Orthonormal vector polynomials in a unit circle, Part I: basis set derived from gradients of Zernike polynomials

Chunyu Zhao and James H. Burge

College of Optical Sciences, the University of Arizona
1630 E. University Blvd, Tucson, AZ 85721
czhao@optics.arizona.edu; jburge@optics.arizona.edu

Abstract: Zernike polynomials provide a well known, orthogonal set of scalar functions over a circular domain, and are commonly used to represent wavefront phase or surface irregularity. A related set of orthogonal functions is given here which represent vector quantities, such as mapping distortion or wavefront gradient. These functions are generated from gradients of Zernike polynomials, made orthonormal using the Gram-Schmidt technique. This set provides a complete basis for representing vector fields that can be defined as a gradient of some scalar function. It is then efficient to transform from the coefficients of the vector functions to the scalar Zernike polynomials that represent the function whose gradient was fit. These new vector functions have immediate application for fitting data from a Shack-Hartmann wavefront sensor or for fitting mapping distortion for optical testing. A subsequent paper gives an additional set of vector functions consisting only of rotational terms with zero divergence. The two sets together provide a complete basis that can represent all vector distributions in a circular domain.

References and links

1. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207-211 (1976).
2. M. Born and E. Wolf, *Principles of Optics,* (Pergamon Press, 1980) pg. 464-468.
3. See http://wyant.opt-sci.arizona.edu/zernikes/zernikes.htm.
4. G. Harbers, P. J. Kunst, and G. W. R. Leibbrandt, "Analysis of lateral shearing interferograms by use of Zernike polynomials," Appl. Opt. **35,** 6162-6172 (1996).
5. R. G. Lane and M. Tallon, "Wave-front reconstruction using a Shack-Hartmann sensor," Appl. Opt. **31,** 6902-6907 (1992).
6. P. C. V. Mallik, C. Zhao, J. H. Burge, "Measurement of a 2-meter flat using a pentaprism scanning system," Opt. Eng. **46**, 023602 (2007).
7. A. Gavrielides, "Vector polynomials orthogonal to the gradient of Zernike polynomials," Opt. Lett, **7**, 526-528 (1982).
8. E. Acosta, S. Bara, M. A. Rama and S. Rios, "Determination of phase mode components in terms of local wave-front slopes: an analytical approach," Opt. Lett. **20**, 1083-1085 (1995).
9. P. E. Murphy, T. G. Brown, and D. T. Moore, "Interference imaging for aspheric surface testing," Appl. Opt. **39**, 2122-2129 (2000).
10. J. H. Burge, *Advanced Techniques for Measuring Primary Mirrors for Astronomical Telescopes,* Ph. D. Dissertation, Optical Sciences, University of Arizona (1993).
11. Durango^{TM} Interferometry Software, Diffraction International, Minnetonka, MN.
12. C. Zhao and J. H. Burge, "Orthonormal vector polynomials in a unit circle, Part II : completing the basis set," to be submitted to Optics Express (2007).
13. T. M. Apostol, *Linear Algebra: A First Course, with Applications to Differential Equations* (John Wiley & Sons, 1997), Page 111-114.
14. R. Upton and B. Ellerbroek, "Gram–Schmidt orthogonalization of the Zernike polynomials on apertures of arbitrary shape," Opt. Lett. **29**, 2840-2842 (2004).

1. Introduction

Zernike polynomials [1-3] are commonly used in optical testing, engineering, and analysis. There are two reasons for this. First of all, Zernike polynomials are orthogonal in a unit circle, which is convenient since many optics are circular in shape. Secondly, the lower order members of Zernike polynomials represent typical optical wavefront aberrations such as power, astigmatism, coma and spherical aberration. Besides direct wavefront measurements, wavefront slopes are often measured as well, e.g. with shearing interferometry [4], Shack-Hartmann sensors [5], or a scanning pentaprism test [6]. Various techniques have been developed to convert measured slope data to a wavefront map expressed in terms of Zernike polynomials. Garvrelides [7] developed a set of vector polynomials that are orthogonal to the gradients of Zernike polynomials but not mutually orthogonal. The coefficient for a specific Zernike polynomial representing the wavefront can then be directly calculated from integration of the dot product of the slope and the corresponding vector polynomial. Acosta [8] *et al,* took a different approach but arrived at similar results. This approach skips the intermediate step of fitting the vector slope data and obtains the wavefront directly. Yet, it is desirable to fit measurement data in the measurement space. In this case, a set of vector polynomials is needed to fit the vector slope data.

Vector polynomials are also used for quantifying mapping distortion, which is important for accurate measurement of optical surfaces [9] and can be severe due to the use of null optics. Typically, polynomial mapping functions are defined and the coefficients are fit to data using least squares techniques. [10, 11]

Although the above problems can be solved using a least squares fit to vector functions that are not orthogonal over the domain, the results are not optimal. The fit to a non-orthogonal basis set can require many more terms than are necessary, and the coefficients themselves may not be meaningful, because the value for any particular coefficients will change as higher order terms are fit. When fitting to real data, the propagation of noise is increased with the use of non-orthogonal basis functions. If the functions are truly orthogonal, the least squares solution is not necessary, coefficients can be determined by a much simpler and computationally efficient inner product. Clearly, an orthonormal basis is desired.

In this paper, we present such a desired set of vector polynomials which are orthonormal in a unit circle. These polynomials are perfect for fitting slope data. Since they are gradients of linear combinations of Zernike polynomials, it is also straightforward to convert the fitted slope map to the wavefront map expressed in terms of Zernike polynomials.

In Section 2, we present the Zernike notations that we adopted from Noll's landmark paper[1] and list the gradients of the Zernikes following the recursion relationships presented there. We then use the Zernike gradients as a basis to obtain an orthonormal set of vector polynomials using the Gram-Schmidt method and present the result in Section 3. The mapping from the orthonormal vector polynomials to gradients of scalar functions represented by standard Zernike polynomials is discussed in Section 4.

