Journal of Modern Optics
Vol. 55, No. 7, 10 April 2008, 1127–1157

Zernike representation and Strehl ratio of optical systems with variable numerical aperture

A.J.E.M. Janssen[a], S. van Haver[b]*, P. Dirksen[a] and J.J.M. Braat[b]

[a]*Philips Research Europe, Eindhoven, The Netherlands;* [b]*Faculty of Applied Sciences, Optics Research Group, Delft University of Technology, Delft, The Netherlands*

(*Received 13 March 2007; final version received 8 August 2007*)

We consider optical systems with variable numerical aperture (NA) on the level of the Zernike coefficients of the correspondingly scalable pupil function. We thus present formulas for the Zernike coefficients and their first two derivatives as a function of the scaling factor $\varepsilon \leq 1$, and we apply this to the Strehl ratio and its derivatives of NA-reduced optical systems. The formulas for the Zernike coefficients of NA-reduced optical systems are also useful for the forward calculation of point-spread functions and aberration retrieval within the Extended Nijboer–Zernike (ENZ) formalism for optical systems with reduced NA or systems that have a central obstruction. Thus, we retrieve a Gaussian, comatic pupil function on an annular set from the intensity point-spread function in the focal region under high-NA conditions.

Keywords: NA reduction; Zernike coefficients; Strehl ratio; central obstruction; aberration retrieval; ENZ theory

1. Introduction and overview

In highly corrected optical systems that operate in or close to the diffraction-limited regime, the residual aberrations are small and the optical design is such that the distribution of aberration over the aperture of the imaging pencils minimizes image degradation. Various criteria to assess image quality are used. The maximum intensity of the point-spread function, the image of a point source in the object plane, is a good measure for image quality. Normalized to unity for the perfect imaging system, its value S for an aberrated system yields useful information on the imaging performance that can be expected. This quantity was defined by Strehl in 1894 and is commonly called the Strehl ratio [1]. Another quality measure for imaging systems is the root mean square value W_{rms} of the wavefront aberration in the exit pupil of a system. For modest aberration values,

*Corresponding author. Email: s.vanhaver@tudelft.nl

ISSN 0950–0340 print/ISSN 1362–3044 online

DOI: 10.1080/09500340701618403
http://www.informaworld.com

smaller than the wavelength λ of the light, a direct relationship can be established between the Strehl ratio S and W_{rms} [2] according to

$$S = 1 - k^2 W_{\mathrm{rms}}^2, \tag{1}$$

with $k = 2\pi/\lambda$. A well-corrected optical system should not produce a Strehl value below 0.80 and the corresponding upper limit for the root mean square wavefront aberration equals $W_{\mathrm{rms}} \leq 0.071\lambda$. The well-corrected optical systems described above are meant to operate at a well-defined and fixed aperture and any change in it compromises the balancing of aberrations that was obtained in the design process. The change in aperture towards higher values generally is excluded because of mechanical constraints. A lower aperture is possible but can lead to rather unexpected effects in, for instance, the root mean square residual aberration or the Strehl ratio of the produced point-spread function of the system. As an example of systems with a variable aperture we quote lithographic projection systems. At some occasions they are used below their maximum aperture value to optimize the imaging on the wafer of a particular mask structure with less demanding features. Another possible change in effective aperture is brought about by a central obstruction. This type of obstruction is found when extra beams of light have to be transported through the imaging system with the aid of auxiliary mirrors or, simply, because the system is meant to be catadioptric by design. Examples of these systems are also found in optical lithography and, of course, in astronomical observation. Another class of optical systems with varying aperture uses a so-called iris diaphragm. In most cases, the iris diaphragm is found in imaging systems that are operating far away from the diffraction limit, like in those for classical photography. However, with the advent of short-focus, image-sensor based photographic lenses, these systems operate close to the diffraction limit. On the other hand, in these modern devices the action of the mechanical iris diaphragm has mostly been replaced by an electronic shutter. One optical imaging system in which the (circular) iris remains fully active is the eye of humans and humanoids. In ophthalmology, the eye doctor carries out measurements on the optical eye with varying iris diameter or at full aperture using an iris-freezing drug. When studying visual perception or when obtaining images of the retina via the eye lens, the aberrations as a function of iris diameter need to be well known and, if necessary, they are scaled from the full diameter to the actual active diameter. Important changes in aberration correction and Strehl intensity can be expected in this case.

Both the point-spread function and, for high-quality systems, the Strehl ratio can be expressed in terms of the Zernike coefficients of the pupil function. The description of optical systems in terms of Zernike expansions of pupil functions is, in fact, a very powerful and widely used method to which a whole chapter (Chapter 9) has been devoted in [3]. It is therefore of considerable interest to quantify how the Zernike coefficients of the pupil function vary when the aperture is reduced to a fraction $\varepsilon < 1$ of the maximum aperture value. Changing the aperture has a major and non-trivial impact on the Zernike coefficients, and in recent years several papers [4–9] have been devoted to this difficult problem. The contribution of the present paper is that it offers a complete mathematical treatment of this problem, with applications to the Strehl ratio and computation of the point-spread function of optical systems as well as to optical system characterization for obstructed systems from through-focus intensity data.

To be more specific, we consider a pupil function

$$P(\rho,\vartheta) = A(\rho,\vartheta)\exp[\mathrm{i}\Phi(\rho,\vartheta)], \quad 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \tag{2}$$

on a unit disk with amplitude $A \ge 0$ and real phase Φ, where we assume that the normalizations are such that the maximum aperture value gives rise to a pupil radius of unity. The pupil function P can be thought of as being represented in the form of a Zernike series according to (polar coordinates)

$$P(\rho,\vartheta) = \sum_{n,m} \beta_n^m Z_n^m(\rho,\vartheta), \quad 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi. \tag{3}$$

Here $Z_n^m(\rho,\vartheta)$ denotes the Zernike function

$$Z_n^m(\rho,\vartheta) = R_n^m(\rho)\begin{cases} \cos m\vartheta, & 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \quad (4a) \\ \sin m\vartheta, & 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \quad (4b) \end{cases}$$

with integer n, $m \ge 0$ such that $n-m$ is even and ≥ 0, and R_n^m is the Zernike polynomial of the azimuthal order m and of degree n, see [3], Appendix VII. For simplicity, we shall only consider the cosine option in (4), the treatment for the sine option in (4) being largely the same.

The through-focus complex-amplitude point-spread function U pertaining to the optical system is expressed in terms of the pupil function P as the diffraction integral

$$U(r,\varphi,f) = \frac{1}{\pi}\int_0^1\int_0^{2\pi} \exp(\mathrm{i}f\rho^2)P(\rho,\vartheta)\ \exp[2\pi\mathrm{i}r\rho\cos(\vartheta-\varphi)]\rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta, \tag{5}$$

where we have used polar coordinates r, φ in the image planes and f denotes the focal variable. Also, the Strehl ratio of the optical system is defined as

$$S = \frac{\left|(1/\pi)\int_0^1\int_0^{2\pi} P(\rho,\vartheta)\rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta\right|^2}{\left|(1/\pi)\int_0^1\int_0^{2\pi} |P(\rho,\vartheta)|\rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta\right|^2}. \tag{6}$$

Using the Zernike expansion (3) of P, the point-spread function U admits the representation

$$U(r,\varphi,f) = 2\sum_{n,m} \mathrm{i}^m \beta_n^m V_n^m(r,f)\cos m\varphi, \tag{7}$$

in which the V_n^m are specific functions that have become available in tractable form recently (Section 4 has details for this). Similarly, under a small-aberration assumption, the Strehl ratio S can be approximated in terms of the Zernike expansion coefficients α of the aberration phase Φ according to

$$\Phi(\rho,\vartheta) = \sum_{n,m} \alpha_n^m Z_n^m(\rho,\vartheta), \quad 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \tag{8}$$

as

$$S_\varepsilon \approx S(\alpha) := 1 - \sum_{n,m} \frac{(\alpha_n^m)^2}{\varepsilon_m(n+1)} \tag{9}$$

(real α's; $\varepsilon_0 = 1$, $\varepsilon_1 = \varepsilon_2 = \cdots = 1$, Neumann's symbol). In the summation in Equation (9) the term with $(n, m) = (0, 0)$ is omitted.

Now reducing the numerical aperture (NA) to a fraction $\varepsilon \le 1$ of its maximum value, means that we set $P(\rho, \vartheta) = 0$ outside the disk $0 \le \rho \le \varepsilon$, $0 \le \vartheta \le 2\pi$, and leave P as it is inside the disk. In order to obtain convenient forms for U and S as in (7) and (9), it is, in principle, possible to expand the new $P = P_\varepsilon$ as a Zernike series on the full disk $0 \le \rho \le 1$, $0 \le \vartheta \le 2\pi$. However, the resulting series has poorly decaying coefficients, due to the discontinuity at $\rho = \varepsilon$ (Appendix 3 is explicit about this). Also, it is not clear what the new phase $\Phi = \Phi_\varepsilon$ and its Zernike expansion are going to be, the amplitude being 0 outside the disk $0 \le \rho \le \varepsilon$. We therefore choose for a different approach in which we observe that the new $U = U_\varepsilon$ is obtained as

$$\begin{aligned} U_\varepsilon(r, \varphi, f) &= \frac{1}{\pi} \int_0^\varepsilon \int_0^{2\pi} \exp(\mathrm{i} f \rho^2) P(\rho, \vartheta)\ \exp[2\pi \mathrm{i} r \rho \cos(\vartheta - \varphi)] \rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \\ &= \frac{\varepsilon^2}{\pi} \int_0^1 \int_0^{2\pi} \exp(\mathrm{i} f \varepsilon^2 \rho^2) P(\varepsilon\rho, \vartheta)\ \exp[2\pi \mathrm{i} r \varepsilon \rho \cos(\vartheta - \varphi)] \rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \end{aligned} \tag{10}$$

in which the last expression in (10) has been obtained from the middle one by changing the variable ρ, $0 \le \rho \le \varepsilon$, into $\varepsilon\rho$, $0 \le \rho \le 1$. By the same variable transformation, we have that the new Strehl ratio $S = S_\varepsilon$ is given by

$$S = \frac{\left| (1/\pi) \int_0^1 \int_0^{2\pi} P(\varepsilon\rho, \vartheta) \rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \right|^2}{\left| (1/\pi) \int_0^1 \int_0^{2\pi} |P(\varepsilon\rho, \vartheta)| \rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \right|^2}. \tag{11}$$

Equations (10) and (11) show that we are in the same position as in (7) and (9), when we would have available the Zernike expansion of the scaled pupil $P(\varepsilon\rho, \vartheta)$, $0 \le \rho \le 1$, $0 \le \vartheta \le 2\pi$, and of the scaled phase $\Phi(\varepsilon, \rho, \vartheta)$, $0 \le \rho \le 1$, $0 \le \vartheta \le 2\pi$. Limiting ourselves here to P (the developments for Φ being the same), we thus seek to find the Zernike expansion

$$P(\varepsilon\rho, \vartheta) = \sum_{n,m} \beta_n^m(\varepsilon) Z_n^m(\rho, \vartheta), \quad 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \tag{12}$$

in which the coefficients $\beta_n^m(\varepsilon)$ should be related to the β_n^m in the Zernike expansion of the unscaled $P(\rho, \vartheta)$, see (3).

The problem of expressing Zernike coefficients of scaled pupil functions into Zernike coefficients of unscaled pupil functions has been considered recently, see [4–9]. The developments in [4–7] eventually led to an explicit series expansion for the matrix elements

$M^m_{nn'}(\varepsilon)$ required to compute $\beta^m_n(\varepsilon)$ from $\beta^m_{n'}$ according to

$$\frac{1}{2(n+1)}\beta^m_n(\varepsilon) = \sum_{n'} M^m_{nn'}(\varepsilon)\beta^m_{n'}, \quad n = m, m+2, \ldots, \tag{13}$$

see [7], Equation (19) (Dai's formula) and Appendix 1 where we give a new proof for it. We note here that the computation scheme decouples per azimuthal order m due to azimuthal orthogonality of the $Z^m_n(\rho, \vartheta)$. The factor $1/2(n+1)$ in front of $\beta^m_n(\varepsilon)$ in (13) is due to the normalization $\int_0^1 (R^m_n(\rho))^2 \rho \, \mathrm{d}\rho = 1/2(n+1)$ of the Zernike polynomials. The matrix elements $M^m_{nn'}(\varepsilon)$ take by orthogonality of the R^m_n the form

$$M^m_{nn'}(\varepsilon) = \int_0^1 R^m_{n'}(\varepsilon\rho) R^m_n(\rho) \rho \, \mathrm{d}\rho, \quad n, n' = m, m+2, \ldots. \tag{14}$$

Indeed, since by (3)

$$P(\varepsilon\rho, \vartheta) = \sum_{n', m} \beta^m_{n'} R^m_{n'}(\varepsilon\rho) \cos m\vartheta, \quad 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \tag{15}$$

we need to find the Zernike^m expansion of $R^m_{n'}(\varepsilon\rho)$:

$$R^m_{n'}(\varepsilon\rho) = \sum_n 2(n+1) M^m_{nn'} R^m_n(\rho), \quad 0 \le \rho \le 1. \tag{16}$$

A more general matrix approach for scaling, rotating and displacing pupils was considered in [8]; this yields, however, not the explicit type of results for $M^m_{nn'}$ that we are interested in here.

