Straightforward path to Zernike polynomials

Anthony Yen*

ASML Technology Development Center, San Jose, California, United States

Abstract. Starting from Weierstrass' approximation theorem, Zernike polynomials are obtained by a few straightforward steps involving only the recast of the aberration function as a double sum in the polar coordinates followed by the weighted orthogonalization of a power series. The origin of the name Fringe Zernike polynomials is also explained. © *2021 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JMM.20.2.020501]

Zernike polynomials are used extensively in microlithography to characterize the imaging optics and in evaluating the resulting images. Yet few lithographers have questioned how these polynomials are obtained. Frits Zernike invented the eponymous circle polynomials as solutions of a self-adjoint differential equation subject to circular boundary conditions.[1–3] The angular parts of his solutions are simply $\cos m\varphi$ and $\sin m\varphi$ with $m \geq 0$, but the general expression for the radial part of the solution for $0 \leq r \leq 1$ looks quite daunting at first:

$$
\begin{aligned}
R_n^m(r) &= \frac{r^{-m}}{\left(\frac{n-m}{2}\right)!} \left(\frac{d}{d(r^2)}\right)^{(n-m)/2} \left\{ r^{n+m}(r^2-1)^{(n-m)/2} \right\} \\
&= \sum_{s=0}^{(n-m)/2} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)!} r^{n-2s}.
\end{aligned}
\tag{1}
$$

An alternate derivation of the above formulae is described in Born and Wolf.[4] To understand fully both derivations, the reader has to be familiar with some specialized topics in mathematical physics. It is the purpose of this letter to lessen the complexity and demonstrate that the Zernike polynomials can be obtained using straight-forward mathematics involving three steps described below.

Let $W(x, y)$ be an aberration function or any function that is continuous within and on the unit circle. According to Weierstrass' approximation theorem, $W(x, y)$ may be expressed in a polynomial to arbitrary degree of accuracy:

$$
W(x, y) = \lim_{N \to \infty} \sum_{p,q=0}^{N} A_{pq} x^p y^q,
\tag{2}
$$

where p and q are integers.[5] Anticipating the split of the polynomials into radial and angular parts, our first step is to express $x^p y^q$ in polar corrdinates. We accomplish this by letting $x = r \cos \theta$ and $y = r \sin \theta$ and making use of Euler's formula and the binomial theorem

*Address all correspondence to Anthony Yen, tony.yen@asml.com

$$x^p y^q = (r \cos \theta)^p (r \sin \theta)^q = r^{p+q} \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right)^p \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^q$$

$$= \frac{r^{p+q}}{2^{p+q} i^q} \sum_{s=0}^{p} \binom{p}{s} (e^{i\theta})^{p-s} (e^{-i\theta})^s \sum_{t=0}^{q} \binom{q}{t} (e^{i\theta})^{q-t} (-e^{-i\theta})^t$$

$$= \frac{r^{p+q}}{2^{p+q} i^q} \sum_{s=0}^{p} \sum_{t=0}^{q} \binom{p}{s} \binom{q}{t} (-1)^t (e^{i\theta})^{p+q-(s+t)} (e^{-i\theta})^{s+t}$$

$$= \frac{r^{p+q}}{2^{p+q} i^q} e^{i(p+q)\theta} \sum_{l=0}^{p+q} C_l e^{-2il\theta},$$

where we have let $l = s + t$ and combined all the exponential terms with fixed l into a single term with coefficient C_l. We then insert the above expression into Eq. (2) and get

$$W(x, y) = W(r, \theta) = \sum_{m=0}^{\infty} r^m \sum_{l=0}^{m} C_{lm} e^{i(m-2l)\theta}, \tag{3}$$

where we have let $m = p + q$ and, just like before, combined all the exponential terms with fixed m into a single term with coefficient C_{lm}.