The vector set is made complete with the addition of a complementary set of vector polynomials with non-zero curl, as presented in a subsequent paper. [12] The addition of this second set of functions provides a complete basis, capable of representing any vector distribution in the circular domain. Applications of the vector polynomials for fitting the slope data taken from Shack-Hartmann sensors or other slope measurement devices, and in correcting mapping distortions for null tests of aspheric surfaces will be presented in subsequent papers as well.

2. Zernike polynomials and their gradients

There are different numbering schemes for Zernike polynomials. In this paper we adopt Noll's notation and numbering scheme which defines the polynomials in polar coordinates as

$$Z_{even\ j} = \sqrt{n+1}R_n^m(r)\sqrt{2}\cos(m\theta) \Big\}$$
$$Z_{odd\ j} = \sqrt{n+1}R_n^m(r)\sqrt{2}\sin(m\theta) \Big\} \qquad m \neq 0 \qquad (1)$$
$$Z_j = \sqrt{n+1}R_n^0(r), \qquad\qquad m = 0$$

where

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s(n-s)!}{s![(n+m)/2-s]![(n-m)/2-s]!}r^{n-2s} \qquad (2)$$

j: the general index of Zernike polynomials
n: the power of the radial coordinate r
m: the multiplication factor of the angular coordinate θ
n and m have the following relations: $m \leq n$ and (n - m) is even

The general index j has no physical meaning, while the indices n and m do. For a given j, there is a unique corresponding pair of (n, m), and the parity of j determines the angle dependence of the polynomial. While for a given pair of (n, m), j is ambiguous when $m \neq 0$. In some relationships given in the subsequent text, n and m are usually known, but the corresponding j (therefore the sine or cosine angle dependence of the polynomial) depends on other factors. For this reason, we choose to use $j(n, m)$ for the general index of a Zernike polynomial to show that n and m are known and the actual j will be determined by other conditions. The first 37 polynomials of this numbering scheme are listed in the Appendix, where the aforementioned relationship between j and (n, m) can be seen as well.

As the first step toward establishing an orthonormal basis of vector polynomials, we derive the gradients of the Zernike polynomials. We take the gradient of each Zernike polynomial and apply the recursion relationships from Noll to represent the gradients as linear combinations of lower order Zernike polynomials. The first 37 gradient terms are presented in Table 1. These functions provide a complete basis to represent gradients, but they require further manipulation to create an orthonormal set.

Table 1. Gradient of Zernike polynomials

$\nabla Z_1 = 0$

$\nabla Z_2 = \hat{i}2Z_1$

$\nabla Z_3 = \hat{j}2Z_1$

$\nabla Z_4 = \hat{i}2\sqrt{3}Z_2 + \hat{j}2\sqrt{3}Z_3$

$\nabla Z_5 = \hat{i}\sqrt{6}Z_3 + \hat{j}\sqrt{6}Z_2$

$\nabla Z_6 = \hat{i}\sqrt{6}Z_2 - \hat{j}\sqrt{6}Z_3$

$\nabla Z_7 = \hat{i}2\sqrt{3}Z_5 + \hat{j}2(\sqrt{2}Z_1 + \sqrt{6}Z_4 - \sqrt{3}Z_6)$

$\nabla Z_8 = \hat{i}2(\sqrt{2}Z_1 + \sqrt{6}Z_4 + \sqrt{3}Z_6) + \hat{j}2\sqrt{3}Z_5$

$\nabla Z_9 = \hat{i}2\sqrt{3}Z_5 + \hat{j}2\sqrt{3}Z_6$

$\nabla Z_{10} = \hat{i}2\sqrt{3}Z_6 - \hat{j}2\sqrt{3}Z_5$

$\nabla Z_{11} = \hat{i}2\sqrt{5}(Z_2 + \sqrt{2}Z_8) + \hat{j}2\sqrt{5}(Z_3 + \sqrt{2}Z_7)$

$\nabla Z_{12} = \hat{i}\sqrt{10}(Z_2 + \sqrt{2}Z_8 + \sqrt{2}Z_{10}) + \hat{j}\sqrt{10}(-Z_3 - \sqrt{2}Z_7 + \sqrt{2}Z_9)$

$\nabla Z_{13} = \hat{i}\sqrt{10}(Z_3 + \sqrt{2}Z_7 + \sqrt{2}Z_9) + \hat{j}\sqrt{10}(Z_2 + \sqrt{2}Z_8 - \sqrt{2}Z_{10})$

$\nabla Z_{14} = \hat{i}2\sqrt{5}Z_{10} - \hat{j}2\sqrt{5}Z_9$

$\nabla Z_{15} = \hat{i}2\sqrt{5}Z_9 + \hat{j}2\sqrt{5}Z_{10}$

$\nabla Z_{16} = \hat{i}(2\sqrt{3}Z_1 + 6Z_4 + 3\sqrt{2}Z_6 + 2\sqrt{15}Z_{11} + \sqrt{30}Z_{12}) + \hat{j}(3\sqrt{2}Z_5 + \sqrt{30}Z_{13})$

$\nabla Z_{17} = \hat{i}(3\sqrt{2}Z_5 + \sqrt{30}Z_{13}) + \hat{j}(2\sqrt{3}Z_1 + 6Z_4 - 3\sqrt{2}Z_6 + 2\sqrt{15}Z_{11} - \sqrt{30}Z_{12})$

$\nabla Z_{18} = \hat{i}(3\sqrt{2}Z_6 + \sqrt{30}Z_{12} + \sqrt{30}Z_{14}) + \hat{j}(-3\sqrt{2}Z_5 - \sqrt{30}Z_{13} + \sqrt{30}Z_{15})$

