Concise formula for the Zernike coefficients of scaled pupils

Augustus J. E. M. Janssena and Peter Dirksenb

aPhilips Research Europe, HTC 36, 5656 AE Eindhoven, The Netherlands
bPhilips 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.

2 Zernike Coefficients of Scaled Pupils

We consider a pupil function

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

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:

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

where \(m\) and \(n\) are integers with \(n-m\) being even and \(\geq 0\), and \(R_n^m(\rho)\) denotes the Zernike radial polynomial of azimuthal order \(m\) and degree \(n\). For simplicity we only consider cosin terms. Scaling to a smaller pupil with relative size \(\varepsilon\) means that we have to expand the scaled phase \(\Phi(\varepsilon\rho, \theta)\) into a Zernike expansion:

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

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

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

\[ \alpha_n^m(\varepsilon) = \sum_{n'} \alpha_{n'}^m [R_n^m(\varepsilon) - R_{n'}^{m+2}(\varepsilon)], \quad n = m, m+2, \ldots, \]

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

3 Examples and Sensitivity Analysis

Figure 1 shows an example where we scale low-order coma \(\alpha_3^1(\varepsilon)\) with \(\alpha_3^1(\varepsilon=2\pi(0.016))\) and \(\alpha_3^1(\varepsilon=2\pi(0.016))\). For this special case, Eq. (4) and Dai’s Eq. (18) become

\[ \alpha_3^1(\varepsilon) = \alpha_3^1(\varepsilon) R_3^1(\varepsilon) + \alpha_3^1(\varepsilon) R_3^3(\varepsilon) \]

and

\[ \alpha_3^1(\varepsilon) = \varepsilon^2 [\alpha_3^1 + 4\alpha_3^1 (\varepsilon^2 - 1)], \]

respectively. Both equations give an identical result, as \(R_3^1(\varepsilon) = \varepsilon^2\) and \(R_3^3(\varepsilon) = 4\varepsilon^4 (\varepsilon^2 - 1)\).

A consequence of Eq. (4) is that

\[ \left. \frac{\partial \alpha_n^m}{\partial \varepsilon} \right|_{\varepsilon=1} = n \alpha_n^m + 2(n+1) \left[ \alpha_{n+2}^m + \alpha_{n+4}^m + \ldots \right]. \]

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

The Strehl ratio \(\Omega^2\) is approximated as:

...
We shall show that
\[ M_m^m(e) = \frac{1}{2(n+1)} [R_n^m(e) - R_{n+2}^m(e)]; \]
in particular, it follows that \( M_m^m(e) = 0 \) when \( n' < n \) and that \( M_m^m(e) \) does not depend on \( m \), except that in Eq. (10) we only use \( n, n' = m, m + 2, \ldots \). For this we use
\[ R_n^0(p) = (-1)^{(n-k)/2} \int_0^\infty J_{n+1}(r) J_k(p r) \, dr, \quad 0 \leq p < 1, \]
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 \( R_n^0 = 0 \). We use Eq. (13) in Eq. (11) to rewrite \( R_n^m(p e) \) and interchange integrals to get
\[ M_m^m(e) = (-1)^{(n'-m)/2} \int_0^\infty J_{n'+1}(r) \times \left[ \int_0^1 R_n^m(p) J_m(p e) \, dp \right] \, dr. \]
To the inner integral we apply the result
\[ \int_0^1 R_n^m(p) J_m(p v) \, dp = (-1)^{(m-n)/2} \left[ J_{n+1}(v) \right] \]
from the Nijboer-Zernike theory, and we get

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

By decoupling per azimuthal order \( m=0, 1, \ldots \), and normalization and orthogonality of \( R_n^m(p), n=m,m+2, \ldots \), the Zernike coefficients \( a^m_n(e) \) of \( \Phi(p e, \theta) \) and the Zernike coefficients \( a^m_n \) of \( \Phi(p, \theta) \) are related by
\[ a^m_n(e) = 2(n+1) \sum_{n'} a^m_{n'} M_{n'}^m(e), \quad n = m, m + 2, \ldots \] (10)
The summation in Eq. (10) is over \( n'=m,m+2,\ldots \), and

\[ S = 1 - \sum_{n,m} \frac{(a^m_n)^2}{\gamma_m(n+1)} \] (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
\[ \frac{dS}{de} = 2 \sum_{n,m} \frac{(a^m_n)^2}{\gamma_m(n+1)} - 2 \sum_{n,m} \frac{1}{\gamma_m} \left( \sum_n a^m_n \right)^2 \]
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 \( e \) of the Strehl ratio in the case of a mixture of low- and high-order spherical aberration: \( a^0_n = 2 \pi(0.02) \) and \( a^1_n = 2 \pi(0.02) \), where the scaled Strehl ratio \( S(e) \) 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 \( \sum_n a^m_n = 0 \) 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 \( e=1 \) is positive.
\[ M_{mn}^n(e) = (-1)^{(n'-n-2m)/2} \int_0^\infty \frac{J_{n+1}(r) J_{n+1}(er)}{er} \, dr. \]  

Next we use the identity \(^7\)

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

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:

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

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.