Orthonormal vector polynomials in a unit circle, Part II : completing the basis set

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. Previously, we have developed a basis of functions generated from gradients of Zernike polynomials. Here, we complete the basis by adding a complementary set of functions with zero divergence – those which are defined locally as a rotation or curl.

OCIS codes: (220.4840) testing; (080.1010) aberrations; (010.1080) active or adoptive optics

References and links

1. C. Zhao and J. H. Burge, "Orthonormal vector polynomials in a unit circle, Part I: basis set derived from gradients of Zernike polynomials," Opt. Express **15**, 18014-18024 (2007).
2. C. Zhao, *et al*, "Figure measurement of a large optical flat with a Fizeau interferometer and stitching technique," Proc. SPIE **6293**, 62930k, (2006).
3. R. J. Noll, "Zernike polynomials and atmospheric turbulence", J. Opt. Soc. Am. **66**, 1976, 207-211.
4. H. F. Davis and A. D. Snider, *Introduction to Vector Analysis*, (Wm. C. Brown Publisher, 1986).

1. Introduction

In a previous paper,[1] we developed an orthonormal set of vector polynomials over a unit circle, which we call $\vec{S}$ polynomials. These polynomials are great for fitting the slope data taken by a Shack-Hartmann sensor. But since they are gradients of linear combinations of Zernike polynomials, they have zero curl, which means they make an incomplete set of vector polynomials such that an arbitrary continuously differentiable vector function defined over a unit circle cannot be represented by linear combinations of these polynomials. Additional vector polynomials must be added to make a complete set. An optical application that needs a complete set of vector polynomials is to fit the mapping distortions of an interferometric null test. See Reference 2 for an example of mapping distortion in a null test and how to correct it. The lowest modes of mapping distortions include translation, scaling and rotation. Polynomial $\vec{S}_2$ and $\vec{S}_3$ represent x and y translation, respectively, and $\vec{S}_4$ represents scaling. But no $\vec{S}$ polynomial represents rotation. The reason is that the rotation vector has non-zero curl, while all $\vec{S}$ polynomials have zero curl. In this paper, we derived a complementary set of vector polynomials which have zero divergence and non-zero curl. This new set combined with the zero-curl set $\vec{S}$ makes a complete set such that it can represent any continuously differentiable vector polynomials defined over a unit circle.

In Section 2, we review the $\vec{S}$ polynomials we derived and reported in a previous paper. We then proceed to derive the complementary $\vec{T}$ polynomials in Section 3.

#92913 - $15.00 USD Received 19 Feb 2008; revised 18 Apr 2008; accepted 19 Apr 2008; published 24 Apr 2008

2. The $\vec{S}$ polynomials

In a previous paper[1], we derived an orthonormal set of vector polynomials over a unit circle. We call this set the $\vec{S}$ polynomials. Each $\vec{S}$ polynomial is the gradient of a scalar function:

$$\vec{S}_j = \nabla \phi_j = \hat{i}\,\frac{\partial \phi_j}{\partial x} + \hat{j}\,\frac{\partial \phi_j}{\partial y}. \tag{1}$$

The scalar functions ϕ_j are linear combination of Zernike polynomials. Following Noll's notation and numbering scheme,[3]

$$\phi_j = \frac{1}{\sqrt{2n(n+1)}} Z_j \text{, for all } j \text{ with } n = m, \tag{2a}$$

and

$$\phi_j = \frac{1}{\sqrt{4n(n+1)}} \left(Z_j - \sqrt{\frac{n+1}{n-1}} Z_{j'(n'=n-2,m'=m)} \right) \text{, for all } j \text{ with } n \neq m. \tag{2b}$$

Each $\vec{S}$ polynomial is then the linear combinations of gradient of Zernike polynomials following (1) and (2). Since gradient of Zernike polynomials can also be represented by Zernike polynomials,[3] $\vec{S}$ polynomials can be written as linear combinations of Zernike polynomials as well. The first 14 non-trivial $\vec{S}$ polynomials are listed in Table 1.

Table 1. List of the first 14 non-trivial $\vec{S}$ polynomials as linear combinations of Zernike polynomials.