$\nabla Z_{19} = \hat{i}(3\sqrt{2}Z_5 + \sqrt{30}Z_{13} + \sqrt{30}Z_{15}) + \hat{j}(3\sqrt{2}Z_6 + \sqrt{30}Z_{12} - \sqrt{30}Z_{14})$

$\nabla Z_{20} = \hat{i}\sqrt{30}Z_{14} - \hat{j}\sqrt{30}Z_{15}$

$\nabla Z_{21} = \hat{i}\sqrt{30}Z_{15} + \hat{j}\sqrt{30}Z_{14}$

$\nabla Z_{22} = \hat{i}2\sqrt{7}(Z_2 + \sqrt{2}Z_8 + \sqrt{3}Z_{16}) + \hat{j}2\sqrt{7}(Z_3 + \sqrt{2}Z_7 + \sqrt{3}Z_{17})$

$\nabla Z_{23} = \hat{i}\sqrt{14}(Z_3 + \sqrt{2}Z_7 + \sqrt{2}Z_9 + \sqrt{3}Z_{17} + \sqrt{3}Z_{19}) + \hat{j}\sqrt{14}(Z_2 + \sqrt{2}Z_8 - \sqrt{2}Z_{10} + \sqrt{3}Z_{16} - \sqrt{3}Z_{18})$

$\nabla Z_{24} = \hat{i}\sqrt{14}(Z_2 + \sqrt{2}Z_8 + \sqrt{2}Z_{10} + \sqrt{3}Z_{16} + \sqrt{3}Z_{18}) + \hat{j}\sqrt{14}(-Z_3 - \sqrt{2}Z_7 + \sqrt{2}Z_9 - \sqrt{3}Z_{17} + \sqrt{3}Z_{19})$

$\nabla Z_{25} = \hat{i}\sqrt{7}(2Z_9 + \sqrt{6}Z_{19} + \sqrt{6}Z_{21}) + \hat{j}\sqrt{7}(2Z_{10} + \sqrt{6}Z_{18} - \sqrt{6}Z_{20})$

$\nabla Z_{26} = \hat{i}\sqrt{7}(2Z_{10} + \sqrt{6}Z_{18} + \sqrt{6}Z_{20}) + \hat{j}\sqrt{7}(-2Z_9 - \sqrt{6}Z_{19} + \sqrt{6}Z_{21})$

$\nabla Z_{27} = \hat{i}\sqrt{42}Z_{21} + \hat{j}\sqrt{42}Z_{20}$

$\nabla Z_{28} = \hat{i}\sqrt{42}Z_{20} - \hat{j}\sqrt{42}Z_{21}$

$\nabla Z_{29} = \hat{i}\sqrt{8}(\sqrt{3}Z_5 + \sqrt{5}Z_{13} + \sqrt{7}Z_{23}) + \hat{j}\sqrt{8}(\sqrt{2}Z_1 + \sqrt{6}Z_4 - \sqrt{3}Z_6 + \sqrt{10}Z_{11} - \sqrt{5}Z_{12} + \sqrt{14}Z_{22} - \sqrt{7}Z_{24})$

$\nabla Z_{30} = \hat{i}\sqrt{8}(\sqrt{2} + \sqrt{6}Z_4 + \sqrt{3}Z_6 + \sqrt{10}Z_{11} + \sqrt{5}Z_{12} + \sqrt{14}Z_{22} + \sqrt{7}Z_{24}) + \hat{j}\sqrt{8}(\sqrt{3}Z_5 + \sqrt{5}Z_{13} + \sqrt{7}Z_{23})$

$\nabla Z_{31} = \hat{i}\sqrt{8}(\sqrt{3}Z_5 + \sqrt{5}Z_{13} + \sqrt{5}Z_{15} + \sqrt{7}Z_{23} + \sqrt{7}Z_{25}) + \hat{j}\sqrt{8}(\sqrt{3}Z_6 + \sqrt{5}Z_{12} - \sqrt{5}Z_{14} + \sqrt{7}Z_{24} - \sqrt{7}Z_{26})$

$\nabla Z_{32} = \hat{i}\sqrt{8}(\sqrt{3}Z_6 + \sqrt{5}Z_{12} + \sqrt{5}Z_{14} + \sqrt{7}Z_{24} + \sqrt{7}Z_{26}) + \hat{j}\sqrt{8}(-\sqrt{3}Z_5 - \sqrt{5}Z_{13} + \sqrt{5}Z_{15} - \sqrt{7}Z_{23} + \sqrt{7}Z_{25})$

$\nabla Z_{33} = \hat{i}\sqrt{8}(\sqrt{5}Z_{15} + \sqrt{7}Z_{25} + \sqrt{7}Z_{27}) + \hat{j}\sqrt{8}(\sqrt{5}Z_{14} + \sqrt{7}Z_{26} - \sqrt{7}Z_{28})$

$\nabla Z_{34} = \hat{i}\sqrt{8}(\sqrt{5}Z_{14} + \sqrt{7}Z_{26} + \sqrt{7}Z_{28}) + \hat{j}\sqrt{8}(-\sqrt{5}Z_{15} - \sqrt{7}Z_{25} + \sqrt{7}Z_{27})$

$\nabla Z_{35} = \hat{i}2\sqrt{14}Z_{27} + \hat{j}2\sqrt{14}Z_{28}$

$\nabla Z_{36} = \hat{i}2\sqrt{14}Z_{28} - \hat{j}2\sqrt{14}Z_{27}$

$\nabla Z_{37} = \hat{i}6(Z_2 + \sqrt{2}Z_8 + \sqrt{3}Z_{16} + 2Z_{30}) + \hat{j}6(Z_3 + \sqrt{2}Z_7 + \sqrt{3}Z_{17} + 2Z_{29})$

3. An orthonormal set of vector polynomials

We use linear combinations of the above terms to create an orthogonal set. We define the inner product of two vector polynomials defined in a unit circle as

$$(\vec{A},\vec{B}) = \frac{1}{\pi}\iint (\vec{A} \cdot \vec{B})dxdy \qquad (3)$$

where the integration is over a unit circle.

