Investigation of the Field Dependence of the Aberration Functions of

Rotationally Nonsymmetric Optical Imaging Systems

by

Robert W. Gray

Submitted in Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

Supervised by Professor Jannick Rolland

The Institute of Optics

Arts, Science and Engineering

Edmund A. Hajim School of Engineering and Applied Sciences

University of Rochester

Rochester, New York

2015

Dedication

This work is dedicated to

my wife, Yevgenia Skorobogatov-Gray,

my parents, Marjorie S. and Howard J. Gray,

and to all the other square pegs forced into round holes.

May your edges always remain sharp.

Biographical Sketch

Mr. Gray entered the Ph.D. program at the Institute of Optics, University of Rochester, Rochester, NY, in 2011. During the period of research described in this dissertation, the following have been published in peer review journals.

1) Gray, R. W., C. Dunn, K. P. Thompson, J. P. Rolland, "An analytic expression for the field dependence of Zernike polynomials in rotationally symmetric optical system," Optics Express, 20(15), 16436-16449, 2012

2) Gray, R. W. and J. P. Rolland, "Wavefront aberration function in terms of R. V. Shack's vector product and Zernike polynomial vectors," *J. Opt. Soc. Am. A*, 32(10), 1836-1847, 2015.

Additionally, the Matlab® software "ZernikeCalc" was developed and has been made available to the optics community at the Mathworks' Matlab® Central File Exchange website:

http://www.mathworks.com/matlabcentral/fileexchange/33330-zernikecalc

Acknowledgments

I would like to thank Prof. Jannick Rolland for providing me with the opportunity to study wavefront aberration theory and for assisting me in optics with both basic concepts in the classroom and advanced concepts throughout the research period. Her kind assistance throughout the years has made it possible for me to study a very interesting area of optics.

I would also like to thank Dr. Kevin Thompson for his time and for answering my questions about nodal aberration theory and for pushing me to use the advanced optical design software for the validation of this research.

I thank Prof. Miguel Alonso for providing helpful feedback to some of the mathematical material presented here.

And I thank Prof. Judith Pipher for participating as my outside reader.

Aaron Bauer and Kyle Fuerschbach have often assisted me when I had questions about CODE V®. They have been willing listeners and have provided helpful feedback to my ideas as I progressed through this research. Thank you.

Finally, I would like to thank the staff of the Department of Optics for their kind assistance throughout the years that I have been a graduate student with the department.

Abstract

It is shown how Shack's vector product (SVP) can be used to express the expansion of the monochromatic wavefront aberration function for rotationally symmetric optical imaging models using the H. H. Hopkins' expansion coefficients. As a result of this research it was discovered that SVP fits naturally within the framework of 2-dimensional Geometric Algebra (GA). It is further shown how SVP can be used to define two-dimensional vectors in terms of the Zernike polynomials. The wavefront aberration function expansion is then expressed in terms of these Zernike vectors. A method for calculating the H. H. Hopkins' expansion coefficients is described, including the calculation of the per surface, sphere/asphere, intrinsic/extrinsic expansion coefficient contributions. Methods for converting the Zernike expansion coefficients into the H. H. Hopkins' expansion coefficients have been developed and are presented in this work. Nodal aberration theory (NAT), now well established in the literature, is presented in terms of these mathematical developments. As a quantitative validation of the NAT predications of an optical system's field dependence, comparisons are made between the full field displays (FFDs) produced by using a double Zernike expansion of the wavefront aberration function and by application of NAT to rotationally symmetric optical imaging systems having one or more surfaces decentered and/or tilted. Validation of these developments is provided using idealized computer models of several space telescopes having one or more optical surfaces decentered and/or tilted. Initial steps toward one approach for extending NAT to include optical imaging systems containing freeform surface shapes defined by Zernike polynomials is also provided.

Contributors and Funding Sources

Table of Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

$\vec{A} \star_\eta \vec{B}$ Shack's vector product of 2D vectors $\vec{A}$ and $\vec{B}$ using η as the reference

axis.