In [9] it is shown that

$$M^m_{nn'}(\varepsilon) = \frac{R^n_{n'}(\varepsilon) - R^{n+2}_{n'}(\varepsilon)}{2(n+1)}, \quad n, n' = m, m+2, \ldots, \tag{17}$$

where it is understood that $R^k_l \equiv 0$ when k, l are integers ≥ 0 such that $l-k$ is even and < 0. A number of consequences of Equation (17) was noted in [9]. Among these is the formula

$$\beta^m_n(\varepsilon) = \sum_{n'} (R^n_{n'}(\varepsilon) - R^{n+2}_{n'}(\varepsilon)) \beta^m_{n'}, \quad n = m, m+2, \ldots, \tag{18}$$

where the summation is over $n' = n, n+2, \ldots$. Furthermore, in [9] an expression for the derivative of $\alpha^m_n(\varepsilon)$ and for $S(\alpha(\varepsilon))$, see Equations (8) and (9), at $\varepsilon = 1$ is given, showing large sensitivity of aberration coefficients and Strehl ratios for values of ε near the maximum 1.

In this paper we expand on the investigations in [9] which was only a brief letter aimed at the lithographic community. Thus, we present formulas for the first two derivatives of $\beta^m_n(\varepsilon)$ and $S(\alpha(\varepsilon))$ at general $\varepsilon \in [0, 1]$, and we consider these results in the context of the semigroup structure governing scaling operations. We also present examples of pupil functions for which $(d/d\varepsilon)\ [S(\alpha(\varepsilon))]$ and $(d/d\varepsilon)^2\ [S(\alpha(\varepsilon))]$ at $\varepsilon = 1$ can have all four combinations of signs. These examples are somewhat counterintuitive, the common opinion being that Strehl ratios of scaled pupils should decrease when NA is increased. We furthermore show how the in recent years developed Extended Nijboer–Zernike (ENZ) formalism, see the ENZ website [10] or [11–16], for the computation of through-focus

optical point-spread functions and the retrieval of optical aberrations from through-focus intensities, has to be modified so as to apply to scaled optical systems and to systems with a central obstruction. As an example, we show a retrieval result for a Gaussian, comatic pupil function on an annular set under high-NA conditions. The proofs of our results are collected in the three appendices.

2. Mathematical results

In this section we present our results in a mathematical form; the application of these results are to be found in the subsequent sections. All proofs are contained in Appendix 1.

2.1 *Basic results*

We start by repeating the basic result (17), extended as

$$M^m_{nn'}(\varepsilon) = \frac{R^n_{n'}(\varepsilon) - R^{n+2}_{n'}(\varepsilon)}{2(n+1)} = \frac{R^{n+1}_{n'+1}(\varepsilon) - R^{n+1}_{n'-1}(\varepsilon)}{2\varepsilon(n'+1)}, \quad n, n' = m, m+2, \ldots, \tag{19}$$

that we discuss and prove in full in Appendix 1. Since we have by convention that $R^k_l = 0$ when $l - k < 0$, we see that

$$M^m_{nn'}(\varepsilon) = 0, \quad n \text{ outside } \{m, m+2, \ldots, n'\}. \tag{20}$$

Furthermore, $M^m_{nn'}(\varepsilon)$ does not depend on m; however, note that in Equation (17) both indices n, n' are restricted to $m, m+2, \ldots$. Evidently, we also have from the two identities in Equation (19) that

$$\beta^m_n(\varepsilon) = \sum_{n'} \frac{n+1}{(n'+1)\varepsilon}(R^{n+1}_{n'+1}(\varepsilon) - R^{n+1}_{n'-1}(\varepsilon))\beta^m_{n'}, \quad n = m, m+2, \ldots. \tag{21}$$

As a consequence of the second identity in Equation (19) we show in Appendix 1 that for n, $n' = m$, $m+2, \ldots$

$$M^m_{nn}(\varepsilon) = \frac{\varepsilon^n}{2(n+1)}, \tag{22a}$$

$$M^m_{nn'}(\varepsilon) = \frac{-1}{2k}(1-\varepsilon^2)\varepsilon^n P^{(1,n+1)}_{k-1}(2\varepsilon^2 - 1), \quad k = \frac{1}{2}(n'-n) = 1, 2, \ldots, \tag{22b}$$

where $P^{(\alpha,\beta)}_l$ denotes the Jacobi polynomial with parameters α, β and of degree l, see [17], Chapter 22. These formulas can be used to show that $M^m_{nn'}(\varepsilon)$ has appropriate behavior as $\varepsilon \downarrow 0$ or $\varepsilon \uparrow 1$. Indeed,

$$M^m_{nn'}(\varepsilon = 0) = 0 \text{ all allowed } m, n, n', \quad (m, n) \neq (0, 0). \tag{23}$$

Therefore, $\beta_n^m(0) \neq 0$ for $m=n=0$ only: for $\varepsilon=0$ the scaled pupil function reduces to a constant. Furthermore, since $P_{k-1}^{(1,n+1)}(1)=k$, see [17], item 22.4.1 in Table 22.4 on p. 777, we have that

$$M_{nn}^m(\varepsilon) = \frac{1}{2(n+1)} - \frac{n}{2(n+1)}(1-\varepsilon) + \cdots, \tag{24a}$$

$$M_{nn'}^m(\varepsilon) = -(1-\varepsilon) + \cdots, \quad \frac{n'-n}{2} = 1, 2, \ldots \tag{24b}$$

when $\varepsilon \uparrow 1$. The case $\varepsilon=1$ corresponds to the full (unscaled) pupil, and thus the formulas (22a) and (22b) for $\varepsilon=1$ correctly show the orthogonality and normalization properties of the $R_n^m(\rho)$, $0 \leq \rho \leq 1$.

As a consequence of the first identity in (19), the orthogonality and the normalization properties of the $R_n^m(\rho)$, $0 \leq \rho \leq 1$, and the definition of $M_{nn'}^m(\varepsilon)$ in (14), we have that

$$R_{n'}^m(\varepsilon\rho) = \sum_n (R_{n'}^n(\varepsilon) - R_{n'}^{n+2}(\varepsilon)) R_n^m(\rho), \quad n' = m, m+2, \ldots, \tag{25}$$

where the summation is over $n=m, m+2, \ldots, n'$. There is a variety of other identities of type (25) that can be obtained by interchanging ε and ρ and/or by reorganizing the series expression at the right-hand side, and/or by using the second identity in Equation (19). We may also observe that all equations presented up to now are valid for all complex values of ε, ρ (not necessarily restricted to $[0, 1]$).

2.2 *Results on derivatives*

2.2.1 Expressions for derivatives of β_n^m

There holds for $m=0, 1, \ldots$ and $n=m, m+2, \ldots$

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}(\beta_n^m(\varepsilon)) = \frac{1}{\varepsilon} \sum_{n'} (n R_{n'}^n(\varepsilon) + (n+2) R_{n'}^{n+2}(\varepsilon)) \beta_{n'}^m, \tag{26}$$

where the summation is over $n'=n, n+2, \ldots$. The particular case $\varepsilon=1$,

$$\begin{aligned}(\beta_n^m)'(1) &= n\beta_n^m + 2(n+1)[\beta_{n+2}^m + \beta_{n+4}^m + \cdots] \\ &= n\beta_n^m + 2(n+1) \sum_{k=1}^{\infty} \beta_{n+2k}^m\end{aligned} \tag{27}$$

was already presented in [9]. Furthermore, we have

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^2 (\beta_n^m(\varepsilon)) = \sum_{n'} \Bigg\{ &\left(\frac{n(n-1)}{\varepsilon^2(1-\varepsilon^2)} + \frac{3n - n'(n'+2)}{1-\varepsilon^2} \right) R_{n'}^n(\varepsilon) \\ &+ \left(-\frac{(n+2)(n+3)}{\varepsilon^2(1-\varepsilon^2)} + \frac{3(n+2) + n'(n'+2)}{1-\varepsilon^2} \right) R_{n'}^{n+2}(\varepsilon) \Bigg\} \beta_{n'}^m,\end{aligned} \tag{28}$$

which shows that compact results for higher derivatives than the first one should not be expected to exist. The explicit result for the case $\varepsilon = 1$,

$$(\beta_n^m)''(1) = n(n-1)\beta_n^m + 4(n+1)\sum_{k=1}^{\infty}\left[k(n+k+1) - \frac{3}{2}\right]\beta_{n+2k}^m \tag{29}$$

is, however, still reasonably tidy.

2.2.2 Expressions for derivatives of $S(\alpha(\varepsilon))$

We consider pure-phase aberrations $P = \exp(\mathrm{i}\Phi)$ in which the real phase function Φ has the Zernike expansion

$$\Phi(\rho, \vartheta) = \sum_{n,m} \alpha_n^m Z_n^m(\rho, \vartheta), \quad 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \tag{30}$$

with real, reasonably small expansion coefficients α. We shall generalize the notion of the Strehl ratio by defining it as

$$\tilde{S} = \max_V |U|^2, \tag{31}$$

where U is the point-spread function given in Equation (5), and V denotes a subset of the focal volume such as

(i) the single point best focus, on axis,
(ii) the x axis at best focus,
(iii) the optical axis (f axis),
(iv) the whole (x, f) plane.

Note that $|P| = 1$ so that the normalization as in Equation (6) disappears in Equation (31). Furthermore, since we have restricted consideration to the cosine option in (3), we have in (ii) and (iv) only the x axis, rather than the whole image plane (x, y).

In this more general situation, the Strehl ratio has the approximation

$$\begin{aligned} S(\alpha) &= 1 - \frac{1}{\pi}\int_0^1\int_0^{2\pi}\left|\sum_{n,m}^{\sim} \alpha_n^m Z_n^m(\rho, \vartheta)\right|^2 \rho \,\mathrm{d}\rho \,\mathrm{d}\vartheta \\ &= 1 - \sum_{n,m}^{\sim} \frac{(\alpha_n^m)^2}{\varepsilon_m(n+1)}, \end{aligned} \tag{32}$$

where the $\sim$ on top of the summation signs means to indicate that in the summation a set of the form

$$\{(n, m) | m = 0, 1, \ldots; \quad n = m, m+2, \ldots, n(m) - 2\}, \tag{33}$$

with integer $n(m) \ge m$ having the same parity as m, has been deleted. In the above four cases the appropriate choice for the set in (33) is

(i) $n(0) = 2$; $n(m) = m$, $m = 1, 2, \ldots$,
(ii) $n(0) = 2$, $n(1) = 3$; $n(m) = m$, $m = 2, 3, \ldots$,
(iii) $n(0) = 4$; $n(m) = m$, $m = 1, 2, \ldots$,
(iv) $n(0) = 4$, $n(1) = 3$; $n(m) = m$, $m = 2, 3, \ldots$.

We then have the following results. We let

$$\tilde{\Phi}(\rho, \vartheta, \varepsilon) = \sum_{n,m}^{\sim} \alpha_n^m(\varepsilon) Z_n^m(\rho, \vartheta) \tag{34}$$

be the $\sim$-reduced aberration phase of the scaled pupil. Then

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} S(\alpha(\varepsilon)) = \frac{2}{\pi\varepsilon} \int_0^1 \int_0^{2\pi} |\tilde{\Phi}(\rho, \vartheta, \varepsilon)|^2 \rho \, \mathrm{d}\rho \, \mathrm{d}\vartheta - \frac{1}{\pi\varepsilon} \int_0^{2\pi} |\tilde{\Phi}(1, \vartheta, \varepsilon)|^2 \, \mathrm{d}\vartheta \tag{35}$$

and

$$\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^2 S(\alpha(\varepsilon)) = \frac{-6}{\varepsilon^2} \left[\frac{1}{\pi} \int_0^1 \int_0^{2\pi} |\tilde{\Phi}(\rho, \vartheta, \varepsilon)|^2 \rho \, \mathrm{d}\rho \, \mathrm{d}\vartheta - \frac{1}{2\pi} \int_0^{2\pi} |\tilde{\Phi}(1, \vartheta, \varepsilon)|^2 \, \mathrm{d}\vartheta \right] - \frac{1}{\pi\varepsilon} \int_0^{2\pi} \frac{\partial}{\partial\varepsilon} |\tilde{\Phi}(\rho, \vartheta, \varepsilon)|^2 \, \mathrm{d}\vartheta. \tag{36}$$

These formulas can be written succinctly as

$$S(\alpha(\varepsilon)) = 1 - \overline{|\tilde{\Phi}(\varepsilon)|^2_{\mathrm{disk}}}, \tag{37}$$

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}(S(\alpha(\varepsilon))) = \frac{2}{\varepsilon} \left(\overline{|\tilde{\Phi}(\varepsilon)|^2_{\mathrm{disk}}} - \overline{|\tilde{\Phi}(\varepsilon)|^2_{\mathrm{rim}}} \right), \tag{38}$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^2 (S(\alpha(\varepsilon))) = \frac{-6}{\varepsilon^2} \left(\overline{|\tilde{\Phi}(\varepsilon)|^2_{\mathrm{disk}}} - \overline{|\tilde{\Phi}(\varepsilon)|^2_{\mathrm{rim}}} \right) - \frac{2}{\varepsilon} \overline{\frac{\partial}{\partial\varepsilon} |\tilde{\Phi}(\varepsilon)|^2_{\mathrm{rim}}}, \tag{39}$$

where the two types of averaging refer to the whole disk $0 \le \rho \le 1$ and the rim $\rho = 1$, respectively. The formulas (38) and (39) give a clue as to how to choose Φ such that the first and second derivative of $S(\alpha(\varepsilon))$ at $\varepsilon = 1$ exhibit a desired sign combination (also see Section 3).

The first two derivatives of $S(\alpha(\varepsilon))$ can also be expressed solely in terms of $\alpha(\varepsilon)$. There holds

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}(S(\alpha(\varepsilon))) = \frac{2}{\varepsilon} \sum_{n,m}^{\sim} \frac{(\alpha_n^m(\varepsilon))^2}{\varepsilon_m(n+1)} - \frac{2}{\varepsilon} \sum_m \frac{1}{\varepsilon_m} \left(\sum_n^{\sim} \alpha_n^m(\varepsilon) \right)^2, \tag{40}$$

and

$$\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^2 (S(\alpha(\varepsilon))) = \frac{-6}{\varepsilon^2} \left[\sum_{n,m}^{\sim} \frac{(\alpha_n^m(\varepsilon))^2}{\varepsilon_m(n+1)} - \sum_m \frac{1}{\varepsilon_m} \left(\sum_n^{\sim} \alpha_n^m(\varepsilon) \right)^2 \right] - \frac{4}{\varepsilon} \sum_m \frac{1}{\varepsilon_m} \left(\sum_n^{\sim} \alpha_n^m(\varepsilon) \right) \left(\sum_n^{\sim} (\alpha_n^m)'(\varepsilon) \right), \tag{41}$$

where the $\sim$ on top of the $\sum_n$ means to indicate summation over $n = n(m), n(m) + 2, \ldots$.

