A concise formula for the Zernike coefficients of scaled pupils

Augustus J.E.M. Janssen, Peter Dirksen

Philips Research Europe

ABSTRACT

Modern steppers and scanners have a projection lens whose numerical aperture 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.

Keywords: Optical lithography, NA-scaling, Zernike coefficients, Strehl ratio

1. INTRODUCTION

It is well known that different lithographic features have different optimal settings of the numerical aperture (NA) and illumination condition.[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.J.E.M.J.: Philips Research Europe, WO-02, NL-5656 AA Eindhoven, The Netherlands

P.D., Philips Research Europe, Kapeldreef 75, B-3001 Leuven, Belgium, peter.dirksen@philips.com

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[3] gives the Zernike coefficients of the scaled pupil in terms of those 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 directly 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 Eq.(4) is given in Appendix A.

2. ZERNIKE COEFFICIENTS OF SCALED PUPILS

We consider a pupil function

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

on a unit disk $0 \le \rho \le 1$ with real phase Φ, and we assume that Φ is expanded as a Zernike series according to:

$$\Phi(\rho,\theta) = \sum_{n,m} \alpha_n^m R_n^m(\rho)\cos(m\theta), \tag{2}$$

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

$$\Phi(\epsilon\rho, \theta) = \sum_{n,m} \alpha_n^m(\epsilon) R_n^m(\rho) \cos(m\theta), \quad 0 \le \rho \le 1, \quad 0 \le \theta \le 2\pi. \tag{3}$$

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

$$\alpha_n^m(\epsilon) = \sum_{n'} \alpha_{n'}^m \left(R_{n'}^n(\epsilon) - R_{n'}^{n+2}(\epsilon) \right), n = m, m+2, \cdots, \tag{4}$$

where the summation is over $n' = n, n+2, \cdots$, and where we use the convention that $R_n^{n+2} \equiv 0$. We note that

1) the entries $R_{n'}^n(\epsilon) - R_{n'}^{n+2}(\epsilon)$ do not depend on m; we consider them only for $n, n' = m, m+2, \cdots$.

2) $\alpha_n^m(\epsilon = 1) = \alpha_n^m$ since all involved Zernike polynomials, except $R_n^{n+2}(\epsilon) \equiv 0$, equal 1 at $\epsilon = 1$.

3. EXAMPLES AND SENSITIVITY ANALYSIS

Figure 1 shows an example where we scale low order coma $\alpha_3^1(\epsilon)$ with $\alpha_3^1 = 2\pi \cdot 0.016$ and $\alpha_5^1 = 2\pi \cdot 0.016$. For this special case, Eq.(4) and Dai's Eq.(18) become

$$\alpha_3^1(\epsilon) = \alpha_3^1 R_3^3(\epsilon) + \alpha_5^1 \left(R_5^3(\epsilon) - R_5^5(\epsilon) \right) \tag{5}$$

and

$$\alpha_3^1(\epsilon) = \epsilon^3 \left[\alpha_3^1 + 4\alpha_5^1(\epsilon^2 - 1) \right], \tag{6}$$

respectively. Both equations give an identical result, as $R_3^3(\epsilon) = \epsilon^3$ and $R_5^3(\epsilon) - R_5^5(\epsilon) = 4\epsilon^3(\epsilon^2 - 1)$.

A consequence of Eq.(4) is that

$$\frac{\partial \alpha_n^m}{\partial \epsilon}|_{\epsilon=1} = n\alpha_n^m + 2(n+1)\left[\alpha_{n+2}^m + \alpha_{n+4}^m + \cdots\right] . \tag{7}$$

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

The Strehl ratio[4] is approximated as:

$$S \approx 1 - \sum_{n,m} \frac{(\alpha_n^m)^2}{\gamma_m(n+1)} \tag{8}$$

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

$$\frac{\partial S}{\partial \epsilon}|_{\epsilon=1} = 2\sum_{n,m} \frac{(\alpha_n^m)^2}{\gamma_m(n+1)} - 2\sum_m \frac{1}{\gamma_m}(\sum_n \alpha_n^m)^2 = 2\left[\overline{|\Phi|^2_{disk}} - \overline{|\Phi|^2_{rim}}\right] , \tag{9}$$

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 ϵ of the Strehl ratio in the case of a mixture of low and high order spherical aberration: $\alpha_4^0 = 2\pi \cdot 0.02$ and $\alpha_6^0 = -2\pi \cdot 0.02$, where the scaled Strehl ratio $S(\epsilon)$ is calculated by combining Eqs.(8) and (4). This result compares well with numerical calculation obtained from the lithography simulator SOLID-C.[5] 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 $\sum_n \alpha_n^m = 0$ 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 $m = 0, 1, \cdots$ and normalization and orthogonality[4] of $R_n^m(\rho)$, $n = m, m+2, \cdots$, the Zernike coefficients $\alpha_n^m(\epsilon)$ of $\Phi(\rho\epsilon, \theta)$ and the Zernike coefficients α_n^m of $\Phi(\rho, \theta)$ are related by

$$\alpha_n^m(\epsilon) = 2(n+1) \sum_{n'} \alpha_{n'}^m \mathrm{M}_{nn'}^m(\epsilon), \quad n = m, m+2, \cdots . \tag{10}$$

