_2} \left\{ \sum_{m=m_1}^{m_2} a_{m,m;n,n} + \right. \\
&\quad \left. \sum_{\mu=1}^{\mu_{max}} \sum_m (a_{m,m+\mu;n,n} + a_{m+\mu,m;n,n}) \right\} \\
&+ \sum_{\nu=1}^{\nu_{max}} \left\{ \sum_n \sum_m (a_{m,m;n,n+\nu} + a_{m,m;n+\nu,n}) \right. \\
&\quad \left. + \sum_{\mu=1}^{\mu_{max}} \sum_n \sum_m \left[a_{m,m+\mu;n,n+\nu} + a_{m+\mu,m;n,n+\nu} \right. \right. \\
&\quad \left. \left. + a_{m,m+\mu;n+\nu,n} + a_{m+\mu,m;n+\nu,n} \right] \right\} , \quad (8)
\end{aligned}$$

where the various summation ranges determined by m_1 , m_2 , n_1 , n_2 , μ_{max} and ν_{max} are derived from the transformation from a rectangular summation scheme to a summation scheme along diagonals. After some rearrangement, the following expression is obtained

$$\begin{aligned}
G_{kl}(\alpha, \beta) &= \exp [i(k-l)\phi] \left[\sum_n \sum_m \alpha_n^m \beta_n^{m*} V_{n,k}^m V_{n,l}^{m*} + \right. \\
&\quad \left. \sum_{\mu=1}^{\mu_{max}} \left\{ \exp [-i\mu\pi/2] \exp [-i\mu\phi] \sum_m \sum_n (\alpha_n^m \beta_n^{m+\mu*} V_{n,k}^m V_{n,l}^{m+\mu*}) \right. \right. \\
&\quad \left. \left. + \exp [i\mu\pi/2] \exp [i\mu\phi] \sum_m \sum_n (\alpha_n^m \beta_n^{m-\mu*} V_{n,k}^m V_{n,l}^{m-\mu*}) \right\} \right] , \quad (9)
\end{aligned}$$

$$\begin{aligned}
& + \exp [i\mu\pi/2] \exp [i\mu\phi] \sum_m \sum_n \left(\alpha_n^{m+\mu} \beta_n^{m*} V_{n,k}^{m+\mu} V_{n,l}^{m*} \right) \Big\} \\
& + \sum_{\nu=1}^{\nu_{max}} \left\{ \sum_n \sum_m \left(\alpha_n^m \beta_{n+\nu}^{m*} V_{n,k}^m V_{n+\nu,l}^{m*} + \alpha_{n+\nu}^m \beta_n^{m*} V_{n+\nu,k}^m V_{n,l}^{m*} \right) \right. \\
& + \sum_{\mu=1}^{\mu_{max}} \left[\exp [-i\mu\pi/2] \exp [-i\mu\phi] \left\{ \sum_n \sum_m \left(\alpha_n^m \beta_{n+\nu}^{m+\mu*} V_{n,k}^m V_{n+\nu,l}^{m+\mu*} \right. \right. \right. \\
& \qquad \qquad \qquad \left. \left. \left. + \alpha_{n+\nu}^m \beta_n^{m+\mu*} V_{n+\nu,k}^m V_{n,l}^{m+\mu*} \right) \right\} \right. \\
& + \exp [+i\mu\pi/2] \exp [+i\mu\phi] \left\{ \sum_n \sum_m \left(\alpha_n^{m+\mu} \beta_{n+\nu}^{m*} V_{n,k}^{m+\mu} V_{n+\nu,l}^{m*} \right. \right. \\
& \qquad \qquad \qquad \left. \left. \left. + \alpha_{n+\nu}^{m+\mu} \beta_n^{m*} V_{n+\nu,k}^{m+\mu} V_{n,l}^{m*} \right) \right\} \right] \Big\} \Big\} \Big\} , \tag{9}
\end{aligned}$$

where we have suppressed the (r, f) -dependence of the V -functions.

With the G -function notation above, the electric energy density is readily written as

$$\begin{aligned}
\langle w_e(r, \phi, f) \rangle = & \frac{\epsilon_0 n_r^2 s_0^4}{4} \left[G_{0,0}(\beta_x, \beta_x) + \right. \\
& s_0^2 \text{Re} \{ G_{0,2}(\beta_x, \beta_x - i\beta_y) + G_{0,-2}(\beta_x, \beta_x + i\beta_y) \} + \\
& \frac{s_0^4}{4} \{ G_{2,2}(\beta_x - i\beta_y, \beta_x - i\beta_y) + G_{-2,-2}(\beta_x + i\beta_y, \beta_x + i\beta_y) \} + \\
& \frac{s_0^4}{2} \text{Re} \{ G_{2,-2}(\beta_x - i\beta_y, \beta_x + i\beta_y) \} + \\
& G_{0,0}(\beta_y, \beta_y) + \\
& -s_0^2 \text{Re} \{ G_{0,2}(\beta_y, i\beta_x + \beta_y) + G_{0,-2}(\beta_y, -i\beta_x + \beta_y) \} + \\
& \frac{s_0^4}{4} \{ G_{2,2}(i\beta_x + \beta_y, i\beta_x + \beta_y) + G_{-2,-2}(-i\beta_x + \beta_y, -i\beta_x + \beta_y) \} + \\
& \frac{s_0^4}{2} \text{Re} \{ G_{2,-2}(i\beta_x + \beta_y, -i\beta_x + \beta_y) \} + \\
& s_0^2 \{ G_{1,1}(i\beta_x + \beta_y, i\beta_x + \beta_y) + G_{-1,-1}(-i\beta_x + \beta_y, -i\beta_x + \beta_y) \} + \\
& \left. 2s_0^2 \text{Re} \{ G_{1,-1}(i\beta_x + \beta_y, -i\beta_x + \beta_y) \} \right] , \tag{10}
\end{aligned}$$

where the indices x, y of β in the arguments of the G_{kl} -functions refer to the sets of Zernike coefficients to be used, corresponding to either x - or y - linearly polarized light (to be denoted by $\beta_{n,x}^m$ and $\beta_{n,y}^m$, respectively).

Using the summation property

$$G_{kl}(\alpha_1 + \alpha_2, \beta_1 + \beta_2) = G_{kl}(\alpha_1, \beta_1) + G_{kl}(\alpha_1, \beta_2) + G_{kl}(\alpha_2, \beta_1) + G_{kl}(\alpha_2, \beta_2) \tag{11}$$

and the property

$$G_{kl}(\alpha, \beta) = G_{lk}^*(\beta, \alpha) , \tag{12}$$

we write for the electric energy density

$$\begin{aligned}
\langle w_e(r, \phi, f) \rangle = & \\
& \frac{\epsilon_0 n_r^2 s_0^4}{4} \left[G_{0,0}(\beta_x, \beta_x) + G_{0,0}(\beta_y, \beta_y) + \right. \\
& s_0^2 \text{Re} \{ G_{0,2}(\beta_x, \beta_x) + iG_{0,2}(\beta_x, \beta_y) + iG_{0,2}(\beta_y, \beta_x) - G_{0,2}(\beta_y, \beta_y) \} + \\
& s_0^2 \text{Re} \{ G_{0,-2}(\beta_x, \beta_x) - iG_{0,-2}(\beta_x, \beta_y) - iG_{0,-2}(\beta_y, \beta_x) - G_{0,-2}(\beta_y, \beta_y) \} + \\
& \frac{s_0^4}{2} \{ G_{2,2}(\beta_x, \beta_x) + iG_{2,2}(\beta_x, \beta_y) - iG_{2,2}(\beta_y, \beta_x) + G_{2,2}(\beta_y, \beta_y) \} + \\
& \frac{s_0^4}{2} \{ G_{-2,-2}(\beta_x, \beta_x) - iG_{-2,-2}(\beta_x, \beta_y) + iG_{-2,-2}(\beta_y, \beta_x) + G_{-2,-2}(\beta_y, \beta_y) \} + \\
& s_0^2 \{ G_{1,1}(\beta_x, \beta_x) + iG_{1,1}(\beta_x, \beta_y) - iG_{1,1}(\beta_y, \beta_x) + G_{1,1}(\beta_y, \beta_y) \} + \\
& s_0^2 \{ G_{-1,-1}(\beta_x, \beta_x) - iG_{-1,-1}(\beta_x, \beta_y) + iG_{-1,-1}(\beta_y, \beta_x) + G_{-1,-1}(\beta_y, \beta_y) \} + \\
& \left. + 2s_0^2 \text{Re} \{ -G_{1,-1}(\beta_x, \beta_x) + iG_{1,-1}(\beta_x, \beta_y) + iG_{1,-1}(\beta_y, \beta_x) + G_{1,-1}(\beta_y, \beta_y) \} \right]. \tag{13}
\end{aligned}$$

3. Approximated G -functions for modest values of the Zernike coefficients

Like in the scalar retrieval procedure, we now make the assumption that the lens defects (amplitude and phase) are sufficiently small and that the Strehl intensity I_S of the imaging system is relatively high. With the (scalar) relationship $I_S = |\beta_0^0|^2$, we have found that in practice it is needed that $\beta_0^0 \geq 0.5$. The basic functions occurring in the energy density function $\langle w_e(r, \phi, f) \rangle$ applying to the vectorial case have been denoted by $G_{kl}(\alpha, \beta)$. In the case of a dominating α_0^0 and β_0^0 term and, consequently, modest values of any of the remaining $\beta_{n,x}^m$ and $\beta_{n,y}^m$ terms, the general expression for $G_{kl}(\alpha, \beta)$ then reduces in good approximation to

$$\begin{aligned}
G_{kl}(\alpha, \beta) = & e^{i(k-l)\phi} \left[\alpha_0^0 \beta_0^{0*} V_{0,k}^0 V_{0,l}^{0*} \right. \\
& + \sum_{\mu=1}^{\mu_{max}} \left[\exp[-i\mu\pi/2] \exp[-i\mu\phi] \left(\alpha_0^0 \beta_0^{\mu*} V_{0,k}^0 V_{0,l}^{\mu*} + \alpha_0^{-\mu} \beta_0^{0*} V_{0,k}^{-\mu} V_{0,l}^{0*} \right) \right. \\
& \left. + \exp[+i\mu\pi/2] \exp[+i\mu\phi] \left(\alpha_0^0 \beta_0^{-\mu*} V_{0,k}^0 V_{0,l}^{-\mu*} + \alpha_0^\mu \beta_0^{0*} V_{0,k}^\mu V_{0,l}^{0*} \right) \right] \\
& + \sum_{\nu=1}^{\nu_{max}} \left\{ \alpha_0^0 \beta_\nu^{0*} V_{0,k}^0 V_{\nu,l}^{0*} + \alpha_\nu^0 \beta_0^{0*} V_{\nu,k}^0 V_{0,l}^{0*} \right. \\
& + \sum_{\mu=1}^{\mu_{max}} \left[\exp[-i\mu\pi/2] \exp[-i\mu\phi] \left(\alpha_0^0 \beta_\nu^{\mu*} V_{0,k}^0 V_{\nu,l}^{\mu*} + \alpha_\nu^{-\mu} \beta_0^{0*} V_{\nu,k}^{-\mu} V_{0,l}^{0*} \right) \right. \\
& \left. \left. + \exp[+i\mu\pi/2] \exp[+i\mu\phi] \left(\alpha_0^0 \beta_\nu^{-\mu*} V_{0,k}^0 V_{\nu,l}^{-\mu*} + \alpha_\nu^\mu \beta_0^{0*} V_{\nu,k}^\mu V_{0,l}^{0*} \right) \right] \right\}. \tag{14}
\end{aligned}$$