$\vec{S}_2 = \hat{i}Z_1$	$\vec{S}_9 = \frac{1}{\sqrt{2}}\left(\hat{i}Z_5 + \hat{j}Z_6\right)$
$\vec{S}_3 = \hat{j}Z_1$	$\vec{S}_{10} = \frac{1}{\sqrt{2}}\left(\hat{i}Z_6 - \hat{j}Z_5\right)$
$\vec{S}_4 = \frac{1}{\sqrt{2}}\left(\hat{i}Z_2 + \hat{j}Z_3\right)$	$\vec{S}_{11} = \frac{1}{\sqrt{2}}\left(\hat{i}Z_8 + \hat{j}Z_7\right)$
$\vec{S}_5 = \frac{1}{\sqrt{2}}\left(\hat{i}Z_3 + \hat{j}Z_2\right)$	$\vec{S}_{12} = \frac{1}{2}\left(\hat{i}(Z_8 + Z_{10}) + \hat{j}(-Z_7 + Z_9)\right)$
$\vec{S}_6 = \frac{1}{\sqrt{2}}\left(\hat{i}Z_2 - \hat{j}Z_3\right)$	$\vec{S}_{13} = \frac{1}{2}\left(\hat{i}(Z_7 + Z_9) + \hat{j}(Z_8 - Z_{10})\right)$
$\vec{S}_7 = \frac{1}{2}\left(\hat{i}Z_5 + \hat{j}(\sqrt{2}Z_4 - Z_6)\right)$	$\vec{S}_{14} = \frac{1}{\sqrt{2}}\left(\hat{i}Z_{10} - \hat{j}Z_9\right)$
$\vec{S}_8 = \frac{1}{2}\left(\hat{i}(\sqrt{2}Z_4 + Z_6) + \hat{j}Z_5\right)$	$\vec{S}_{15} = \frac{1}{\sqrt{2}}\left(\hat{i}Z_9 + \hat{j}Z_{10}\right)$

If $\vec{A}$ and $\vec{B}$ are two vector polynomials defined over a unit circle, we define their inner product as

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

where integration is over unit circle.

#92913 - $15.00 USD Received 19 Feb 2008; revised 18 Apr 2008; accepted 19 Apr 2008; published 24 Apr 2008

$\vec{S}$ polynomials are orthonormal, which means

$$(\vec{S}_i, \vec{S}_j) = \frac{1}{\pi} \iint \left((\nabla \phi_i) \bullet (\nabla \phi_j) \right) dxdy = \delta_{ij} . \tag{4}$$

3. Derivation of a complementary set of vector polynomials

Any vector field can be written as[4]

$$\vec{v} = \nabla \phi + \nabla \times \vec{P} , \tag{5}$$

where ϕ is a scalar and $\vec{P}$ is a vector. The divergence of $\vec{v}$ is then

$$\nabla \bullet \vec{v} = \nabla^2 \phi + \nabla \bullet (\nabla \times \vec{P}) = \nabla^2 \phi ,$$

and the curl of $\vec{v}$ is

$$\nabla \times \vec{v} = \nabla \times (\nabla \times \vec{P}) = \nabla (\nabla \bullet \vec{P}) - \nabla^2 \vec{P} .$$

The $\vec{S}$ polynomials presented in the previous paper[1] were defined as gradients of scalar functions, so have no curl component and $\vec{P} = 0$. We complete the basis by adding a second set that has zero divergence, therefore zero ϕ, but non zero $\vec{P}$, such that

$$\vec{T} = \nabla \times \vec{P} = \begin{bmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P_x & P_y & P_z \end{bmatrix} . \tag{6}$$

This set has to be mutually orthogonal as well.

Like the $\vec{S}$ polynomials, $\vec{T}$ polynomials are vectors defined in *x*-*y* plane only. A convenient choice of $\vec{P}$ is vectors along *z* axis only, i.e. $P_x = P_y = 0$. We can use a scalar ψ instead to represent $\vec{P}$:

$$\vec{P} = \psi \hat{k} , \tag{7}$$

where ψ is a function of *x* and *y*: $\psi = \psi(x, y)$. It follows that

$$\vec{T}_i = \nabla \times (\psi_i \hat{k}) = \hat{i} \frac{\partial \psi_i}{\partial y} - \hat{j} \frac{\partial \psi_i}{\partial x} . \tag{8}$$