The case that $\varepsilon = 1$ deserves special attention since often the NA is reduced only a small fraction below its maximum value. We have $\alpha_n^m(\varepsilon = 1) = \alpha_n^m$ and $\tilde{\Phi}(\varepsilon = 1) = \tilde{\Phi}$ is the $\sim$-reduced aberration phase of the unscaled pupil. Then we get, for instance,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}(S(\alpha(\varepsilon)))\Big|_{\varepsilon=1} &= 2\left(\overline{|\tilde{\Phi}|^2_{\mathrm{disk}}} - \overline{|\tilde{\Phi}|^2_{\mathrm{rim}}}\right) \\ &= 2\sum_{n,m}^{\sim}\frac{(\alpha_n^m)^2}{\varepsilon_m(n+1)} - 2\sum_m \frac{1}{\varepsilon_m}\left(\sum_n^{\sim}\alpha_n^m\right)^2, \end{aligned} \tag{42}$$

and

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^2 S(\alpha(\varepsilon))\Big|_{\varepsilon=1} &= -6\left[\sum_{n,m}^{\sim}\frac{(\alpha_n^m)^2}{\varepsilon_m(n+1)} - \sum_m \frac{1}{\varepsilon_m}\left(\sum_n^{\sim}\alpha_n^m\right)^2\right] \\ &\quad - 4\sum_m \frac{1}{\varepsilon_m}\left(\sum_n^{\sim}\alpha_n^m\right)\left(\sum_n^{\sim}\alpha_n^m\left[n + \frac{1}{2}(n^2 - n^2(m))\right]\right). \end{aligned} \tag{43}$$

2.2.3 Semigroup structure of the scaling operation

We set for $0 \le \varepsilon \le 1$ and $m = 0, 1, \ldots$

$$\beta^m = (\beta_n^m)_{n=m,m+2,\ldots}, \quad \beta^m(\varepsilon) = (\beta_n^m(\varepsilon))_{n=m,m+2,\ldots}, \tag{44}$$

and

$$N^m(\varepsilon) = (N^m_{nn'}(\varepsilon))_{n,n'=m,m+2,\ldots} = (R^n_{n'}(\varepsilon) - R^{n+2}_{n'}(\varepsilon))_{n,n'=m,m+2,\ldots}. \tag{45}$$

Thus, $N^m_{nn'}(\varepsilon)$ denotes the matrix element of $N^m(\varepsilon)$ with row index n and column index n'. From Equation (18) we have for $m = 0, 1, \ldots$

$$\beta^m(\varepsilon) = N^m(\varepsilon)\beta^m. \tag{46}$$

We shall below omit superscripts m by which we mean to say that we are considering the whole aggregate of coefficients and matrices in (44)–(45) with $m = 0, 1, \ldots$.

The property that scaling the pupil function by factors ε_1 and subsequently ε_2 is the same as scaling by the factor $\varepsilon_1\varepsilon_2 = \varepsilon_2\varepsilon_1$ is reflected in terms of β's and N's as

$$\begin{aligned}\beta(\varepsilon_2\varepsilon_1) &= N(\varepsilon_2)\beta(\varepsilon_1) = N(\varepsilon_2)N(\varepsilon_1)\beta \\ &= \beta(\varepsilon_1\varepsilon_2) = N(\varepsilon_1)\beta(\varepsilon_2) = N(\varepsilon_1)N(\varepsilon_2)\beta \\ &= N(\varepsilon_2\varepsilon_1)\beta = N(\varepsilon_1\varepsilon_2)\beta. \end{aligned} \tag{47}$$

Hence, the $(N(\varepsilon))_{0 \le \varepsilon \le 1}$ have a commutative semigroup structure (for a group structure we would also need the existence of inverses). Note that

$$N(\varepsilon = 1) = I(\text{identity matrix}). \tag{48}$$

The semigroup property

$$N(\varepsilon_2\varepsilon_1) = N(\varepsilon_2)N(\varepsilon_1) = N(\varepsilon_1)N(\varepsilon_2) \tag{49}$$

can also be verified directly from Equation (25). Indeed, using (25) for $R_n^{n''}(\varepsilon\rho)$ and $R_n^{n''+2}(\varepsilon\rho)$, we get for any $m=0,1,\ldots,m=n'',\ n''-2,\ldots$ from Equation (45) that

$$\begin{aligned} N_{n''n}^{m}(\varepsilon\rho) &= R_n^{n''}(\varepsilon\rho) - R_n^{n''+2}(\varepsilon\rho) \\ &= \sum_{n'} (R_n^{n'}(\varepsilon) - R_n^{n'+2}(\varepsilon))(R_{n'}^{n''}(\rho) - R_{n'}^{n''+2}(\rho)) \\ &= \sum_{n'} N_{n'n}^{m}(\varepsilon) N_{n''n'}^{m}(\rho) = (N^m(\rho)N^m(\varepsilon))_{n''n}. \end{aligned} \tag{50}$$

The infinitesimal generator B of the semigroup $(N(\varepsilon))_{0\le\varepsilon\le 1}$ is defined as

$$B = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} N(\varepsilon)\Big|_{\varepsilon=1}. \tag{51}$$

By Equation (46) we have for $m=0,1,\ldots$

$$(\beta^m)'(1) = (N^m)'(1)\beta^m = B\beta^m. \tag{52}$$

Hence, from Equation (27) we see that the row of B^m with index $n=m,\ m+2,\ldots$ is given by

$$\underbrace{0\ \ 0\ \cdots\ 0}_{\frac{n-m}{2}}\ \ n\ \ 2(n+1)\ \ 2(n+1)\ldots. \tag{53}$$

The infinitesimal generator is useful for compact representation of results of computations. This is based on the formula

$$N(\varepsilon) = \exp(B\ \ln\ \varepsilon) = \sum_{k=0}^{\infty} \frac{(\ln\varepsilon)^k}{k!} B^k. \tag{54}$$

For instance, we have for $l=0,1,\ldots$

$$\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^l \beta(\varepsilon) = \left(\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^l \exp(B\ \ln\ \varepsilon)\right)\beta. \tag{55}$$

For $l=1$ this yields

$$\beta'(\varepsilon) = \frac{1}{\varepsilon} BN(\varepsilon)\beta, \tag{56}$$

and this gives (26) when using (53) and (45). Similarly,

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}[S(\alpha(\varepsilon))] = (\nabla S)(\alpha(\varepsilon))\cdot\alpha'(\varepsilon) = \frac{1}{\varepsilon}(\nabla S)(\alpha(\varepsilon))\cdot B\alpha(\varepsilon). \tag{57}$$

This shows how the general results (38) and (40) follow from the $\varepsilon=1$ result in (42).

3. Quality assessment of NA-reduced systems by Strehl ratios

We consider in this section pure-phase aberrations so that the pupil function P is given as $\exp(\mathrm{i}\Phi)$ with Φ having the Zernike expansion (8) with real and sufficiently small α's. We can then replace the Strehl ratio in (6) by the quantity $S(\alpha)$ in (9) where the summation over (n,m) omits the term $(n,m)=(0,0)$. Thus, in the general setting of Subsection 2.2.2 we have the case

(i) $n(0)=2$; $n(m)=m$, $m=1,2,\ldots$.

Figure 1 shows an example where we scale an aberrated pupil with $\alpha_3^1=0.1$ and $\alpha_5^1=0.1$ while all other $\alpha_n^m=0$. We have in this case by Equation (18)

$$\alpha_3^1(\varepsilon)=\alpha_3^1R_3^3(\varepsilon)+\alpha_5^1(R_5^3(\varepsilon)-R_5^5(\varepsilon))=\varepsilon^3\alpha_3^1+(4\varepsilon^5-4\varepsilon^3)\alpha_5^1, \tag{58}$$

$$\alpha_5^1(\varepsilon)=\alpha_5^1R_5^5(\varepsilon)=\varepsilon^5\alpha_5^1. \tag{59}$$

Furthermore, in agreement with Equations (26) and (28)

$$(\alpha_3^1)'(\varepsilon)=3\varepsilon^2\alpha_3^1+(20\varepsilon^4-12\varepsilon^2)\alpha_5^1,\quad (\alpha_5^1)'(\varepsilon)=5\varepsilon^4\alpha_5^1, \tag{60}$$

and

$$(\alpha_3^1)''(\varepsilon)=6\varepsilon\alpha_3^1+(80\varepsilon^3-24\varepsilon)\alpha_5^1,\quad (\alpha_5^1)''(\varepsilon)=20\varepsilon^3\alpha_5^1. \tag{61}$$

In the case $\varepsilon=1$ we then find $\alpha_3^1(1)=\alpha_3^1$, $\alpha_5^1(1)=\alpha_5^1$, and, in agreement with (27) and (29)

$$(\alpha_3^1)'(1)=3\alpha_3^1+8\alpha_5^1,\quad (\alpha_5^1)'(1)=5\alpha_5^1, \tag{62}$$

and

$$(\alpha_3^1)''(1)=6\alpha_3^1+56\alpha_5^1,\quad (\alpha_5^1)''(1)=20\alpha_5^1. \tag{63}$$

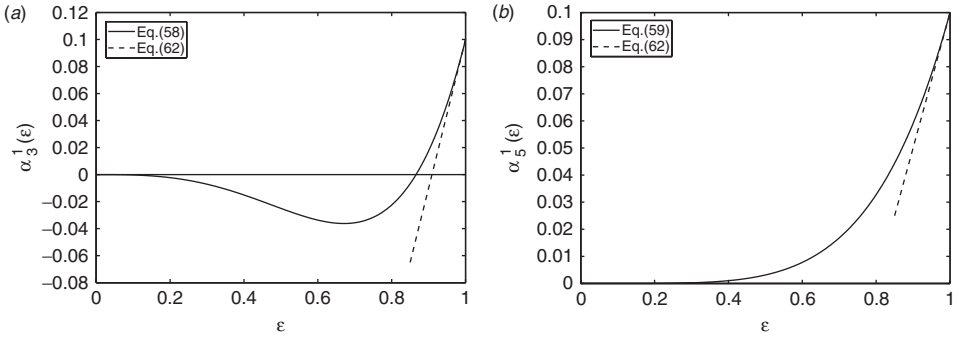

Figure 1. Scaling an aberrated pupil with $\alpha_3^1=\alpha_5^1=0.1$ to relative size $\varepsilon\le 1$. In (*a*) we show $\alpha_3^1(\varepsilon)$ as a function of ε, see (58), where the tangent line at $\varepsilon=1$ has slope given in accordance with (62). In (*b*) we do the same for $\alpha_5^1(\varepsilon)$.

We note from these formulas that the relative sensitivity $(\alpha_n^m)^{-1}\, \mathrm{d}\alpha_n^m/\mathrm{d}\varepsilon$ increases sharply when ε approaches 1. In the case of Figure 1 we have

$$\frac{1}{\alpha_3^1}(\alpha_3^1)'(1) = 11, \quad \frac{1}{\alpha_5^1}(\alpha_5^1)'(1) = 5. \tag{64}$$

We next consider the Strehl ratio for two pupils containing a variety of aberrations (low-to-medium-high order spherical, coma, astigmatism and trefoil). In the first example, see Figure 2, the Strehl ratio as a function of the scaling parameter ε behaves as one expects: it decreases with increasing NA, and it does so faster at higher NA. A somewhat more complicated behavior of the Strehl ratio as a function of NA occurs in the second example, see Figure 3. For this second example, we have displayed in Figure 4 the plots of eight functions $\alpha_n^m(\varepsilon)$ with scaling parameter ε between 0 and 1.

The example in Figure 3 already shows that the behavior of the Strehl ratio as a function of NA can be more complicated than one is inclined to expect. The next example illustrates how complicated things can get, even in the simple case that only two spherical aberration terms are present. In Figure 5 we have plotted

$$S(\alpha(\varepsilon)) = 1 - \sum_{(n,m)\neq(0,0)} \frac{(\alpha_n^m(\varepsilon))^2}{\varepsilon_m(n+1)} \tag{65}$$

for the case of the radially symmetric aberration phase

$$\Phi(\rho, \vartheta) = \Phi(\rho) = \alpha_4^0 R_4^0(\rho) + \alpha_6^0 R_6^0(\rho) \tag{66}$$

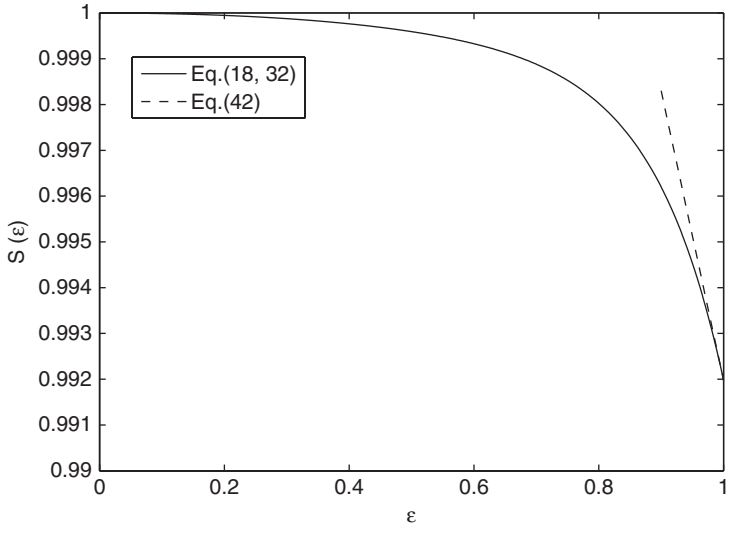

Figure 2. Strehl ratio $S(\alpha(\varepsilon))$ as a function of ε for a pupil containing the following cocktail of non-zero aberrations: $\alpha_2^0 = 0.1$, $\alpha_4^0 = 0.03$, $\alpha_6^0 = 0.01$ (spherical); $\alpha_1^1 = 0.1$, $\alpha_3^1 = 0.03$, $\alpha_5^1 = 0.01$ (coma); $\alpha_2^2 = 0.1$, $\alpha_4^2 = 0.03$, $\alpha_6^2 = 0.01$ (astigmatism); $\alpha_3^3 = 0.1$, $\alpha_5^3 = 0.03$ (trefoil). The drawn line shows $S(\alpha(\varepsilon))$, see (18), (32) with α's instead of β's, and the tangent line at $\varepsilon = 1$ has slope given in accordance with (42).