A more compact notation is possible according to

$$G_{kl}(\alpha, \beta) = \exp [i(k-l)\phi] \sum_{\nu=0}^{\nu_{max}} \sum_{\mu=-\mu_{max}(\nu)}^{+\mu_{max}(\nu)} \left\{ \exp [-i\mu\pi/2] \exp [-i\mu\phi] \right. \\ \left. \times \alpha_0^0 \beta_\nu^{\mu*} V_{0,k}^0 V_{\nu,l}^{\mu*} + (1 - \epsilon_{\nu\mu}) \exp [i\mu\pi/2] \exp [i\mu\phi] \alpha_\nu^\mu \beta_0^{0*} V_{\nu,k}^\mu V_{0,l}^{0*} \right\}, \quad (15)$$

where $\epsilon_{\nu\mu}$ equals unity for $\nu = \mu = 0$ and zero for any other combination of (ν, μ) -values.

A special case arises when the coefficients α and β are equal and $k = l$. We then obtain

$$G_{kk}(\alpha, \alpha) = |\alpha_0^0|^2 |V_{0,k}^0|^2 + \\ 2 \sum_{\nu=0}^{\nu_{max}} \sum_{\mu=-\mu_{max}(\nu)}^{+\mu_{max}(\nu)} \operatorname{Re} \left\{ \exp [i\mu\pi/2] \exp [i\mu\phi] \alpha_\nu^\mu \alpha_0^{0*} V_{\nu,k}^\mu V_{0,k}^{0*} \right\}, \quad (16)$$

with the value $\mu = 0$ excluded in the double summation.

Using the simplified expression for the functions $G_{kl}(\alpha, \beta)$, we subsequently analyze a general state of polarization that is incident on the optical system by putting

$$\beta_{n,x}^m = a\beta_n^m, \quad \beta_{n,y}^m = b\beta_n^m, \quad (17)$$

with $|a|^2 + |b|^2 = 1$ for normalization purposes. The generally complex quantities (a, b) allow us to specify the initial state of polarization. Using the result of Eq.(13), we obtain

$$\langle w_e(r, \phi, f) \rangle_0 = \frac{\epsilon_0 n_r^2 s_0^4}{4} \left[G_{0,0}(\beta, \beta) + \right. \\ s_0^2 \left\{ [|a|^2 - |b|^2] \operatorname{Re}\{G_{0,2}(\beta, \beta)\} - 2\operatorname{Re}(ab^*) \operatorname{Im}\{G_{0,2}(\beta, \beta)\} \right\} + \\ s_0^2 \left\{ [|a|^2 - |b|^2] \operatorname{Re}\{G_{0,-2}(\beta, \beta)\} + 2\operatorname{Re}(ab^*) \operatorname{Im}\{G_{0,-2}(\beta, \beta)\} \right\} + \\ \frac{s_0^4}{2} [\{1 - 2\operatorname{Im}(ab^*)\} G_{2,2}(\beta, \beta) + \{1 + 2\operatorname{Im}(ab^*)\} G_{-2,-2}(\beta, \beta)] + \\ s_0^2 [\{1 - 2\operatorname{Im}(ab^*)\} G_{1,1}(\beta, \beta) + \{1 + 2\operatorname{Im}(ab^*)\} G_{-1,-1}(\beta, \beta)] + \\ \left. -2s_0^2 \left\{ [|a|^2 - |b|^2] \operatorname{Re}\{G_{+1,-1}(\beta, \beta)\} + \right. \right. \\ \left. \left. 2\operatorname{Re}(ab^*) \operatorname{Im}\{G_{+1,-1}(\beta, \beta)\} \right\} \right]. \quad (18)$$

The index zero has been added to $\langle w_e(r, \phi, f) \rangle$ to indicate that no spatially varying birefringence was present in the optical system.

Several special cases for the energy density function arise when we choose special values for (a, b) :

- linear polarization in the x - or y -direction, resp. $(a, b) = (1, 0)$ and $(a, b) = (0, 1)$, and in the diagonal directions, resp. $(a, b) = (1/\sqrt{2}, 1/\sqrt{2})$ and $(a, b) = (1/\sqrt{2}, -1/\sqrt{2})$

- left- or right-handed circular polarization, resp. $a = 1/\sqrt{2}$, $b = i/\sqrt{2}$ and $a = 1/\sqrt{2}$, $b = -i/\sqrt{2}$,
- unpolarized or natural light, a summation in intensity of the above mentioned orthogonal linear or circular states
- radial or tangential (azimuthal) states of polarization

A. *Linear polarization in the x-direction, $(a,b)=(1,0)$*

Note that in the absence of birefringence the coefficients $\beta_{n,x}^m$ equal the corresponding $\beta_{n,y}^m$ and with this assumption the energy density becomes

$$\begin{aligned} \langle w_e^x(r, \phi, f) \rangle_0 &\propto G_{0,0}(\beta, \beta) + s_0^2 \text{Re} [G_{0,2}(\beta, \beta) + G_{0,-2}(\beta, \beta)] \\ &\quad + \frac{s_0^4}{2} [G_{2,2}(\beta, \beta) + G_{-2,-2}(\beta, \beta)] \\ &\quad + s_0^2 [G_{1,1}(\beta, \beta) + G_{-1,-1}(\beta, \beta)] \\ &\quad - 2s_0^2 \text{Re} \{G_{+1,-1}(\beta, \beta)\} \quad . \end{aligned} \quad (19)$$

B. *Linear polarization in the y-direction, $(a,b)=(0,1)$*

The energy is proportional to

$$\begin{aligned} \langle w_e^y(r, \phi, f) \rangle_0 &\propto G_{0,0}(\beta, \beta) - s_0^2 \text{Re} [G_{0,2}(\beta, \beta) + G_{0,-2}(\beta, \beta)] \\ &\quad + \frac{s_0^4}{2} [G_{2,2}(\beta, \beta) + G_{-2,-2}(\beta, \beta)] \\ &\quad + s_0^2 [G_{1,1}(\beta, \beta) + G_{-1,-1}(\beta, \beta)] \\ &\quad + 2s_0^2 \text{Re} \{G_{+1,-1}(\beta, \beta)\} \quad . \end{aligned} \quad (20)$$

Again supposing the absence of (linear) birefringence effects, the subtraction of the two exposure patterns in (19) and (20) yields

$$\begin{aligned} \Delta w_{l,0} &= \langle w_e^x(r, \phi, f) \rangle_0 - \langle w_e^y(r, \phi, f) \rangle_0 = \\ &= 2s_0^2 \text{Re} [G_{0,2}(\beta, \beta) + G_{0,-2}(\beta, \beta) - 2G_{+1,-1}(\beta, \beta)] \quad . \end{aligned} \quad (21)$$

The subtraction of two exposure patterns resulting from orthogonal diagonal linear polarization states yields

$$\begin{aligned} \Delta w_{l,\pi/4} &= \langle w_e^x(r, \phi, f) \rangle_{3\pi/4} - \langle w_e^y(r, \phi, f) \rangle_{\pi/4} = \\ &= 2s_0^2 \text{Im} [G_{0,2}(\beta, \beta) - G_{0,-2}(\beta, \beta) + 2G_{+1,-1}(\beta, \beta)] \quad . \end{aligned} \quad (22)$$

C. *Circular polarization (LC and RC)*

We follow a similar procedure to calculate the exposure patterns in the case of circular polarization. With our convention $\exp\{i(kz - \omega t)\}$ for an outgoing plane wave in the positive z -direction we find the relations $b = ia$ for left-handed circular polarization

(*LC*) and $b = -ia$ for the right-handed case (*RC*). A straightforward calculation shows

$$\begin{aligned} \langle w_e^{RC}(r, \phi, f) \rangle_0 &\propto G_{0,0}(\beta, \beta) + s_0^4 G_{2,2}(\beta, \beta) \\ &\quad + 2s_0^2 G_{1,1}(\beta, \beta) \end{aligned} \quad (23)$$

and, in a corresponding way,

$$\begin{aligned} \langle w_e^{LC}(r, \phi, f) \rangle_0 &\propto G_{0,0}(\beta, \beta) + s_0^4 G_{-2,-2}(\beta, \beta) \\ &\quad + 2s_0^2 G_{-1,-1}(\beta, \beta) \quad . \end{aligned} \quad (24)$$

The difference between right- and left-handed polarization exposure distributions thus equals

$$\begin{aligned} \Delta w_{C,0} &= s_0^4 [G_{2,2}(\beta, \beta) - G_{-2,-2}(\beta, \beta)] \\ &\quad + 2s_0^2 [G_{1,1}(\beta, \beta) - G_{-1,-1}(\beta, \beta)] \quad . \end{aligned} \quad (25)$$

D. Natural light

Finally, the exposure with natural light gives rise to the summation of either both linear orthogonal or circular orthogonal polarization states and this yields, apart from a factor of 2 due to the normalization of the circular eigenstates,

$$\begin{aligned} \langle w_e^N(r, \phi, f) \rangle &= G_{0,0}(\beta, \beta) + s_0^2 [G_{1,1}(\beta, \beta) + G_{-1,-1}(\beta, \beta)] \\ &\quad + \frac{s_0^4}{2} [G_{2,2}(\beta, \beta) + G_{-2,-2}(\beta, \beta)] \quad . \end{aligned} \quad (26)$$

Recall again that in this analysis we have excluded any birefringence effect, which implies that $\beta_{n,x}^m$ and $\beta_{n,y}^m$ are related by a simple factor of proportionality to the coefficients β_n^m that are uniquely defined by the geometrical lens properties.

The general case of partially polarized light can be accounted for by defining a total exposure that is a weighted sum of a fully polarized exposure pattern and an unpolarized pattern according to the degree of partial polarization.

