Concise formula for the Zernike coefficients of scaled pupils

Augustus J. E. M. Janssen\textsuperscript{a} and Peter Dirksen\textsuperscript{b}
\textsuperscript{a}Philips Research Europe, HTC 36, 5656 AE Eindhoven, The Netherlands
\textsuperscript{b}Philips Research Europe, HTC 4, 5656 AE Eindhoven, The Netherlands
E-mail: peter.dirksen@philips.com

Abstract. Modern steppers and scanners have a projection lens whose numerical aperture (NA) can be varied so as to optimize the image performance for certain lithographic features. Thus a variable fraction of the aberrations is actually involved in the imaging process. In this letter, we present a concise formula for the NA scaling of the Zernike coefficients. In addition, we apply our results to the Strehl ratio.

1 Introduction

It is well known that different lithographic features have different optimal settings of the numerical aperture (NA) and illumination condition.\textsuperscript{1,2} Typically the depth of focus or the process latitudes are optimized. To achieve the optimal settings, modern steppers and scanners have a projection lens with a variable numerical aperture, where the NA value can be reduced to about 70% of its maximum value. As a consequence, a smaller fraction of the lens pupil and a corresponding smaller fraction of the phase aberrations is actually involved in the imaging process.

In the literature (see Ref. 3 for a survey) the problem of computing the aberration Zernike coefficients of scaled pupils has been discussed at various places. The recent result of Dai\textsuperscript{3} gives the Zernike coefficients of the scaled pupil in terms of the unscaled pupil in analytic form [see Eq. (18)].

In this letter we give an alternative expression for Dai’s result. Our main result in Eq. (4) agrees mathematically with Dai’s Eq. (18) but has the advantage that it is very simple and direct in terms of Zernike polynomials. Thus it leads to the simple and explicit results in Eqs. (7) and (9) from which the sensitivity of the aberration coefficients and Strehl ratios can be assessed when NA is close to its maximum value. The short and elegant proof of our main result in Eq. (4) is given in Appendix A.

2 Zernike Coefficients of Scaled Pupils

We consider a pupil function

\[ P(\rho, \theta) = \exp(i\Phi(m, \theta)), \]

de on a unit disk \(0 \leq \rho \leq 1\) with real phase \(\Phi\), and we assume that \(\Phi\) is expanded as a Zernike series according to:

\[ \Phi(m, \theta) = \sum_{n,m} a_n^m R_n^m(\rho) \cos(m\theta), \]

where \(m\) and \(n\) are integers with \(n-m\) being even and \(n \geq 0\), and \(R_n^m(\rho)\) denotes the Zernike radial polynomial\textsuperscript{4} of azimuthal order \(m\) and degree \(n\). For simplicity we only consider cos terms. Scaling to a smaller pupil with relative size \(\epsilon=\text{NA}/\text{NA}_{\max}\leq 1\) means that we have to expand the scaled phase \(\Phi(\rho, \theta)\) into a Zernike expansion:

\[ \Phi(\rho, \theta) = \sum_{n,m} a_n^m(\epsilon) R_n^m(\rho) \cos(m\theta), \]

\[ 0 \leq \rho \leq 1, \quad 0 \leq \theta \leq 2\pi. \]

The problem is how to express the Zernike coefficients \(a_n^m(\epsilon)\) of the scaled pupil function in terms of the Zernike coefficients \(a_n^m\) of the unscaled pupil function. In the Appendix A we prove our main result:

\[ a_n^m(\epsilon) = \sum_{n'} a_n^{m}(R_{n'}^m(\epsilon) - R_{n'}^{m+2}(\epsilon)], \quad n = m + 2, \ldots, \]

where the summation is over \(n'=n,n+2,\ldots\), and we use the convention that \(R_{n'}^{m+2}=0\). We note that (1) the entries \(R_{n'}^m(\epsilon) - R_{n'}^{m+2}(\epsilon)\) do not depend on \(m\); we consider them only for \(n,n'=m,m+2,\ldots\), and (2) \(a_n^m(\epsilon=1)=a_n^m\) since all involved Zernike polynomials, except \(R_{n'}^{m+2}(\epsilon)=0\), equal 1 at \(\epsilon=1\).

3 Examples and Sensitivity Analysis

Figure 1 shows an example where we scale low-order coma \(a_3^1(\epsilon)\) with \(a_3^1=2\pi(0.016)\) and \(a_3^1=2\pi(0.016)\). For this special case, Eq. (4) and Dai’s Eq. (18) become

\[ a_3^1(\epsilon) = a_3^1 R_3^1(\epsilon) + a_3^1 [R_3^1(\epsilon) - R_3^1(\epsilon)] \]

and

\[ a_3^1(\epsilon) = \epsilon^2 [a_3^1 + 4 a_3^1 (\epsilon^2 -1)], \]

respectively. Both equations give an identical result, as \(R_3^1(\epsilon)=\epsilon^2\) and \(R_3^1(\epsilon)=4\epsilon^2(\epsilon^2 - 1)\).

A consequence of Eq. (4) is that

\[ \frac{\partial a_n^m}{\partial \epsilon}_{\epsilon=1} = n a_n^m + 2(n+1) [a_{n+2}^m + a_{n+4}^m + \ldots]. \]

Equation (7) shows a significant sensitivity of \(a_n^m(\epsilon)\) when \(\epsilon\) is slightly below its nominal value of 1, especially when high-order aberrations are present. In our example we obtain a relative sensitivity \(1/a_n(\partial a_n/\partial \epsilon)\) of 11, indicating that the NA value should be specified to \(10^{-3}\) accuracy when the required coefficient accuracy is 1%.