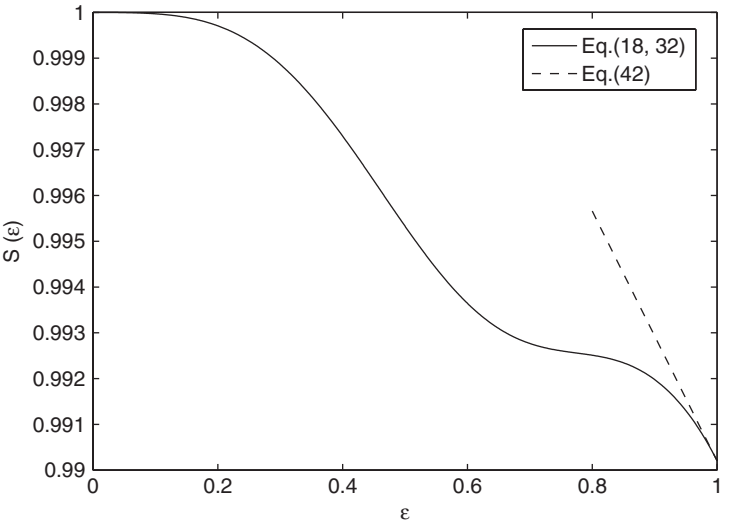

Figure 3. Same caption as in Figure 2, with a different cocktail of aberrations: $\alpha_2^0 = 0.1$, $\alpha_4^0 = -0.075$, $\alpha_6^0 = 0.05$ (spherical); $\alpha_1^1 = 0.1$, $\alpha_3^1 = 0.05$, $\alpha_5^1 = -0.025$ (coma); $\alpha_2^2 = 0.05$, $\alpha_4^2 = 0.1$, $\alpha_6^2 = -0.05$ (astigmatism); $\alpha_3^3 = 0.05$, $\alpha_5^3 = 0.05$ (trefoil).

with $\alpha_4^0 = -\alpha_6^0 = 0.04\pi$. We see that $S(\alpha(\varepsilon))$ exhibits the expected behaviour only to $\varepsilon = 0.6$, after which $S(\alpha(\varepsilon))$ increases.

Let us examine what happens at $\varepsilon = 1$ in the case that $S(\alpha(\varepsilon))$ and Φ are as in (65)–(66) with general α_4^0 and α_6^0. We get from formulas (42)–(43) in this case

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}(S(\alpha(\varepsilon)))\Big|_{\varepsilon=1} = 2\left(\frac{1}{7}(\alpha_6^0)^2 + \frac{1}{5}(\alpha_4^0)^2\right) - 2(\alpha_6^0 + \alpha_4^0)^2, \tag{67}$$

and

$$\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^2(S(\alpha(\varepsilon)))\Big|_{\varepsilon=1} = -6\left(\frac{1}{7}(\alpha_6^0)^2 + \frac{1}{5}(\alpha_4^0)^2 - (\alpha_6^0 + \alpha_4^0)^2\right) - 4(\alpha_6^0 + \alpha_4^0)(22\alpha_6^0 + 10\alpha_4^0). \tag{68}$$

We set $t = \alpha_6^0/\alpha_4^0$, and we consider the quadratics

$$Q_1(t) = (\alpha_4^0)^2 \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(S(\alpha(\varepsilon)))\Big|_{\varepsilon=1} = 2\left(\frac{1}{7}t^2 + \frac{1}{5}\right) - 2(t+1)^2, \tag{69}$$

$$\begin{aligned} Q_2(t) &= (\alpha_4^0)^2\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^2(S(\alpha(\varepsilon)))\Big|_{\varepsilon=1} \\ &= -6\left(\frac{1}{7}t^2 + \frac{1}{5} - (t+1)^2\right) - 4(t+1)(22t+10). \end{aligned} \tag{70}$$

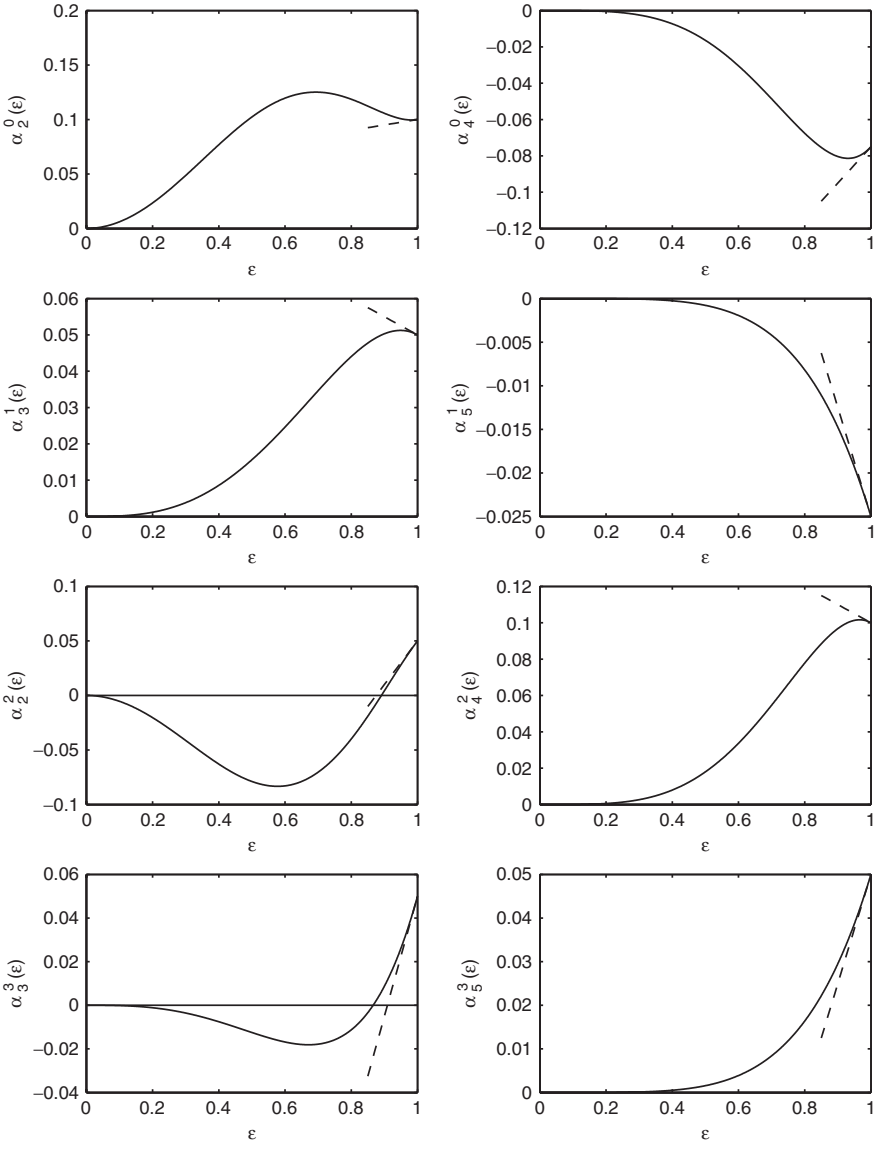

Figure 4. Scaling the aberrated pupil of Figure 3 to relative size $\varepsilon \leq 1$. The drawn lines in the separate figures show $\alpha_n^m(\varepsilon)$, as indicated along the vertical axis, according to (18) with α's instead of β's. The tangent lines at $\varepsilon = 1$ have slope given in accordance with (27) with α's instead of β's.

There holds

$$Q_1(t) = -\frac{12}{7}\left(t^2 + \frac{7}{3}t + \frac{14}{15}\right) = -\frac{12}{7}(t - t_{1,+})(t - t_{1,-}), \tag{71}$$

$$Q_2(t) = -\frac{580}{7}\left(t^2 + \frac{7}{5}t + \frac{308}{725}\right) = -\frac{580}{7}(t - t_{2,+})(t - t_{2,-}), \tag{72}$$

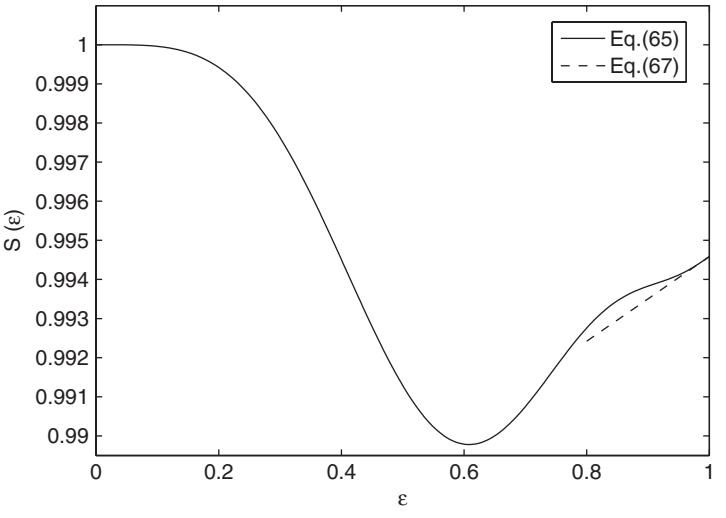

Figure 5. Strehl ratio $S(\alpha(\varepsilon))$ as a function of ε for a pupil with two non-zero aberration coefficients, $\alpha_4^0 = -\alpha_6^0 = 0.04\pi$. The drawn line shows $S(\alpha(\varepsilon))$ as given by (65) while the tangent line at $\varepsilon = 1$ has slope given in accordance with (67).

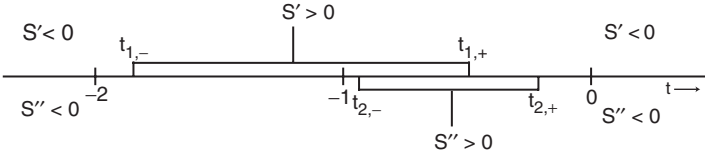

Figure 6. Sign of S', $S'' = (\mathrm{d}/\mathrm{d}\varepsilon)[S(\alpha(\varepsilon))]$, $(\mathrm{d}/\mathrm{d}\varepsilon)^2[S(\alpha(\varepsilon))]$ at $\varepsilon = 1$ for the case that the aberrations are given in (66) and $t = \alpha_4^0/\alpha_6^0$. The special points $t_{1,\pm}$ and $t_{2,\pm}$ are given in (73)–(74).

where

$$t_{1,+} = -\frac{7}{6} + \left(\frac{77}{180}\right)^{1/2} = -0.512619438, \quad t_{1,-} = -\frac{7}{6} - \left(\frac{77}{180}\right)^{1/2} = -1.820713896, \tag{73}$$

$$t_{2,+} = -\frac{7}{10} + \left(\frac{189}{2900}\right)^{1/2} = -0.444711117, \quad t_{2,-} = -\frac{7}{10} - \left(\frac{189}{2900}\right)^{1/2} = -0.955288883. \tag{74}$$

Since $t_{1,-} < t_{2,-} < t_{1,+} < t_{2,+}$ it is seen that all sign combinations for the first and second derivative of $S(\alpha(\varepsilon))$ at $\varepsilon = 1$ occur. Also see Figure 6.

4. ENZ-theory for NA-reduced optical systems and for systems with a central obstruction

In the so-called Extended Nijboer–Zernike (ENZ) theory of diffraction, a general pupil function P, defined on the full disk $0 \le \rho \le 1$, is expanded as a Zernike series as in (3), and

the complex-amplitude point-spread function is obtained as

$$U(r,\varphi,f) = \frac{1}{\pi}\int_0^1\int_0^{2\pi} \exp(\mathrm{i}f\rho^2)P(\rho,\vartheta)\ \exp[2\pi\mathrm{i}\rho r\cos(\vartheta-\varphi)]\rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta$$
$$= 2\sum_{n,m}\mathrm{i}^m\beta_n^m V_n^m(r,f)\cos m\varphi. \tag{75}$$

In (75) we have normalized polar coordinates r, φ in the image planes, and f is the normalized focal variable (under low-to-medium-high NA conditions it is permitted to represent the defocus factor by exp $(\mathrm{i}f\rho^2)$). The V_n^m in (75) are given in integral form as

$$V_n^m(r,f) = \int_0^1 \exp(\mathrm{i}f\rho^2)R_n^m(\rho)J_m(2\pi r\rho)\rho\ \mathrm{d}\rho, \tag{76}$$

where J_m is the Bessel function of the first kind and of order m. The V_n^m can be computed in the form of a well-convergent power-Bessel series for all values of r and values of $|f|$ up to 20, or, alternatively, in the form of a somewhat more complicated Bessel–Bessel series that converges virtually without loss-of-digits for all values of r and f, see [11,15]. For basic ENZ theory, we refer to [11,12]; for retrieval of aberrations from intensity point-spread functions in the focal region under low-to-medium-high NA conditions within the ENZ framework, we refer to [14,16]; for point-spread function computation under high-NA conditions (including vector diffraction theory and polarization), we refer to [13]; for aberration and birefringence retrieval under high-NA conditions, we refer to [18,19].