4. The incorporation of polarization *and* birefringence effects

The propagation of a polarized wave through a general optical system leads to a perturbation of the initial state of polarization due to the polarization-dependent amplitude and phase changes on transmission through the (coated) air-glass interfaces. On top of this, anisotropy of the lens materials, induced by structural properties or e.g. by residual stresses in the lens materials, leads to a gradual change of the state of polarization on propagation. In this section we first describe the cumulative effect of the birefringence on the polarization state in the exit pupil. In the next step, we incorporate the amplitude and phase effects due to the birefringence in our description of the complex exit pupil function for x and y -polarization using an extra set of complex Zernike coefficients and we point out how a set of exposures with different polarization states at the entrance of the optical system yield the unknown birefringence data in the exit pupil.

A. The field components E_x and E_y in the exit pupil

We approximate the cumulative effects of birefringence in the optical system by defining at each point in the exit pupil a certain value of the retardation due to the birefringence, Δ , and the azimuths of the orthogonal principal axes, e.g. by defining the angle α of the 'slow' axis of the birefringence. In practical cases, it is allowed to neglect the dichroism introduced by an optical system and for this reason we will limit ourselves to the influence of the phase retardation on the strength of the x - and y -field components in the exit pupil of the optical system.

To analyze the state of polarization in the exit pupil in the presence of a sequence of birefringent elements and surfaces effects we use the Jones matrix analysis³⁰. The matrix relation between the x - and y -components of the input and output electric fields is in general given by³¹

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} a_j \\ b_j \end{pmatrix} \quad (27)$$

where the complex amplitudes of the x - and y components of the incident electric field at the j -th exposure have been denoted by (a_j, b_j) . Note that the field components (E_x, E_y) are a function of the position in the pupil; the functions describing their locally varying complex amplitude can be expanded with the aid of Zernike polynomials. The field components (E_x, E_y) , affected by the birefringence of the optical system, formally replace the original components (a_j, b_j) related to the entrance pupil. The Zernike expansion corresponding to (E_x, E_y) is used to determine the vector components of the field on the exit pupil sphere; these are then used to evaluate the field in the focal region, see Eqs.(2)-(3).

B. Procedure for evaluation of the birefringence of the optical system

Basically, we need to evaluate the four complex matrix elements m_{ij} for each sample point in the exit pupil, leading to eight independent quantities to be determined. But since we have excluded dichroism, the matrix above has a special structure³²⁻³⁶ and can be written as

$$M = \begin{pmatrix} m_{11} & m_{12} \\ -m_{12}^* & m_{11}^* \end{pmatrix}, \quad (28)$$

with the property $|m_{11}|^2 + |m_{12}|^2 = 1$. The eigenstates of this matrix are elliptical in general. Once the eigenvalues and eigenstates have been found, the orientation α of the slow and fast axes and the value Δ of the phase birefringence are known.

Because of the special structure of the unitary matrix M , three independent quantities need to be determined on top of the geometrical wavefront aberration and transmission defects of the system. We thus need four retrieval operations to determine the complex quantities m_{11} and m_{12} plus the polarization-independent geometrical defects of the system. Preferred polarization states (a_j, b_j) are two orthogonal linear polarization states, e.g. $(1, 0)$ and $(0, 1)$ and the circular ones, viz. $(1, i)/\sqrt{2}$ for left-circularly polarized and $(1, -i)/\sqrt{2}$ for right-circularly polarized light. The four exposures with the preferred polarization states lead, after retrieval, to four different sets of β -coefficients, $\beta_{n;(1,0)}^m$, $\beta_{n;(0,1)}^m$, $\beta_{n;(1,i)}^m$ and $\beta_{n;(1,-i)}^m$. Note that the coefficients

$\beta_{n,x}^m$ and $\beta_{n,y}^m$, used previously in Eqs.(2)-(3), correspond to the first two sets that we discuss here in the framework of birefringence retrieval. From the four sets of β -coefficients we obtain the complex amplitude in a general point of the exit pupil for four different polarization states. This is basically sufficient to uniquely determine the size and the orientation of the cumulative birefringence of the optical system in that specific point of the exit pupil. In addition, we obtain the geometrical defects of the system that are independent of the state of polarization of the incident light.

5. Outline of the basic retrieval scheme

The various expressions that have been obtained for the electric field density constitute the intensity pattern detected by a sensor or the exposure profile in a storage layer (e.g. a photoresist layer in lithography). These measured data, collected from a set of axially displaced (defocused) planes, serve as the input for the retrieval scheme that will yield the complex $\beta_{n;(a_j,b_j)}^m$ -coefficients that describe the high-NA imaging system. The basic term that appears in the expressions is the real or imaginary part of $G_{k,l}(\beta, \beta)$; if $k = l$ the function is real anyhow. We now want to sketch the basic approach for retrieval of the complex β -coefficients.

Following our retrieval approach for the low-NA scalar case¹⁸ we first detect the azimuthal periodicities in the measured intensity patterns according to

$$\Psi^m(r, f) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} I(r, \phi, f) \exp(im\phi) d\phi, \quad (29)$$

where $I(r, \phi, f)$ is the measured intensity function in the focal volume.

In our retrieval scheme, the measured through-focus intensity pattern will be matched by the linearized intensity distribution according to Eq.(18). In compact notation we write this approximated analytical expression as

$$w_{an}(r, \phi, f) = \sum_{k,l} F_{k,l}(r, \phi, f) \quad (30)$$

and the various functions $F_{k,l}$ are, apart from a constant factor, given by

$$\begin{aligned} F_{0,0} &= G_{0,0}(\beta, \beta) \\ F_{0,2} &= s_0^2 \left[\left\{ |a|^2 - |b|^2 \right\} \operatorname{Re} \{ G_{0,2}(\beta, \beta) \} \right. \\ &\quad \left. - 2\operatorname{Re}(ab^*) \operatorname{Im} \{ G_{0,2}(\beta, \beta) \} \right] \\ F_{0,-2} &= s_0^2 \left[\left\{ |a|^2 - |b|^2 \right\} \operatorname{Re} \{ G_{0,-2}(\beta, \beta) \} \right. \\ &\quad \left. + 2\operatorname{Re}(ab^*) \operatorname{Im} \{ G_{0,-2}(\beta, \beta) \} \right] \\ F_{+1,-1} &= -2s_0^2 \left[\left\{ |a|^2 - |b|^2 \right\} \operatorname{Re} \{ G_{+1,-1}(\beta, \beta) \} \right. \\ &\quad \left. + 2\operatorname{Re}(ab^*) \operatorname{Im} \{ G_{+1,-1}(\beta, \beta) \} \right] \\ F_{1,1} &= s_0^2 \{ 1 - 2\operatorname{Im}(ab^*) \} G_{1,1}(\beta, \beta) \\ F_{-1,-1} &= s_0^2 \{ 1 + 2\operatorname{Im}(ab^*) \} G_{-1,-1}(\beta, \beta) \\ F_{2,2} &= \frac{s_0^4}{2} \{ 1 - 2\operatorname{Im}(ab^*) \} G_{2,2}(\beta, \beta) \\ F_{-2,-2} &= \frac{s_0^4}{2} \{ 1 + 2\operatorname{Im}(ab^*) \} G_{-2,-2}(\beta, \beta). \end{aligned} \quad (31)$$

The general expression for $G_{k,l}$ (linearized approximation for dominating β_0^0) reads

$$G_{k,l}(\beta, \beta) = \beta_0^0 \exp\{i(k-l)\phi\} \sum_{\nu} \sum_{\mu} \left\{ \beta_{\nu}^{\mu*} \Psi_{\nu;k,l}^{\mu*}(r, f) \exp(-i\mu\phi) \right. \\ \left. + (1 - \epsilon_{\nu\mu}) \beta_{\nu}^{\mu} \Psi_{\nu;l,k}^{\mu}(r, f) \exp(+i\mu\phi) \right\} , \quad (32)$$

where $\epsilon_{\nu\mu}$ has been defined in Eq.(15) and where we also introduced the shorthand notation

$$\Psi_{\nu;k,l}^{\mu}(r, f) = (+i)^{\mu} V_{0,k}^{0*}(r, f) V_{\nu,l}^{\mu}(r, f) . \quad (33)$$

A harmonic decomposition applied to the general terms $G_{k,l}$ of this analytic function yields the result

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} G_{k,l}(r, \phi, f) \exp(im\phi) d\phi = \\ \beta_0^0 \sum_{\nu} \left[\beta_{\nu}^{(+k-l+m)*} \Psi_{\nu;k,l}^{(+k-l+m)*}(r, f) \right. \\ \left. (1 - \epsilon_{\nu, -k+l-m}) \beta_{\nu}^{(-k+l-m)} \Psi_{\nu;l,k}^{(-k+l-m)}(r, f) \right] , \quad (34)$$

where we have used the property that, without loss of generality, β_0^0 can be taken to be real (reference phase of the pupil function is zero). We ultimately need the harmonic decomposition of the full analytic expression $w_{an}(r, \phi, f)$; because of the rather lengthy expression involved, we give the result for the various harmonic functions $\Psi_{an}^m(r, f)$ that arise from this decomposition in the Appendix.

Having available now the harmonic azimuthal dependence of order m both from the measurement data ($\Psi^m(r, f)$) and from the analytical functions ($\Psi_{an}^m(r, f)$) with which the measurement data have to be matched, the equation to be solved for each azimuthal order number m reads

$$\Psi_{an}^m(r, f) \approx \Psi^m(r, f) . \quad (35)$$

Here, the right-hand side function has been obtained via measurement values in a large number of lateral and axial positions in the focal volume. The left-hand side contains the unknown β -coefficients that have to be calculated and the \approx -sign expresses that the linearized version of the analytical intensity distribution that has been used.

The approximate equality in Eq.(35) can be solved for the unknown β -coefficients in various ways. Our preferred method uses the fact that the functions $\Psi_{\nu;k,l}^{\mu}(r, f)$ that implicitly appear at the right-hand side of Eq.(35) are close to being orthogonal. By applying inner products with the involved $\Psi_{\nu;k,l}^{\mu}$ on either side of (35), we obtain an approximate linear system in the β -coefficients that produces estimates of these β -coefficients upon solving it. The inner product that we choose here is defined for functions $\Psi(r, f)$ and $\Phi(r, f)$ as

$$(\Psi, \Phi) = \int_0^R \int_{-F}^{+F} \Psi(r, f) \Phi^*(r, f) r dr df . \quad (36)$$

The integration limits R and $\pm F$ formally should be infinitely large but, in practice, they are determined by the lateral and axial extent of the measured data set. The

inner products of the form $\left(\Psi_{\nu;k,l}^m, \Psi_{\nu';k',l'}^{m'}\right)$ are calculated just once and their numerical values serve to fill the matrix corresponding to the system of linear equations. In general, the procedure is to first calculate the best fit β -values without the birefringence included. Given the residual error of this solution for the various polarization states, the full set of $\beta_{n;(a_j,b_j)}^m$ -coefficients is then taken into account to evaluate the birefringence effects of the optical system.