The inner product of two $\vec{T}$ polynomials is then

$$\begin{aligned} (\vec{T}_i, \vec{T}_j) &= \iint \left(\left(\frac{\partial \psi_i}{\partial y} \right) \left(\frac{\partial \psi_j}{\partial y} \right) + \left(- \frac{\partial \psi_i}{\partial x} \right) \left(- \frac{\partial \psi_j}{\partial x} \right) \right) dxdy \\ &= \iint \left((\nabla \psi_i) \bullet (\nabla \psi_j) \right) dxdy . \end{aligned} \tag{9}$$

#92913 - $15.00 USD Received 19 Feb 2008; revised 18 Apr 2008; accepted 19 Apr 2008; published 24 Apr 2008

We choose a basis of functions $\{\psi_i\}$ that we use to generate the $\vec{T}$ polynomials to be the same basis as we used to generate the $\vec{S}$ polynomials, $\{\phi_i\}$. By Eq. (4), we know that their choice will create $\vec{T}$ polynomials that are mutually orthogonal.

From Eqs. (1) and (8), and with $\psi_i = \phi_i$, we know that $\vec{S}_i(x,y)$ and $\vec{T}_i(x,y)$ have same magnitude and are orthogonal to each other at any point in a unit circle, therefore $(\vec{S}_i, \vec{T}_i) = 0$. But the sets $\vec{S}$ and $\vec{T}$ are not fully independent. For all the *j* with *m* = *n*, we can show that

$$\nabla \times \vec{T}_j = -\hat{k}\nabla^2 \phi_j ,$$

and

$$\nabla^2 \phi_j \propto \nabla^2 Z_j \propto \left(\frac{1}{r}\frac{\partial}{\partial r}(r\frac{\partial}{\partial r}) + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} \right) \left[r^n \begin{pmatrix} \cos n\theta \\ \sin n\theta \end{pmatrix} \right] = 0, \quad (10)$$

which means $\vec{T}_j$ has 0 curl and is therefore not linearly independent of $\vec{S}$ polynomials. For example, when j = 9 or 10, m = n = 3: $\vec{T}_9 = \frac{1}{\sqrt{2}}(\hat{i}Z_6 - \hat{j}Z_5) = \vec{S}_{10}$ and $\vec{T}_{10} = \frac{1}{\sqrt{2}}(\hat{i}Z_5 - \hat{j}Z_6) = -\vec{S}_9$.

For any other pair of *i* and *j*, $(\vec{S}_i, \vec{T}_j) = 0$.

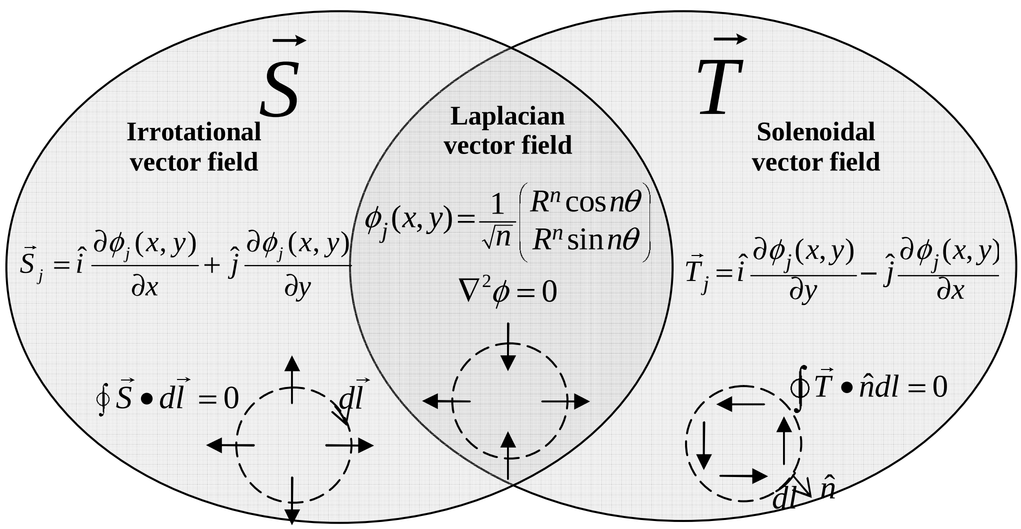


Fig. 1. Relations between the $\vec{S}$ and $\vec{T}$ polynomials. The Laplacian vector fields are the overlap between $\vec{S}$ and $\vec{T}$. The dashed circles and associated solid arrows illustrate the local behaviors of the vectors in different sets after subtracting the local constant vector.