4.1 *ENZ point-spread function calculation for scaled pupils*

We shall first consider the scaling issue in the basic ENZ setting as presented above. So assume that we have a smooth pupil function $P(\rho,\vartheta)$, $0\le\rho\le1$, $0\le\vartheta\le2\pi$. Setting the NA value to a fraction $\varepsilon\le1$ of its maximum amounts to hard thresholding of $P(\rho,\vartheta)$ to the value 0 for $\varepsilon\le\rho\le1$. A direct use of the ENZ formalism, in which the thresholded pupil function is developed as a Zernike series on the full pupil $0\le\rho\le1$, is cumbersome since an unacceptable number of terms in such a series is required for a reasonable accuracy. In Appendix 3 we show that for the case that $P(\rho,\vartheta)=1$, $0\le\rho\le1$, $0\le\vartheta\le2\pi$, the thresholded pupil has Zernike coefficients β_{2n}^0 that decay as slow as $n^{-1/2}$. Instead, we proceed as in (10), where the point-spread function of the NA-reduced optical system is written in the form

$$U_\varepsilon(r,\varphi,f) = \frac{\varepsilon^2}{\pi}\int_0^1\int_0^{2\pi} \exp(\mathrm{i}f\varepsilon^2\rho^2)P(\varepsilon\rho,\vartheta)\ \exp[2\pi\mathrm{i}r\varepsilon\rho\cos(\vartheta-\varphi)]\rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta. \tag{77}$$

From the Zernike expansion

$$P(\varepsilon\rho,\vartheta) = \sum_{n,m}\beta_n^m(\varepsilon)Z_n^m(\rho,\vartheta) \tag{78}$$

of the scaled pupil function, with $\beta_n^m(\varepsilon)$ given by Equation (18), we then get

$$U_\varepsilon(r,\varphi,f) = 2\varepsilon^2 \sum_{n,m} \mathrm{i}^m \beta_n^m(\varepsilon) V_n^m(\varepsilon r, \varepsilon^2 f)\cos m\varphi. \tag{79}$$

An alternative approach, leading to the same computation scheme for U_ε, is to insert the Zernike expansion (3) of P into the first double-integral expression in Equation (10). Using $Z_n^m(\rho,\vartheta) = R_n^m(\rho)\cos m\vartheta$, this leads to

$$U_\varepsilon(r,\varphi,f) = 2\sum_{n,m} \mathrm{i}^m \beta_n^m V_n^m(r,f;\varepsilon)\cos m\varphi, \tag{80}$$

where

$$V_n^m(r,f;\varepsilon) = \int_0^\varepsilon \exp(\mathrm{i}f\rho^2) R_n^m(\rho) J_m(2\pi r\rho)\rho \,\mathrm{d}\rho. \tag{81}$$

With the substitution $\rho = \varepsilon\rho_1$, $0 \le \rho_1 \le 1$, in the latter integral and using Equation (25) to write $R_n^m(\varepsilon\rho_1)$ as a linear combination of the $R_n^m(\rho_1)$, we get from Equation (76)

$$V_n^m(r,f;\varepsilon) = \varepsilon^2 \sum_{n'} (R_n^{n'}(\varepsilon) - R_n^{n'+2}(\varepsilon)) V_{n'}^m(r\varepsilon, f\varepsilon^2). \tag{82}$$

The advantage of (80) and (82) over (79) is that (80) is directly in terms of the Zernike coefficients of the unscaled pupil function and that the scaling operation is completely represented by the modification of V_n^m functions as in (82).

In a similar fashion we can compute point-spread functions pertaining to a pupil function P that vanishes for $0 \le \rho < \varepsilon$ and that admits in $\varepsilon \le \rho \le 1$ a Zernike expansion

$$P(\rho,\vartheta) = \sum_{n,m} \gamma_n^m Z_n^m(\rho,\vartheta), \quad \varepsilon \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi. \tag{83}$$

In Appendix 2 we shall address the problem of how to obtain a feasible Zernike approximation as in (83) for a well-behaved P in an annulus $\varepsilon \le \rho \le 1$. Now the point-spread function U corresponding to this P is given by

$$U(r,\varphi,f) = 2\sum_{n,m} \mathrm{i}^m \gamma_n^m W_n^m(r,f;\varepsilon)\cos m\varphi, \tag{84}$$

where

$$\begin{aligned} W_n^m(r,f;\varepsilon) &= \int_\varepsilon^1 \exp(\mathrm{i}f\rho^2) R_n^m(\rho) J_m(2\pi r\rho)\rho \,\mathrm{d}\rho \\ &= V_n^m(r,f) - V_n^m(r,f;\varepsilon) \end{aligned} \tag{85}$$

with $V_n^m(r,f)$ and $V_n^m(r,f;\varepsilon)$ given in (76) and (82), respectively.

A related result concerns the computation of point-spread functions for certain multi-ring systems. Assume we have a pupil function $P(\rho,\vartheta)$, $0 \le \rho \le 1$, $0 \le \vartheta \le 2\pi$, with a Zernike expansion as in (3), and numbers

$$0 = \varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_J = 1\ ; \quad a_1, a_2, \ldots, a_J \in \mathbb{C}, \tag{86}$$

and consider as pupil function

$$\tilde{P}(\rho, \vartheta) = a_j P(\rho, \vartheta), \quad \varepsilon_{j-1} \leq \rho < \varepsilon_j, \quad 0 \leq \vartheta \leq 2\pi, \tag{87}$$

with j running from 1 to J. Compare [20] where the case $P = 1$ is considered. Then the point-spread function $\tilde{U}$ corresponding to $\tilde{P}$ is given by

$$\tilde{U}(r, \varphi, f) = 2 \sum_{n,m} \mathrm{i}^m \beta_n^m \tilde{V}_n^m(r, f) \cos m\varphi \tag{88}$$

in which

$$\begin{aligned} \tilde{V}_n^m(r, f) &= \sum_{j=1}^{J} a_j \int_{\varepsilon_{j-1}}^{\varepsilon_j} \exp(\mathrm{i} f \rho^2) R_n^m(\rho) J_m(2\pi r \rho) \rho \, \mathrm{d}\rho \\ &= \sum_{j=1}^{J} a_j [V_n^m(r, f; \varepsilon_j) - V_n^m(r, f; \varepsilon_{j-1})] \end{aligned} \tag{89}$$

with $V_n^m(r, f; \varepsilon)$ given in (82).

4.2 *ENZ aberration retrieval for optical systems with a central obstruction*

Assume that we have an optical system with an unknown pupil function $P(\rho, \vartheta)$, $0 \leq \rho \leq 1$, $0 \leq \vartheta \leq 2\pi$. In [14,16] it has been shown how one can estimate P from the through-focus intensity point-spread function $I = |U|^2$ of the optical system. The key step is to choose the unknown Zernike coefficients β_n^m of P such that the match between the recorded intensity I and the theoretical intensity

$$|U(r, \varphi, f)|^2 = \left| 2 \sum_{n,m} \mathrm{i}^m \beta_n^m V_n^m(r, f) \cos m\varphi \right|^2 \tag{90}$$

is maximal. This procedure, and sophisticated variants of it, is remarkably accurate in estimating pupil functions with aberrations as large as twice the diffraction limit.

In the case that the pupil function is known to be obstructed in the region $0 \leq \rho \leq \varepsilon$, aberration retrieval can be still practiced with the above sketched approach by appropriately modifying the V_n^m functions involved in it. We thus propose P of the form

$$P(\rho, \vartheta) = \sum_{n,m} \gamma_n^m Z_n^m(\rho, \vartheta), \quad \varepsilon \leq \rho \leq 1, \quad 0 \leq \vartheta \leq 2\pi, \tag{91}$$

in which the γ_n^m are unknowns that are to be found by matching the recorded intensity and the theoretical intensity. This theoretical intensity is given in this case, see (84), as

$$I(r, \varphi, f) = |U(r, \varphi, f)|^2 = \left| 2 \sum_{n,m} \mathrm{i}^m \gamma_n^m W_n^m(r, f; \varepsilon) \cos m\varphi \right|^2, \tag{92}$$

with W_n^m given in (85). In Subsection 4.3 we shall illustrate this procedure with an example (Gaussian, comatic pupil function obstructed in the disk $0 \le \rho \le 1/2 = \varepsilon$), with the extra complication that the optical system has a high NA.

4.3 *Extension to high-NA systems*

In [13] the extension to high-NA optical systems of the ENZ approach to the calculation of optical point-spread functions has been given, and in [18,19] the retrieval procedure has been extended to high-NA systems. These extensions, though involved, are still based on the availability of computational schemes for certain basic integrals. We consider now for integer n, m, with $n - |m| \ge 0$ and even, and for integer j the integral

$$V_{n,j}^m(r, f, s_0) = \int_0^1 \frac{\rho^{|j|}(1 + (1 - s_0^2\rho^2)^{1/2})^{-|j|+1}}{(1 - s_0^2\rho^2)^{1/4}} \times \exp\left[\frac{\mathrm{i}f}{u_0}(1 - (1 - s_0^2\rho^2)^{1/2})\right] R_n^{|m|}(\rho) J_{m+j}(2\pi r\rho)\rho \,\mathrm{d}\rho \tag{93}$$

in which

$$s_0 \text{ is the NA value} \in [0, 1], \quad u_0 = 1 - (1 - s_0^2)^{1/2}. \tag{94}$$

In the case that the pupil function is thresholded to 0 in a set $\varepsilon \le \rho < 1$, the integration ranges of the integrals in (93) have to be changed from [0, 1] to $[0, \varepsilon]$, yielding the high-NA versions of the $V_n^m(r, f; \varepsilon)$ in (82). As in (81) the substitution $\rho = \varepsilon\rho$, $0 \le \rho_1 \le 1$, combined with the result of Equation (25) works wonders, and we get

$$V_{n,j}^m(r, f, s_0; \varepsilon) = \sum_{n'} \varepsilon^{|j|+2}(R_n^{n'}(\varepsilon) - R_n^{n'+2}(\varepsilon)) V_{n',j}^m\left(r\varepsilon, f\frac{u_0(\varepsilon)}{u_0}, s_0\varepsilon\right), \tag{95}$$

in which $u_0(\varepsilon) = 1 - [1 - (s_0\varepsilon)^2]^{1/2}$. Hence, the calculation of point-spread functions and aberration retrieval for high-NA systems with thresholded pupil functions can still be done on the level of the Zernike coefficients of the pupil function. Also, in the case of a pupil function obstructed in the disk $0 \le \rho < \varepsilon$ as in (83), we can do ENZ computation and retrieval under high-NA conditions by replacing the W_n^m in (85) by

$$W_{n,j}^m(r, f, s_0; \varepsilon) = V_{n,j}^m(r, f, s_0) - V_{n,j}^m(r, f, s_0; \varepsilon). \tag{96}$$

As an example, we consider retrieval of the Gaussian, comatic pupil function (with $s_0 = 0.95$)

$$P(\rho, \vartheta) = \exp[-\gamma\rho^2 + \mathrm{i}\alpha R_3^1(\rho)\cos\vartheta], \quad 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \tag{97}$$

which is obstructed in the disk $0 \le \rho < \varepsilon = 1/2$ and we take $\alpha = \gamma = 0.1$. For this we apply the procedures as given in [18,19] in which the $V_{n,j}^m$ of (93) are to be replaced throughout by the $W_{n,j}^m$ of (96). Since this example can at the present stage only be considered in a simulation environment, we should generate the 'recorded' intensity in the focal region

in simulation. This we do as follows. In [16], (A19) on p. 1726, the Zernike expansion of P on the full disk $0 \le \rho \le 1$,

$$P(\rho, \vartheta) = \sum_{m=0}^{\infty} \sum_{p=0}^{\infty} \beta_{m+2p}^{m} Z_{m+2p}^{m}(\rho, \vartheta), \quad 0 \le \rho \le 1, \quad 0 \le \vartheta \le 2\pi, \tag{98}$$

has been given explicitly with β's in the form of a triple series with good convergence properties when γ and α in (97) are of order unity. The representation (98) holds naturally also on the annular set $\varepsilon \le \rho \le 1$, $0 \le \vartheta \le 2\pi$, and thus the point-spread function computation scheme of [18,19], using this representation of P and with $W_{n,j}^{m}$ instead of $V_{n,j}^{m}$ as described above, can be applied. In theory, when infinite series in (98) are used, the coefficients retrieved with this procedure are not identical to the β's used in (98) (see Appendix 2 for more details). However, in the experiment we input and retrieve only β's with $0 \le m, n \le 10$, and thereby avoid this non-uniqueness problem. For this finite set of aberration coefficients, we use the iterative version (predictor–corrector approach) of the ENZ-retrieval method as described in [18], Section 4.1 and [21], Appendix A. In Figure 7(*a*) we show modulus and phase of $P_{\text{analytic}} - P_{\text{beta}}$ where P_{analytic} is the P of (97) and P_{beta} is the P of (98) in which the summations have been restricted to the terms