6. Graphical illustration of the basic functions $V_{n,j}^m(r, f)$, $G_{kl}(\beta, \beta)$ and a high-NA retrieval example

In this section we present some typical examples of the amplitude function $V_{n,j}^m(r, f)$ that plays a basic role in the calculation of the complex amplitude of the cartesian electric field components in the focal region. We also present some examples of the analytical function $G_{kl}(\beta, \beta)$ that plays an important role when evaluating the energy density in the focal region and when the inverse problem is solved for retrieving aberrational lens properties. Some characteristic aberrations like coma and astigmatism will be treated in more detail and the subtle interplay between the state of polarization in the exit pupil and the azimuth of a non-circularly symmetrical aberration will be discussed.

A. The aberration-free $V_{n,j}^m$ -functions

In Fig.2 we have plotted the functions $V_{0,0}^0(r, 0)$, $V_{0,+1}^0(r, 0)$, $V_{0,-1}^0(r, 0)$ and $V_{0,\pm 2}^0(r, 0)$ in the upper row. The same functions have been represented in the middle row for a value of the defocus parameter equal to 2π . In both cases, the numerical aperture of the imaging system was 0.95. These functions with $n = m = 0$ are generally the dominant ones that determine the complex amplitude of the electric field components. If there are no aberrations at all, they are the only ones needed for the calculation of the intensity in the focal region (the coefficient β_0^0 equals unity and all other β -coefficients are zero). Some remarks follow from the inspection of the $V_{0,j}^0$ -functions for the aberration-free case. The amplitude at the central point of the diffraction image ($r = f = 0$) is given by the value of $V_{0,0}^0$ only. We also note that for equal values of $|j|$ the V -functions change sign for odd j . In the defocused case, we see that the on-axis amplitude does not vanish for a value of the defocusing parameter equal to 2π . In the scalar diffraction case at low numerical aperture, the V_0^0 -function is zero precisely at this defocusing value. At high numerical aperture, the on-axis amplitude is non-vanishing because of the non-quadratic defocus phase and the non-uniform amplitude distribution over the exit pupil (radiometric effect). In the lower row of Fig.2 we have produced contour plots of the absolute value of the three electric field components and of the resulting total energy density (or intensity) in optimum focus. The incident state of polarization was linear and oriented along the x -axis (see Eq.(2) for the expressions for the field components in the focal region). The E_x field is the sum of the dominant $V_{0,0}^0$ -function and the smaller contributions from the $V_{0,\pm 2}^0$ -functions (with zero-azimuth off-set) which tend to reduce the field value along the azimuths $\phi = 0, \pi$ (x -axis) and lead to an increase along the azimuths $\phi = \pi/2, 3\pi/2$. The E_y -component only consists of a 2ϕ -component with a phase shift of $\pi/4$ and is

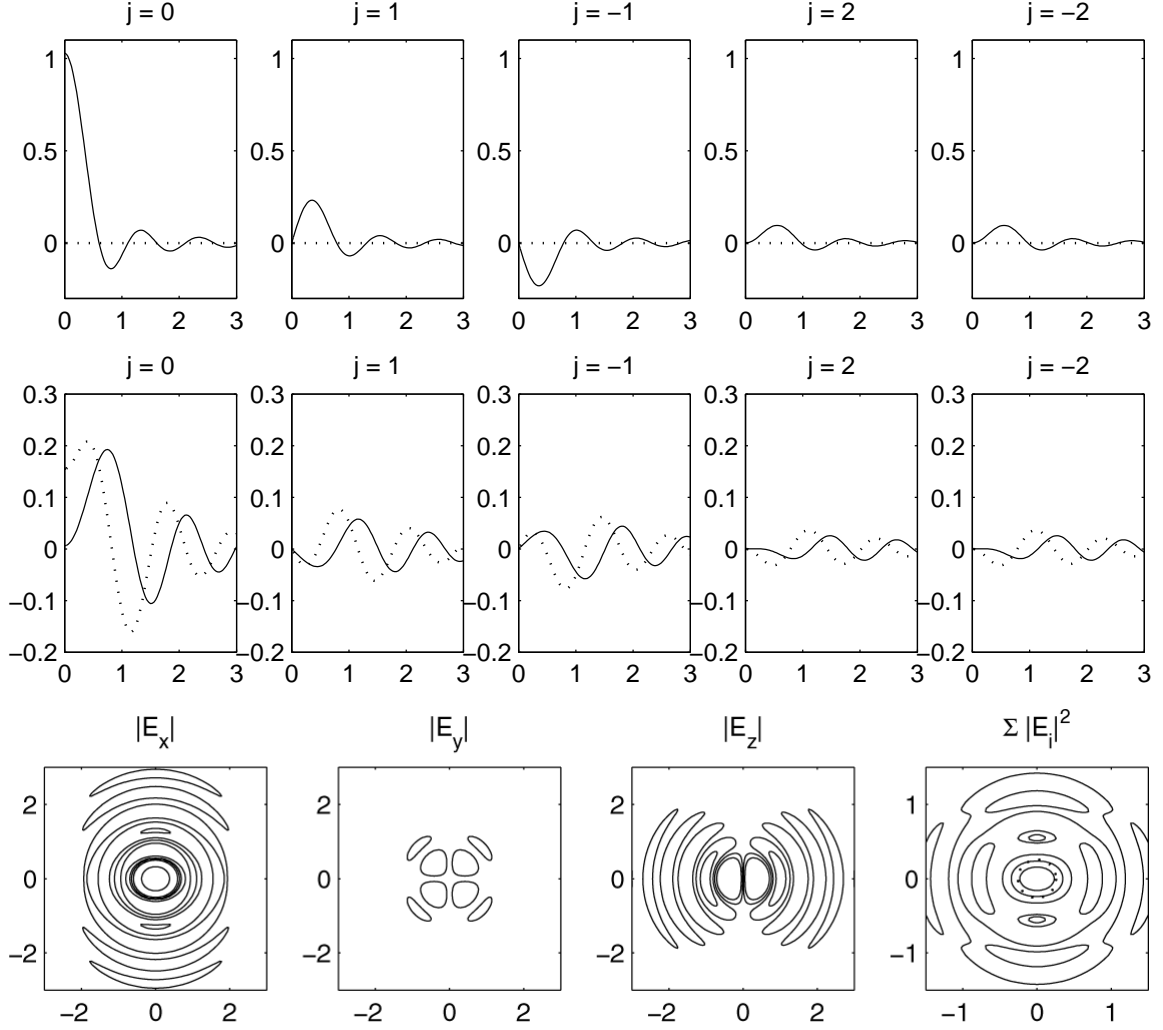


Fig. 2. The functions $V_{0,j}^0(r, f)$ (upper row: $f = 0$, middle row: $f = 2\pi$) for the aberration-free case (NA=0.95, linear polarization along the x -direction). The horizontal coordinate r is expressed in the diffraction unit λ/s_0 with s_0 the numerical aperture of the imaging system. The solid and dotted lines in the first and second row apply to, respectively, the real and imaginary part of the $V_{0,j}^0$ -functions. Lower row: contour plots of the three electric field components, $|E_x|$, $|E_y|$ and $|E_z|$, and of the electric energy density $\sum |E_i|^2$. The contour lines for the electric field components have been chosen at 0.5, 0.09 and 0.025; for the electric energy density the levels are 0.75, 0.50, 0.25, 0.017 and 0.005. In the latter contour plot, the dotted circle indicates the circular 0.50-contour of the hypothetical in-focus scalar intensity distribution.

the weakest of all three. In the intensity plot we have also given the 50%-level of the scalar Airy distribution with the transverse position expressed in units of λ/NA . It is

clearly visible from the contour plot that the full width at half maximum (FWHM) of the high-NA intensity distribution is slightly smaller in the y -cross-section but drastically increased along the x -cross-section. From the contour plots of the three field components, it can be seen that this effect is created by the E_z -field component that constitutes the most prominent extra feature at high-NA values with respect to the scalar case.

B. The definition of Strehl-intensity at high NA

A close inspection of the upper-left graph of Fig.2, representing the function $V_{0,0}^0(r, 0)$ at NA=0.95, reveals that the on-axis amplitude in the nominal focal plane is not equal to unity but slightly larger because of the vector addition and the radiometric effect. In the absence of aberrations we define the Strehl intensity by

$$I_S = \frac{\left| \frac{1}{\pi} \int_0^{2\pi} \int_0^1 A(\rho, \theta) \rho d\rho d\theta \right|^2}{\frac{1}{\pi} \int_0^{2\pi} \int_0^1 |A(\rho, \theta)|^2 \rho d\rho d\theta} . \quad (37)$$

With the amplitude function

$$A(\rho, \theta) = \frac{1}{(1 - s_0^2 \rho^2)^{1/4}} \left[\frac{1 + (1 - s_0^2 \rho^2)^{1/2}}{2} \right] , \quad (38)$$

we find the following expression for the Strehl-intensity

$$I_S = \left(\frac{8}{75s_0^2} \right) \frac{[8 - 5(1 - s_0^2)^{3/4} - 3(1 - s_0^2)^{5/4}]^2}{4 + 3s_0^2 - (1 - s_0^2)^{1/2} \{4 - s_0^2\}} \quad (39)$$

For the value of $s_0=0.95$, the on-axis intensity is 1.05856 (numerator of Eq.(37)) and, after normalisation to the incident power (1.060075), we find a Strehl intensity I_S equal to 0.99857. All calculated intensity values in the remainder of this paper have been normalized according to this definition.