The inner product is taken of the above gradient functions, and some results are shown in Table 2 (the table is symmetric about the diagonal, but only non-zero elements under the diagonal are shown). These Zernike gradient polynomials are not orthogonal, as the matrix of inner products listed in Table 2 is not diagonal.

Table 2. List of the inner products of the first 13 Zernike gradients

Inner Product ∇Z_j	∇Z_i												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0												
2	0	4											
3	0	0	4	0									
4	.			24									
5	.				12								
6						12							
7			$4\sqrt{2}$				56						
8		$4\sqrt{2}$						56					
9									24				
10										24			
11				$8\sqrt{15}$							120		
12						$4\sqrt{15}$						120	
13					$4\sqrt{15}$								100

$(\nabla Z_4, \nabla Z_3) = 0$

$(\nabla Z_6, \nabla Z_6) = 12$

$(\nabla Z_6, \nabla Z_{12}) = 4\sqrt{15}$

3.1 Orthogonalization of gradient functions

Using the Gram-Schmidt orthogonalization method [13, 14] (general description for the method can be found in Reference 13, and an optical application can be found in Reference 14), we construct a new set of vector polynomials with Zernike gradient polynomials as basis. The gradient of Z_1 is zero, therefore it is not used in the construction of the new set. We choose to index the first polynomial of this new set as 2 to maintain its correspondence with Zernike polynomials. The first 36 such polynomials are listed in Table 3.

Table 3. List of first 36 orthonormal vector polynomials $\vec{S}_i$ as functions of Zernike gradients.

$\vec{S}_2 = \dfrac{1}{\sqrt{4}}\nabla Z_2$	$\vec{S}_{14} = \dfrac{1}{\sqrt{40}}\nabla Z_{14}$	$\vec{S}_{26} = \dfrac{1}{\sqrt{168}}\left(\nabla Z_{26} - \sqrt{\dfrac{7}{5}}\nabla Z_{14}\right)$
$\vec{S}_3 = \dfrac{1}{\sqrt{4}}\nabla Z_3$	$\vec{S}_{15} = \dfrac{1}{\sqrt{40}}\nabla Z_{15}$	$\vec{S}_{27} = \dfrac{1}{\sqrt{84}}\nabla Z_{27}$
$\vec{S}_4 = \dfrac{1}{\sqrt{24}}\nabla Z_4$	$\vec{S}_{16} = \dfrac{1}{\sqrt{120}}\left(\nabla Z_{16} - \sqrt{\dfrac{3}{2}}\nabla Z_8\right)$	$\vec{S}_{28} = \dfrac{1}{\sqrt{84}}\nabla Z_{28}$
$\vec{S}_5 = \dfrac{1}{\sqrt{12}}\nabla Z_5$	$\vec{S}_{17} = \dfrac{1}{\sqrt{120}}\left(\nabla Z_{17} - \sqrt{\dfrac{3}{2}}\nabla Z_7\right)$	$\vec{S}_{29} = \dfrac{1}{\sqrt{224}}\left(\nabla Z_{29} - \sqrt{\dfrac{4}{3}}\nabla Z_{17}\right)$
$\vec{S}_6 = \dfrac{1}{\sqrt{12}}\nabla Z_6$	$\vec{S}_{18} = \dfrac{1}{\sqrt{120}}\left(\nabla Z_{18} - \sqrt{\dfrac{3}{2}}\nabla Z_{10}\right)$	$\vec{S}_{30} = \dfrac{1}{\sqrt{224}}\left(\nabla Z_{30} - \sqrt{\dfrac{4}{3}}\nabla Z_{16}\right)$
$\vec{S}_7 = \dfrac{1}{\sqrt{48}}\left(\nabla Z_7 - \sqrt{2}\nabla Z_3\right)$	$\vec{S}_{19} = \dfrac{1}{\sqrt{120}}\left(\nabla Z_{19} - \sqrt{\dfrac{3}{2}}\nabla Z_9\right)$	$\vec{S}_{31} = \dfrac{1}{\sqrt{224}}\left(\nabla Z_{31} - \sqrt{\dfrac{4}{3}}\nabla Z_{19}\right)$
$\vec{S}_8 = \dfrac{1}{\sqrt{48}}\left(\nabla Z_8 - \sqrt{2}\nabla Z_2\right)$	$\vec{S}_{20} = \dfrac{1}{\sqrt{60}}\nabla Z_{20}$	$\vec{S}_{32} = \dfrac{1}{\sqrt{224}}\left(\nabla Z_{32} - \sqrt{\dfrac{4}{3}}\nabla Z_{18}\right)$
$\vec{S}_9 = \dfrac{1}{\sqrt{24}}\nabla Z_9$	$\vec{S}_{21} = \dfrac{1}{\sqrt{60}}\nabla Z_{21}$	$\vec{S}_{33} = \dfrac{1}{\sqrt{224}}\left(\nabla Z_{33} - \sqrt{\dfrac{4}{3}}\nabla Z_{21}\right)$
$\vec{S}_{10} = \dfrac{1}{\sqrt{24}}\nabla Z_{10}$	$\vec{S}_{22} = \dfrac{1}{\sqrt{168}}\left(\nabla Z_{22} - \sqrt{\dfrac{7}{5}}\nabla Z_{11}\right)$	$\vec{S}_{34} = \dfrac{1}{\sqrt{224}}\left(\nabla Z_{34} - \sqrt{\dfrac{4}{3}}\nabla Z_{20}\right)$
$\vec{S}_{11} = \dfrac{1}{\sqrt{80}}\left(\nabla Z_{11} - \sqrt{\dfrac{5}{3}}\nabla Z_4\right)$	$\vec{S}_{23} = \dfrac{1}{\sqrt{168}}\left(\nabla Z_{23} - \sqrt{\dfrac{7}{5}}\nabla Z_{13}\right)$	$\vec{S}_{35} = \dfrac{1}{\sqrt{112}}\nabla Z_{35}$
$\vec{S}_{12} = \dfrac{1}{\sqrt{80}}\left(\nabla Z_{12} - \sqrt{\dfrac{5}{3}}\nabla Z_6\right)$	$\vec{S}_{24} = \dfrac{1}{\sqrt{168}}\left(\nabla Z_{24} - \sqrt{\dfrac{7}{5}}\nabla Z_{12}\right)$	$\vec{S}_{36} = \dfrac{1}{\sqrt{112}}\nabla Z_{36}$
$\vec{S}_{13} = \dfrac{1}{\sqrt{80}}\left(\nabla Z_{13} - \sqrt{\dfrac{5}{3}}\nabla Z_5\right)$	$\vec{S}_{25} = \dfrac{1}{\sqrt{168}}\left(\nabla Z_{25} - \sqrt{\dfrac{7}{5}}\nabla Z_{15}\right)$	$\vec{S}_{37} = \dfrac{1}{\sqrt{288}}\left(\nabla Z_{37} - \sqrt{\dfrac{9}{7}}\nabla Z_{22}\right)$