Our second step is to re-arrange the terms in Eq. (3). To start, let us write down its first few terms, say from $m = 0$ to $m = 4$. They are

$$C_{00} + r(C_{01} e^{i\theta} + C_{11} e^{-i\theta}) + r^2 (C_{02} e^{i2\theta} + C_{12} + C_{22} e^{-i2\theta})$$
$$+ r^3 (C_{03} e^{i3\theta} + C_{13} e^{i\theta} + C_{23} e^{-i\theta} + C_{33} e^{-i3\theta})$$
$$+ r^4 (C_{04} e^{i4\theta} + C_{14} e^{i2\theta} + C_{24} + C_{34} e^{-i2\theta} + C_{44} e^{-i4\theta}).$$

These terms can be rearranged so that the following pattern can be seen:

$$\{ C_{00} + r(C_{01} e^{i\theta} + C_{11} e^{-i\theta}) + r^2 (C_{02} e^{i2\theta} + C_{22} e^{-i2\theta}) + r^3 (C_{03} e^{i3\theta} + C_{33} e^{-i3\theta})$$
$$+ r^4 (C_{04} e^{i4\theta} + C_{44} e^{-i4\theta}) \} + \{ r^2 C_{12} + r^3 (C_{13} e^{i\theta} + C_{23} e^{-i\theta})$$
$$+ r^4 (C_{14} e^{i2\theta} + C_{34} e^{-i2\theta}) \} + \{ r^4 C_{24} \}.$$

Continuing this process and making use of Euler's formula, Eq. (3) can be expressed as

$$W(r, \theta) = \sum_{m=0}^{\infty} r^m (A_m \cos m\theta + B_m \sin m\theta)$$

$$+ \sum_{m=0}^{\infty} r^2 r^m (A_m' \cos m\theta + B_m' \sin m\theta)$$

$$+ \sum_{m=0}^{\infty} r^4 r^m (A_m'' \cos m\theta + B_m'' \sin m\theta) + \cdots$$

Grouping all the cosine and sine terms together, we have

$$W(r, \theta) = \sum_{m=0}^{\infty} \cos m\theta \, r^m (A_m + A_m' r^2 + A_m' r^4 + \cdots)$$

$$+ \sum_{m=0}^{\infty} \sin m\theta \, r^m (B_m + B_m' r^2 + B_m' r^4 + \cdots)$$

$$= \sum_{m=0}^{\infty} \cos m\theta \, r^m \sum_{k=0}^{\infty} A_{mk} r^{2k} + \sum_{m=0}^{\infty} \sin m\theta \, r^m \sum_{k=0}^{\infty} B_{mk} r^{2k}. \tag{4}$$

In reaching the above expression, no requirement of rotational symmetry about an axis had to be imposed.

Our third and last step involves expressing r^{2k} as a linear combination of orthogonal polynomials satisfying the orthogonal relation over the interval [0, 1]. Once this is accomplished, both summations over k can be expressed as linear combinations of these polynomials. Therefore, the first thing to do is to obtain these orthogonal radial polynomials (actually the Zernike radial polynomials) by orthogonalizing the set $\{1, r^2, r^4 \ldots r^{2k} \ldots\}$. We do this by first letting $r^2 = u$ so that the orthogonalization process becomes for the set $\{1, u, u^2, \ldots u^k \ldots\}$. We may then associate the orthogonalization of this set with shifted Legendre polynomials $P_k(u)$. $P_k(u)$'s can be obtained through the Gram-Schmidt orthogonalization process on the above set or simply by making use of the formula[6]

$$P_k(u) = \frac{1}{k!} \frac{d^k}{du^k} \{u^k(u-1)^k\}.$$

(See also Ref. 5, pp. 233–239. The Legendre polynomials discussed in the text are defined on the interval $[-1, 1]$ and the associated formula is called Rodrigues' formula.) The first three shifted Legendre polynomials are $P_0(u) = 1$, $P_1(u) = 2u - 1$, $P_2(u) = 6u^2 - 6u + 1$. Therefore, we may express 1 as $P_0(u)$, u as $\frac{P_1(u)}{2} + \frac{P_0(u)}{2}$, u^2 as $\frac{P_2(u)}{6} + \frac{P_1(u)}{2} + \frac{P_0(u)}{3}$, and so on. Hence any linear combination of powers of u can be expressed as a linear combination of $P_k(u)$'s. There is only one catch, however. We have to include the common factor $r^m = u^{m/2}$ in Eq. (4) in the orthogonalization process, so if $G_k^m(u)$'s are the resulting polynomials, our orthogonalization relation has to be, instead of $\int_0^1 P_k(u) \cdot P_{k'}(u) du = Const.\delta_{kk'}$,

$$\int_0^1 u^{m/2} G_k^m(u) \cdot u^{m/2} G_{k'}^m(u) du = \int_0^1 G_k^m(u) G_{k'}^m(u) u^m du = Const.\delta_{kk'}. \tag{5}$$