C. Field components and intensity belonging to an aberrated system (coma)

In Fig.3 we have plotted the set of graphs belonging to a wavefront in the exit pupil (linear polarization along the x -direction) that has a comatic aberration of the lowest order. The wavefront aberration is given by $2\pi W^x(\rho, \theta) = \Phi(\rho, \theta) = \alpha_3^1 R_3^1(\rho) \cos \theta$ and the corresponding coefficients $\beta_{n,x}^m$ have been calculated by the insertion of this expression in Eq.(1). In the upper row we have now plotted the radial cross-sections of functions $V_{3,j}^1(r, 0)$; in the graphs of the middle row, the same functions appear with a defocus value of $f = 2\pi$. The functions now show a more general behavior and there is no identity or change of sign, respectively, for indices j that are even or odd. The field components (moduli) and the intensity pattern are given in the lower row and they show the typical $\cos(\phi)$ -asymmetry. But on top of this basic azimuthal frequency, we expect higher frequency components because of the presence of azimuthal components with ranges from -3ϕ to $+3\phi$ for the E_x - and E_y -components and from -2ϕ to $+2\phi$ for the E_z -component. After the squaring operation, we thus can expect a highest

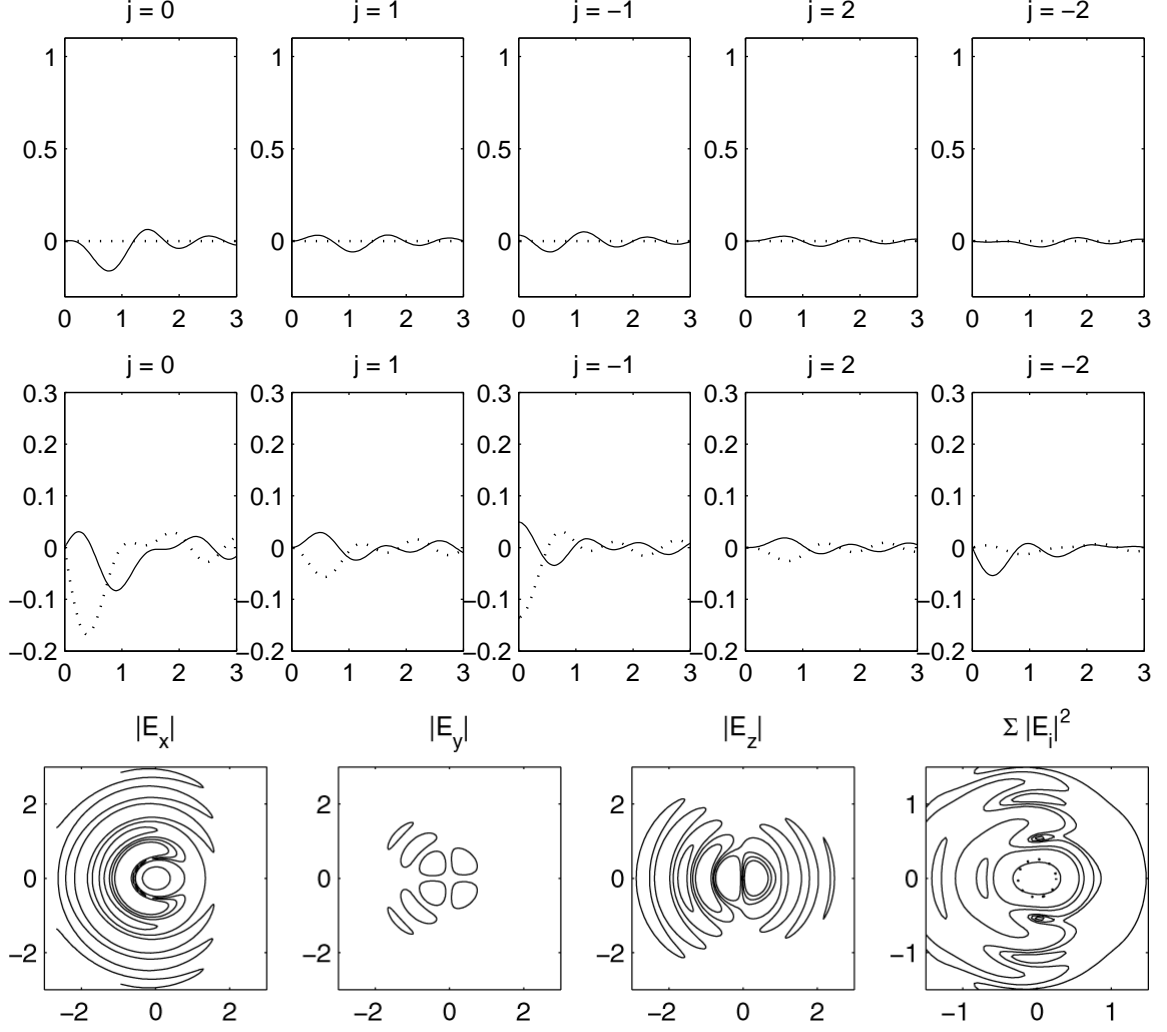


Fig. 3. Same legend as in Fig.2 but now with comatic aberration of lowest order ($\alpha_3^1 = 1$). The values of the relevant $\beta_{n,x}^m$ -coefficients (second-order approximation of the phase aberration function) are: $\beta_{0,x}^0 = 15/16$, $\beta_{2,x}^0 = -1/80$, $\beta_{4,x}^0 = -1/16$, $\beta_{6,x}^0 = -9/80$, $\beta_{3,x}^1 = \beta_{3,x}^{-1} = i/2$, $\beta_{2,x}^2 = \beta_{2,x}^{-2} = -1/20$, $\beta_{6,x}^2 = \beta_{6,x}^{-2} = -3/40$; all $\beta_{n,y}^m$ identical zero. In the contour plot of the energy density (lower row, right-hand figure) the contour levels have been chosen 0.5, 0.1, 0.05, 0.01, 0.005 and 0.002. For reasons of comparison, we have also included the dotted contour plot in the center corresponding to the 50% relative height for the hypothetical scalar diffraction image (same comatic aberration value).

azimuthal frequency of 6ϕ in the intensity pattern of the lower-right graph. In this graph, for comparison, we have again plotted the FWHM-contour (dotted) of the hypothetical scalar diffraction pattern at the same value of the numerical aperture (aberration-free case).

D. Construction of the image intensity from the G -functions

In the retrieval procedure that was presented in Section 5, the functions $G_{kl}(r, f)$ play a key role. The ranges of values of (k, l) are limited by the condition $|k|, |l| \leq 2$ and there are in total eight G -functions that occur in the expression for the energy density in the focal region, see Eq.(13). In Fig.4 we have plotted these eight functions

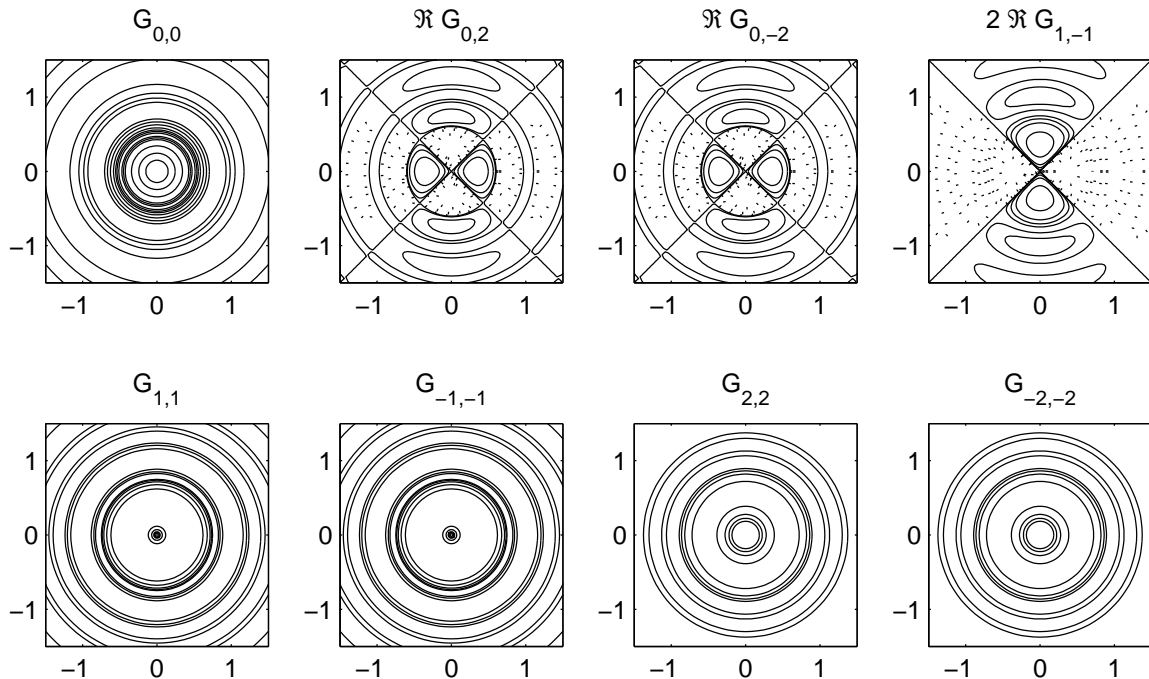


Fig. 4. The eight $G_{kl}(r, f)$ -functions that contribute to the energy density in the focal volume and that are used in the aberration retrieval scheme. The unit along the axes is the diffraction unit. The contour plots apply to the aberration-free case in the nominal focal plane ($f=0$, $\text{NA}=0.95$). To visualize the features of the various functions, the contour levels have been changed from plot to plot. $G_{0,0}$: 0.75, 0.5, 0.25, 0.10, 0.07, 0.05, 0.02, 0.01, 0.005 and 0.001 ; $\text{Re}[G_{0,2}]$, $\text{Re}[G_{0,-2}]$ and $2\text{Re}[G_{1,-1}]$: 0.055, 0.015, 0.005, 0.001, 0, -0.001, -0.005, -0.015, -0.055 (contours with negative values are dotted); $G_{1,1}$, $G_{-1,-1}$, $G_{2,2}$ and $G_{-2,-2}$: 0.12, 0.06, 0.005, 0.002, 0.001, 0.0005. Note that the functions $G_{1,1}$, $G_{-1,-1}$, $G_{2,2}$ and $G_{-2,-2}$ all have a dough-nut shape with a zero on axis.

according to their relative importance in the expression for the energy density (weight factor containing s_0); the imaging system is free of any aberration. The dominating G -function is $G_{0,0}$ and this is the only one that remains in the limiting case of very small numerical aperture ($s_0 \rightarrow 0$). In the nominal aberration-free focus, this function equals $|V_{0,0}^0(r, f)|^2$ and it closely resembles the basic Airy diffraction pattern if we neglect the influence of the radiometric effect and an average increase in lateral size due to the vectorial image formation. The functions in the lower row with equal k - and l -

indices have a central zero in the nominal focal plane. They lead to a further 'blurring' of the diffraction image because of the vectorial effects. The functions in the upper row with $|k - l| = 2$ lead to the absence of rotational symmetry in image formation with e.g. linearly polarized light. These functions are not everywhere positive as were the functions with $k = l$. The contributions to the focal plane intensity of the G -functions with $|k - l| = 2$ is proportional to $\text{Re}[G_{0,2} + G_{0,-2} - 2G_{1,-1}]$ if we have linearly polarized light along the x -direction in the entrance pupil (see Eq.(13) with $\beta_y = 0$). From Fig.4 we see that the contributions from the three non-circularly symmetric G -functions lead to a broadening of the central lobe along the x -cross-section (FWHM is 34% larger than that of the hypothetical scalar profile). Their negative contributions along the y -cross-section lead to a narrowing of the intensity profile and a reduced FWHM-value (-7% with respect to the 'scalar' profile). The asymmetry effect leads to an elliptic 50%-intensity contour and the ratio of the long and short axis amounts to 1.44. In Fig.5 we have plotted the G -functions in the same

