# Strehl ratio and optimum focus of high-numerical-aperture beams

**Augustus J.E.M. Janssen** Philips Research Europe, HTC 36, NL-5656 AE Eindhoven, The Netherlands

**Sven van Haver** Optics Research Group, Faculty of Applied Sciences, Technical University Delft, Lorentzweg 1, NL-2628 CJ Delft, The Netherlands

**Joseph J.M. Braat** Optics Research Group, Faculty of Applied Sciences, Technical University Delft, Lorentzweg 1, NL-2628 CJ Delft, The Netherlands

**Peter Dirksen** Philips Research Europe, HTC 4, NL-5656 AE Eindhoven, The Netherlands

We analytically calculate the focus setting for a beam with a high numerical aperture (NA) that optimizes its Strehl ratio in the case of small aberrations up to the 'just' diffraction-limited value (Strehl ratio $\geq$0.80). The optimum focus setting deviates from the one that follows from a minimization of the wavefront aberration with the aid of the Zernike aberration coefficients. This deviation stems largely from the fact that the common quadratic approximation of the focus term becomes inadequate in the high-NA case. Fundamental high-NA amplitude nonuniformity in the exit pupil of an optical system in the case of a linearly polarized incident beam also influences the optimum focus setting. Results for spherical aberration and astigmatism are presented for an NA-value of 0.95. [DOI: 10.2971/jeos.2007.07008]

**Keywords:** vector diffraction, aberrations, Strehl ratio, high-numerical-aperture beams

## 1 INTRODUCTION

Strehl ratio plays an important role in the assessment of optical systems as it is a generally accepted number to characterize the quality of the intensity impulse response (point-spread function) of an imaging system. In microscopy, optical data storage and optical projection lithography [1], very high values of numerical aperture are required for the imaging or read-out of the finest details. Obtaining the highest possible Strehl ratio and the corresponding optimum focus-setting, both in design and operation, is crucial for optimum printing and reading-out of sub-wavelength features. The expression for Strehl ratio and the best focus setting for these high-quality imaging systems with a very high numerical aperture is the subject of this paper.

The original definition by Strehl for optical system quality is given by the ratio of the central intensity in the diffraction image of an aberrated system and the theoretical maximum intensity in the unaberrated case. It is usually applied to systems with a uniform amplitude transmittance function but can be extended to more general systems [2, 3], where the exit pupil function $g(\rho,\theta)$ is given by

$$g(\rho,\theta) = A(\rho,\theta)\exp\{i\Phi(\rho,\theta)\}, \tag{1}$$

with $(\rho,\theta)$ the normalized polar coordinates on the exit pupil sphere, $A$ the normalized amplitude transmission function (normally taken equal to unity in the center of the exit pupil) and $\Phi$ the phase departure due to the wavefront aberration function $W$ in the exit pupil of the optical system ($\Phi = 2\pi W/\lambda$ with $\lambda$ the wavelength of the light). The Strehl ratio is then given by

$$S = \frac{\left|\frac{1}{\pi}\int_0^{2\pi}\int_0^1 A(\rho,\theta)\exp\{i\Phi(\rho,\theta)\}\rho d\rho d\theta\right|^2}{\left|\frac{1}{\pi}\int_0^{2\pi}\int_0^1 A(\rho,\theta)\rho d\rho d\theta\right|^2}, \tag{2}$$

leading to unity in the aberration-free case ($\Phi \equiv 0$).

The amplitude function can be partly measured in the entrance pupil of the optical system as far as it is related to the structure of the incident beam, $A_E(\rho,\theta)$, with the index $E$ referring to the entrance pupil of the system. In many test cases, this function will be rather uniform and can be put equal to unity. The function $A_E$ can be further modified on its way through the optical system and then has to be measured in the exit pupil of the imaging system, using $A(\rho,\theta) = \sqrt{I(\rho,\theta)}$, with $I(\rho,\theta)$ the intensity function over the beam cross-section on the exit pupil sphere. If the exit pupil is not accessible to measurements, $I(\rho,\theta)$ has to be measured in the optical 'far field'. In this paper we include a complication that arises in high-numerical-aperture imaging systems where the mapping of the amplitude from entrance to exit pupil is not obeying the simple scaling law of low-NA systems. Two effects need to be considered:

- Design of the optical system
  The mapping of the complex amplitude distribution from entrance to exit pupil is determined by the design choice. For large field systems, the sine condition by Abbe has to be satisfied [4]. In the case of an incident parallel beam,

Received January 8, 2007; published February 23, 2007 ISSN 1990-2573

the corresponding portions of the axial imaging pencils in object and image space are transformed in such a way that the amplitude in the high-NA image space is given by

$$A_N \propto \frac{1}{(1-s_0^2\rho^2)^{1/4}}\,, \tag{3}$$

with $s_0 = \sin\alpha = NA/n$ and $n$ the refractive index of the image space of the optical system (radiometric effect). The corresponding intensity is obtained from, for instance, a far field measurement.

- Vector diffraction
  When we take into account the vector character of the focusing beam and assume, for instance, an $x$-linearly polarized light beam in the entrance pupil, the on-axis amplitude in focus is only built up from the $x$-component of the field on the exit pupil sphere ($I(0,0) \propto E_x E_x^*$; see Figure 1 for the definition of the field vectors in the high-numerical-aperture image space. The vector diffraction effects in focus have been studied in Ref. [5]; later applications of this theory can be found in [6]-[11]. It follows that the contribution to the $x$-component in focus by the field components on the exit pupil is proportional to

$$A_x(\rho,\theta) \propto \frac{1+\sqrt{1-s_0^2\rho^2}-[1-\sqrt{1-s_0^2\rho^2}]\cos 2\theta}{2}. \tag{4}$$

  This distribution can be experimentally verified by a far-field intensity measurement using an $x$-analyzer that is positioned perpendicularly to the $z$-axis.

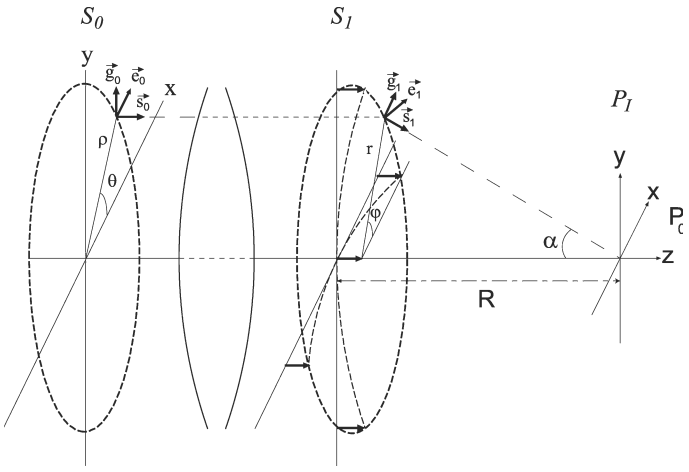


FIG. 1 The propagation of the incident wave from the entrance pupil $S_0$ through the optical system towards the exit pupil $S_1$ and the focal region at the image plane $P_I$. The orthogonal unit vectors for the field components and the wave vector are indicated in object and image space by $(\vec{e}_0,\vec{g}_0,\vec{s}_0)$ and $(\vec{e}_1,\vec{g}_1,\vec{s}_1)$, respectively. A point on the exit pupil sphere is defined by means of the polar coordinates $(\rho,\theta)$. The aperture (NA) of the imaging pencil is given by $s_0 = \sin\alpha_{max}$.