$\vec{A}\vec{B}$ Geometric algebra's geometric vector product of vectors $\vec{A}$ and $\vec{B}$.

$\vec{A} \wedge \vec{B}$ Geometric algebra's wedge (a.k.a. outer) product of vectors $\vec{A}$ and $\vec{B}$.

$\vec{A} \bullet \vec{B}$ Vector dot (a.k.a. inner) product of vectors $\vec{A}$ and $\vec{B}$.

$\vec{A}^*$ Reflection of 2D vector about the x-axis, called the vector conjugate.

$\hat{e}_1$, $\hat{e}_2$ Unit vectors along the positive x- and y-axis directions, respectively.

η Parameter for the selection of either x- or y- axis.

$\lfloor k,l \rfloor$ The smaller of the two numbers k and l.

$\lfloor r \rfloor$ The integer part of the positive real number r.

$\vec{\Sigma}$ Unnormalized 2D offset vector.

$\vec{\sigma}_j$ Nodal Aberration Theory's normalized 2D offset vector.

$U^{m_H, m_\rho}_{n_H, n\rho}$ Expansion coefficient for the double Zernike expansion of the wavefront

aberration function.

$V^m_{n_H, n_\rho}$ Expansion coefficient for the single Zernike expansion of the wavefront

aberration function.

$\left\{ \vec{V} \right\}_{\eta}^{n}$ Shack's vector product of the 2D vector $\vec{V}$ to the nth power using η as the reference axis.

$W_{klm;j}$ Wavefront aberration function expansion coefficient for surface j.

$W_{klm\{M\}}$ Wavefront aberration function expansion coefficients with astigmatism related terms associated with the medial focal surface.

$W_{klm\{S\}}$ Wavefront aberration function expansion coefficients with astigmatism related terms associated with the sagittal focal surface.

$Z_n^{\pm m}$ A Zernike polynomial having n as the highest radial power and $+m$ or $-m$ azimuthal factor. n and m positive integers or zero.

$Z_{n\{\eta\}}^{\pm m}$ A Zernike polynomial using η as the reference axis.

$\vec{Z}_{n\{\eta\}}^{m}$ A Zernike vector composed of a Zernike polynomial for the x- and y-components.

0-P   Zero-to-Peak.

2D   Two Dimensional

3D   Three Dimensional

AFA   Aberration Field Axis

ES   Extrinsic Sphere

EA   Extrinsic Asphere

FFD   Full Field Display

FOV   Field of View

GA   Geometric Algebra

GQ   Gaussian Quadrature

HFOV   Half Field of View

IS   Intrinsic Sphere

IA   Intrinsic Asphere

MCA   Mechanical Coordinate Axis

NAT   Nodal Aberration Theory

OAR   Optical Axis Ray

OPD   Optical Path Difference

OPL   Optical Path Length

RMS   Root Mean Square

RWRS   Real incident Wave shape and Real Surface shape

RWSS   Real incident Wave shape and Spherical Surface shape

SVP   Shack's vector product

SWRS   Sphere incident Wave shape and Real Surface shape

SWSS   Sphere incident Wave shape and Spherical Surface shape

TMA   Three mirror anastigmat

Chapter 1 Introduction

Portions of this Chapter have been taken from the author's article [2] and appear here in a slightly revised form.

This research concerns the development of a bridge between the optical design community's utilization of the H. H. Hopkins expansion of the monochromatic wavefront aberration function and the optical testing community's utilization of a Zernike expansion of the wavefront aberration function. Although the field dependence of the wavefront aberrations has been noted and utilized in the optical design community for many years (hence the development of the Fringe Zernike ordering of the Zernike polynomials), it is less well known or utilized in the optical testing community. Equations are developed in this work that explicitly provide the field dependence of the Zernike polynomial expansion of the wavefront aberration function as well as equations for expressing the Zernike expansion coefficients in terms of the H. H. Hopkins expansion coefficients.