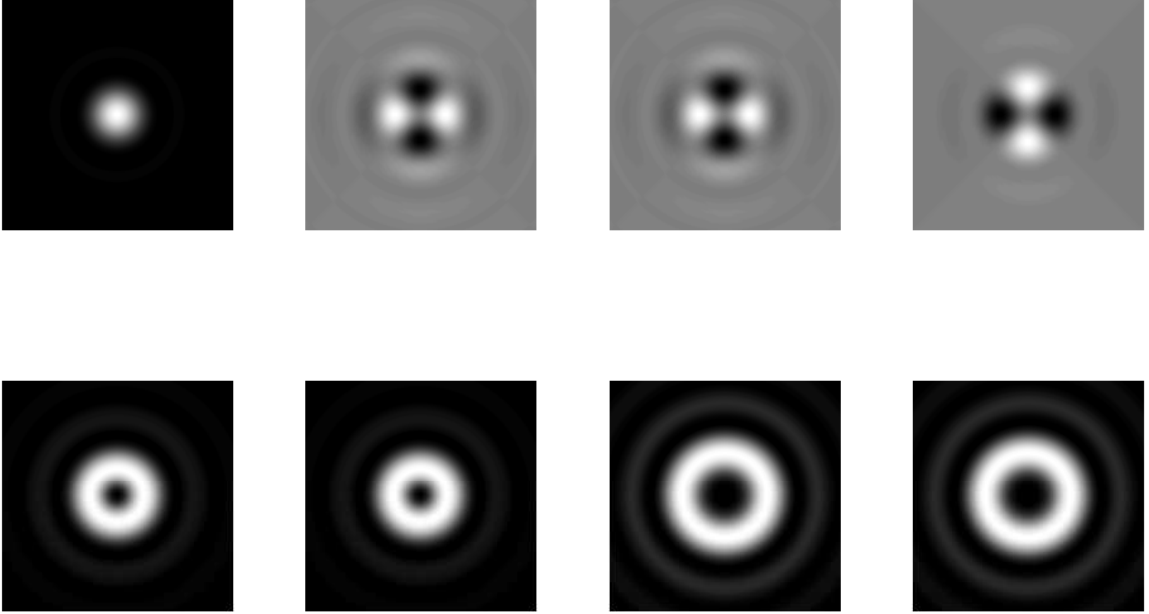


Fig. 5. Gray-level plots of the G -functions for the aberration-free case in the nominal focal plane. The order of representation is the same as in Fig.4. The plots of the functions $\text{Re}[G_{0,2}]$, $\text{Re}[G_{0,-2}]$ and $2\text{Re}[G_{1,-1}]$ have been coded with gray for zero level and with white and black shades for positive and negative values, respectively. Note the dough-nut shape of the functions $G_{1,1}$, $G_{-1,-1}$, $G_{2,2}$ and $G_{-2,-2}$ in the lower row. There is no relationship between the gray levels in the various graphs; all levels are relative with respect to the local maximum or the zero level.

order but now using gray scale levels to represent the intensity contributions. This is especially useful for a comparison of the intensity levels of the G_{kk} -functions and

to make clear the doughnut-shape of the G -functions with $k = l \neq 0$. Note that in image formation with natural light or circularly polarized light we only encounter the G_{kk} -functions in the expression for the energy density. This means that for these cases the vectorial image formation will always lead to larger FWHM-values than those given by the scalar extrapolation.

E. Application of the G -functions to a general aberrated system

In this subsection we discuss the G -functions for a general aberrated case with linearly polarized illumination in the x -direction. As the aberration function we choose lowest order astigmatism with $2\pi W^x(\rho, \theta) = \Phi(\rho, \theta) = \alpha_2^2 R_2^2(\rho) \cos 2(\theta - \theta_0)$. For the off-set angle we take $\theta_0 = \pi/6$ and the amplitude of the phase aberration is given by $\alpha_2^2 = 1$. A first approximation of the main aberration coefficients is given by $\beta_{2,x}^2 = i(1/4 - i\sqrt{3}/4)$ and $\beta_{2,x}^{-2} = i(1/4 + i\sqrt{3}/4)$ and the remaining β_x -coefficients are found from the Zernike expansion of the function $\exp\{i\Phi(\rho, \theta)\}$, approximated up to the second order; all β_y -coefficients are identical zero. In Fig.(6) we have plotted the corresponding G_{kl} -functions for a defocus value $f = 0$. We see that the G_{00} -function resembles the intensity profile we would expect in the case of astigmatic aberration with an azimuth of $\pi/6$. The G_{kk} -functions with $k \neq 0$ in the middle row also show the astigmatic behaviour with maximum azimuthal periodicities of 4ϕ . The functions $G_{1,1}$ and $G_{-1,-1}$, as well as $G_{2,2}$ and $G_{-2,-2}$, show a mutual rotation of $\pi/2$. This follows from e.g. Eq.(16) for the astigmatic case because the intensity contribution with periodicity 2ϕ in these functions changes sign when the k -index of the G_{kk} -functions changes sign. The three G -functions ($\text{Re}[G_{0,2}]$, $\text{Re}[G_{0,2}]$ and $\text{Re}[G_{1,-1}]$), that introduce by themselves an extra periodicity of 2ϕ in the intensity pattern due to the linear state of polarization, show the most complicated patterns because of the summation of periodicities with an azimuthal off-set of $\pi/6$. The intensity distributions can show periodicities up to a frequency of 8ϕ . In the retrieval approach we have chosen to select the azimuthal periodicities from the through-focus intensity distribution because these periodicities have a virtually straightforward relationship with the azimuthal periodicities we encounter in the aberration function. Finally, in the lower row of Fig.6, we have plotted the intensity distribution in three focal planes, resp. with f -parameters of $-\pi/3$, 0 and $+\pi/3$. The axial settings $f = \pm\pi/3$ correspond well with the positions of the astigmatic focal lines according to scalar diffraction theory. The focal lines are clearly visible as well as their distortion due to the nonparallelism of the linear state of polarization and the astigmatic principal cross-sections. The orientation of the elliptic shape in focus due to the linear state of polarization is preserved on both sides of focus; the two astigmatic focal lines are perpendicular to each other. The combined effect at an angular off-set of $\pi/6$ leads to the typical pattern in the last row of the picture.

F. Retrieval at high-NA with unpolarized light

Finally, we present in this section a relatively simple numerical example that shows the importance of including the high-NA imaging effects in the retrieval scheme. We consider the frequently occurring situation, e.g. like in a lithographic projection

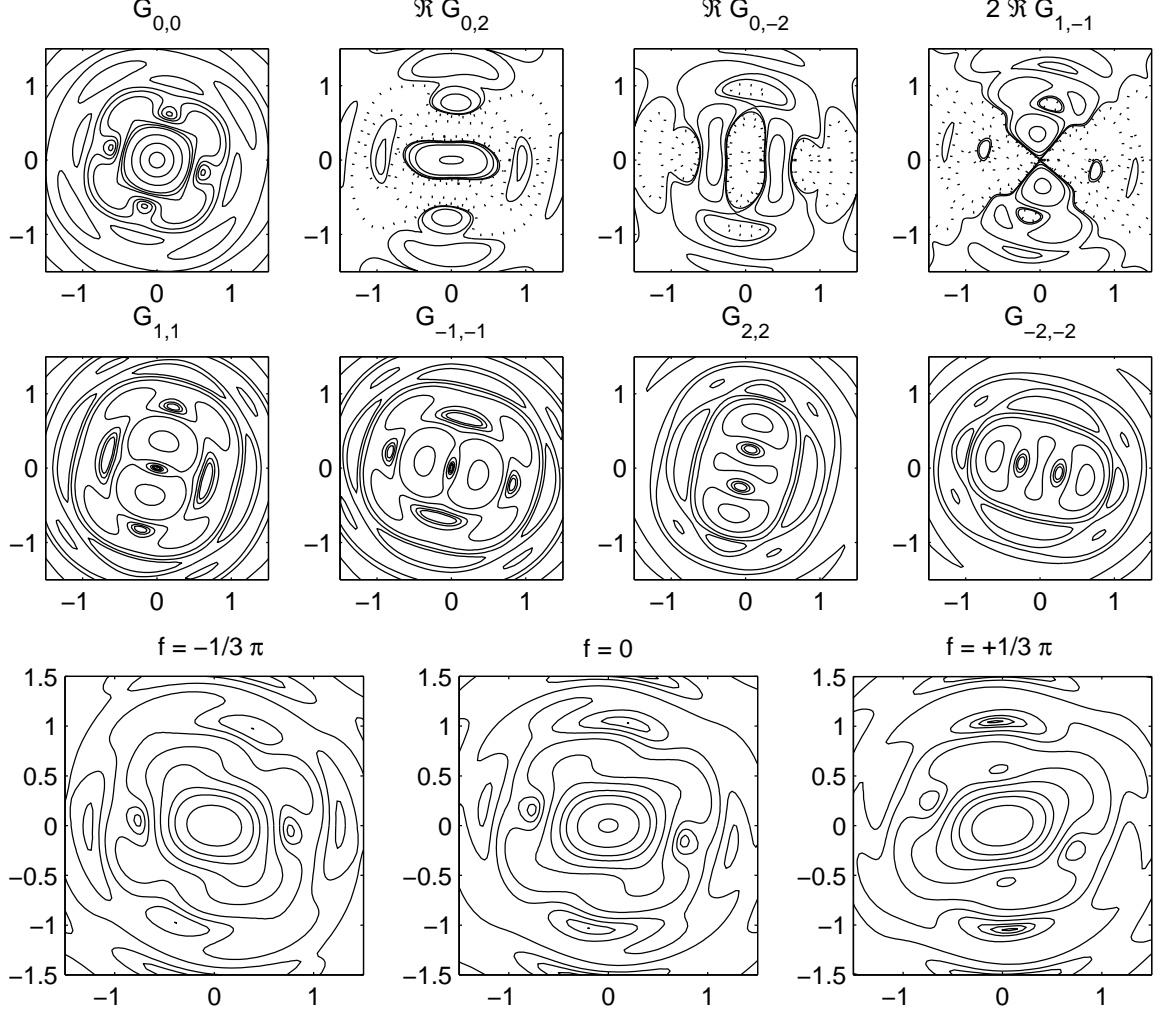


Fig. 6. Same legend as in Fig.4 for the upper and middle row. The graphs apply to a system with astigmatic wavefront aberration and the β -coefficients (second order approximation of the aberration function) are given by the following values: $\beta_{0,x}^0=11/12$, $\beta_{2,x}^0=-1/8$, $\beta_{4,x}^0=-1/24$, $\beta_{2,x}^2=1/4(\sqrt{3}+i)$, $\beta_{2,x}^{-2}=1/4(-\sqrt{3}+i)$, $\beta_{4,x}^4=1/16(1+i\sqrt{3})$, $\beta_{4,x}^{-4}=1/16(1-i\sqrt{3})$. In the lower row, contour plots are given for the astigmatic focal distribution with defocus values f of $-\pi/3$, 0 and $+\pi/3$, respectively, where the defocus values $f = \pm\pi/3$ approximately correspond to the image positions of the two focal lines of the astigmatic pencil. The choice of the various contour levels is identical to that in Fig.4.