The $\vec{S}$ and $\vec{T}$ polynomials can be thought of as vector fields in a unit circle. In vector calculus, $\vec{S}$ is known as irrotational vector fields which have zero curl everywhere, and $\vec{T}$ is known as solenoidal vector fields which have zero divergence everywhere. The two types of

#92913 - $15.00 USD Received 19 Feb 2008; revised 18 Apr 2008; accepted 19 Apr 2008; published 24 Apr 2008

vector fields have some overlap where both divergence and curls are everywhere zero, which is known as Laplacian vector field. The overall relationship between $\vec{S}$ and $\vec{T}$ vector fields is illustrated in Figure 1. The overlapped area contains terms derived from corresponding scalar ϕ polynomials whose Laplacian is 0. If ϕ represents wavefront, these terms correspond to a wavefront that has zero net curvature at any point in the pupil.

It is useful to compare the different types of functions defined here. The $\vec{S}$ functions are generated from gradients, thus have no curl. Since $\vec{S}$ functions are 2-d vectors defined in a plane, mathematically, we can express the curl as line integral along a closed path in the plane:

$$\oint \vec{S} \bullet d\vec{l} = 0 . \tag{11}$$

The $\vec{T}$ functions have no divergence. Again they are 2-d vectors defined in a plane. Mathematically, we express divergence of a 2-d vector as a line integral over a closed path:

$$\oint \vec{T} \bullet \hat{n} dl = 0 , \tag{12}$$

where $\hat{n}$ is the unit normal vector pointing out of the closed path.

The intersection, which includes both $\vec{S}$ and $\vec{T}$, is of the form that fits both Eqs. (11) and (12), having both zero divergence and zero curl. Graphical depictions of the local behavior of the functions are included in Fig. 1: dashed circles represent infinitesimal region and solid arrows represent local vectors (after a constant vector is subtracted.)

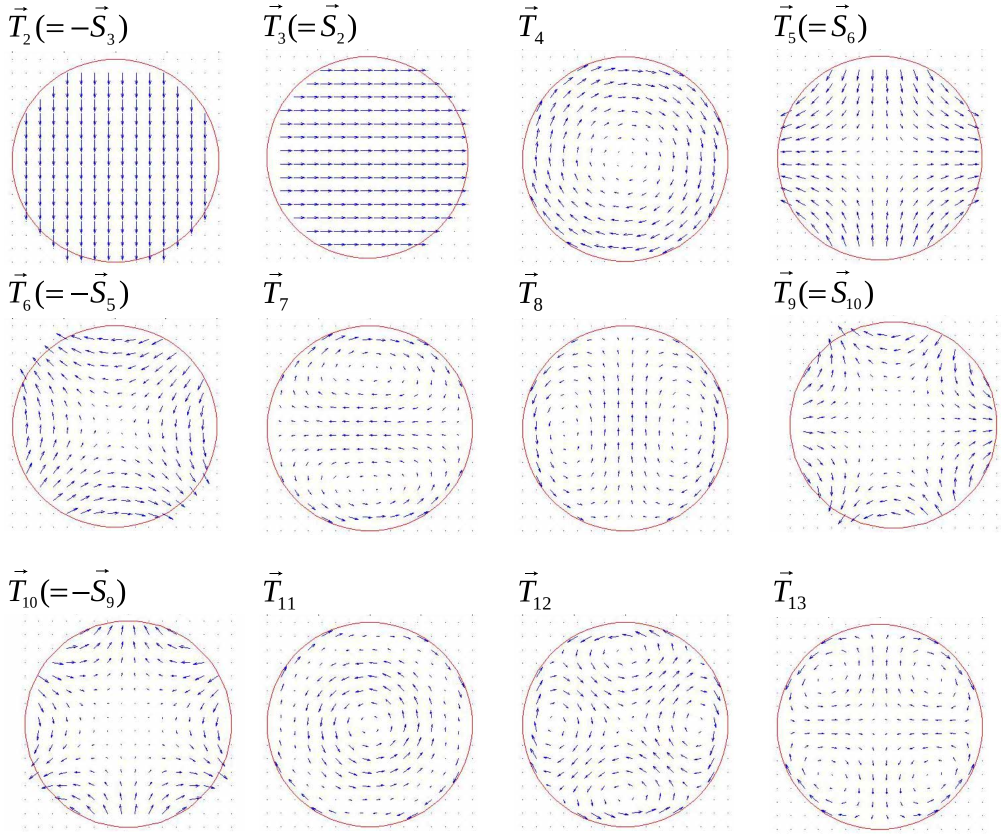
Table 2 lists expressions for the first 15 $\vec{T}$ polynomials. The plots of first 12 non-trivial $\vec{T}$ polynomials are shown in Table 3. The complete set of orthogonal vector polynomials that fully spans the circular domain can be written as the combined set of $\vec{S}$ polynomials and independent $\vec{T}$ polynomials since the Laplacian type functions are included in both sets. Care must be taken to ensure that the common functions are not counted twice.