In general, the S polynomials can be simply expressed in terms of Zernike gradient polynomials:

For all j with $n = m$,

$$\vec{S}_j = \frac{1}{\sqrt{2n(n+1)}}\nabla Z_j. \tag{4}$$

For all j with $n \neq m$,

$$\vec{S}_j = \frac{1}{\sqrt{4n(n+1)}}\left(\nabla Z_j - \sqrt{\frac{n+1}{n-1}}\nabla Z_{j'(n'=n-2,m'=m)}\right) \tag{5}$$

where $j - j'$ is even when $m \neq 0$.

3.3 S as linear combinations of Zernike polynomials

Given that the vector polynomials S are functions of Zernike gradient polynomials and the Zernike gradient polynomials are functions of Zernike polynomials, we can express S in terms of Zernike polynomials as listed in Table 4.

Table 4. List of *S* polynomials expressed as linear combinations of Zernike polynomials.

$\vec{S}_2 = \hat{i}Z_1$

$\vec{S}_3 = \hat{j}Z_1$

$\vec{S}_4 = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_2 + \hat{j}Z_3\right)$

$\vec{S}_5 = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_3 + \hat{j}Z_2\right)$

$\vec{S}_6 = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_2 - \hat{j}Z_3\right)$

$\vec{S}_7 = \dfrac{1}{2}\left(\hat{i}Z_5 + \hat{j}(\sqrt{2}Z_4 - Z_6)\right)$

$\vec{S}_8 = \dfrac{1}{2}\left(\hat{i}(\sqrt{2}Z_4 + Z_6) + \hat{j}Z_5\right)$

$\vec{S}_9 = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_5 + \hat{j}Z_6\right)$

$\vec{S}_{10} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_6 - \hat{j}Z_5\right)$

$\vec{S}_{11} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_8 + \hat{j}Z_7\right)$

$\vec{S}_{12} = \dfrac{1}{2}\left(\hat{i}(Z_8 + Z_{10}) + \hat{j}(-Z_7 + Z_9)\right)$

$\vec{S}_{13} = \dfrac{1}{2}\left(\hat{i}(Z_7 + Z_9) + \hat{j}(Z_8 - Z_{10})\right)$

$\vec{S}_{14} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{10} - \hat{j}Z_9\right)$

$\vec{S}_{15} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_9 + \hat{j}Z_{10}\right)$

$\vec{S}_{16} = \dfrac{1}{2}\left(\hat{i}(\sqrt{2}Z_{11} + Z_{12}) + \hat{j}Z_3\right)$

$\vec{S}_{17} = \dfrac{1}{2}\left(\hat{i}Z_3 + \hat{j}(\sqrt{2}Z_{11} - Z_{12})\right)$

$\vec{S}_{18} = \dfrac{1}{2}\left(\hat{i}(Z_{12} + Z_{14}) + \hat{j}(-Z_{13} + Z_{15})\right)$

$\vec{S}_{19} = \dfrac{1}{2}\left(\hat{i}(Z_{13} + Z_{15}) + \hat{j}(Z_{12} - Z_{14})\right)$

$\vec{S}_{20} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{14} - \hat{j}Z_{15}\right)$

$\vec{S}_{21} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{15} + \hat{j}Z_{14}\right)$

$\vec{S}_{22} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{16} + \hat{j}Z_{17}\right)$

$\vec{S}_{23} = \dfrac{1}{2}\left(\hat{i}(Z_{17} + Z_{19}) + \hat{j}(Z_{16} - Z_{18})\right)$

$\vec{S}_{24} = \dfrac{1}{2}\left(\hat{i}(Z_{16} + Z_{18}) + \hat{j}(-Z_{17} + Z_{19})\right)$

$\vec{S}_{25} = \dfrac{1}{2}\left(\hat{i}(Z_{19} + Z_{21}) + \hat{j}(Z_{18} - Z_{20})\right)$

$\vec{S}_{26} = \dfrac{1}{2}\left(\hat{i}(Z_{18} + Z_{20}) + \hat{j}(-Z_{19} + Z_{21})\right)$

$\vec{S}_{27} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{21} + \hat{j}Z_{20}\right)$

$\vec{S}_{28} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{20} - \hat{j}Z_{21}\right)$

$\vec{S}_{29} = \dfrac{1}{2}\left(\hat{i}Z_{23} + \hat{j}(\sqrt{2}Z_{22} - Z_{24})\right)$