The second integral in Eq. (5) suggests that the presence of the factor $u^{m/2}$ may be regarded as orthogonalizing the set $\{1, u, u^2, \ldots u^k \ldots\}$ with the weight u^m. The formula for the polynomials obtained by orthogonalizing the set $\{1, u, u^2, \ldots\}$ with the weight equal to u^m instead of 1 (which would result in shifted Legendre polynomials) is given as

$$G_k^m(u) = \frac{1}{k!} \frac{1}{u^m} \frac{d^k}{du^k} \{u^m u^k(u-1)^k\}. \tag{6}$$

The validity of Eq. (6) can be established as follows. First, the polynomial so generated is of order k because the term of the highest power inside the brackets to be differentiated is u^{m+2k}. Second, the following integral is valid (Ref. 6, p. 324):

$$\int_0^1 G_k^m(u) u^{k'} u^m du = 0, \quad 0 \le k' < k.$$

Since $G_{k'}^m(u)$ is a linear combination of powers of u with $u^{k'}$ being the term of the highest power, Eq. (5) therefore stands.

Now with powers of u represented by $G_k^m(u)$'s, Eq. (4) can be recast as

$$W(r, \theta) = \sum_{m,k=0}^{\infty} C_k^m r^m G_k^m(r^2) \cos m\theta + \sum_{m,k=0}^{\infty} D_k^m r^m G_k^m(r^2) \sin m\theta,$$

where C_k^m and D_k^m are the new coefficients.

The Zernike polynomial is simply $Z_k^m(r, \theta) = r^m G_k^m(r^2) \begin{Bmatrix} \cos m\theta \\ \sin m\theta \end{Bmatrix} = R_k^m(r) \begin{Bmatrix} \cos m\theta \\ \sin m\theta \end{Bmatrix}$, where $R_k^m(r) = r^m G_k^m(r^2)$ is called the Zernike radial polynomial. Since the angular parts are already orthogonal, as

$$\int_0^{2\pi} \cos m\theta \, \sin m'\theta \, d\theta = 0,$$

$$\int_0^{2\pi} \cos m\theta \, \cos m'\theta \, d\theta = \pi(1 + \delta_{m0})\delta_{mm'},$$

$$\int_0^{2\pi} \sin m\theta \, \sin m'\theta \, d\theta = \pi(1 - \delta_{m0})\delta_{mm'},$$

and since $R_k^m(r)$'s satisfy

$$\int_0^1 R_k^m(r)R_{k'}^m(r)dr^2 = \int_0^1 G_k^m(r^2)G_{k'}^m(r^2)r^{2\,m}dr^2 = Const.\delta_{kk'}$$

because of Eq. (5), $Z_k^m(r,\theta)$'s therefore satisfy the orthogonal relation over an area bounded by the unit circle, as

$$\int_0^1 \int_0^{2\pi} Z_k^m(r,\theta)Z_{k'}^{m'}(r,\theta)rdr\,d\theta = Const.\delta_{kk'}\delta_{mm'}.$$

The explicit expression for the Zernike radial polynomials can now be written down immediately as

$$R_k^m(r) = r^m G_k^m(r^2) = \frac{r^{-m}}{k!}\frac{d^k}{d(r^2)^k}\{r^{2(m+k)}(r^2-1)^k\}.$$

Defining $n = m + 2k$ brings us to Eq. (1) put forth originally by Frits Zernike.