For aberration theorists working in optical design, a common starting point for evaluation of aberrations is a form of the wavefront aberration function expansion written out by H. H. Hopkins [3] in the 1950s. R. V. Shack later wrote the wavefront aberration function expansion in terms of 2-dimensional (2D) pupil and field vectors [4]. This led Shack to develop a vector multiplication that has become known in the optics literature as Shack's vector multiplication and Shack's vector product (SVP) [5]. This vector product is similar to the multiplication of complex numbers [6]. It has also been defined in terms of the Cartesian components of the two vectors involved. SVP of the two vectors has

then been defined by converting the vectors into complex numbers, performing the complex number multiplication, and then converting the result back into a 2D vector. Using geometric algebra (GA) [7, 8, 9, 10] as the 2D vector algebra eliminates any need to convert 2D vectors into complex numbers and vice versa. This is because, as the research to be described in Chapter 2 has demonstrated, SVP is a fundamental construct of 2D GA.

Zernike polynomials [11, 12, 13] have been adopted by both the optical testing and the optical design communities and the supporting fields of optical engineering and optical alignment following the introduction of commercial laser-based interferometers. They are orthogonal and complete over a unit radius circular pupil and they present the balance between multiple orders of the aberrations of H. H. Hopkins in the context of minimizing the RMS wavefront error for optical design. Zernike polynomials provide an excellent metric basis for describing and understanding errors in the shape of an optical surface. In the initial application of Zernike polynomials to the testing of individual optical surfaces there was no motivation to consider field dependence, only aperture dependence was being sought. With the research described here, the field dependence is made explicit and may then find useful application in the testing community.

Researchers have used Zernike polynomial fits to the wavefront computed at a sparse set of field-of-view points in an attempt to characterize the performance of misaligned optical systems. Specifically, McLeod [14] used Zernike polynomials to describe characteristics of 4th order astigmatism within the field of view resulting from misalignment between the primary and secondary mirrors of a Ritchey-Chrétien

telescope. Rakich [15] describes the field dependence of the aberrations using the Burch plate diagram method [16] for simplifying the fourth-order analysis of optical systems and uses Zernike polynomials to resolve individual plate contributions to the system aberration into the coefficients for coma and astigmatism in Zernike terms that arise from misalignments. Noethe and Guisard [17] present measurements of the astigmatic field for the European Southern Observatory (ESO) Very Large Telescope (VLT), using measurements of the coefficients of their Zernike polynomial astigmatic components Z4 and Z5 (Fringe ordering Z5 and Z6). Matsuyama and Ujike [18] have developed "functions that are orthogonal to each other and expressed by a simple combination of Zernike function(s) of pupil coordinates and Zernike function(s) of field coordinates." Kim, et al. [19] used Zernike coefficients to develop a merit function for a telescope alignment scheme. Lee, et al. [20] used Zernike polynomials as an orthogonal basis for decomposing alignment influence functions. Schechter and Levinson [21] studied the fourth- and sixth-order aberration patterns that arise when small misalignments are present in a rotationally symmetric telescope system.

In Chapter 3 of this work the real number form (as opposed to the complex number form) of the Zernike polynomials is defined and it is shown how they can be written in terms of SVP. It is then shown how to define Zernike vectors that will be used to express the wavefront aberration function expansion, a new result of this research. An Appendix (Appendix III) to this Chapter is provided giving the details for the discrete orthogonality relation of the real number Zernike polynomials. The relations developed are used in a Gaussian Quadrature (GQ) procedure (together with optical path length

difference data obtained by real ray tracing for an optical model under consideration) to calculate the wavefront aberration function expansion coefficients. An additional Appendix (Appendix IV) reports on a discovery of this research that the Zernike polynomials can be written as a sum of decentered, rotationally symmetric terms.