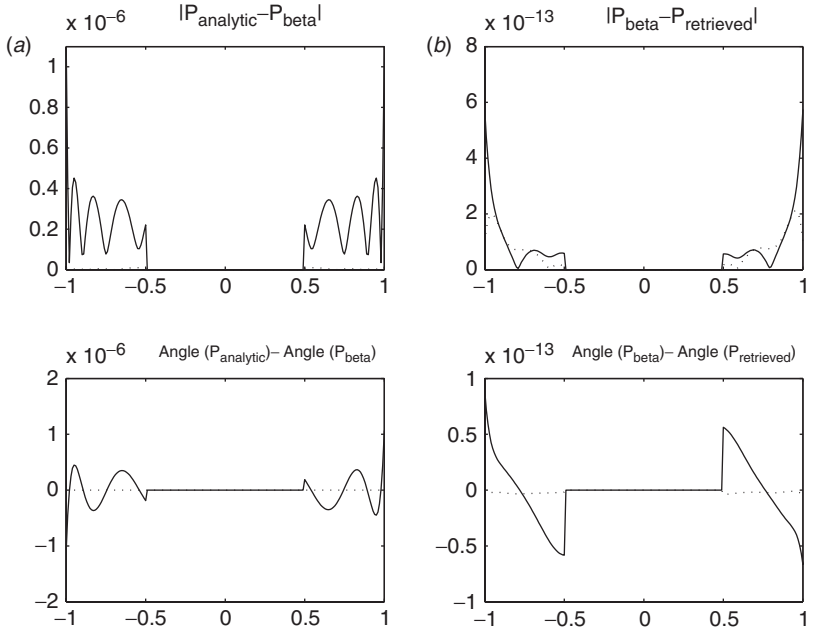

Figure 7. Modulus and phase of (*a*) $P_{\text{analytic}} - P_{\text{beta}}$ and (*b*) $P_{\text{beta}} - P_{\text{retrieved}}$ of two perpendicular cross-sections on the annular set $\varepsilon = \frac{1}{2} \le \rho \le 1$. Here P_{analytic} is given in (97) with $\alpha = \gamma = 0.1$ and P_{beta} is given by (98) in which the summations are restricted to m, p with $0 \le m, m + 2p \le 10$. Finally, $P_{\text{retrieved}}$ is obtained as the linear combination of Zernike terms Z_{m+2p}^{m}, with m, p such that $0 \le m, m + 2p \le 10$, where the coefficients β are obtained by applying 60 iterations of the high-NA ($s_0 = 0.95$), predictor–corrector version of the ENZ-retrieval method with $|U_{\text{beta}}|^2$ as 'recorded' intensity in the focal region.

m, p with $0 \le m$, $m + 2p \le 10$. In Figure 7(*b*) we show modulus and phase of $P_{\text{beta}} - P_{\text{retrieved}}$. Here P_{beta} is the same as in Figure 7(*a*), and $P_{\text{retrieved}}$ is the pupil function resulting from taking the same linear combination of Zernike terms that constitute P_{beta} but now with β's that are obtained by applying 60 iterations of the predictor–corrector procedure using $|U_{\text{beta}}|^2$ as 'recorded' intensity in the focal region. Note that the error levels in Figure 7(*a*) are well below those in Figure 7(*b*), whence a Figure 7(*c*), showing modulus and phase of $P_{\text{analytic}} - P_{\text{retrieved}}$, would practically coincide with Figure 7(*a*) and is therefore omitted.

5. Conclusion

In this paper we discussed applications of a recently derived mathematical result for obtaining the modified Zernike aberration coefficients of a scaled pupil function that comprises both the transmission variation and wavefront aberration of the optical system. The formulae that lead to the Zernike coefficients of the scaled pupil function have been further developed to yield also the derivatives of the separate Zernike coefficients with respect to the scaling ratio. Using our general scaling results we were able to evaluate the behavior of a typical quality factor of an optical system like the Strehl ratio under radial scaling. Surprisingly enough, the first and second derivatives of the Strehl ratio with respect to the radial scaling parameter do not always exhibit the behavior that follows from a simple intuitive picture, for instance a decrease in Strehl ratio with increasing aperture. Both negative and positive values are possible for the first and second derivatives. This means that the initial Zernike coefficients of the full pupil could be tailored in such a way that desired derivative values are realized under radial scaling.

The mathematical method for obtaining the Zernike coefficients of a radially scaled pupil can be implemented in the semi-analytic expression that was previously developed by us to evaluate the diffraction integral when using the Zernike coefficients of the pupil function. With the modified Zernike coefficients of the scaled pupil function, we easily obtain its intensity point-spread function in the focal region. We have also shown that it is possible to accommodate other modified pupil functions, for instance a centrally obstructed pupil function that is common for catadioptric imaging systems. Using the analytic tools available for the forward diffraction calculation from exit pupil to focal region, we have also addressed the inverse problem, namely, the reconstruction of the amplitude and phase of the pupil function of the optical system. We have demonstrated that the pupil function can be reconstructed in a wide range of amplitude and phase defects by using the information from several defocused intensity distributions; this method remains applicable for radially scaled pupil functions and pupil functions with a central obstruction. It is sufficient to adapt the forwardly calculated amplitude distributions in the focal region, that are associated with a typical Zernike aberration term, to the modified pupil shape. It was also shown in this paper that not only the scalar diffraction formalism can be handled, but also the complete vector diffraction case which is needed at high values of the numerical aperture of the imaging system. With the mathematical tools that have been made available in this paper, we have extended the range of application of our previous methods (analytic forward calculation and inverse aberration retrieval) to more

general pupil shapes. These more general shapes like radially down-scaled pupils and apertures with a central obstruction are frequently encountered in astronomical imaging, ophthalmologic metrology and in instrumentation for eye surgery.

References

[1] Strehl, K. *Theorie des Fernrohrs*; Barth: Leipzig, 1894.
[2] Maréchal, A. *Rev. Opt. (Théor. Instrum.)* **1947,** *26,* 257–277; *J. Opt. Soc. Am.* **1947,** *37*, 982–982.
[3] Born, M.; Wolf, E. *Principles of Optics*, 7th ed.; Cambridge University Press: Cambridge, UK, 2002.
[4] Goldberg, K.A.; Geary, K. *J. Opt. Soc. Am. A* **2001,** *18,* 2146–2152.
[5] Schwiegerling, J. *J. Opt. Soc. Am. A* **2002,** *19,* 1937–1945.
[6] Campbell, C.E. *J. Opt. Soc. Am. A* **2003,** *20,* 209–217.
[7] Dai, G.-M. *J. Opt. Soc. Am. A* **2006,** *23,* 539–543.
[8] Bará, S.; Arines, J.; Ares, J.; Prado, P. *J. Opt. Soc. Am. A* **2006,** *23,* 2061–2066.
[9] Janssen, A.J.E.M.; Dirksen, P. *J. Microlithogr. Microfabr. Microsyst.* **2006,** *5,* 030501, 1–3.
[10] The Extended Nijboer-Zernike (ENZ) analysis website. http:\\www.nijboerzernike.nl (accessed March, 2007).
[11] Janssen, A.J.E.M. *J. Opt. Soc. Am. A* **2002,** *19,* 849–857.
[12] Braat, J.J.M.; Dirksen, P.; Janssen, A.J.E.M. *J. Opt. Soc. Am. A* **2002,** *19,* 858–870.
[13] Braat, J.J.M.; Dirksen, P.; Janssen, A.J.E.M.; van de Nes, A.S. *J. Opt. Soc. Am. A* **2003,** *20,* 2281–2292.
[14] Dirksen, P.; Braat, J.J.M.; Janssen, A.J.E.M.; Juffermans, C. *J. Microlithogr. Microfabr. Microsyst.* **2003,** *2,* 61–68.
[15] Janssen, A.J.E.M.; Braat, J.J.M.; Dirksen, P. *J. Mod. Opt.* **2004,** *51,* 687–703.
[16] van der Avoort, C.; Braat, J.J.M.; Dirksen, P.; Janssen, A.J.E.M. *J. Mod. Opt.* **2005,** *52,* 1695–1728.
[17] Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions*, 9th ed.; Dover: New York, 1970.
[18] Braat, J.J.M.; Dirksen, P.; Janssen, A.J.E.M.; van Haver, S.; van de Nes, A.S. *J. Opt. Soc. Am. A* **2005,** *22,* 2635–2650.
[19] van Haver, S.; Braat, J.J.M.; Dirksen, P.; Janssen, A.J.E.M. *J. Eur. Opt. Soc.-RP* **2006,** *1,* 06004, 1–8.
[20] Boivin, A. *J. Opt. Soc. Am.* **1952,** *42,* 60–64.
[21] Noll, R.J. *J. Opt. Soc. Am.* **1976,** *66,* 207–211.
[22] Nijboer, B.R.A. The Diffraction Theory of Aberrations. Ph.D. Thesis, University of Groningen, Groningen, The Netherlands, 1942. http:\\www.nijboerzernike.nl (accessed March, 2007).
[23] Wang, J.Y.; Silva, D.E. *Appl. Opt.* **1980,** *19,* 1510–1518.
[24] Tatian, B. *J. Opt. Soc. Am.* **1974,** *64,* 1083–1091.
[25] Tricomi, F.G. *Vorlesungen über Orthogonalreihen*; Springer: Berlin, 1955.

Appendix 1

A1. Proof of the mathematical properties in Section 2

A1.1 Proof of and comments on the basic results

We start by showing the result in Equation (19). To that end we recall the representation (14),

$$M^m_{nn'}(\varepsilon) = \int_0^1 R^m_n(\varepsilon\rho) R^m_n(\rho)\rho \, \mathrm{d}\rho, \quad n, n' = m, m+2, \ldots, \tag{99}$$

and use the result

$$R_l^k(\rho) = (-1)^{(l-k)/2} \int_0^\infty J_{l+1}(r) J_k(\rho r)\, \mathrm{d}r, \quad 0 \le \rho < 1, \tag{100}$$

that we shall discuss in more detail below. For now it is relevant to note that (100) holds for integer k, $l \ge 0$ with same parity, with the right-hand side of (100) vanishing when $l < k$. Using (100) in (99) with $k = m$, $l = n'$ and $\varepsilon\rho$ instead of ρ, we get by interchanging integrals

$$\begin{aligned} M_{nn'}^m(\varepsilon) &= (-1)^{(n'-m)/2} \int_0^1 R_n^m(\rho) \left(\int_0^\infty J_{n'+1}(r) J_m(\varepsilon\rho r)\, \mathrm{d}r \right) \rho\, \mathrm{d}\rho \\ &= (-1)^{(n'-m)/2} \int_0^\infty J_{n'+1}(r) \left(\int_0^1 R_n^m(\rho) J_m(\rho\varepsilon r) \rho\, \mathrm{d}\rho \right) \mathrm{d}r. \end{aligned} \tag{101}$$

We next use the basic result

$$\int_0^1 R_n^m(\rho) J_m(\rho v) \rho\, \mathrm{d}\rho = (-1)^{(n-m)/2} \frac{J_{n+1}(v)}{v} \tag{102}$$

from the classical Nijboer–Zernike theory, see [3], formula (39) on p. 910. There results

$$M_{nn'}^m(\varepsilon) = (-1)^{(n'+n-2m)/2} \int_0^\infty \frac{J_{n'+1}(r) J_{n+1}(\varepsilon r)}{\varepsilon r}\, \mathrm{d}r. \tag{103}$$

We then apply the result

$$\frac{J_\nu(z)}{z} = \frac{1}{2\nu} (J_{\nu-1}(z) + J_{\nu+1}(z)), \tag{104}$$

see [17], 9.1.27 on p. 361, to rewrite $J_{n+1}(\varepsilon r)/\varepsilon r$ and $J_{n'+1}(r)/r$ for either instance of Equation (19). Thus, $M_{nn'}^m(\varepsilon)$ is written in two ways as a sum containing two terms of the type that occurs at the right-hand side of (100) (with ε instead of ρ). This then yields both statements in Equation (19).

We make some comments on the result (100). It is often attributed to Noll, see [21], formula (9) on p. 207, but it can already be found in Nijboer's thesis [22], formula (2.22) on p. 26, where it is discussed in connection with the proof of the basic result (102), see [22], (2.20) on p. 25. The right-hand side of (100) vanishes when $l - k < 0$, which is consistent with the convention that $R_l^k(\rho) \equiv 0$ in that case. The result (100) is in fact a special case of the Weber–Schafheitlin formula, see [17], 11.4.33 on p. 487, for the integral

$$\int_0^\infty \frac{J_\mu(at) J_\nu(bt)}{t^\lambda}\, \mathrm{d}t, \quad 0 \le b < a, \tag{105}$$

that is expressed in terms of the hypergeometric function ${}_2F_1$. In the case of formula (100) we have

$$\mu = l + 1, \quad \nu = k, \quad a = 1, \quad b = \varepsilon, \quad \lambda = 0, \tag{106}$$

and the ${}_2F_1$ reduces to a terminating series (polynomial in ε^2). This series can then be expressed, by [17], 15.4.6 on p. 561, in terms of Jacobi polynomials ($\alpha = 0$, $\beta = k$, degree $= 1/2(l - k)$ and thus vanishing when $l - k < 0$, argument $2\varepsilon^2 - 1$), and then one obtains the result (100) by the relation

$$R_l^k(\rho) = \rho^k P_{(l-k)/2}^{(0,k)}(2\rho^2 - 1). \tag{107}$$

As a consequence of Equation (19) we have that

$$\rho R_n^m(\rho) - \rho R_n^{m+2}(\rho) = \frac{m+1}{n+1} (R_{n-1}^{m+1}(\rho) - R_{n+1}^{m+1}(\rho)), \tag{108a}$$

which is reminiscent of Noll's result, see [21], formula (13) on p. 208,

$$(R_{n+1}^{m+1})'(\rho) - (R_{n-1}^{m+1})'(\rho) = (n+1)(R_n^m(\rho) + R_n^{m+2}(\rho)) \tag{108b}$$

that we shall use in the sequel.