The combination of the various contributions leads to an amplitude function according to

$$A_I(\rho,\theta) = A_E(\rho,\theta)\left[\frac{1+\sqrt{1-s_0^2\rho^2}-[1-\sqrt{1-s_0^2\rho^2}]\cos 2\theta}{2(1-s_0^2\rho^2)^{1/4}}\right] \tag{5}$$

with $A_E(\rho,\theta)$ the incident field in the entrance pupil. We have added the index $I$ to emphasize that the amplitude function corresponds to an incident beam that is linearly polarized, in this particular case along the $x$-axis ($\theta = 0$).
In all cases of interest, $A_I(\rho,\theta)$ will be a smooth function but its deviation from unity can become important. This is illustrated in Figure 2 where we have plotted the amplitude function $A_I(\rho,\theta)$ of Eq. (5) for several values of the numerical aperture $s_0$ and for two values of the azimuth ($\theta = 0$ and $\pi/2$) where the deviation from unity is maximum. The function $A_E$ in the entrance pupil has been set equal to unity.

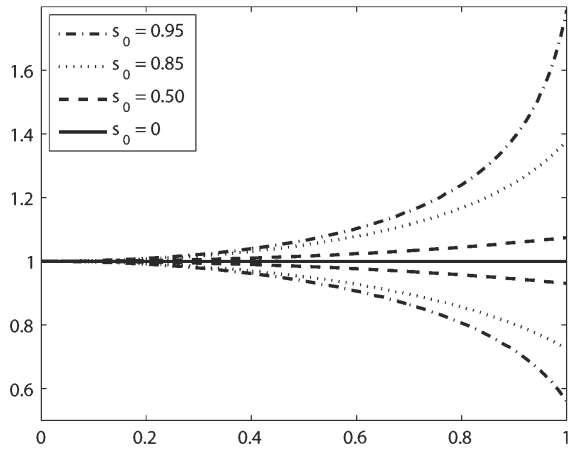


FIG. 2 Plot of the function $A_I(\rho,\theta)$, the contribution to the $E_x$-component in the focal region, as a function of the radial coordinate $\rho$, see Eq. (5), for two cross-sections: $\theta = 0$ with $A_I \le 1$ and $\theta = \pi/2$ ($A_I \ge 1$). The values of the numerical aperture $s_0$ are 0.0, 0.50, 0.85 and 0.95, respectively.

The curves in Figure 2 show that for the high value of $s_0$=0.95 there is, at the rim of the exit pupil, an apparent amplitude nonuniformity of more than a factor of three. We use the adjective 'apparent' to stress that it is the contribution to the $E_x$-field in the focal region that has been plotted in Figure 2, each point on the exit pupil corresponding to a particular wave vector direction in the focusing field.

The phase function $\Phi$ is also supposed to be a smooth function over the exit pupil. Moreover, if the quality of the optical system is within the 'just' diffraction-limited range, the aberration is weak and the amplitude of $\Phi$ certainly does not exceed a range of typically $\pi$. To allow a focal shift, we now extend the phase function with a defocusing term $\Phi_d(\rho)$. At low numerical aperture, the defocusing term is well represented by a simple quadratic exponential function of the normalized radial coordinate $\rho$ according to $\exp\{if\rho^2\}$. With our sign convention, the defocus parameter $f$ is related to an axial shift $z$ of the focal plane by $z = -\lambda f/(\pi n s_0^2)$. The focal depth $\delta_f$ is defined by $|f| = \pi/2$, leading to $\delta = \lambda/2ns_0^2$. At high values of $s_0$, the parabolic approximation is insufficient and the correct expression, see [11, 12], reads

$$\Phi_d(\rho) = f\,\frac{1-\sqrt{1-s_0^2\rho^2}}{1-\sqrt{1-s_0^2}} = f\Psi(\rho), \tag{6}$$

with $\Psi(\rho)$ a defocus function that will be further analyzed in Section 2 of this paper. The relationship between the focal

shift $z$ and the defocus parameter $f$ for the high-numerical-aperture case is given by

$$z = -\frac{\lambda}{2\pi n\left\{1-\sqrt{1-s_0^2}\right\}} f \tag{7}$$

and the focal depth $\delta_f$ now amounts to $\lambda/(4n\{1-\sqrt{1-s_0^2}\})$.

Inserting the high-NA amplitude function $A_l$ in Eq. (2) and adding the defocus term to the phase function, we obtain the following expression for the Strehl ratio,

$$S = \frac{\left|\frac{1}{\pi}\int_0^{2\pi}\int_0^1 A_l(\rho,\theta)\exp\{i[\Phi(\rho,\theta)+f\Psi(\rho)]\}\rho d\rho d\theta\right|^2}{\left|\frac{1}{\pi}\int_0^{2\pi}\int_0^1 A_l(\rho,\theta)\rho d\rho d\theta\right|^2}. \tag{8}$$

Under the above mentioned conditions (smooth functions, limited phase excursion), we are allowed, as it is customary in evaluating the Strehl ratio, to apply a series expansion up to the second order of the integrand in the numerator of the expression in Eq. (8). To this goal we will apply series expansions using Zernike polynomials to the functions $\Phi(\rho,\theta)$ and $\Psi(\rho)$. Strehl ratio analysis at low numerical aperture is already based on such expansions and here we extend this analysis to cover the particularities of the high-numerical-aperture case. It has turned out to be important to also have an adequate series expansion of the function $A_l(\rho,\theta)$. Although moderate amplitude variation over the beam cross-section in the exit pupil is commonly believed to be of minor influence on the Strehl ratio, our high-NA case ($s_0 = 0.95$) gives rise to such a modulation that it is absolutely necessary to include this effect.

In the remainder of this paper we first present the Zernike expansions of the various functions encountered in expression (8) for the Strehl ratio. The next step is to establsih the second order expansions of the integrands, to evaluate the various integrals and to find the expression for the optimum focus setting at high numerical aperture. At this optimum focus setting, we then analytically evaluate the Strehl ratio using the second order approximation. These results will then be compared to an analytic result where we have put $A_l \equiv 1$ and to numerical evaluations of the same quantities (focus setting and Strehl ratio) using (8) without any approximation. As a typical test cases, we will introduce spherical aberration and astigmatism of lowest order in the phase function $\Phi$ as these aberrations show the most pronounced effect on the optimum focus setting.

## 2 ANALYTIC EXPRESSION FOR STREHL RATIO AND OPTIMUM FOCUS

Having analyzed the Strehl ratio definition for the high-numerical-aperture case, we now introduce the Zernike expansions of the various functions encountered in the integrands of Eq. (8). Regarding the exponential phase function, in line with standard aberration analysis of $\Phi$, we write

$$\begin{aligned}
\Phi(\rho,\theta) &= \sum_{n,m}\alpha_n^m R_n^m(\rho)\cos m\theta\,,\\
\Phi_1(\rho,\theta) &= \left(\sum_{n,m}\alpha_n^m R_n^m(\rho)\cos m\theta\right) - \alpha_0^0 R_0^0(\rho)\,,\\
\Psi(\rho,\theta) &= \frac{1-\sqrt{1-s_0^2\rho_0^2}}{1-\sqrt{1-s_0^2}} = \sum_{n,m}\gamma_{2n}^0 R_n^m(\rho)\,,\\
\Psi_1(\rho,\theta) &= \left(\sum_{n,m}\gamma_{2n}^0 R_n^m(\rho)\right) - \gamma_0^0 R_0^0(\rho)\,.
\end{aligned} \tag{9}$$

The defocus phase term $\Psi$ has been developed into a radially symmetric Zernike expansion with even index coefficients $\gamma_{2n}^0$; the values of these coefficients are basically obtained by evaluating the inner products of $\Psi$ with the relevant radial Zernike polynomial. The results are given in Appendix A and have been earlier derived in Appendix B of Ref. [11]. We have introduced the functions $\Phi_1$ and $\Psi_1$ to split off the constant phase terms that are irrelevant for the determination of optimum focus and Strehl ratio. Note that we have limited ourselves in the aberration analysis to $\cos m\theta$-dependent aberration terms, but an extension of the analysis to the general case including aberration terms with arbitrary azimuthal dependence is straightforward.