Chapter 4 develops and presents the core results of this research. The H. H. Hopkins formulation of the wavefront aberration function used by aberration theorists working in the field of optical design is written in terms of SVP formulated in the context of GA. In this form, the field dependence and the pupil dependence can be separated, and the vector nature of the field dependence is made explicit. It is shown that a subset of the Zernike vectors introduced in Chapter 3 can also be written in terms of SVP. This subset of Zernike vectors is exactly the collection of Zernike vectors needed for expressing the wavefront aberration function of rotationally symmetric optical imaging systems in terms of Zernike vectors. New equations for the wavefront aberration function expansion having the field and pupil parameters separated, referenced to the Sagittal and the Medial focal surfaces, are presented. A double Zernike expansion (in field and in pupil parameters) is also developed. Equations for obtaining the H. H. Hopkins' expansion coefficients from the double Zernike expansion coefficients are determined.

Chapter 5 provides a brief review of nodal aberration theory (NAT) [3]. NAT is used to investigate the field dependence of the wavefront aberration function expansion of rotationally symmetric optical imaging systems that have one or more of its surfaces perturbed (decentered and/or tilted).

Chapter 6 presents validation of the mathematical development developed in the preceding Chapters using 2 and 3-mirror telescope models. This Chapter provides examples of utilizing the developed field dependent equations for rotationally symmetric optical imaging systems that have at least one surface that is decentered or tilted. Quantitative and qualitative comparisons between the field dependent display plots generated by real ray trace data and the double Zernike expansion, and the single Zernike expansion using NAT's sigma offset vectors are presented.

Chapter 7 proposes a new method for utilizing NAT to obtain the field contributions due to freeform optical elements defined using *xy* polynomials and Zernike polynomials. Examples of utilizing the new approach are also provided.

Finally, Chapter 8 provides conclusions and comments on future research directions.

Chapter 2 Shack's Vector Product (SVP) in the Context of Geometric Algebra

In this work, geometric algebra (GA) is used primarily as an alternative to using the complex number algebra for 2D vector calculations. GA is an algebraic system that includes many other algebraic systems that are traditionally considered separately, such as vectors, complex numbers, quaternions, differential forms, Dirac and Pauli spinors, tensors, etc. GA also provides real geometric interpretations of its elements (as opposed to "imaginary" elements). Additionally, GA does not suffer from the deficiencies inherent in the traditional vector algebra. For example, in the traditional vector algebra of Gibbs and Heaviside, the vector cross product is only valid in a 3-dimensional vector space, it makes no notational distinction between vectors and pseudo-vectors (for example, torque, angular momentum and the magnetic field pseudo-vectors), and "does not include a consistent way to represent vector rotations" [22]. Typically, in the vector algebra approach to vector rotations, the vector would first be converted into a complex number representation, then multiplied by a phasor, then converted back into a vector. The matrix formulation of vector rotations again requires objects and an algebra beyond the vector algebra (the matrix algebra) to accomplish vector rotations. This is what the authors mean when they say that the vector algebra does not have a consistent way to represent vector rotations: An additional algebra (the complex number algebra, or matrix algebra) is required. GA provides a single algebra which overcomes all these

deficiencies. Finally, since GA is being used in many disciplines, from computer graphics to theoretical physics, it has the potential to provide a common mathematical system that will aid cross disciplinary studies.

The goals of this Chapter are to provide a brief introduction to GA at the level necessary to understand this work, show how Shack's vector product (SVP) of two 2D vectors, which is central to this work, can be expressed using the GA vector products and to illustrate how calculations involving SVP can be performed. It does not assumed that the reader has been previously exposed to GA. In the next Chapter, it is shown how SVP in the GA form can be used to define a special set of Zernike polynomial vectors. This set of Zernike vectors is exactly the collection of Zernike vectors needed for expressing the expansion of the monochromatic wavefront aberration function for rotationally symmetric optical imaging systems in terms of Zernike vectors. This will be demonstrated in a later Chapter.