We furthermore see that m has disappeared altogether from the two right-hand side members of Equation (19). As a consequence we have for all integer n, n', $m \geq 0$ of same parity and such that $n' \geq n \geq m$ that

$$M_{nn'}^m(\varepsilon) = \int_0^1 \rho^n R_{n'}^n(\varepsilon\rho)\rho \, \mathrm{d}\rho. \tag{109}$$

We shall next show Dai's formula

$$M_{nn'}^m(\varepsilon) = \begin{cases} 0, & n' - n < 0, \qquad (110a) \\ \dfrac{1}{2}\varepsilon^n \displaystyle\sum_{j=0}^{i} \frac{(-1)^{i-j}(n+i-j)!\varepsilon^{2j}}{(n+j+1)!(i-j)!j!}, & i = \dfrac{n'-n}{2} \geq 0, \qquad (110b) \end{cases}$$

when the integers n, n', $m \geq 0$, of same parity and $n \geq m$, $n' \geq m$. This result was given in [7], formula (19), but the proof given in [7] is somewhat cumbersome. Of course, (110) follows directly from either instance of Equation (19) and the explicit expression

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(n-s)!(-1)^s \rho^{n-2s}}{((n-m)/2-s)!((n+m)/2-s)!} \tag{111}$$

for the Zernike polynomials. An alternative proof of Equation (110) is based upon the representation (99) of $M_{nn'}^m(\varepsilon)$, the explicit formula (111) used with $R_{n'}^m(\varepsilon\rho)$ instead of $R_n^m(\rho)$, the result

$$\int_0^1 \rho^\alpha R_n^m(\rho)\rho \, \mathrm{d}\rho = (-1)^{(n-m)/2} \frac{(m-\alpha)(m-\alpha+2)\cdots(n-\alpha-2)}{(m+\alpha+2)(m+\alpha+4)\cdots(n+\alpha+2)} \tag{112}$$

and some administration involving factorials. The result (112) was presented and proved in [13], Appendix A; it also occurs implicitly in Nijboer's thesis [22], p. 25, where it is used to prove the basic result (102).

We next show the results (22a) and (22b). The result in (22a) follows immediately from the first identity in Equation (19) and the fact that $R_n^n(\varepsilon) = \varepsilon^n$, $R_n^{n+2}(\varepsilon) = 0$. To show the result (22b), we start from (107) and we consider the second identity in Equation (19). Now using the contiguity property, see [17], 22.7.15 on p. 782,

$$P_l^{(0,\beta)}(x) - P_{l+1}^{(0,\beta)}(x) = \frac{1}{2}(1-x)(2l+\beta+1)P_l^{(1,\beta)}(x), \tag{113}$$

we obtain (22b).

The other properties noted in Subsection 2.1 are immediate consequences of the orthogonality relation

$$\int_0^1 R_n^m(\rho)R_{n'}^m(\rho)\rho d\rho = \frac{\delta_{nn'}}{2(n+1)}, \tag{114}$$

where $\delta_{nn'}$ is Kronecker's delta.

A1.2 Proof of the results on derivatives

A1.2.1 Derivatives of Zernike coefficients

We shall first show the result (26) on $(\beta_n^m)'$. To that end we start with Equation (21) so that

$$\beta_n^m(\varepsilon) = \sum_{n'} \frac{n+1}{n'+1} \frac{1}{\varepsilon} (R_{n'+1}^{n+1}(\varepsilon) - R_{n'-1}^{n+1}(\varepsilon)) \beta_{n'}^m. \tag{115}$$

Now

$$\begin{aligned} \left(\frac{1}{\varepsilon} (R_{n'+1}^{n+1}(\varepsilon) - R_{n'-1}^{n+1}(\varepsilon)) \right)' &= \frac{-1}{\varepsilon^2} (R_{n'+1}^{n+1}(\varepsilon) - R_{n'-1}^{n+1}(\varepsilon)) + \frac{1}{\varepsilon} ((R_{n'+1}^{n+1})'(\varepsilon) - (R_{n'-1}^{n+1})'(\varepsilon)) \\ &= \frac{-1}{\varepsilon} \frac{n'+1}{n+1} (R_{n'}^n(\varepsilon) - R_{n'}^{n+2}(\varepsilon)) + \frac{1}{\varepsilon} ((R_{n'+1}^{n+1})'(\varepsilon) - (R_{n'-1}^{n+1})'(\varepsilon)), \end{aligned} \tag{116}$$

where the second identity in Equation (19) has been used. We next use Noll's result in (108b), and we get

$$\begin{aligned} \left(\frac{1}{\varepsilon} (R_{n'+1}^{n+1}(\varepsilon) - R_{n'-1}^{n+1}(\varepsilon)) \right)' &= \frac{-1}{\varepsilon} \frac{n'+1}{n+1} (R_{n'}^n(\varepsilon) - R_{n'}^{n+2}(\varepsilon)) + \frac{1}{\varepsilon} (n'+1)(R_{n'}^n(\varepsilon) + R_{n'}^{n+2}(\varepsilon)) \\ &= \frac{1}{\varepsilon} \frac{n'+1}{n+1} (n R_{n'}^n(\varepsilon) + (n+2) R_{n'}^{n+2}(\varepsilon)). \end{aligned} \tag{117}$$

Inserting the latter result into (115) we get Equation (26).

Equation (27) is an immediate consequence from Equation (26) and the fact that $R_n^{n+2}(1) = 0$, $R_{n'}^n(1) = R_{n'}^{n+2}(1) = 1$, $n' = n+2, n+4, \ldots$.

We next show Equation (28). To that end we start with Equation (18),

$$\beta_n^m(\varepsilon) = \sum_{n'} (R_{n'}^n(\varepsilon) - R_{n'}^{n+2}(\varepsilon)) \beta_{n'}^m, \tag{118}$$

and we use [22], formula (2.11) on p. 21,

$$\rho(1-\rho^2)(R_n^m)''(\rho) + (1-3\rho^2)(R_n^m)'(\rho) + \left(n(n+2) - \frac{m^2}{\rho} \right) R_n^m(\rho) = 0. \tag{119}$$

Thus, we get

$$\begin{aligned} \varepsilon(1-\varepsilon^2)[(R_{n'}^n)''(\varepsilon) - (R_{n'}^{n+2})''(\varepsilon)] &= (3\varepsilon^2 - 1)[(R_{n'}^n)'(\varepsilon) - (R_{n'}^{n+2})'(\varepsilon)] \\ &+ \left(\frac{n^2}{\varepsilon} - n'(n'+2)\varepsilon \right) R_{n'}^n(\varepsilon) - \left(\frac{(n+2)^2}{\varepsilon} - n'(n'+2) R_{n'}^{n+2}(\varepsilon) \right). \end{aligned} \tag{120}$$

Now by (108a) and (120)

$$\begin{aligned} (R_{n'}^n)'(\varepsilon) - (R_{n'}^{n+2})'(\varepsilon) &= \frac{n+1}{n'+1} \left(\frac{1}{\varepsilon} (R_{n'-1}^{n+1}(\varepsilon) - R_{n'+1}^{n+1}(\varepsilon)) \right)' \\ &= \frac{1}{\varepsilon} (n R_{n'}^n(\varepsilon) + (n+2) R_{n'}^{n+2}(\varepsilon)). \end{aligned} \tag{121}$$

This then yields

$$\varepsilon(1-\varepsilon^2)[(R^n_{n'})''(\varepsilon)-(R^{n+2}_{n'})''(\varepsilon)]=\frac{1}{\varepsilon}(n(n-1)R^n_{n'}(\varepsilon)-(n+2)(n+3)R^{n+2}_{n'}(\varepsilon))$$
$$+\varepsilon((3n-n'(n'+2))R^n_{n'}(\varepsilon)+(3(n+2)$$
$$+n'(n'+2))R^{n+2}_{n'}(\varepsilon)), \tag{122}$$

and from this Equation (28) follows.

We finally verify Equation (29). To that end we use Equation (26) which we differentiate once more and obtain

$$(\beta^m_n)''(\varepsilon)=\left(\frac{n}{\varepsilon}\beta^m_n R^n_{n'}(\varepsilon)+\frac{1}{\varepsilon}\sum_{n'}(nR^n_{n'}(\varepsilon)+(n+2)R^{n+2}_{n'}(\varepsilon))\beta^m_{n'}\right)', \tag{123}$$

where the summation is over $n'=n+2, n+4, \ldots$. Carrying through the differentiation in (123) while using, see (119),

$$R^m_n(1)=1, \quad (R^m_n)'(1)=\frac{1}{2}(n(n+2)-m^2), \tag{124}$$

we get

$$(\beta^m_n)''(1)=n(n-1)\beta^m_n+\sum_{n'}[-(2n+2)+\frac{1}{2}(n'(n'+2)-n^2)$$
$$+\frac{1}{2}(n+2)(n'(n'+2)-(n+2)^2)]\beta^m_{n'}$$
$$=n(n-1)\beta^m_n+\sum_{n'}(n+1)[n'(n'+2)-n(n+2)-6]\beta^m_{n'}, \tag{125}$$

where the summations are over $n'=n+2, n+4, \ldots$. Now writing $n'=n+2k$, $k=1,2,\ldots$ in the last member of (125) we get Equation (29).

A1.2.2 Derivatives of Strehl ratios

We consider the approximating $S(\alpha)$ as given in Equation (32),

$$S(\alpha)=1-\frac{1}{\pi}\int_0^1\int_0^{2\pi}\left|\sum_{n,m}^{\sim}\alpha^m_n Z^m_n(\rho,\vartheta)\right|^2\rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta$$
$$=1-\sum_{n,m}^{\sim}\frac{(\alpha^m_n)^2}{\varepsilon_m(n+1)}, \tag{126}$$

where the $\sim$ on top of the summation signs refers to deletion of the terms in a set

$$\{(n,m)|m=0,1,\ldots; \quad n=m,m+2,\ldots,n(m)-2\}, \tag{127}$$

with integers $n(m)\geq m$ of same parity as m. We shall also write $\Sigma^*_{n,m}$ to indicate summation over the set in (127). Thus, $\sum^{\sim}_{n,m}=\Sigma_{n,m}-\Sigma^*_{n,m}$, etc. With these conventions we have

$$\Phi(\varepsilon\rho,\vartheta)=\sum_{n,m}\alpha^m_n(\varepsilon)Z^m_n(\rho,\vartheta)$$
$$=\sum_{n,m}^{\sim}\alpha^m_n(\varepsilon)Z^m_n(\rho,\vartheta)+\sum_{n,m}^{*}\alpha^m_n(\varepsilon)Z^m_n(\rho,\vartheta). \tag{128}$$

The first term in the last member of (128) is what we have abbreviated in Equation (130) as $\tilde{\Phi}(\rho, \vartheta, \varepsilon)$.

We shall now prove Equation (35). We have from (126) and (128)

$$S(\alpha(\varepsilon)) = 1 - \frac{1}{\pi}\int_0^1\int_0^{2\pi}\left|\Phi(\varepsilon\rho, \vartheta) - \sum_{n,m}^{*}\alpha_n^m(\varepsilon)Z_n^m(\rho, \vartheta)\right|^2 \rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta$$

$$= 1 - \frac{1}{\pi\varepsilon^2}\int_0^{\varepsilon}\int_0^{2\pi}\left|\Phi(\rho, \vartheta) - \sum_{n,m}^{*}\alpha_n^m(\varepsilon)Z_n^m(\rho\varepsilon^{-1}, \vartheta)\right|^2 \rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta. \tag{129}$$

Then we get upon differentiating

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}(S(\alpha(\varepsilon))) = \frac{2}{\pi\varepsilon^3}\int_0^{\varepsilon}\int_0^{2\pi}\left|\Phi(\rho, \vartheta) - \sum_{n,m}^{*}\alpha_n^m(\varepsilon)Z_n^m(\rho\varepsilon^{-1}, \vartheta)\right|^2 \rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta$$

$$- \frac{1}{\pi\varepsilon^2}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\int_0^{\varepsilon}\int_0^{2\pi}\left|\Phi(\rho, \vartheta) - \sum_{n,m}^{*}\alpha_n^m(\varepsilon)Z_n^m(\rho\varepsilon^{-1}, \vartheta)\right|^2 \rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta\right]. \tag{130}$$

The first term in the second member of (130) equals, see (129),

$$\frac{2}{\pi\varepsilon}\int_0^1\int_0^{2\pi}\left|\Phi(\varepsilon\rho, \vartheta) - \sum_{n,m}^{*}\alpha_n^m(\varepsilon)Z_n^m(\rho, \vartheta)\right|^2 \rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta = \frac{2}{\varepsilon}(1 - S(\alpha(\varepsilon))), \tag{131}$$

and so, see Equations (34) and (128), there only remains to be considered the second term in the second member of (130). For this second term we use the result

$$\frac{\mathrm{d}}{\mathrm{d}y}\left[\int_0^y f(x, y)\,\mathrm{d}x\right] = f(y, y) + \int_0^y \frac{\partial f}{\partial y}(x, y)\mathrm{d}x, \tag{132}$$

and we get

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\int_0^{\varepsilon}\int_0^{2\pi}\left|\Phi(\rho, \vartheta) - \sum_{n,m}^{*}\alpha_n^m(\varepsilon)Z_n^m(\rho\varepsilon^{-1}, \vartheta)\right|^2 \rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta\right]$$

$$= \varepsilon\int_0^{2\pi}\left|\Phi(\varepsilon, \vartheta) - \sum_{n,m}^{*}\alpha_n^m(\varepsilon)Z_n^m(1, \vartheta)\right|^2 \mathrm{d}\vartheta - 2\int_0^{\varepsilon}\int_0^{2\pi}\left(\Phi(\rho, \vartheta) - \sum_{n,m}^{*}\alpha_n^m(\varepsilon)Z_n^m(\rho\varepsilon^{-1}, \vartheta)\right)$$

$$\times \sum_{n,m}^{*}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}[\alpha_n^m(\varepsilon)Z_n^m(\rho\varepsilon^{-1}, \vartheta)]\rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta. \tag{133}$$

From (128) and Equation (34) it is seen that the first term in the second member of (133) yields the second term in the second member of (35) when inserted into (130). Thus, it is enough to show that the second term in the second member of (133) vanishes altogether. To that end we compute first