$\vec{S}_{30} = \dfrac{1}{2}\left(\hat{i}(\sqrt{2}Z_{22} + Z_{24}) + \hat{j}Z_{23}\right)$

$\vec{S}_{31} = \dfrac{1}{2}\left(\hat{i}(Z_{23} + Z_{25}) + \hat{j}(Z_{24} - Z_{26})\right)$

$\vec{S}_{32} = \dfrac{1}{2}\left(\hat{i}(Z_{24} + Z_{26}) + \hat{j}(-Z_{23} + Z_{25})\right)$

$\vec{S}_{33} = \dfrac{1}{2}\left(\hat{i}(Z_{25} + Z_{27}) + \hat{j}(Z_{26} - Z_{28})\right)$

$\vec{S}_{34} = \dfrac{1}{2}\left(\hat{i}(Z_{26} + Z_{28}) + \hat{j}(-Z_{25} + Z_{27})\right)$

$\vec{S}_{35} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{27} + \hat{j}Z_{28}\right)$

$\vec{S}_{36} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{28} - \hat{j}Z_{27}\right)$

$\vec{S}_{37} = \dfrac{1}{\sqrt{2}}\left(\hat{i}Z_{30} + \hat{j}Z_{29}\right)$

For a given S_j with corresponding indices $j(n, m)$, we define its x and y components as S_{jx} and S_{jy}, respectively, i.e.

$$\vec{S}_j = \hat{i} S_{jx} + \hat{j} S_{jy} \tag{6}$$

From observation of the first 37 S polynomials, we found that both S_{jx} and S_{jy} are linear combinations of at most two Zernike polynomials with corresponding indices $j'(n-1, m\pm 1)$ which may or may not exist binding by the rules $n \geq m \geq 0$:

$$\begin{aligned} S_{jx} &= C_a Z_{ja(n-1,m-1)} + C_{a'} Z_{ja'(n-1,m+1)}, \\ S_{jy} &= C_b Z_{jb(n-1,m-1)} + C_{b'} Z_{jb'(n-1,m+1)}. \end{aligned} \tag{7}$$

For a given $j(n, m)$, a set of rules can be used to determine all the parameters in Eq. (7) to express S_j as linear combinations of Zernike polynomials. These rules are summarized in Table 5. These rules are useful for obtaining analytical expression of any S polynomial by programming. They are complex since we have to deal with different cases of j, n and m combinations. The complexity mostly comes from the numbering scheme. In Noll's numbering scheme, even j correspond to cosine angle terms and odd j to sine angle terms and these terms swap order each time after an $m = 0$ term. The rules will be simpler if we just use the sine/cosine dependence of the terms. Basically, if an S polynomial has the same j index of a Zernike polynomial, its x-component is the linear combination of the Zernikes with same sine or cosine angle dependence, and the y-component has the opposite angle dependence. For example, for S_{32}, the corresponding Z_{32} has cosine angle dependence, so the x-component of S_{32} has Z_{24} and Z_{26} terms which both have cosine angle dependence, while the y-component of S_{32} has Z_{23} and Z_{25} terms which both have sine angle dependence.

Table 5. The rules for writing S in terms of linear combinations of Zernikes.

$S_{j(n,m)}$ $=\hat{i}S_{jx}+\hat{j}S_{jy}$			Rules						
						S_{jx}		S_{jy}	
m	n	j	ja, ja'	jb, jb'	ja-jb	C_a $(Z_{ja(n-1,m-1)})$	$C_{a'}$ $(Z_{ja'(n-1,m+1)})$	C_b $(Z_{jb(n-1,m-1)})$	$C_{b'}$ $(Z_{jb'(n-1,m+1)})$
1	1	2	ja=1, no ja'	jb=1, no jb'	0	1	NA	0	NA
		3	ja=1, no ja'	jb=1, no jb'	0	0	NA	1	NA
0	even ≥2		no ja, ja' even	no jb, jb' odd		NA	$\sqrt{1/2}$	NA	$\sqrt{1/2}$
1	=2s+1, s odd	even	ja=jb, ja'-ja=2	jb=ja, jb'-jb=1	0	$\sqrt{1/2}$	1/2	0	1/2
		odd	ja=jb, ja'-ja=1	jb=ja, jb'-jb=2	0	0	1/2	$\sqrt{1/2}$	-1/2
	=2s+1, s even	even	ja=jb, ja'-ja=1	jb=ja, jb'-jb=2	0	$\sqrt{1/2}$	1/2	0	1/2
		odd	ja=jb, ja'-ja=2	jb=ja, jb'-jb=1	0	0	1/2	$\sqrt{1/2}$	-1/2
m = n > 1	=2s or 2s-1, s odd	even	no ja'	no jb'	-1	$\sqrt{1/2}$	NA	$-\sqrt{1/2}$	NA
		odd	no ja'	no jb'	1	$\sqrt{1/2}$	NA	$\sqrt{1/2}$	NA
	=2s or 2s-1, s even	even	no ja'	no jb'	1	$\sqrt{1/2}$	NA	$-\sqrt{1/2}$	NA
		odd	no ja'	no jb'	-1	$\sqrt{1/2}$	NA	$\sqrt{1/2}$	NA
m>1 m<n	=2s or 2s-1, s odd	even	ja'-ja=2	jb'-jb=2	-1	1/2	1/2	-1/2	1/2
		odd	ja'-ja=2	jb'-jb=2	1	1/2	1/2	1/2	-1/2
	=2s or 2s-1, s even	even	ja'-ja=2	jb'-jb=2	1	1/2	1/2	-1/2	1/2
		odd	ja'-ja=2	jb'-jb=2	-1	1/2	1/2	1/2	-1/2

3.4 Plots of vector polynomial functions

The plots of first 12 *S* vector polynomials are shown in Table 6.