With $A_E(\rho,\theta)\equiv 1$, the amplitude function $A_l$ is split into

$$\begin{aligned}
A_l(\rho,\theta) &= \left[\frac{1+\sqrt{1-s_0^2\rho^2}-[1-\sqrt{1-s_0^2\rho^2}]\cos 2\theta}{2(1-s_0^2\rho^2)^{1/4}}\right]\\
&= A^0(\rho) - A^2(\rho)\cos 2\theta\,,
\end{aligned} \tag{10}$$

with

$$A^0(\rho) = \sum_{n=0}^{\infty} a_{2n}^0 R_{2n}^0(\rho)\,, A^2(\rho) = \sum_{n=1}^{\infty} a_{2n}^2 R_{2n}^2(\rho)\,. \tag{11}$$

The expressions for the coefficients $a_{2n}^0$ and $a_{2n}^2$ as a function of $s_0$ can be found in Appendix A.

To calculate the approximated Strehl ratio we start by expanding the exponential of the integrand in the numerator of Eq. (8) up to second order according to $\exp(ix) = 1 + ix - \frac{1}{2}x^2$. The complex amplitude $U$ in the numerator (equal to $|U|^2$) is given by

$$\begin{aligned}
U &= \frac{1}{\pi}\int_0^{2\pi}\int_0^1 A_l(\rho,\theta)\exp\{i[\Phi_1(\rho,\theta)+f\Psi_1(\rho)]\}\rho d\rho d\theta\\
&\approx \frac{1}{\pi}\int_0^{2\pi}\int_0^1 \left(A^0(\rho)-A^2(\rho)\cos 2\theta\right)\times\\
&\left[1+i\{\Phi_1(\rho,\theta)+f\Psi_1(\rho)\}-\frac{1}{2}\{\Phi_1(\rho,\theta)+f\Psi_1(\rho)\}^2\right]\rho d\rho d\theta.
\end{aligned} \tag{12}$$

Using the notation

$$\Phi_1^m(\rho) = \sum_{0\neq n=m,m+2,\cdots}\alpha_n^m R_n^m(\rho) = \frac{\epsilon_m}{2\pi}\int_0^{2\pi}\Phi_1(\rho,\theta)\cos m\theta\, d\theta, \tag{13}$$

for $m = 0, 2$ and with $\epsilon_0 = 1,\ \epsilon_2 = 2$ , we write

$$\begin{aligned} U \approx\ & 2\int_0^1 A^0(\rho)\rho d\rho + 2i\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho) + f\Psi_1(\rho)\right\}\rho d\rho \\ & - \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 A^0(\rho)\left\{\Phi_1(\rho,\theta) + f\Psi_1(\rho)\right\}^2\rho d\rho d\theta \\ & - i\int_0^1 A^2(\rho)\Phi_1^2(\rho)\rho d\rho \\ & + \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 A^2(\rho)\left\{\Phi_1(\rho,\theta) + f\Psi_1(\rho)\right\}^2\cos 2\theta\rho d\rho d\theta, \end{aligned} \tag{14}$$

where it may be noted that the first integral above equals $a_0^0$.

Deleting $4^{th}$ order terms, expanding squares and carrying out integrations (using Eq. (13)), we obtain

$$\begin{aligned} |U|^2 \approx\ & (a_0^0)^2 - \frac{a_0^0}{\pi}\int_0^{2\pi}\int_0^1 A^0(\rho)\left\{\Phi_1(\rho,\theta)\right\}^2\rho d\rho d\theta \\ & - 4a_0^0 f\int_0^1 A^0(\rho)\Psi_1(\rho)\Phi_1^0(\rho)\rho d\rho \\ & - 2a_0^0 f^2\int_0^1 A^0(\rho)\left\{\Psi_1(\rho)\right\}^2\rho d\rho \\ & + \frac{a_0^0}{\pi}\int_0^{2\pi}\int_0^1 A^2(\rho)\left\{\Phi_1(\rho,\theta)\right\}^2\cos 2\theta\rho d\rho d\theta \\ & + 2a_0^0 f\int_0^1 A^2(\rho)\Psi_1(\rho)\Phi_1^2(\rho)\rho d\rho \\ & + \left(\int_0^1 A^2(\rho)\Phi_1^2(\rho)\rho d\rho\right)^2 \\ & - \left(2\int_0^1 A^2(\rho)\Phi_1^2(\rho)\rho d\rho\right) \\ & \times\left(2\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho) + f\Psi_1(\rho)\right\}\rho d\rho\right) \\ & + \left(2\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho) + f\Psi_1(\rho)\right\}\rho d\rho\right)^2. \end{aligned} \tag{15}$$

We shall argue now that we can ignore the last two terms of Eq. (15), and to that end we consider orders of magnitudes guided by the numerical results of the pilot case $s_0 = 0.95$, see Table A1 in Appendix A. Thus, by orthogonality and normalization of the Zernike polynomials

$$2\int_0^1 A^2(\rho)\Phi_1^2(\rho)\rho d\rho \approx \frac{1}{3}a_2^2\,\alpha_2^2, \tag{16}$$

$$2\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho) + f\Psi_1(\rho)\right\}\rho d\rho \approx \frac{1}{3}a_2^0\left(\alpha_2^2 + f\gamma_2^0\right). \tag{17}$$

This should be compared with, for instance,

$$\frac{a_0^0}{\pi}\int_0^{2\pi}\int_0^1 A^0(\rho)\left\{\Phi_1(\rho,\theta)\right\}^2\rho d\rho d\theta \approx \sum_{(n,m)\neq(0,0)}\frac{(\alpha_n^m)^2}{\epsilon_m(n+1)}, \tag{18}$$

where $\epsilon_0 = 1$ and $\epsilon_1 = \epsilon_2 = ... = 2$, and with

$$4a_0^0 f\int_0^1 A^0(\rho)\Psi_1(\rho)\Phi_1^0(\rho)\rho d\rho \approx \frac{2}{3}f\gamma_2^0\alpha_2^0, \tag{19}$$

$$2a_0^0 f^2\int_0^1 A^0(\rho)\left\{\Psi_1(\rho)\right\}^2\rho d\rho \approx \frac{1}{3}f^2\left(\gamma_2^0\right)^2. \tag{20}$$

With $a_0^0 = 0.056$, $a_2^2 = 0.414$, $\gamma_2^0 = 0.473$, it is then easily seen that the terms in Eqs. (16)-(17) can be ignored compared to those in Eqs. (18)-(20).

Ignoring the last two terms in Eq. (15) and differentiating we find

$$\begin{aligned} \frac{\partial|U|^2}{\partial f} =\ & -2a_0^0\left[2\int_0^1 A^0(\rho)\Psi_1(\rho)\Phi_1^0(\rho)\rho d\rho\right. \\ & +2f\int_0^1 A^0(\rho)\left\{\Psi_1(\rho)\right\}^2\rho d\rho \\ & \left. -\int_0^1 A^2(\rho)\Psi_1(\rho)\Phi_1^2(\rho)\rho d\rho\right]. \end{aligned} \tag{21}$$

Setting this equal to 0, we then find

$$f = -\frac{\int_0^1 A^0(\rho)\Psi_1(\rho)\Phi_1^0(\rho)\rho d\rho - \frac{1}{2}\int_0^1 A^2(\rho)\Psi_1(\rho)\Phi_1^2(\rho)\rho d\rho}{\int_0^1 A^0(\rho)\left\{\Psi_1(\rho)\right\}^2\rho d\rho}. \tag{22}$$

Using the results from Appendix A,

$$A^0(\rho)\Psi_1(\rho) = \sum_{n=0}^{\infty} C_{2n}^0 R_{2n}^0(\rho),\ A^2(\rho)\Psi_1(\rho) = \sum_{n=1}^{\infty} E_{2n}^2 R_{2n}^2(\rho), \tag{23}$$

we get the final expression for the optimum focus setting

$$f = -\frac{\sum_{n=1}^{\infty}\frac{C_{2n}^0\alpha_{2n}^0}{2(2n+1)} - \frac{1}{2}\sum_{n=1}^{\infty}\frac{E_{2n}^2\alpha_{2n}^2}{2(2n+1)}}{\sum_{n=1}^{\infty}\frac{C_{2n}^0\gamma_{2n}^0}{2(2n+1)}} \tag{24}$$

from orthogonality of the Zernike polynomials, normalized according to $\int_0^1\left\{R_n^m(\rho)\right\}^2\rho d\rho = [2(n+1)]^{-1}$.