Incidentally, using k instead of n to index the Zernike polynomials is not a bad thing. One advantage is that k is independent of m. The ordering sequence of the Zernike polynomials used by Zeiss and ASML is a modified version of the indexing scheme originated at the University of Arizona. We can learn the origin of this Fringe indexing scheme from Katherine Creath and Robert E. Parks' article:[7] "The first program for analyzing interferograms was written by Jim Rancourt, PhD 1974 (Fig. 11),[19]. . . Later, Loomis, PhD 1980, wrote a FRINGE MANUAL, and updated the program to output the 37 "FRINGE" Zernike polynomials,[20] and the beginning of the confusion about whose numbering of the polynomials one might be using." Citation [19] in their article is: Optical Sciences Center, "FRINGE Software Program," OSC Newsletter 8(12), 29 (1974). Citation [20] refers to John S. Loomis, FRINGE User's Manual, Optical Sciences
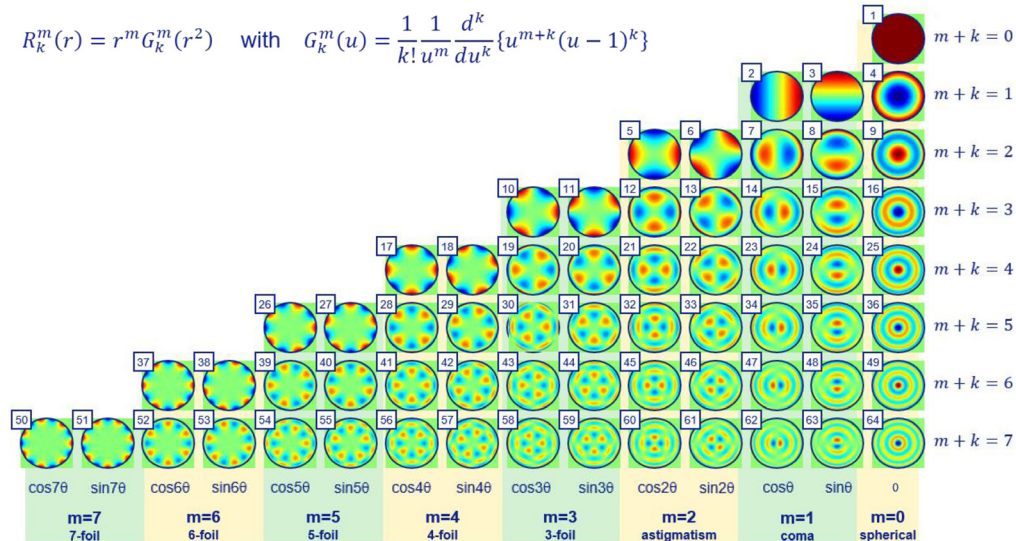

Fig. 1 Indexing scheme of Zernike polynomials used by Zeiss and ASML. These plots were originally generated by Marco Moers.

Table 1 Explicit expressions of the first 36 Zernike polynomials.

Index	Mathematical expression	Name	m (period)	$m + k$	$n = m + 2k$ (power)
1	1	Piston	0	0	0
2	$r \cos \theta$	Tilt x	1	1	1
3	$r \sin \theta$	Tilt y	1	1	1
4	$2r^2 - 1$	Focus	0	1	2
5	$r^2 \cos 2\theta$	Astigmatism x	2	2	2
6	$r^2 \sin 2\theta$	Astigmatism y	2	2	2
7	$(3r^3 - 2r) \cos \theta$	Coma x	1	2	3
8	$(3r^3 - 2r) \sin \theta$	Coma y	1	2	3
9	$6r^4 - 6r^2 + 1$	Spherical aberration	0	2	4
10	$r^3 \cos 3\theta$	Three-fold x	3	3	3
11	$r^3 \sin 3\theta$	Three-fold y	3	3	3
12	$(4r^4 - 3r^2) \cos 2\theta$	Astigmatism x	2	3	4
13	$(4r^4 - 3r^2) \sin 2\theta$	Astigmatism y	2	3	4
14	$(10r^5 - 12r^3 + 3r) \cos \theta$	Coma x	1	3	5
15	$(10r^5 - 12r^3 + 3r) \sin \theta$	Coma y	1	3	5
16	$20r^6 - 30r^4 + 12r^2 - 1$	Spherical aberration	0	3	6
17	$r^4 \cos 4\theta$	Four-fold x	4	4	4
18	$r^4 \sin 4\theta$	Four-fold y	4	4	4
19	$(5r^5 - 4r^3) \cos 3\theta$	Three-fold x	3	4	5
20	$(5r^5 - 4r^3) \sin 3\theta$	Three-fold y	3	4	5
21	$(15r^6 - 20r^4 + 6r^2) \cos 2\theta$	Astigmatism x	2	4	6
22	$(15r^6 - 20r^4 + 6r^2) \sin 2\theta$	Astigmatism y	2	4	6
23	$(35r^7 - 60r^5 + 30r^3 - 4r) \cos \theta$	Coma x	1	4	7
24	$(35r^7 - 60r^5 + 30r^3 - 4r) \sin \theta$	Coma y	1	4	7
25	$70r^8 - 140r^6 + 90r^4 - 20r^2 + 1$	Spherical aberration	0	4	8
26	$r^5 \cos 5\theta$	Five-fold x	5	5	5
27	$r^5 \sin 5\theta$	Five-fold y	5	5	5
28	$(6r^6 - 5r^4) \cos 4\theta$	Four-fold x	4	5	6
29	$(6r^6 - 5r^4) \sin 4\theta$	Four-fold y	4	5	6
30	$(21r^7 - 30r^5 + 10r^3) \cos 3\theta$	Three-fold x	3	5	7
31	$(21r^7 - 30r^5 + 10r^3) \sin 3\theta$	Three-fold y	3	5	7
32	$(56r^8 - 105r^6 + 60r^4 - 10r^2) \cos 2\theta$	Astigmatism x	2	5	8