Table 6. Plots of first 12 *S* polynomials in a unit circle.
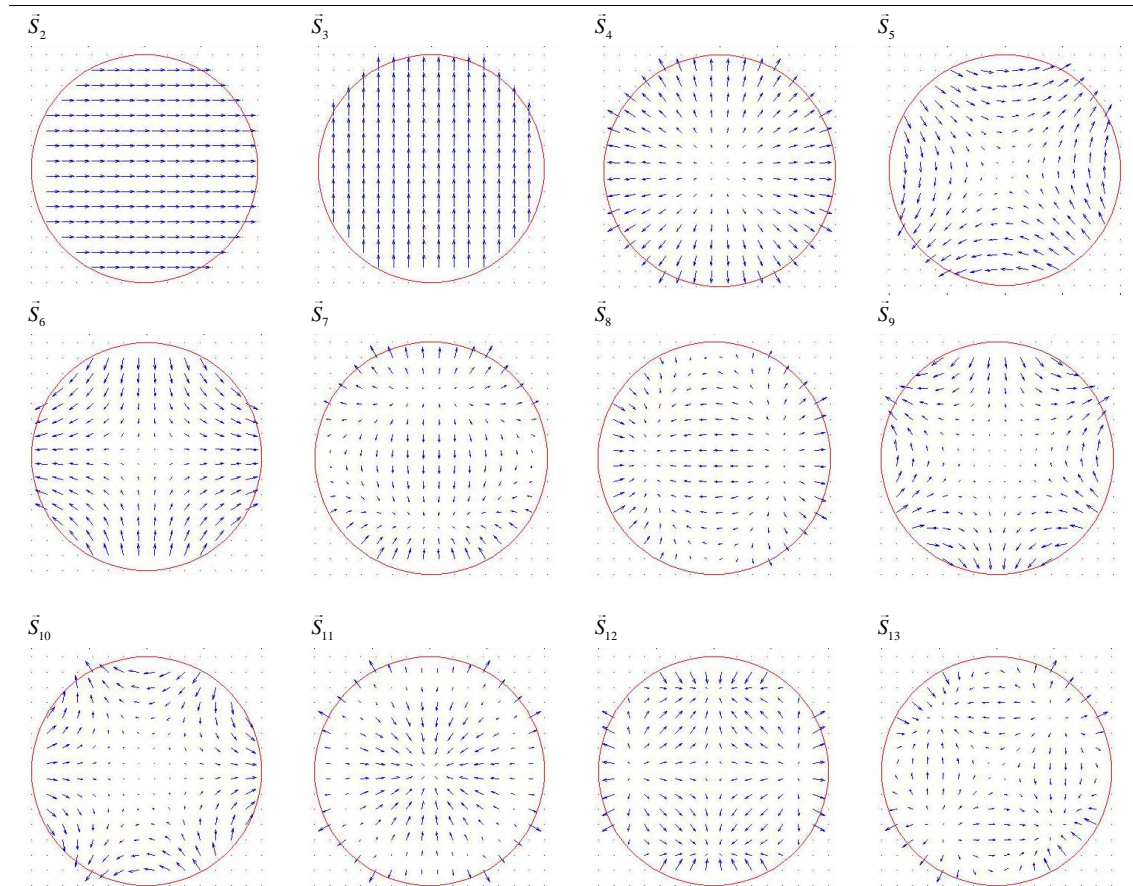
4. Relating the vector polynomials to gradients of scalar functions

The set of *S* polynomials fully spans the space of vector distributions $\vec{V}(x, y)$ over the unit circle where a scalar function $\Phi(x,y)$ exists such that $\vec{V}(x, y) = \nabla\Phi(x, y)$. It is useful to represent the vector data using the vector polynomials *S* and relate to a scalar functions ϕ that are defined as $\vec{S}_i = \nabla\phi_i$.

Applying the rules listed in (4) and (5), the scalar functions can be calculated as

For all *j* with $n = m$,

$$\phi_j = \frac{1}{\sqrt{2n(n+1)}} Z_j .$$ (8)

For all *j* with $n \neq m$,

$$\phi_j = \frac{1}{\sqrt{4n(n+1)}} \left(Z_j - \sqrt{\frac{n+1}{n-1}} Z_{j'(n'=n-2,m'=m)} \right), \tag{9}$$

where $j - j'$ is even when $m \neq 0$.

These relationships match those demonstrated for the vector functions listed in Table 3. For example, $\vec{S}_7 = \frac{1}{\sqrt{48}} \left(\nabla Z_7 - \sqrt{2} \nabla Z_3 \right)$, which leads to $\phi_7 = \frac{1}{\sqrt{48}} \left(Z_7 - \sqrt{2} Z_3 \right)$.

Applying these relations, the vector data $\vec{V}(x, y)$ is decomposed into a linear combination of the orthonormal S polynomials as

$$\vec{V} = \sum \alpha_i \vec{S}_i. \tag{10}$$

Using the definitions of the scalar functions Φ and ϕ_i ($\vec{V} = \nabla \Phi$, $\vec{S}_i = \nabla \phi_i$), we have

$$\Phi = \sum \alpha_i \phi_i, \tag{11}$$

where the coefficients α_i were found from the vector decomposition in Eq. (10). Then the scalar function Φ can in turn be represented as linear combinations of standard Zernike polynomials:

$$\Phi = \sum \alpha_i \phi_i = \sum \gamma_i Z_i \tag{12}$$

The coefficients of these standard Zernike polynomials can be found by

$$\gamma_j = \frac{\alpha_{j(n,m)}}{\sqrt{2n(n+1)}} \qquad n = m$$

$$\gamma_j = \frac{\alpha_{j(n,m)}}{\sqrt{4n(n+1)}} - \frac{\alpha_{j'(n+2,m)}}{\sqrt{4(n+1)(n+2)}} \quad n \neq m \tag{13}$$

where $j - j'$ is even when $m \neq 0$.