As an incidental note we observe that

$$\begin{aligned} \sum_{n=1}^{\infty}\frac{C_{2n}^0\,\alpha_{2n}^0}{2(2n+1)} &= \int_0^1 A^0(\rho)\Psi_1(\rho)\Phi_1^0(\rho)\rho d\rho\,, \\ \frac{1}{2}\sum_{n=1}^{\infty}\frac{E_{2n}^2\,\alpha_{2n}^2}{2(2n+1)} &= \frac{1}{2}\int_0^1 A^2(\rho)\Psi_1(\rho)\Phi_1^2(\rho)\rho d\rho\,, \\ \sum_{n=1}^{\infty}\frac{C_{2n}^0\gamma_{2n}^0}{2(2n+1)} &= \int_0^1 A^0(\rho)\left(\Psi_1(\rho)\right)^2\rho d\rho\,. \end{aligned} \tag{25}$$

When substituting the identities above in Eq. (24) we observe that the inner product ($\rho d\rho$ on $0 \le \rho \le 1$) of the functions $(fA^0\Psi_1 + A^0\Phi_1^0 - \frac{1}{2}A^2\Phi_1^2)$ and $\Psi_1$ has been made zero by the particular choice of $f$. This means that the aberration function corresponding to 'best' focus does not contain a high-NA defocus term of the form $\Psi_1 = (\Psi(\rho) - \gamma_0^0)$. The appearance of the high-NA amplitude functions $A^0$ and $A^2$ in the aberration function $(fA^0\Psi_1 + A^0\Phi_1^0 - \frac{1}{2}A^2\Phi_1^2)$ means that the true phase aberration $f\Psi_1 + \Phi_1$ has been automatically weighted with the high-NA amplitude functions in obtaining the optimum focus setting.

The expression for Strehl ratio of Eq. (15) holds for the on-axis intensity. To gather information on the off-axis intensity at a certain focus setting $f$, we can introduce a wavefront tilt in, for instance, the $x$- or the $y$-direction. Wavefront tilt is represented by the coefficient $\alpha_1^1$, multiplied with $\cos\theta$ for an $x$-excursion in the focal volume and $\sin\theta$ for the $y$-direction. We have introduced such a wavefront tilt in the expression for $U$ in the case that the wavefront aberration itself is limited to the circularly symmetric terms ($\alpha_{2n}^0 \neq 0$). Carrying out the integrations

of the relevant terms of Eq. (15), see Appendix B.1, we obtain

$$|U(r,\phi,f)|^2 \approx \left(a_0^0\right)^2 - 2a_0^0\left[\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho)\right\}^2\rho d\rho \right.$$
$$+2f\int_0^1 A^0(\rho)\Psi_1(\rho)\Phi_1^0(\rho)\rho d\rho + f^2\int_0^1 A^0(\rho)\left\{\Psi_1(\rho)\right\}^2\rho d\rho$$
$$\left.+\frac{(2\pi r)^2}{8}\,a_0^0\left(a_0^0+\frac{1}{3}a_2^0\right)-\frac{(2\pi r)^2\cos 2\phi}{24}\,a_0^0\,a_2^2\right], \qquad (26)$$

where we have used cylindrical coordinates $(r,\phi,f)$ in the focal region with the origin in the center of the nominal image plane. The expression is quadratic in the lateral field coordinate $r$ and provides us with the principal curvatures in the $x$- and $y$-cross-sections of the intensity distribution in the chosen focal plane. It follows from the expression of Eq. (26) that, when focusing a high-NA beam, the intensity distribution has two principal curvatures leading to the well-known elliptical profile of the point-spread function if the incident state of polarization is linear. The analytic expression also confirms that the major axis of the elliptic intensity profile is found along the polarization direction of the incident light. The method using a wavefront tilt coefficient to obtain the off-axis intensity can be extended to non-circularly-symmetric aberration, for instance lowest order astigmatism.

The Strehl ratio following from Eq. (26) is given by

$$S \approx 1-\frac{2}{a_0^0}\left[\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho)\right\}^2\rho d\rho \right.$$
$$\left.+2f\int_0^1 A^0(\rho)\Psi_1(\rho)\Phi_1^0(\rho)\rho d\rho + f^2\int_0^1 A^0(\rho)\left\{\Psi_1(\rho)\right\}^2\rho d\rho\right]$$
$$=1-\frac{2}{a_0^0}\left[\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho)\right\}^2\rho d\rho \right.$$
$$\left.+f\sum_{n=1}^{\infty}\frac{C_{2n}^0\alpha_{2n}^0}{2n+1}+f^2\sum_{n=1}^{\infty}\frac{C_{2n}^0\gamma_{2n}^0}{2(2n+1)}\right], \qquad (27)$$

where the optimum $f$-value has to be used to find the focus setting with maximum Strehl ratio.

The evaluation of the remaining integral in Eq. (27) in terms of Zernike coefficients $a_{2n}^0$ and $\alpha_{2n}^0$ can be done in principle by working out the Zernike coefficients of $\left\{\Phi_1^0(\rho)\right\}^2$. In the examples in Section 3 with $\alpha_2^0$ and $\alpha_4^0$ as the only nonzero coefficients, this is easily done (see Appendix B). In Section 3 we also present an analysis of $S$ in the presence of lowest order astigmatism ($\alpha_2^2$ and $\alpha_2^0$). The detailed derivation of $S$ for this case is also found in Appendix B.

## 3 NUMERICAL RESULTS

The effects of a high NA-value (NA=0.95) on the Strehl ratio and optimum focus setting are illustrated in Figures 3, 4 and 5. To check the foregoing analysis we start by introducing spherical aberration of lowest order accompanied by a focus off-set, represented by their Zernike coefficients $\alpha_4^0$ and $\alpha_2^0$. We then first apply our high-NA analysis with an exact treatment of the defocus effect according to Eq. (6) but neglecting the amplitude effects at high NA that are given by Eq. (5). This approach is in line with the common opinion that phase defects are more influential on a quantity like Strehl ratio than amplitude nonuniformities. Figures 3 presents the paraxially approximated Strehl ratio using Eq. (27) with $s_0 \to 0$, the value according to Eq. (27) for $s_0 = 0.95$, and the result from a numerical calculation using, for instance, Eq. (8).

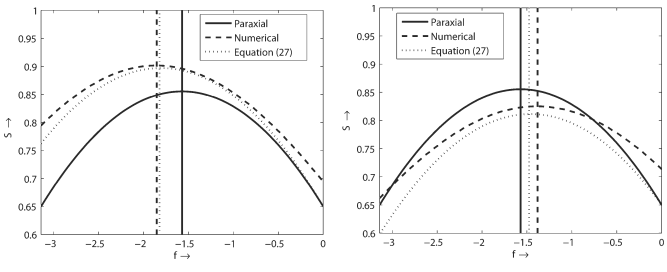


FIG. 3 Analytic calculation of the Strehl ratio as a function of the defocus parameter $f$ with a uniform amplitude distribution on the exit pupil sphere. The aberration is lowest order spherical aberration with a focus off-set (NA=0.95; $\alpha_2^0 = \pi/4$, $\alpha_4^0 = \pm 0.85$). The drawn line represents the low-NA approximation (paraxial), the dotted curve is obtained with the aid of Eq. (27) including the uniform amplitude function. Finally, the dashed curve, denoted by 'numerical', follows from Eq. (8) where the correct amplitude function at high-NA is taken into account. Lefthand figure: $\alpha_4^0 = +0.85$, righthand figure: $\alpha_4^0 = -0.85$. The respective maximum points of the three curves are indicated with a corresponding vertical line.