system, that the illumination is unpolarized ('natural' light). Unpolarized light is represented in our analysis by adding incoherently two orthogonal polarization states; for the sake of simplicity, we select linearly polarized light along the x - and y -direction. If we turn to Eq.(A1) and carry out the summation of both orthogonally polarized

contributions with, respectively, $(a, b) = (1, 0)$ and $(a, b) = (0, 1)$, we find

$$\begin{aligned} \Psi_{an}^m(r, f) = & \\ \beta_0^0 \sum_{\nu} \left\{ \beta_{\nu}^{m*} \left[\Psi_{\nu;0,0}^{m*} + s_0^2 \left\{ \Psi_{\nu;1,1}^{m*} + \Psi_{\nu;-1,-1}^{m*} \right\} + \frac{s_0^4}{2} \left\{ \Psi_{\nu;2,2}^{m*} + \Psi_{\nu;-2,-2}^{m*} \right\} \right. \right. & \\ \left. \left. + \beta_{\nu}^{-m} (1 - \epsilon_{\nu,m}) \left[\Psi_{\nu;0,0}^{-m} + s_0^2 \left\{ \Psi_{\nu;1,1}^{-m} + \Psi_{\nu;-1,-1}^{-m} \right\} + \frac{s_0^4}{2} \left\{ \Psi_{\nu;2,2}^{-m} + \Psi_{\nu;-2,-2}^{-m} \right\} \right] \right\} & \end{aligned} \quad (40)$$

and this function is then used to construct the linear system of equations as indicated in Section 5.

In this numerical experiment, we used an asymmetric set of Zernike coefficients according to: $\beta_3^{+1} = \beta_3^{-1} = 0.1i$, $\beta_5^{+1} = \beta_5^{-1} = -0.02i$ and $\beta_7^{+1} = \beta_7^{-1} = -0.02i$. The fact that the (small) coefficients are purely imaginary implies that the lens defect can be contributed to wavefront aberration, of comatic nature in this case. The through-focus intensity distribution has been calculated using the basic result from Eqs.(2)-(4) with a value of the numerical aperture equal to 0.95 (refractive index n equals unity). The retrieval scheme is then applied using values of the numerical aperture in the range from 0.001 to 0.95. In Fig. 7 we show that the correct values are retrieved only if the exact value of the numerical aperture is used and thus the influence of the vectorial effects is correctly included. The scalar case ($NA \rightarrow 0$) shows a substantial deviation from the correct lens values of the order of 10 to even 100%, especially for the higher order β -coefficients.

Data sets taken from intensity patterns in the focal volume at different incident states of polarization could not be studied yet. These data sets at high NA are not easily created in a lithographic projection system since the standard illumination system has not been provided with special polarization means. For this reason, a complete check of the vectorial retrieval scheme, including the birefringence effects of the projection lens, could not be carried out yet. With respect to the sensitivity of the method and the range of aberration that can be covered, we refer to two recent publications³⁷⁻³⁸ for the scalar case. Here it is first shown that retrieval operations remain stable down to Strehl intensity levels as small as 0.30; or, equivalently, *rms* wavefront aberration values are allowed up to twice the diffraction limit ($150 \text{ m}\lambda$). The robustness of the retrieval method in the presence of noise and off-sets has also been studied in these references. When using a basically identical retrieval scheme as the one described in the preceding chapter, signal to noise ratios in the intensity patterns as low as 10 to 5 can be allowed without compromising the retrieved lens data.

7. Conclusion and outlook

We have presented a high-numerical-aperture analysis of the intensity distribution in the focal volume of an imaging system using the vectorial version of the extended Nijboer-Zernike approach. The three-dimensional intensity distribution has been obtained by means of a series expansion of basic functions in the Nijboer-Zernike theory

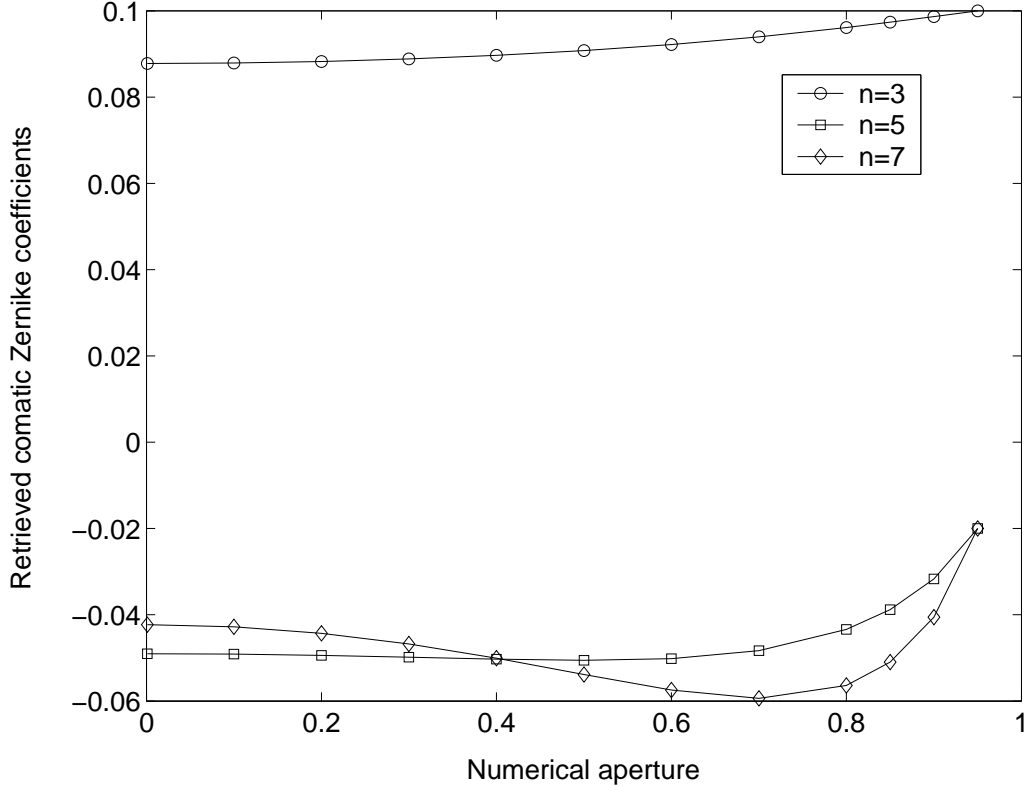


Fig. 7. Variation in the retrieved value of optical aberration coefficients when an incorrect value of the numerical aperture is used and the vectorial imaging effects are not correctly applied. Forward calculation at an NA-value of 0.95. Retrieval of the β -aberration coefficients (comatic wavefront aberration of 3rd, 5th and 7th order, resp. with Zernike coefficients of +0.1, -0.02 and -0.02 radians) at various values of NA. The correct aberration values are only retrieved when the NA-value at retrieval is chosen identical to the value used in the forward calculation scheme. The retrieved values for the scalar scheme, +0.088, -0.050 and -0.042, respectively, are found in the graph at the abscissa-value NA=0.

using generalized aberration coefficients related to both the amplitude and the phase of the complex pupil function. For high-NA imaging systems, the original scalar theory has been extended to the vectorial case and an extended set of aberration coefficients has been introduced, describing the behavior of the optical system as a function of the state of polarization of the incident light. It has been shown that the intensity distribution in the focal region can be constructed from a set of elementary functions that give rise to the basically non-circularly symmetric intensity profile in the focal region and to the relative increase in spot size at high NA as compared to the scalar prediction. Our theoretical approach has also shown that the collection of focal intensity data from four exposures with well-selected polarization states of the incident

light enables the retrieval of the 'polarization' aberration coefficients. They represent the geometrical aberrations and the spatial distribution of birefringence (azimuth and size) in the exit pupil of the optical system. A first numerical exercise to illustrate our theoretical approach has shown that, even in the case of illumination with unpolarized light, the correct aberration coefficients are only retrieved when the vectorial formulation of image formation is correctly included. Future experimental work should concentrate on the collection of through-focus intensity data for various incident polarization states in a high-NA imaging system; with such experimental data, the retrieval of polarization aberrations becomes possible along the lines described in this paper.

Acknowledgement

The authors would like to acknowledge stimulating discussions with Dr. S. Stallinga of Philips Research Laboratories, Eindhoven. Sven van Haver from Delft University has been very effective in producing the numerical results and graphical representations.