This procedure is useful for applications such as processing data from a Shack Hartmann sensor. The centroid data, which is proportional to wavefront slopes, can be fit to the vector S polynomials to give a set of coefficients α_i. These are converted directly to a standard Zernike polynomial representation of the wavefront, with coefficients γ_i.

A reverse problem is: given a scalar function Φ and its Zernike decomposition coefficients γ_i, we can find α_i from Eq. (13). When Φ is a wavefront, the rms spot radius is $r = 2f \sqrt{\left(\sum \alpha_i^2 \right)}$, where f is the system F number.

5. Summary

We derived an orthonormal set of vector polynomials in a unit circle. It has many potential applications, one of which is fitting slope data in optical testing. These polynomials are linear combinations of at most two Zernike polynomial's gradients. They can be expressed as linear combinations of at most four scalar Zernike polynomials as well. After wavefront slope data, e.g. data taken with a Shack-Hartmann sensor, is fit with the vector polynomials, it is straightforward to convert the fitted slope map to the wavefront map expressed in terms of Zernike polynomials.

Appendix:

The first 37 Zernike polynomials according to Noll's numbering:

j	n, m	Zernikes	j	n, m	Zernikes
1	$n=0$ $m=0$	$Z_1 = 1$	20	$n=5$ $m=5$	$Z_{20} = \sqrt{12}\, r^5 \cos 5\theta$
2	$n=1$ $m=1$	$Z_2 = 2r\cos\theta$	21	$n=5$ $m=5$	$Z_{21} = \sqrt{12}\, r^5 \sin 5\theta$
3	$n=1$ $m=1$	$Z_3 = 2r\sin\theta$	22	$n=6$ $m=0$	$Z_{22} = \sqrt{7}(20r^6 - 30r^4 + 12r^2 - 1)$
4	$n=2$ $m=0$	$Z_4 = \sqrt{3}(2r^2 - 1)$	23	$n=6$ $m=2$	$Z_{23} = \sqrt{14}(15r^6 - 20r^4 + 6r^2)\sin 2\theta$
5	$n=2$ $m=2$	$Z_5 = \sqrt{6}\, r^2 \sin 2\theta$	24	$n=6$ $m=2$	$Z_{24} = \sqrt{14}(15r^6 - 20r^4 + 6r^2)\cos 2\theta$
6	$n=2$ $m=2$	$Z_6 = \sqrt{6}\, r^2 \cos 2\theta$	25	$n=6$ $m=4$	$Z_{25} = \sqrt{14}(6r^6 - 5r^4)\sin 4\theta$
7	$n=3$ $m=1$	$Z_7 = \sqrt{8}(3r^3 - 2r)\sin\theta$	26	$n=6$ $m=4$	$Z_{26} = \sqrt{14}(6r^6 - 5r^4)\cos 4\theta$
8	$n=3$ $m=1$	$Z_8 = \sqrt{8}(3r^3 - 2r)\cos\theta$	27	$n=6$ $m=6$	$Z_{27} = \sqrt{14}\, r^6 \sin 6\theta$
9	$n=3$ $m=3$	$Z_9 = \sqrt{8}\, r^3 \sin 3\theta$	28	$n=6$ $m=6$	$Z_{28} = \sqrt{14}\, r^6 \cos 6\theta$
10	$n=3$ $m=3$	$Z_{10} = \sqrt{8}\, r^3 \cos 3\theta$	29	$n=7$ $m=1$	$Z_{29} = \sqrt{16}(35r^7 - 60r^5 + 30r^3 - 4r)\sin\theta$
11	$n=4$ $m=0$	$Z_{11} = \sqrt{5}(6r^4 - 6r^2 + 1)$	30	$n=7$ $m=1$	$Z_{30} = \sqrt{16}(35r^7 - 60r^5 + 30r^3 - 4r)\cos\theta$
12	$n=4$ $m=2$	$Z_{12} = \sqrt{10}(4r^4 - 3r^2)\cos 2\theta$	31	$n=7$ $m=3$	$Z_{31} = \sqrt{16}(21r^7 - 30r^5 + 10r^3)\sin 3\theta$
13	$n=4$ $m=2$	$Z_{13} = \sqrt{10}(4r^4 - 3r^2)\sin 2\theta$	32	$n=7$ $m=3$	$Z_{32} = \sqrt{16}(21r^7 - 30r^5 + 10r^3)\cos 3\theta$
14	$n=4$ $m=4$	$Z_{14} = \sqrt{10}\, r^4 \cos 4\theta$	33	$n=7$ $m=5$	$Z_{33} = \sqrt{16}(7r^7 - 6r^5)\sin 5\theta$
15	$n=4$ $m=4$	$Z_{15} = \sqrt{10}\, r^4 \sin 4\theta$	34	$n=7$ $m=5$	$Z_{34} = \sqrt{16}(7r^7 - 6r^5)\cos 5\theta$
16	$n=5$ $m=1$	$Z_{16} = \sqrt{12}(10r^5 - 12r^3 + 3r)\cos\theta$	35	$n=7$ $m=7$	$Z_{35} = \sqrt{16}\, r^7 \sin 7\theta$
17	$n=5$ $m=1$	$Z_{17} = \sqrt{12}(10r^5 - 12r^3 + 3r)\sin\theta$	36	$n=7$ $m=7$	$Z_{36} = \sqrt{16}\, r^7 \cos 7\theta$
18	$n=5$ $m=3$	$Z_{18} = \sqrt{12}(5r^5 - 4r^3)\cos 3\theta$	37	$n=8$ $m=0$	$Z_{37} = \sqrt{9}(70r^8 - 140r^6 + 90r^4 - 20r^2 + 1)$
19	$n=5$ $m=3$	$Z_{19} = \sqrt{12}(5r^5 - 4r^3)\sin 3\theta$			