Table 2. Analytical expressions of the first 15 $\vec{T}$ polynomials.

$\vec{T}_1 = 0$, trivia	
$\vec{T}_2 = -\hat{j}Z_1 = -\vec{S}_3$	$\vec{T}_9 = \frac{1}{\sqrt{2}}(\hat{i}Z_6 - \hat{j}Z_5) = \vec{S}_{10}$
$\vec{T}_3 = \hat{i}Z_1 = \vec{S}_2$	$\vec{T}_{10} = \frac{1}{\sqrt{2}}(-\hat{i}Z_5 - \hat{j}Z_6) = -\vec{S}_9$
$\vec{T}_4 = \frac{1}{\sqrt{2}}(\hat{i}Z_3 - \hat{j}Z_2)$	$\vec{T}_{11} = \frac{1}{\sqrt{2}}(\hat{i}Z_7 - \hat{j}Z_8)$
$\vec{T}_5 = \frac{1}{\sqrt{2}}(\hat{i}Z_2 - \hat{j}Z_3) = \vec{S}_6$	$\vec{T}_{12} = \frac{1}{2}\left(\hat{i}(-Z_7 + Z_9) - \hat{j}(Z_8 + Z_{10})\right)$
$\vec{T}_6 = \frac{1}{\sqrt{2}}(-\hat{i}Z_3 - \hat{j}Z_2) = -\vec{S}_5$	$\vec{T}_{13} = \frac{1}{2}\left(\hat{i}(Z_8 - Z_{10}) - \hat{j}(Z_7 + Z_9)\right)$
$\vec{T}_7 = \frac{1}{2}\left(\hat{i}(\sqrt{2}Z_4 - Z_6) - \hat{j}Z_5\right)$	$\vec{T}_{14} = \frac{1}{\sqrt{2}}(-\hat{i}Z_9 - \hat{j}Z_{10}) = -\vec{S}_{15}$
$\vec{T}_8 = \frac{1}{2}\left(\hat{i}Z_5 - \hat{j}(\sqrt{2}Z_4 + Z_6)\right)$	$\vec{T}_{15} = \frac{1}{\sqrt{2}}(\hat{i}Z_{10} - \hat{j}Z_9) = \vec{S}_{14}$

#92913 - $15.00 USD Received 19 Feb 2008; revised 18 Apr 2008; accepted 19 Apr 2008; published 24 Apr 2008

Table 3. Plots of the first 12 non-trivial $\vec{T}$ polynomials.



4. Summary

We derived a set of vector polynomials defined over a unit circle which complements the set we presented in a previous paper. Each set of the vector polynomials is orthonormal over a unit circle. But there are some overlap between the two sets. We can combine the two sets in such a way that the overlapped subset is counted only once, then we obtain a complete set of vector polynomials defined over a unit circle. The combined set is useful for fitting any continuously differentiable vector functions in a circular domain. In particular, it is useful in fitting the mapping distortions often seen in an interferometric null test. We will explore and report its applications in subsequent papers.

If interested, you can request the MATLAB codes for calculating the $\vec{S}$ and $\vec{T}$ polynomials from Dr. Chunyu Zhao, czhao@optics.arizona.edu.

#92913 - $15.00 USD Received 19 Feb 2008; revised 18 Apr 2008; accepted 19 Apr 2008; published 24 Apr 2008