Note that the numerical calculation does include the basic high-NA amplitude nonuniformity $A_l(\rho,\theta)$ according to Eq. (5) with $A_E \equiv 1$. We observe from the figure that our analytic predictions for the optimum focus setting approach the exact numerically calculated values, but a significant difference is still there. The same holds for the predicted maximum Strehl ratios according to the exact numerical calculation and our predicted values from Eq. (27). The conclusion from the foregoing is that the amplitude nonuniformity at a value $s_0 = 0.95$ is such that it can not be neglected.

In Figure 4 we produce the results of our analytic treatment including the high-NA amplitude function $A_l(\rho,\theta)$ of Eq. (5) with $A_E \equiv 1$. The curves apply to the same aberration and defocus settings as in Figure 3 and present the paraxial approximation, the analytic results from our analysis and the numerically obtained data from Eq. (8). We remark that, like in Figure 3, this latter result could also have been obtained by a numerical evaluation of the vector diffraction integral [5] with the appropriate normalization according to Eq. (2); both results were found to be in perfect correspondence. In contrast with Figure 3 we now observe an almost perfect correspondence between the analytic results and the numerically obtained values. They show a pronounced difference with the predictions from paraxial theory, both with respect to the position of the optimum focal plane and the maximum obtainable Strehl ratio. The focus offset with respect to the paraxial case at the NA-value of 0.95 is seen to be approximately 20% of one focal depth (which corresponds to $\Delta f = \pi/2$). This deviation is significant with respect to the paraxial low-NA prediction and has to be taken into account in the design and manufacturing of optical systems. Even when no spherical aberration is present, the paraxially predicted focus setting is not correct (see Figure 4, third graph). This effect again has to be attributed to the 'apparent' spherical aberration term that is introduced by a defocusing at high NA-value. The focus offset in the third graph also explains the unequal focus offset with respect to 'paraxial' in the two upper graphs of Figure 4.

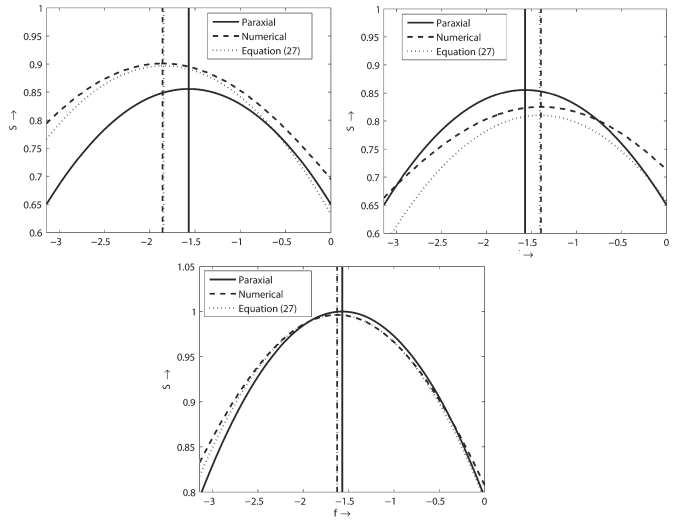


FIG. 4 Legend as in Figure 3, but now the analytic calculations have included the amplitude nonuniformity. The amplitude function $A_l(\rho,\theta)$ is given by Eq. (5) with $A_E \equiv 1$. Upper lefthand figure: $\alpha_4^0 = +0.85$, upper righthand figure: $\alpha_4^0 = -0.85$). An extra figure has been added where $\alpha_4^0 = 0$. The vertical lines for the optimum focus setting now virtually coincide for the analytic and numerical calculation. We remark that, even in the aberration-free case (lower figure), a focus off-set is observed with respect to the paraxial prediction.

The aberration-free optimum focus is found well in between the two focus settings for focused beams with spherical aberration of opposite sign. Although a strong defocus like in the first two graphs of Figure 4) will generally not be the final focus setting of a high-quality optical instrument, these settings are encountered during initial measurement and quality tuning of the instrument. The convergence process to optimum quality is improved when the correct focus settings at high-NA are taken into account instead of the paraxial predictions.

In Figure 5 we show the focus offset that is found according to the three approaches when the spherical aberration coefficient $\alpha_0^4$ is varied. The focus setting according to the paraxial approximation, as it was to be expected, does not show any dependence on the presence of spherical aberration and the focus setting is found at the fixed value $f = -2\alpha_0^2$.

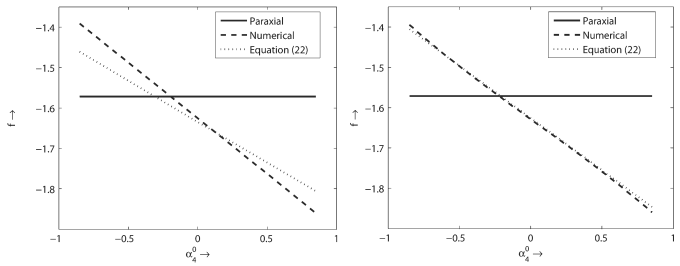


FIG. 5 Best-focus setting $f$ as a function of the Zernike coefficient of spherical aberration $\alpha_4^0$ in three different cases: paraxial approximation, numerical evaluation and our analysis of best focus setting according to Eq. (22). The initial defocusing coefficient $\alpha_2^0$ is $\pi/4$. The paraxial approximation predicts a constant best focus setting at $f = -\pi/2$, the numerical evaluation shows a deviation from this value, even when $\alpha_4^0 = 0$. In both graphs, we have also presented the focus setting according to Eq. (27), in the lefthand graph without taking into account the high-NA amplitude nonuniformity, in the righthand graph this effect has been included.

In the lefthand graph of Figure 5 we have shown the behaviour of $f$ according to our analytic treatment in the case that the amplitude nonuniformity according to Eq. (5) was omitted from the analysis (dotted curve, labeled 'Eq. (22)'). The difference with the exact, numerically obtained data that include the high-NA amplitude effects (dashed curve, labeled 'numerical') is still appreciable although there is a considerable positive correlation with Eq. (22). The righthand figure applies to the same cases with the exception that the amplitude nonuniformity now has been included in our analytic results (dotted curve). The correspondence with the non-approximated numerical calculations (dashed curved) becomes very satisfactory, showing the nonnegligible role played by the amplitude nonuniformity in the evaluation of optimum focus for a high-numerical-aperture beam.

A special case for the optimum focus setting $f$ arises when the phase function $\Phi_1$ only comprises a second order 'aberration' term with coefficient $\alpha_2^0 \neq 0$. Then Eq. (24) reduces to

$$f = -\frac{\frac{1}{6}C_2^0}{\sum_{n=1}^{\infty}\frac{C_{2n}^0\gamma_{2n}^0}{2(2n+1)}}\alpha_2^0 = -\frac{\frac{1}{6}C_2^0}{\int_0^1 A^0(\rho)\{\Psi_1(\rho)\}^2\rho d\rho}\alpha_2^0. \tag{28}$$

In the paraxial approximation, the constant in front of $\alpha_2^0$ in Eq. (28) equals $-2$. When $s_0 = 0.95$ and the amplitude effects are ignored ($A_l \equiv 1$), this coefficient equals $-2.0807$, and when the $A_l$ of Eq. (10) is used, this coefficient equals -2.0697. Finally, when we consider the limiting case $s_0 \uparrow 1$, the coefficients equal $-2$ (paraxial), -12/5=-2.4 (Eq. (24), $A_l \equiv 1$), and -203/88=-2.3068 (Eq. (24), $A_l$ as in Eq. (10)), respectively.