The Strehl ratio\textsuperscript{4} is approximated as:
omitted. A consequence of Eq. (9) is that

\[
S = 1 - \sum_{n,m} \frac{(a_n^m)^2}{\gamma_m(n+1)}
\]  

(8)

in which \(\gamma_0 = 1, \gamma_1 = \gamma_2 = \ldots = 2\) and the term with \(n=m=0\) is omitted. A consequence of Eq. (7) is that

\[
\frac{\partial S}{\partial \epsilon} \bigg|_{\epsilon=1} = 2\sum_{n,m} \frac{(a_n^m)^2}{\gamma_m(n+1)} - 2 \sum_{m} \frac{1}{\gamma_m} \left( \sum_{n} a_n^m \right)^2
\]

(9)

where the two averaging operations are over the whole disk \(0 \leq \rho \leq 1\) and the rim \(\rho = 1\), respectively. Figure 2 shows the dependence on \(\epsilon\) of the Strehl ratio in the case of a mixture of low- and high-order spherical aberation: \(a_1^0 = 2\pi(0.02)\) and \(a_6^0 = -2\pi(0.02)\), where the scaled Strehl ratio \(S(\epsilon)\) is calculated by combining Eqs. (8) and (4). This result compares well with the numerical calculation obtained from the lithography simulator SOLID-C. Intuitively it is expected that the Strehl ratio increases when the NA is decreased. However, in general this is not true. In the special case shown in Fig. 2, the Strehl ratio decreases when the NA is decreased from its maximum value. This result can be seen as follows: as in our example \(\Sigma_{n} a_n^m = 0\) for each \(m\) value, the phase aberration \(\Phi(1, \theta)\) at the rim of the pupil equals 0. From Eq. (9) it then follows that the slope at \(\epsilon = 1\) is positive.

**Appendix A: Proof of the Main Result, Eq. (4)**

By decoupling per azimuthal order \(m=0,1,\ldots\), and normalization and orthogonality of \(R_m^r(\rho)\), \(n=m,m+2,\ldots\), the Zernike coefficients \(a_n^m(e)\) of \(\Phi(\rho, \theta)\) and the Zernike coefficients \(a_n^m\) of \(\Phi(\rho, \theta)\) are related by

\[
a_n^m(e) = 2(n+1) \sum_{n'} a_{n'}^m M_{n'}^m(e), \quad n=m,m+2,\ldots\n\]

(10)

The summation in Eq. (10) is over \(n'=m,m+2,\ldots\), and

\[
M_{n'}^m(e) = \int_0^1 R_n^m(\rho) R_{n'}^m(\rho e) \rho \, d\rho, \quad n,n'=m,m+2,\ldots
\]

(11)

We shall show that

\[
M_{n'}^m(e) = \frac{1}{2(n+1)} [R_{n'}^m(e) - R_{n'+2}^m(e)];
\]

(12)

in particular, it follows that \(M_{n'}^m(e)=0\) when \(n' < n\) and that \(M_{n'}^m(e)\) does not depend on \(m\), except that in Eq. (10) we only use \(n,n'=m,m+2,\ldots\). For this we use

\[
R_l^k(\rho) = (-1)^{(l-k)/2} \int_0^\infty J_{l+1}(r) J_k(\rho r) \, dr, \quad 0 \leq \rho < 1,
\]

(13)

when \(k, l\) are integers \(\geq 0\) with same parity; in the case of \(k-l>0\), the right-hand side of Eq. (13) vanishes, which is consistent with the convention that then \(R_l^k = 0\). We use Eq. (13) in Eq. (11) to rewrite \(R_{n'}^m(\rho e)\) and interchange integrals to get

\[
M_{n'}^m(e) = (-1)^{(n'-m)/2} \int_0^\infty J_{n'+1}(r)
\]

\[
\times \left[ \int_0^1 R_n^m(\rho) J_m(\rho e) \rho \, d\rho \right] dr.
\]

(14)

To the inner integral we apply the result

\[
\int_0^1 R_n^m(\rho) J_m(\rho v) \rho \, d\rho = (-1)^{(n-m)/2} \left[ J_{n+1}(v) \right] / v
\]

(15)

from the Nijboer-Zernike theory, and we get...
\[ M_{mn}^{\prime}(e) = (-1)^{(\alpha^2 + \pi - 2\pi n)/2} \int_0^\infty \frac{J_{n+1}(r) J_n(r)}{er} \, dr. \quad (16) \]

Next we use the identity\(^7\)

\[ \frac{J_{n+1}(er)}{er} = \frac{J_n(er) + J_{n+2}(er)}{2(n+1)}, \quad (17) \]

and use Eq. (13) to rewrite the resulting two integrals in terms of Zernike polynomials. This gives Eq. (12).

**Appendix B: Dai’s Formula**

We reproduce Dai’s formula\(^3\) for the scaling of Zernike coefficients:

\[ \alpha_n^m(e) = \epsilon^2 \left[ \alpha_n^m + (n + 1) \right] \times \sum_{i=1}^{(N-n)/2} \alpha_n^{m+2i} \sum_{j=0}^i \frac{(-1)^i (n+i+j)!}{(n+j+1)!(i-j)!j!} \epsilon^{2j} \]

with \(N\) being the maximum \(n\) value.

**References**

5. SOLID-C (release 6.3.0), SIGMA-C GmbH, Thomas-Dehlerstrasse 9, D-81737 Munich, Germany.