Table 1 (*Continued*).

Index	Mathematical expression	Name	m (period)	$m + k$	$n = m + 2k$ (power)
33	$(56r^8 - 105r^6 + 60r^4 - 10r^2)\sin 2\theta$	Astigmatism y	2	5	8
34	$(126r^9 - 280r^7 + 210r^5 - 60r^3 + 5r)\cos \theta$	Coma x	1	5	9
35	$(126r^9 - 280r^7 + 210r^5 - 60r^3 + 5r)\sin \theta$	Coma y	1	5	9
36	$252r^{10} - 630r^8 + 560r^6 - 210r^4 + 30r^2 - 1$	Spherical aberration	0	5	10

Center, University of Arizona, Tucson, AZ, November 1976. Hence we believe that it was John Loomis who invented this indexing scheme in conjunction with the wavefront-fitting program called FRINGE, originally written by Jim Rancourt. It is therefore a gross misnomer that the Zernike polynomials we lithographers use are often referred to as Fringe Zernike polynomials, as if there are various sets of such polynomials; it is the "Fringe" indexing scheme of the one and only set of Zernike polynomials!

The indexing scheme used by Zeiss and ASML is shown in Fig. 1. As one can see, rows are arranged by the ascending order of $m + k$. Since the power of every radial polynomial is $n = m + 2k$ and since $(m + 2k) + m = 2(m + k)$ is fixed for every row, the rightmost entry of every row, with $m = 0$, has the highest power. Table 1 lists explicitly the Zernike polynomials according to this indexing scheme.

If the pupil function is rather roughly behaved, it may be necessary to include Zernike polynomials of very high orders. For numerical computations involving Zernike radial polynomials of $n \geq 40$, Janssen and Dirksen suggested an alternate form of Eq. (1) with advantages in computation time, accuracy and ease of implementation.[8] Based on Janssen and Dirksen's integral expression, Shakibaei and Paramesran found a concise recursive relation for $R_n^m(r)$ leading to a reduction in computational complexity.[9]

Acknowledgments

References

1. F. Zernike, "Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode," *Physica* **1**, 689–704 (1934).
2. F. Zernike, "Diffraction theory of the knife-edge test and its improved form, the phase-contrast method," *J. Micro/Nanolithogr. MEMS MOEMS* **1**(2), 87–94 (2002).
3. B. R. A. Nijboer, "The diffraction theory of aberrations," Doctoral Thesis, University of Groningen (1942).
4. M. Born and E. Wolf, *Principle of Optics*, 7th ed., Sec. 9.2.1 and Appendix VII, Cambridge University Press, Cambridge (1999).
5. F. W. Byron and R. W. Fuller, *Mathematics of Classical and Quantum Physics*, p. 239, Dover Publications, New York (1992).
6. A. Yen and S.-S. Yu, *Optical Physics for Nanolithography*, pp. 316–320, SPIE Press, Bellingham, Washington, DC (2018).
7. K. Creath and R. E. Parks, "Optical metrology at the optical sciences center: a historical review," *Proc. SPIE* **9186**, 91860T (2014).

8. A. J. E. M. Janssen and P. Dirksen, "Computing Zernike polynomials of arbitrary degree using the discrete Fourier transform," *J. Eur. Opt. Soc. Rapid Publ.* **2**, 07012 (2007).

9. B. H. Shakibaei and R. Paramesran, "Recursive formula to compute Zernike radial polynomials," *Opt. Lett.* **38**, 2487 (2013).

10. J. Braat and P. Török, *Imaging Optics*, pp. 888–890, Cambridge University Press, Cambridge (2019).