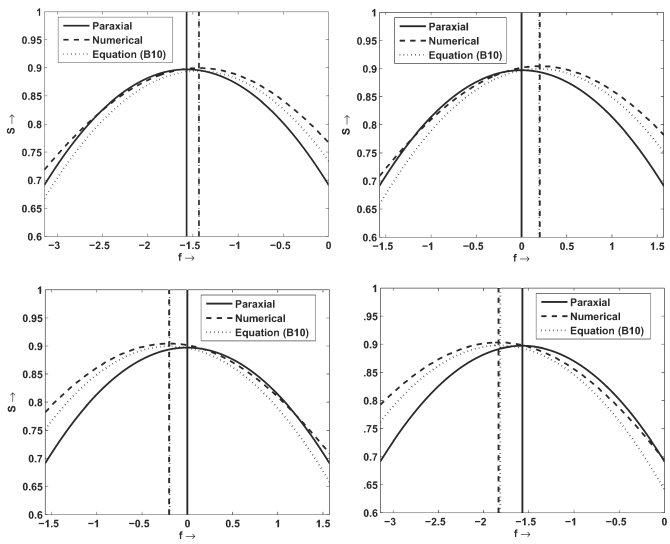


FIG. 6 Strehl ratio as a function of focus setting in the presence of astigmatic aberration with coefficient $\alpha_2^2 \neq 0$. The paraxially obtained curve ('Paraxial', $s_0 \to 0$) and the numerically calculated curve using the exact integral expression ('numerical') are given together with the result of our analytic treatment, see Eq. (B.10). In the lefthand graphs, the coefficient $\alpha_2^2$ is $+\pi/4$ (upper left graph) and $-\pi/4$ (lower left graph). The defocus coefficient $\alpha_2^0$ equals zero in both cases. In the righthand graphs, the same astigmatic coefficients are used for the upper and lower graphs, but now combined with a focus offset of one paraxial focal depth ($\alpha_2^0 = \pi/4$.

In Figure 6 we present the results for Strehl intensity as a function of focus setting in the presence of lowest order astigma-

tism ($\alpha_2^2 \neq 0$). The various curves again apply to the paraxial approximation, the numerical evaluation of Eq. (5) and our analytic evaluation of $S$ according to Eq. (B.10). In Figure 6 we have permanently included the high-NA amplitude nonuniformity in our analytic calculations, the reason why the results for the numerical and analytic results correspond very well. In the upper left graph, the astigmatic coefficient $\alpha_2^2$ equals $+\pi/4$, leading to a lesser curved wavefront cross-section for the radial section at $\theta = 0$. As it was the case throughout this paper, the incident beam in the entrance pupil was linearly polarized along the $x$-axis ($\theta = 0$). Because of the smaller amplitude in the exit pupil in this cross-section, the stronger curvature of the wavefront in the radial section with $\theta = \pi/2$ will dominate in determining the best focus, in this case closer to the exit pupil than the paraxial best-focus. To optimise Strehl ratio, we now have to introduce a nonzero defocus value $f_a$ such that the dominant wavefont curvature in the cross-section $\theta = \pi/2$ is partly compensated according to $f_a + w\alpha_2^2 \cos(\pi) = 0$ with $w$ a weighting factor. From the Figure we see that $f_a \approx +0.2$, so that the $w$-value is close to 0.25 at our $NA$-value of 0.95. The inverse effect is produced when the sign of $\alpha_2^2$ is opposite (lower left graph). In the two righthand graphs we have added a focus offset by making $\alpha_2^0 = +\pi/4$. The paraxially best focus is found at $f = -\pi/2$. In the upper right graph with $\alpha_2^2 = +\pi/4$, an off-set is found for the high-NA calculations to a more positive $f$-value, but less pronounced than in the upper left graph. This is caused by the 'apparent' spherical aberration that is introduced by the focus offset and that counteracts the offset introduced by astigmatism. A comparable effect is present in the lower right figure ($\alpha_2^2 = -\pi/4$), but here we observe an addition of the $f$-shifts due to astigmatism and spherical aberration.

Regarding the approximated Strehl ratios according to our second-order analysis and the exact numerically obtained values in all the examples, these remain very close as long as we stay in the range $S \geq 0.8$. The small difference in maximum value stems from the approximation to second order of the exponential phase function in the integrand of Eq. (2). The maximum Strehl ratio is either higher or lower than the one following from the low-NA expression, see Figure 4. This effect can again be attributed to the influence of defocusing on spherical aberration in the high-NA case. A defocusing is accompanied by wavefront deformation of orders $2n$ with a significant contribution at $2n = 4$. This contribution can enhance the already existing spherical aberration and lower the Strehl ratio like in Figure 4a or produce the opposite effect like in Figure 4b, depending on the sign of the coefficient $\alpha_4^0$. Accordingly, the approximation quality of the analytic formula, which has been devised neglecting higher orders, is affected in a similar fashion. Figure 4 also shows that at high-NA values the decrease in Strehl ratio with defocusing according to the exact formula is less pronounced than in the approximated expressions, both for high and low (paraxial) numerical aperture. This effect can again be attributed to the expansion of the phase function $\Phi$ up to only second order and the neglect of the higher orders.

## 4 CONCLUSION

We have evaluated the Strehl ratio for high-numerical-aperture imaging systems. Maximum Strehl ratio is an important criterion in the design and experimental optimization of an optical system. Its simple relationship with minimum quadratic wavefront deviation is not maintained in the high numerical aperture case. For finding the focus setting with maximum Strehl ratio, it is essential to use the rigorous expression for defocusing in a high-NA system instead of the paraxial quadratic approximation. Apart from using the correct expression for the phase departure of the pupil function, it is also needed to take into account the apparent amplitude nonuniformity due to the vector effects in high-NA image formation. Our analysis shows that the focal setting that is commonly derived from a Zernike expansion of the aberration function needs to be adapted at high NA values to find the image plane with the highest possible Strehl ratio. The amplitude nonuniformity in the exit pupil that is inherent to high-NA imaging has a non-negligible influence in determining this optimum focus setting and calculating the maximum Strehl ratio. The Zernike coefficients of the optical system are only correct when they are determined in the best-focus position.

## A ANALYTIC RESULTS FOR THE VARIOUS ZERNIKE EXPANSION COEFFICIENTS

In this Appendix we present the various Zernike expansions of the amplitude and phase functions that are encountered in the high-NA Strehl ratio analysis. The derivations are based on the evaluation of inner products with the radial Zernike polynomials. Only the results are given, the intricate derivations that were regularly encountered have been omitted.
We first present some basic quantities, related to high-NA imaging and then produce the expressions for the Zernike coefficients of the relevant aperture functions. At the end of the Appendix, numerical values of coefficients are listed pertaining to an NA-value of 0.95. These numerical values are useful in evaluating the relative importance of terms contributing to the Strehl intensity and they support the reasoning why some of these terms have been deleted from the analysis.