The summation in (10) is over $n' = m, m+2, \cdots$, and

$$\mathrm{M}_{nn'}^m(\epsilon) = \int_0^1 R_n^m(\rho) R_{n'}^m(\rho\epsilon)\rho d\rho, \quad n, n' = m, m+2, \cdots . \tag{11}$$

We shall show that

$$\mathrm{M}_{nn'}^m(\epsilon) = \frac{1}{2(n+1)} \left(R_{n'}^n(\epsilon) - R_{n'}^{n+2}(\epsilon) \right); \tag{12}$$

in particular, it follows that $\mathrm{M}_{nn'}^m(\epsilon) = 0$ when $n' < n$ and that $\mathrm{M}_{nn'}^m(\epsilon)$ does not depend on m, except that in Eq.(10) we only use $n, n' = m, m+2, \cdots$. For this we use that[6]

$$R_l^k(\rho) = (-1)^{\frac{k-l}{2}} \int_0^\infty J_{k+1}(r) J_l(\rho r) dr, \ 0 \le \rho < 1, \tag{13}$$

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

$$\mathrm{M}_{nn'}^m(\epsilon) = (-1)^{\frac{n'-m}{2}} \int_0^\infty J_{n'+1}(r) (\int_0^1 R_n^m(\rho) J_m(\rho\epsilon r)\rho d\rho) dr. \tag{14}$$

To the inner integral we apply the result

$$\int_0^1 R_n^m(\rho) J_m(\rho v)\rho d\rho = (-1)^{\frac{n-m}{2}} \frac{J_{n+1}(v)}{v} \tag{15}$$

from the Nijboer-Zernike theory,[4] and we get

$$\mathrm{M}^{m}_{nn'}(\epsilon) = (-1)^{\frac{n'+n-2m}{2}} \int_0^\infty \frac{J_{n'+1}(r)J_{n+1}(\epsilon r)}{\epsilon r} dr. \tag{16}$$

Next we use the identity[7]

$$\frac{J_{n+1}(\epsilon r)}{\epsilon r} = \frac{J_n(\epsilon r) + J_{n+2}(\epsilon r)}{2(n+1)}, \tag{17}$$

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:

$$\alpha^m_n(\epsilon) = \epsilon^n \left[\alpha^m_n + (n+1) \sum_{i=1}^{(N-n)/2} \alpha^m_{n+2i} \sum_{j=0}^{i} \frac{(-1)^{i+j}(n+i+j)!}{(n+j+1)!(i-j)!j!} \epsilon^{2j} \right] \tag{18}$$

with N the maximum $n-$value.

REFERENCES

1. B.J. Lin, "Signamization of resist images", Proc. SPIE **3051**, 620-628 (1997).
2. A. Engelen, R. Socha, E. Hendrickx, W. Scheepers, F. Nowak, M. van Dam, A. Liebchen, and D. Faas, "Optimization on Step & Scan Systems", Proc. SPIE **5377**, 1323-1333 (2004).
3. G.M. Dai., "Scaling Zernike expansion coefficients to smaller pupil sizes: a simpler formula," J. Opt. Soc. Am. A **23**, 539 (2006).
4. M. Born and E. Wolf, *Principles of Optics*, 7th rev. ed., Sec. 9.2, Cambridge University Press, Cambridge, MA (2001).
5. SOLID-C (release 6.3.0), SIGMA-C GmbH, Thomas-Dehlerstrasze 9, D-81737 Munich, Germany.
6. R.J. Noll, " Zernike polynomials and atmospheric turbulence", J. Opt. Soc. Am. **66**, 207 (1976).
7. M. Abramowitz and I.A. Stegun, "Handbook of Mathematical Functions", first item of 9.1.27 on p. 361, Dover, New York (1970).

Captions

Figure [1]

Scaling low order coma $\alpha_3^1(\epsilon)$ *when* $\alpha_3^1 = 2\pi \cdot 0.016$ *and* $\alpha_5^1 = 2\pi \cdot 0.016$. *The solid line represents Eq.(4), the dashed line the tangent line given by Eq.(7), and Eq.(18) is represented by the circle-symbol.*

Figure [2]

Scaling the Strehl ratio when $\alpha_4^0 = -\alpha_6^0 = 2\pi \cdot 0.02$. *The solid line represents the combination of Eqs.(8) and (4), the circles represent SOLID-C numerical calculation, and the dashed line represents the tangent line given by Eq.(9).*

Figures [1] and [2]