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}[\alpha_n^m(\varepsilon)Z_n^m(\rho\varepsilon^{-1}, \vartheta)] = (\alpha_n^m)'(\varepsilon)Z_n^m(\rho\varepsilon^{-1}, \vartheta) - \varepsilon^{-2}\alpha_n^m(\varepsilon)\rho\left(\frac{\partial}{\partial\rho}Z_n^m\right)(\rho\varepsilon^{-1}, \vartheta), \tag{134}$$

insert this into the term under consideration, change variables in which ρ is replaced by $\varepsilon\rho$ and use (128) to get

$$-2\int_0^1\int_0^{2\pi}\sum_{n,m}^{\sim}\alpha_n^m(\varepsilon)Z_n^m(\rho, \vartheta)\sum_{n,m}^{*}\left[\varepsilon^2(\alpha_n^m)'(\varepsilon)Z_n^m(\rho, \vartheta) - \varepsilon\alpha_n^m(\varepsilon)\rho\left(\frac{\partial}{\partial\rho}Z_n^m\right)(\rho, \vartheta)\right]\rho\,\mathrm{d}\rho\,\mathrm{d}\vartheta \tag{135}$$

for this term in (133). We now note that for each (n, m) in the set (127)

$$\varepsilon^2(\alpha_n^m)'(\varepsilon)Z_n^m(\rho,\vartheta) - \varepsilon\alpha_n^m(\varepsilon)\rho\left(\frac{\partial}{\partial\rho}Z_n^m\right)(\rho,\vartheta) \tag{136}$$

is a linear combination of Zernike functions $Z_{n'}^m(\rho,\vartheta)$ with $n' = m, m+2, \ldots, n(m)-2$. Indeed, since $Z_n^m(\rho,\vartheta) = R_n^m(\rho)\cos m\vartheta$, we can write (136) as a polynomial in ρ with terms $\rho^m, \rho^{m+2}, \ldots, \rho^n$ times $\cos m\vartheta$. Now $n \le n(m)-2$, and the $R_{n'}^m(\rho)$, $n' = m,\ m+2, \ldots, n(m)-2$, span the space of linear combinations of $\rho^m, \rho^{m+2}, \ldots, \rho^{n(m)-2}$, from which the claim follows. Therefore, all terms occurring in the $\sum^*$ in (135) are orthogonal to all terms in the $\tilde{\sum}$, and thus (135) vanishes. This yields the result (35) that can also be written in the form

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}(S(\alpha(\varepsilon))) = \frac{2}{\varepsilon}(1 - S(\alpha(\varepsilon))) - \frac{1}{\pi\varepsilon}\int_0^{2\pi}|\tilde{\Phi}(1,\vartheta,\varepsilon)|^2\,\mathrm{d}\vartheta. \tag{137}$$

The result (36) is an immediate consequence of (137) and (34), (128) and (131).

We shall now establish the results (40) and (41). The proof of (40) simply consists of using (126) with $\alpha_n^m(\varepsilon)$ instead of α_n^m in (137) together with the fact that

$$Z_n^m(1,\vartheta) = \cos m\vartheta, \quad m = 0, 1, \ldots, \quad n = m, m+2, \ldots \tag{138}$$

and Parseval's theorem. As to the result (41), we find in a similar fashion from Equation (36) that

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right)^2(S(\alpha(\varepsilon))) = &-\frac{6}{\varepsilon^2}\left[\tilde{\sum_{n,m}}\frac{(\alpha_n^m(\varepsilon))^2}{\varepsilon_m(n+1)} - \sum_m\frac{1}{\varepsilon_m}\left(\tilde{\sum_n}\alpha_n^m(\varepsilon)\right)^2\right]\\ &-\frac{2}{\pi\varepsilon}\int_0^{2\pi}\left(\tilde{\sum_{n,m}}\alpha_n^m(\varepsilon)\cos m\vartheta\right)\left(\tilde{\sum_{n,m}}(\alpha_n^m)'(\varepsilon)\cos m\vartheta\right)\mathrm{d}\vartheta,\end{aligned} \tag{139}$$

and the last expression of the second member in (139) can easily be seen to be the same as the last expression of the second member in Equation (41) by Parseval's theorem.

The special cases $\varepsilon = 1$ in Equations (42) and (43) follow from $\alpha_n^m(1) = \alpha_n^m$ and from, see Equation (27),

$$(\alpha_n^m)'(1) = n\alpha_n^m + 2(n+1)[\alpha_{n+2}^m + \alpha_{n+4}^m + \cdots], \tag{140}$$

and some administration required to rewrite

$$\sum_n(\alpha_n^m)'(1) \quad \text{as} \quad \sum_n\alpha_n^m[n + \frac{1}{2}(n^2 - n^2(m))]. \tag{141}$$

Appendix 2

A2. Zernike expansion of pupil functions on an annulus

Assume that we have a smooth pupil function $P(\rho,\vartheta)$ defined on an annulus $\varepsilon \le \rho \le 1$, $0 \le \vartheta \le 2\pi$. Such a P can be represented on the annulus as an infinite Zernike series in many ways. Indeed, we can interpolate P into the disk $0 \le \rho < \varepsilon$, $0 \le \vartheta \le 2\pi$, in one way or another, and expand the completed P as a Zernike series using that the Z_n^m form a complete orthogonal system for functions defined on the unit disk. Thus, the linear independence of the infinite set of Z_n^m's restricted to the annulus is an issue. However, any finite set of these Z_n^m's is linearly independent on $\varepsilon \le \rho \le 1$, $0 \le \vartheta \le 2\pi$. Hence, for any finite index set I of (n, m), there are unique

coefficients β_n^m, $(n,m) \in I$, such that

$$\int_\varepsilon^1 \int_0^{2\pi} \left| P(\rho, \vartheta) - \sum_{(n,m)\in I} \beta_n^m Z_n^m(\rho, \vartheta) \right|^2 \rho \, \mathrm{d}\rho \, \mathrm{d}\vartheta \tag{142}$$

is minimal.

We shall describe the actual computation of the optimal β_n^m in (142) for the case that the index set I is of the form

$$I = \{(n,m) | m = 0, 1, \ldots, M; \quad n = m, m+2, \ldots, N(m)\}, \tag{143}$$

where the $N(m)$ are integers $\geq m$ having the same parity as m. Let $m = 0, 1, \ldots,$ and set

$$P^m(\rho) = \frac{1}{2\pi} \int_0^{2\pi} P(\rho, \vartheta) \cos m\vartheta \, \mathrm{d}\vartheta, \quad \varepsilon \leq \rho \leq 1, \tag{144}$$

the mth azimuthal average of P. Thus, we have

$$P(\rho, \vartheta) = \sum_{m=0}^{\infty} \varepsilon_m P^m(\rho) \cos m\vartheta, \quad \varepsilon \leq \rho \leq 1, \quad 0 \leq \vartheta \leq 2\pi. \tag{145}$$

Recall that $Z_n^m(\rho, \vartheta) = R_n^m(\rho) \cos m\vartheta$. Hence, for fixed $m = 0, 1, \ldots, M$, the optimal β_n^m, $n = m$, $m+2, \ldots, N(m)$, minimize

$$\int_\varepsilon^1 \left| \varepsilon_m P^m(\rho) - \sum_n \beta_n^m R_n^m(\rho) \right|^2 \rho \, \mathrm{d}\rho, \tag{146}$$

where the summation is over $n = m, m+2, \ldots, N(m)$. The minimum is assumed by

$$[\beta_m^m, \beta_{m+2}^m, \ldots, \beta_{N(m)}^m]^{\mathrm{T}} = \varepsilon_m (\mathcal{M}^m(\varepsilon))^{-1} p^m(\varepsilon), \tag{147}$$

where $p^m(\varepsilon) = [p_m^m(\varepsilon), p_{m+2}^m(\varepsilon), \ldots, p_{N(m)}^m]^{\mathrm{T}}$ and

$$p_{n'}^m(\varepsilon) = \int_\varepsilon^1 P^m(\rho) R_{n'}^m(\rho) \rho \, \mathrm{d}\rho, \quad n' = m, m+2, \ldots, N(m), \tag{148}$$

and

$$\mathcal{M}^m(\varepsilon) = \left(\int_\varepsilon^1 R_n^m(\rho) R_{n'}^m(\rho) \rho \, \mathrm{d}\rho \right)_{n,n' = m, m+2, \ldots, N(m)}. \tag{149}$$

It is a consequence of our results on scaled Zernike expansions that the matrix elements $\mathcal{M}_{nn'}^m(\varepsilon)$ can be found explicitly. We have by (114) that

$$\begin{aligned} \mathcal{M}_{nn'}^m(\varepsilon) &= \frac{\delta_{nn'}}{2(n+1)} - \int_0^\varepsilon R_n^m(\rho) R_{n'}^m(\rho) \rho \mathrm{d}\, \rho \\ &= \frac{\delta_{nn'}}{2(n+1)} - \varepsilon^2 \int_0^1 R_n^m(\varepsilon\rho) R_{n'}^m(\varepsilon\rho) \rho \, \mathrm{d}\rho \\ &= \frac{\delta_{nn'}}{2(n+1)} - \varepsilon^2 \sum_{n''} \frac{(R_n^{n''}(\varepsilon) - R_n^{n''+2}(\varepsilon))(R_{n'}^{n''}(\varepsilon) - R_{n'}^{n''+2}(\varepsilon))}{2(n''+1)}. \end{aligned} \tag{150}$$

For the last identity in (150) we have used Equation (25), which gives the Zernike^m expansions of $R_n^m(\varepsilon\rho)$ and $R_{n'}^m(\varepsilon\rho)$, and (114) once more. Note that the summation in the series in the last member of (150) is over $n'' = m, m+2, \ldots, \min(n, n')$. Observe also that m has disappeared. This latter fact has

advantages in the case that all $N(m) = M$ are equal, so that all $\mathcal{M}^m(\varepsilon)$ are bordered or extended versions of one another, allowing recursive computation of $(\mathcal{M}^m(\varepsilon))^{-1}$, $m = M,\ M-2, \ldots$ and $m = M-1,\ M-3, \ldots$. The matrices $\mathcal{M}^m(\varepsilon)$ also occur when performing Gram–Schmidt orthogonalization of the Zernike polynomials $R_n^m(\rho)$, $n = m,\ m+2, \ldots$ on $[\varepsilon, 1]$ with weight $\rho\,\mathrm{d}\rho$, see [23], that lead to the orthogonal annulus polynomials of Tatian, see [24].

We have to address the issue of well-conditionedness of the matrices $\mathcal{M}^m(\varepsilon)$ that have to be inverted per Equation (147). To that end we consider the normalized system

$$\mathcal{N}_n^m(\rho) := \frac{R_n^m(\rho)}{\|R_n^m\|_\varepsilon}, \quad n = m, m+2, \ldots, N(m), \tag{151}$$

where $\|R_n^m\|_\varepsilon = (\int_\varepsilon^1 |R_n^m(\rho)|^2 \rho\ \mathrm{d}\rho)^{1/2}$. A preliminary experiment, with $\varepsilon = \frac{1}{2}$, $m = 0$ and $N(0) = 24$, has shown that for any $a_0, a_2, \ldots, a_{24}$ we have

$$\left\| \sum_n a_n \mathcal{N}_n^0(\rho) \right\|_\varepsilon \geq c \left(\sum_n |a_n|^2 \right)^{1/2}, \tag{152}$$

where c is of the order 0.01. Since the maximum value of the lower index n in Zernike expansions rarely exceeds 10, it appears that the inversion of the $\mathcal{M}^m(\varepsilon)$ in the relevant cases present no problems.

Appendix 3

A3. Zernike expansion of hard-thresholded pupil

We consider the Zernike expansion of the pupil function $P \equiv 1$, hard thresholded on the annulus $\varepsilon \leq \rho < 1$ to 0, so that

$$P_\varepsilon(\rho, \vartheta) = \begin{cases} 1, & 0 \leq \rho < \varepsilon\ ,\ 0 \leq \vartheta \leq 2\pi, \quad (153a) \\ 0, & \varepsilon \leq \rho \leq 1\ ,\ 0 \leq \vartheta \leq 2\pi \quad (153b) \end{cases}.$$

There holds

$$P_\varepsilon(\rho, \vartheta) = \sum_{n=0}^{\infty} \beta_{2n,\varepsilon}^0 R_{2n}^0(\rho), \quad 0 \leq \rho \leq 1, \quad 0 \leq \vartheta \leq 2\pi, \tag{154}$$

where $\beta_{0,\varepsilon}^0 = \varepsilon^2$, and, for $n = 1, 2, \ldots$,

$$\beta_{2n,\varepsilon}^0 = 2(2n+1) \int_0^\varepsilon R_{2n}^0(\rho) \rho\ \mathrm{d}\rho = \frac{1}{2}(R_{2n+2}^0(\varepsilon) - R_{2n-2}^0(\varepsilon)). \tag{155}$$

Here [25], item (10.10) on p. 190 has been used. Now, with $\varepsilon = \cos x$ and $x \in (0, \pi/2)$, and P_p the Legendre polynomials, we have

$$\begin{aligned} R_{2p}^0(\varepsilon) &= P_p(2\varepsilon^2 - 1) = P_p(\cos\ 2x) \\ &= \left(\frac{2}{\pi p \sin 2x} \right)^{1/2} \cos\left((2p+1)x - \frac{1}{4}\pi \right) + O(p^{-3/2}), \quad p \to \infty, \end{aligned} \tag{156}$$

where we have used [25], item (10.22) on p. 194. Then it follows from (155) and (156) that

$$\beta_{2n,\varepsilon}^0 = -2 \left(\frac{\sin\ 2x}{2\pi n} \right)^{1/2} \sin\left((2n+1)x - \frac{1}{4}\pi \right) + O(n^{-3/2}), \quad n \to \infty. \tag{157}$$