- Definition of some constants

$$c_0 = (1 - s_0^2)^{1/2}, \qquad d_0 = \left(\frac{1 - c_0}{s_0}\right)^2 \tag{A.1}$$

- Expansion coefficients of $(1 - s_0^2\rho^2)^\alpha$

$$(1 - s_0^2\rho^2)^\alpha = \sum_{n=0}^{\infty} D_{2n}^0(\alpha) R_{2n}^0(\rho)\,, \quad n = 0, 1, \ldots \tag{A.2}$$

$$D_{2n}^0(\alpha) = \frac{2n+1}{n+1} \sum_{k=n}^{\infty} \frac{(-1)^k \binom{\alpha}{k}\binom{k}{n}}{\binom{n+k+1}{k}} s_0^{2k}\,, \quad n = 0, 1, \ldots \tag{A.3}$$

- Expansion coefficients of $1 - (1 - s_0^2\rho^2)^\alpha$

$$1-(1-s_0^2\rho^2)^{\alpha}=\sum_{n=1}^{\infty}G_{2n}^2(\alpha)R_{2n}^2(\rho)\ ,n=1,2,... \quad \text{(A.4)}$$

$$G_{2n}^2(\alpha)=-\frac{2n+1}{n}\sum_{k=n}^{\infty}\frac{(-1)^k\binom{\alpha}{k}\binom{k-1}{n-1}}{\binom{n+k+1}{k+1}}s_0^{2k}\ ,\quad n=1,2,... \quad \text{(A.5)}$$

- Expansion coefficients of $\Psi(\rho)$

$$\Psi(\rho)=\frac{1-\sqrt{1-s_0^2\rho^2}}{1-c_0}=\sum_{n=0}^{\infty}\gamma_{2n}^0R_{2n}^0(\rho);\ \Psi_1(\rho)=\Psi(\rho)-\gamma_0^0 \quad \text{(A.6)}$$

$$\gamma_0^0=\frac{1+2c_0}{3(1+c_0)},\gamma_{2n}^0=\frac{1}{2}\left(\frac{d_0^{n-1}}{2n-1}-\frac{d_0^{n+1}}{2n+3}\right),n=1,2,... \quad \text{(A.7)}$$

- Expansion coefficients of $A^0(\rho)$

$$A^0(\rho)=\frac{1+\sqrt{1-s_0^2\rho^2}}{2(1-s_0^2\rho^2)^{1/4}}=\sum_{n=0}^{\infty}a_{2n}^0R_{2n}^0(\rho)\ ; \quad \text{(A.8)}$$

$$a_{2n}^0=\frac{1}{2}\left(D_{2n}^0(-1/4)+D_{2n}^0(1/4)\right),\ \ n=0,1,... \quad \text{(A.9)}$$

- Expansion coefficients of $A^2(\rho)$

$$A^2(\rho)=\frac{1-\sqrt{1-s_0^2\rho^2}}{2(1-s_0^2\rho^2)^{1/4}}=\sum_{n=0}^{\infty}a_{2n}^2R_{2n}^2(\rho)\ ; \quad \text{(A.10)}$$

$$a_{2n}^2=\frac{1}{2}\left(G_{2n}^2(1/4)-G_{2n}^2(-1/4)\right),\ \ n=1,2,... \quad \text{(A.11)}$$

- Expansion coefficients of $A^0(\rho)\Psi_1(\rho)$

$$A^0(\rho)\Psi_1(\rho)=\frac{1+\sqrt{1-s_0^2\rho^2}}{2(1-s_0^2\rho^2)^{1/4}}\left[\frac{1-\sqrt{1-s_0^2\rho^2}}{1-c_0}-\gamma_0^0\right]=\sum_{n=0}^{\infty}C_{2n}^0R_{2n}^0(\rho)\ ; \quad \text{(A.12)}$$

$$C_{2n}^0=\frac{1}{2(1-c_0)}\left[\{1-\gamma_0^0(1-c_0)\}D_{2n}^0(-1/4)-\gamma_0^0(1-c_0)D_{2n}^0(1/4)-D_{2n}^0(3/4)\right],\ \ n=0,1,... \quad \text{(A.13)}$$

- Expansion coefficients of $A^2(\rho)\Psi_1(\rho)$

$$A^2(\rho)\Psi_1(\rho)=\frac{1-\sqrt{1-s_0^2\rho^2}}{2(1-s_0^2\rho^2)^{1/4}}\left[\frac{1-\sqrt{1-s_0^2\rho^2}}{1-c_0}-\gamma_0^0\right]=\sum_{n=1}^{\infty}E_{2n}^2R_{2n}^2(\rho)\ ; \quad \text{(A.14)}$$

$$E_{2n}^2=\frac{-1}{2(1-c_0)}\left[\{1-\gamma_0^0(1-c_0)\}G_{2n}^2(-1/4)-\{2-\gamma_0^0(1-c_0)\}G_{2n}^2(1/4)+G_{2n}^2(3/4)\right],n=1,2,... \quad \text{(A.15)}$$

To conclude this Appendix, we give the numerical values of the most important coefficients that are encountered in the Strehl ratio analysis ($s_0=0.95$).

| n | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $c_0$ | 0.312250 | — | — | — | — | — |
| $d_0$ | 0.524100 | — | — | — | — | — |
| $a_{2n}^0$ | 1.028866 | 0.056333 | 0.042380 | 0.022633 | 0.011679 | 0.005987 |
| $\gamma_{2n}^0$ | 0.412650 | 0.472532 | 0.077067 | 0.023276 | 0.008485 | 0.003395 |
| $C_{2n}^0$ | 0.009615 | 0.496726 | 0.103650 | 0.042274 | 0.019268 | 0.009182 |
| $a_{2n}^2$ | — | 0.413719 | 0.115181 | 0.046428 | 0.020838 | 0.009813 |
| $E_{2n}^2$ | — | 0.121137 | 0.135165 | 0.056529 | 0.025408 | 0.011910 |

TABLE A1 Table of the numerical values

# B ANALYTIC RESULTS FOR STREHL INTENSITY FOR SOME SPECIFIC ABERRATION TYPES

## B.1 Introduction of wavefront tilt to obtain off-axis intensity values

In this subsection we introduce a wavefront tilt and calculate the corresponding on-axis Strehl intensity. We thus obtain information on the off-axis intensity distribution of the non-tilted focused beam. We demonstrate this method for a beam that is affected by circularly symmetric aberration. A wavefront comprising a circularly symmetric component plus a wavefront tilt is represented by

$$\Phi_1(\rho,\theta)=\alpha_1^1\rho\cos(\theta-\phi)+\Phi_1^0(\rho), \quad \text{(B.1)}$$

where $\alpha_1^1$ is the wavefront tilt expressed in radians, $\phi$ determines the azimuth of the wavefront tilt and $\Phi_1^0(\rho)$ is the circularly symmetric aberration term. The substitution of this particular wavefront aberration in the first six integral terms occurring in Eq. (15) leads to the following results. The first integral yields two contributions according to

$$\begin{aligned}&-\frac{a_0^0}{\pi}\int_0^{2\pi}\int_0^1A^0(\rho)\left\{\Phi_1(\rho,\theta)\right\}^2\rho d\rho d\theta\\&=-\frac{a_0^0}{\pi}\int_0^{2\pi}\int_0^1A^0(\rho)\left[\left(\alpha_1^1\right)^2\rho^2\cos^2(\theta-\phi)\right.\\&\left.+2\alpha_1^1\rho\cos(\theta-\phi)\Phi_1^0(\rho)+\left\{\Phi_1^0(\rho)\right\}^2\right]\rho d\rho d\theta\\&=-a_0^0\left(\alpha_1^1\right)^2\int_0^1A^0(\rho)\rho^2\rho d\rho+2a_0^0\int_0^1A^0(\rho)\left\{\Phi_1^0(\rho)\right\}^2\rho d\rho\\&=-\frac{1}{4}a_0^0\left(\alpha_1^1\right)^2\left(\frac{1}{3}a_2^0+a_0^0\right)+2a_0^0\int_0^1A^0(\rho)\left\{\Phi_1^0(\rho)\right\}^2\rho d\rho.\end{aligned} \quad \text{(B.2)}$$

The second and third integral of Eq. (15) are unaffected by the wavefront tilt. The substitution of (B.1) in the fourth integral

yields as only nonzero contribution

$$\frac{a_0^0}{\pi}\int_0^{2\pi}\int_0^1 A^2(\rho)\,\left(\alpha_1^1\right)^2\rho^2\left\{\frac{1}{2}+\frac{1}{2}\cos(2\theta-2\phi)\right\}\cos 2\theta\rho d\rho d\theta$$
$$=\frac{a_0^0}{2}\left(\alpha_1^1\right)^2\cos 2\phi\int_0^1 A^2(\rho)\rho^2\,\rho d\rho=\frac{1}{12}a_0^0a_2^2\left(\alpha_1^1\right)^2\cos 2\phi. \tag{B.3}$$