E-mail address corresponding author: j.j.m.braat@tnw.tudelft.nl

References

- [1] D. Malacara, *Optical Shop Testing* (2nd edition, Wiley, Hoboken NJ, USA, 1992).
- [2] R. W. Gerchberg, and W. O. Saxton, "Phase determination from image and diffraction plane pictures in electron-microscope," *Optik* **34**, 277-286 (1971)
- [3] R. W. Gerchberg, and W. O. Saxton, "Practical algorithm for determination of phase from image and diffraction pictures," *Optik* **35**, 237-246 (1972)
- [4] B. R. Frieden, "Restoring with maximum likelihood and maximum entropy," *J. Opt. Soc. Am.* **62**, 511-518 (1972)
- [5] W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.* **62**, 55-59 (1972).
- [6] J. R. Fienup, "Phase retrieval algorithms - a comparison," *Appl. Opt.* **21**, 2758-2769 (1982).
- [7] J. R. Fienup, J.C. Marron, T. J. Schultz, and J. H. Seldin, "Hubble space telescope characterized by using phase-retrieval algorithms," *Appl. Opt.* **32**, 1747-1767 (1993).
- [8] J. G. Walker, "The phase retrieval problem - a solution based on zero location by exponential apodization," *Opt. Acta* **28**, 735-738 (1981).
- [9] R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Opt. Eng.* **21**, 829-832 (1982).
- [10] J. R. Fienup, "Phase retrieval for undersampled broadband images," *J. Opt. Soc. Am.* **A16**, 1831-1837 (1999).
- [11] D. Van Dyck, and W. Coene, "A new procedure for wavefunction restoration in high-resolution electron-microscopy," *Optik* **77**, 125-128 (1987).
- [12] B. R. Frieden, and C. Oh, "Multiple filter approach to phase retrieval from modulus data," *Appl. Opt.* **31**, 1103-1108 (1992).
- [13] E. S. Angel, and A. K. Jain, "Restoration of images degraded by spatially varying point-spread functions by a conjugate gradient method," *Appl. Opt.* **17**, 2186-2190 (1978).
- [14] R. Barakat, and B. H. Sandler, "Determination of the wavefront aberration function from measured values of the point-spread function - a 2-dimensional phase retrieval problem (least squares method)," *J. Opt. Soc. Am.* **A 9**, 1715-1723 (1992).
- [15] I. Iglesias, "Parametric wave-aberration retrieval from point-spread function data by use of pyramidal recursive algorithm," *Appl. Opt.* **37**, 5427-5430 (1998).
- [16] B. M. Hanser, M. G. L. Gustafsson, D. A. Agrad, and J.W. Sedat, "Phase retrieval for high-numerical-aperture optical systems," *Opt. Lett.* **28**, 801-803 (2003).
- [17] J. Wesner, J. Heil and T. Sure, "Reconstructing the pupil function of microscope objectives from the intensity PSF", in: *Current Developments in Lens Design and Optical Engineering III*, eds. R.E. Fischer, W.J. Smith, R.B. Johnson, Proc. SPIE **4767**, pp. 32-43 (2003).
- [18] P. Dirksen, J. J. M. Braat, A. J. E. M. Janssen, and C. Juffermans, "Aberration retrieval using the extended Nijboer-Zernike approach," *J. Microlithogr. Microfabr. Microsyst.* **2**, 61-68 (2003).

- [19] A. J. E. M. Janssen, "Extended Nijboer-Zernike approach for the computation of optical point-spread functions," *J. Opt. Soc. Am.* **A 19**, 849-857 (2002).
- [20] J. J. M. Braat, P. Dirksen, A. J. E. M. Janssen, "Assessment of an extended Nijboer-Zernike approach for the computation of optical point-spread functions," *J. Opt. Soc. Am.* **A 19**, 858-870 (2002).
- [21] W. Ignatowsky, "Diffraction by a lens of arbitrary aperture," *Trans. Opt. Inst* **1(4)**, 1-36 (1919).
- [22] B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems II. Structure of the image field in an aplanatic system," *Proc. Roy. Soc. A* **253**, 358-379 (1959).
- [23] M. Mansuripur, "Distribution of light at and near the focus of high-numerical-aperture objectives," *J. Opt. Soc. Am.* **A 3**, 2086-2093 (1986).
- [24] M. Mansuripur, "Certain computational aspects of vector diffraction problems," *J. Opt. Soc. Am.* **A 6**, 786-805 (1989).
- [25] R. Kant, "An analytical solution of vector diffraction for focusing optical systems," *J. Mod. Opt.* **40**, 337-348 (1993).
- [26] R. Kant, "An analytical solution of vector diffraction for focusing optical systems with Seidel aberrations," *J. Mod. Opt.* **40**, 2293-2311 (1993).
- [27] C. J. R. Sheppard and P. Török, "Efficient calculation of electromagnetic diffraction in optical systems using a multipole expansion," *J. Mod. Opt.* **44**, 803-819 (1997).
- [28] J. J. M. Braat, P. Dirksen, A. J. E. M. Janssen, and A. S. van de Nes, "Extended Nijboer-Zernike representation of the vector field in the focal region of an aberrated high-aperture optical system," *J. Opt. Soc. Am.* **A 20**, 2281-2292 (2003).
- [29] J. D. Jackson, *Classical Electrodynamics* (3rd edition, Wiley, Hoboken NJ, USA, 1998).
- [30] R. S. Longhurst, *Geometrical and Physical Optics* (3rd edition, Longman, London, UK, 1974).
- [31] J. P. McGuire and R. A. Chipman, "Diffraction image formation in optical systems with polarization aberrations. I: Formulation and example," *J. Opt. Soc. Am.* **A 7**, 1614-1626 (1990).
- [32] S.-Y. Lu and R. A. Chipman, "Homogeneous and inhomogeneous Jones matrices," *J. Opt. Soc. Am.* **A 11**, 766-773 (1994).
- [33] G. D. VanWiggeren and R. Roy, "Transmission of linearly polarized light through a single-mode fiber with random fluctuations of birefringence," *Appl. Opt.* **38**, 3888-3892 (1999).
- [34] S. Stallinga, "Axial birefringence in high-numerical-aperture optical systems and the light distribution close to focus," *J. Opt. Soc. Am.* **A 18**, 2846-2859 (2001).
- [35] S. Stallinga, "Light distribution close to focus in biaxially birefringent media," *J. Opt. Soc. Am.* **A 21**, 1785-1798 (2004).
- [36] S. Stallinga, "Strehl ratio for focusing into biaxially birefringent media," *J. Opt. Soc. Am.* **A 21**, 2406-2413 (2004).
- [37] C. van der Avoort, J.J.M. Braat, P. Dirksen, A.J.E.M. Janssen, "Aberration

retrieval from the intensity point-spread function in the focal region using the extended Nijboer-Zernike approach,” to appear in *J. Mod. Opt.*, 2005.

- [38] P. Dirksen, J.J.M. Braat, A.J.E.M. Janssen, A. Leeuwestein, ”Aberration retrieval for high-NA optical systems using the extended Nijboer-Zernike theory,” to appear in *Proc. SPIE 5754*, Conference on Microlithography 2005, San Jose, USA, February 26 - March 4, 2005.

Appendix A

Expression for the azimuthal harmonic components of the analytically derived intensity distribution in the focal volume

To solve the basic 'retrieval' problem as represented by Eq.(35), we need analytic expressions for the azimuthal harmonic components that are present in the linearized intensity distribution through the focal volume of an aberrated imaging system. The incident illumination is a linear superposition of two orthogonal linear states of polarization (with complex amplitudes a and b , respectively, for the x - and y -polarized states). After some straightforward manipulation one obtains for the component with upper index m

$$\begin{aligned}
\Psi_{an}^m(r, f) = & \beta_0^0 \sum_{\nu} \left\{ \beta_{\nu}^{m*} \left[\Psi_{\nu;0,0}^{m*} + s_0^2 \left\{ \Psi_{\nu;1,1}^{m*} + \Psi_{\nu;-1,-1}^{m*} \right\} + \frac{s_0^4}{2} \left\{ \Psi_{\nu;2,2}^{m*} + \Psi_{\nu;-2,-2}^{m*} \right\} \right. \right. \\
& \left. \left. - 2\text{Im}(ab^*) \left\{ s_0^2 \left\{ \Psi_{\nu;1,1}^{m*} - \Psi_{\nu;-1,-1}^{m*} \right\} - \frac{s_0^4}{2} \left\{ \Psi_{\nu;2,2}^{m*} - \Psi_{\nu;-2,-2}^{m*} \right\} \right\} \right] \right. \\
& + \beta_{\nu}^{-m} (1 - \epsilon_{\nu,m}) \left[\Psi_{\nu;0,0}^{-m} + s_0^2 \left\{ \Psi_{\nu;1,1}^{-m} + \Psi_{\nu;-1,-1}^{-m} \right\} + \frac{s_0^4}{2} \left\{ \Psi_{\nu;2,2}^{-m} + \Psi_{\nu;-2,-2}^{-m} \right\} \right. \\
& \left. \left. - 2\text{Im}(ab^*) \left\{ s_0^2 \left\{ \Psi_{\nu;1,1}^{-m} - \Psi_{\nu;-1,-1}^{-m} \right\} - \frac{s_0^4}{2} \left\{ \Psi_{\nu;2,2}^{-m} - \Psi_{\nu;-2,-2}^{-m} \right\} \right\} \right] \\
& + s_0^2 \left\{ (|a|^2 - |b|^2) \text{Re} \left[\beta_{\nu}^{(m-2)*} \Psi_{\nu;0,2}^{(m-2)*} \right] - 2\text{Re}(ab^*) \text{Im} \left[\beta_{\nu}^{(m-2)*} \Psi_{\nu;0,2}^{(m-2)*} \right] \right\} \\
& + s_0^2 \left\{ (|a|^2 - |b|^2) \text{Re} \left[\beta_{\nu}^{(m+2)*} \Psi_{\nu;0,-2}^{(m+2)*} \right] + 2\text{Re}(ab^*) \text{Im} \left[\beta_{\nu}^{(m+2)*} \Psi_{\nu;0,-2}^{(m+2)*} \right] \right\} \\
& - 2s_0^2 \left\{ (|a|^2 - |b|^2) \text{Re} \left[\beta_{\nu}^{(m+2)*} \Psi_{\nu;+1,-1}^{(m+2)*} \right] + 2\text{Re}(ab^*) \text{Im} \left[\beta_{\nu}^{(m+2)*} \Psi_{\nu;+1,-1}^{(m+2)*} \right] \right\} \\
& + s_0^2 (1 - \epsilon_{\nu,-m+2}) \left\{ (|a|^2 - |b|^2) \text{Re} \left[\beta_{\nu}^{(-m+2)} \Psi_{\nu;2,0}^{(-m+2)} \right] \right. \\
& \quad \left. - 2\text{Re}(ab^*) \text{Im} \left[\beta_{\nu}^{(-m+2)} \Psi_{\nu;2,0}^{(-m+2)} \right] \right\} \\
& + s_0^2 (1 - \epsilon_{\nu,-m-2}) \left\{ (|a|^2 - |b|^2) \text{Re} \left[\beta_{\nu}^{(-m-2)} \Psi_{\nu;-2,0}^{(-m-2)} \right] \right. \\
& \quad \left. + 2\text{Re}(ab^*) \text{Im} \left[\beta_{\nu}^{(-m-2)} \Psi_{\nu;-2,0}^{(-m-2)} \right] \right\} \\
& - 2s_0^2 (1 - \epsilon_{\nu,-m-2}) \left\{ (|a|^2 - |b|^2) \text{Re} \left[\beta_{\nu}^{(-m-2)} \Psi_{\nu;-1,+1}^{(-m-2)} \right] \right. \\
& \quad \left. + 2\text{Re}(ab^*) \text{Im} \left[\beta_{\nu}^{(-m-2)} \Psi_{\nu;-1,+1}^{(-m-2)} \right] \right\} , \tag{A1}
\end{aligned}$$

where the function Ψ has been defined in Eq.(33).

Note that in the special case treated in the numerical example of Section 6 ('natural' light), the only non-zero contributions to $\Psi_{an}^m(r, f)$ are those found on the second and fourth line of the equation above.