The fifth and sixth integral are zero in the circularly symmetric case and the substitution of the above results in Eq. (15) then yields the expression of Eq. (26) where we have used the relationship $\alpha_1^1/2\pi = r$ with $r$ the normalized radial image plane coordinate. In the special case that $\Phi_1^0(\rho)=\alpha_2^0R_2^0(\rho)$, the only remaining integral, see Eq. (B.2), can be evaluated as

$$\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho)\right\}^2\rho d\rho=\left(\alpha_2^0\right)^2\left(\frac{1}{6}a_0^0+\frac{1}{15}a_4^0\right). \tag{B.4}$$

## B.2 Spherical aberration

For some other basic aberration types, analytic results for the Strehl ratio can be obtained. In the case of circularly symmetric aberration terms, the expression for the Strehl ratio of Eq. (27) can be evaluated and at the optimum $f$-value this yields

$$S_{opt}=1-\frac{1}{a_0^0}\left[2\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho)\right\}^2\rho d\rho-\frac{\left(\sum_{n=1}^{\infty}\frac{C_{2n}^0\alpha_{2n}^0}{2n+1}\right)^2}{\sum_{n=1}^{\infty}\frac{C_{2n}^0c_{2n}^0}{2n+1}}\right]. \tag{B.5}$$

The integral occurring in Eq.(27) and Eq.(B.5) can be written as

$$I=2\int_0^1 A^0(\rho)\left\{\Phi_1^0(\rho)\right\}^2\rho d\rho=\sum_{n=0}^{\infty}a_{2n}^0\frac{C_{R_{2n}^0}\left[\left(\Phi_1^0\right)^2\right]}{2n+1}, \tag{B.6}$$

where we have indicated between brackets that the coefficients $C$ are those of the Zernike polynomial $R_{2n}^0$ in the expansion of $\left(\Phi_1^0\right)^2$. The evaluation of the Strehl intensity in the special case that only $\alpha_2^0$ and $\alpha_4^0$ (lowest order 'spherical') are nonzero leads to the following value of the integral above

$$I=a_0^0\left[\frac{(\alpha_2^0)^2}{3}+\frac{(\alpha_4^0)^2}{5}\right]+\frac{4}{15}a_2^0\alpha_2^0\alpha_4^0+$$
$$\frac{a_4^0}{5}\left[\frac{2(\alpha_2^0)^2}{3}+\frac{2(\alpha_4^0)^2}{7}\right]+\frac{6}{35}a_6^0\alpha_2^0\alpha_4^0+\frac{2}{35}a_8^0(\alpha_4^0)^2. \tag{B.7}$$

## B.3 Second order astigmatism and defocus

The wavefront aberration is given by

$$\Phi_1(\rho,\theta)=\alpha_2^0R_2^0(\rho)+\alpha_2^2R_2^2(\rho)\cos 2\theta, \tag{B.8}$$

where $\alpha_2^2$ now is the astigmatic coefficient with the principal curvatures of the astigmatic wavefront oriented along the $x$- and $y$-axes ($\phi = 0$). A somewhat longer derivation is needed here to calculate the on-axis intenstity $I$ and the Strehl ratio using the best focus setting $f_{opt}$, but the analysis basically proceeds along the same lines as in (B.1). We present below the analytic results for the six relevant integrals of Eq. (15) when setting $\Phi_1^0(\rho)=\alpha_2^0R_2^0(\rho)$ and $\Phi_1^2(\rho)=\alpha_2^2R_2^2(\rho)$. We have

$$\int_0^{2\pi}\int_0^1 A^0(\rho)\left\{\Phi_1(\rho,\theta)\right\}^2\rho d\rho d\theta$$
$$=\pi\left(\alpha_2^0\right)^2\left[\frac{2}{15}a_4^0+\frac{1}{3}a_0^0\right]+\frac{1}{2}\pi\left(\alpha_2^2\right)^2\left[\frac{1}{30}a_4^0+\frac{1}{6}a_2^0+\frac{1}{3}a_0^0\right],$$
$$\int_0^1 A^0(\rho)\Psi_1(\rho)\Phi_1^0(\rho)\rho d\rho=\frac{1}{6}C_2^0\alpha_2^0,$$
$$\int_0^1 A^0(\rho)\left\{\Psi_1(\rho)\right\}^2\rho d\rho=\sum_{n=0}^{\infty}\frac{C_{2n}^0\gamma_{2n}^0}{2(2n+1)},$$
$$\int_0^{2\pi}\int_0^1 A^2(\rho)\left\{\Phi_1(\rho,\theta)\right\}^2\cos 2\theta\rho d\rho d\theta=\frac{1}{2}\pi\alpha_2^0\alpha_2^2\left(\frac{1}{3}a_2^2+\frac{1}{5}a_4^2\right),$$
$$\int_0^1 A^2(\rho)\Psi_1(\rho)\Phi_1^2(\rho)\rho d\rho=\frac{1}{6}E_2^2\alpha_2^2,$$
$$\int_0^1 A^2(\rho)\Phi_1^2(\rho)\rho d\rho=\frac{1}{6}a_2^2\alpha_2^2\,. \tag{B.9}$$

With the analytically obtained integral values, the optimum focus setting is obtained from Eq. (22). Finally, the expression for the Strehl ratio as a function of $f$ is given by

$$S\approx 1-\frac{(\alpha_2^0)^2}{a_0^0}\left[\frac{2}{15}a_4^0+\frac{1}{3}a_0^0\right]-\frac{(\alpha_2^2)^2}{a_0^0}\left[\frac{1}{60}a_4^0+\frac{1}{12}a_2^0+\frac{1}{6}a_0^0\right]$$
$$-f\left[\frac{2C_2^0\alpha_2^0-E_2^2\alpha_2^2}{3a_0^0}\right]-f^2\frac{1}{a_0^0}\sum_{n=0}^{\infty}\frac{C_{2n}^0\gamma_{2n}^0}{2n+1}$$
$$+\frac{\alpha_2^0\alpha_2^2}{a_0^0}\left[\frac{1}{6}a_2^2+\frac{1}{10}a_4^2\right]+\frac{(a_2^2)^2(\alpha_2^2)^2}{36(a_0^0)^2}\,. \tag{B.10}$$

## References


[1] N. Bobroff and A.E. Rosenbluth, Appl. Opt. **31**, 1523–1536 (1992).

[2] V. N. Mahajan, J. Opt. Soc. Am. **72**, 1258–1266 (1982).

[3] M. Born and E. Wolf, *Principles of Optics* (4th rev. ed., Pergamon Press, New York, 1970).

[4] W. Welford, *Aberrations of optical systems* (Adam Hilger, Bristol, 1986).

[5] B. Richards and E. Wolf, P. Roy. Soc. A:Math Phy **253**, 358–379 (1959).

[6] S.M. Mansfield and G.S. Kino, Appl. Phys. Lett. **57**, 2615–2616 (1990).

[7] K. Svoboda and S.M. Block, Annu. Rev. Bioph. Biom. **23**, 247–285 (1994).

[8] P. Török, P. Varga, Z. Laczik and G.R. Booker, J. Opt. Soc. Am. A **12**, 325–332 (1995).

[9] I. Ichimura, S. Hayashi and G.S. Kino, Appl. Opt. **36**, 4339–4348 (1997).

[10] S. Quabis, R. Dorn, M. Eberler, O. Glockl and G. Leuchs, Opt. Commun. **179**, 1–6 (2000).

[11] J. J. M. Braat, P. Dirksen, A. J. E. M. Janssen, and A.S. van de Nes, J. Opt. Soc. Am. A **20**, 2281–2292 (2003).

[12] B.J. Lin, J. Microlith. Microfab. **1**, 7–12 (2002).

[13] S. Stallinga, Appl. Opt. **44**, 849–858 (2005).