2.1 Historical Background

The concept of a vector is now so important and basic to the sciences that it is hard to imagine that there was a time in which vectors, vector algebra and vector calculus were not known. Vectors had to be invented. This is also true of complex numbers, complex number algebra, and complex analysis. A fascinating account of the history and development of vectors can be found in Crowe's book [23]. In 1545, Gerolamo Cardano "…acknowledged the existence of what are now called imaginary numbers" in his book *Ars Magna* [24]. However, it was not until the work of Cauchy (early to mid-1800's) that complex numbers were defined as a pair of real numbers [25]. Hamilton, attempting to
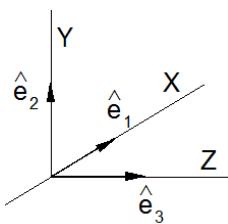
extend the complex number algebra to two and then to three dimensions, discovered and developed his quaternions in 1843 [23, 26]. Although well known, but not universally adopted at the time, quaternions have seen a recent increase in interest largely due to the fact that it is easier and more efficient to program rotations in 3D Euclidean space about an arbitrary axis using quaternions than using the Euler angle formalism [27]. Maxwell used Hamilton's quaternions, and components, in writing out his equations of Electrodynamics. It was principally Gibbs and Heaviside who changed Maxwell's equations into the vector form that we know today. Maxwell never wrote his equations in modern vector form. Vectors were being developed at that time as a reduction of Hamilton's quaternions.

While this development of quaternions and vectors was underway, Hermann Grassmann developed and published an alternative algebra in his 1844 book *Lineale Ausdehnungslehre*. This algebra includes the complex number algebra, Hamilton's quaternion algebra and vector algebra, although this was not appreciated at that time. Grassmann's book was not well received so another, refined, edition was published in 1862 called *Ausdehnungslehre* [28]. But once again, although known, it was not adopted. William K. Clifford further developed Grassmann's work in the 1870s and introduced the term "geometric algebra". "In 1878, one year after Grassmann's death, William K. Clifford (1845-1879) published his 'Applications of Grassmann's extensive algebra'…, in which he successfully unified Grassmann's extensive algebra with Hamilton's quaternion … description of rotations. This was the birth of (Clifford) Geometric Algebra…" [29].

Although additional work in this area of algebra slowly progressed, present day interest did not significantly increase until David Hestenes saw that the algebras used in quantum mechanics (the spinor algebra of Pauli and Dirac) were exactly those of Clifford (Grassmann) algebra [30]. Hestenes continued to develop Clifford (Grassmann) Geometric Algebra, which is now seen to unify complex numbers, quaternions, vectors, differential forms, spinors, as well as Pluker coordinates, homogeneous coordinates, Euclidean geometry, differential geometry, projective geometry, and more, into a unified geometric algebra and calculus for physics and engineering [7]. Recently, a plethora of work in the application of GA has occurred in many areas including: rigid body dynamics (including robotics), computer vision, computer graphics, theoretical and applied physics, computational engineering, and other areas.

2.2 Introduction to Geometric Algebra

For a Euclidean three dimensional vector space, the orthonormal basis vectors are usually denoted in GA as $\hat{e}_1$, $\hat{e}_2$, $\hat{e}_3$.
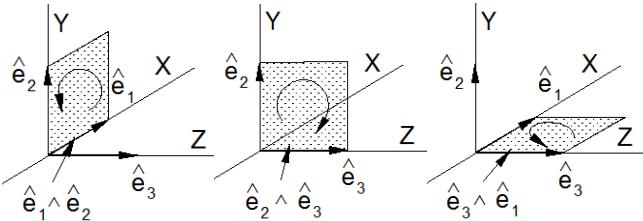
Figure 2.1 Basis vector notation.

The wedge product (also called the outer or exterior product), denoted by "$\wedge$", of two basis vectors is defined by the relations

$$\hat{e}_i \wedge \hat{e}_j = -\hat{e}_j \wedge \hat{e}_i, \ i \neq j, \tag{2.1}$$

$$\hat{e}_i \wedge \hat{e}_j = 0, \ i = j, \tag{2.2}$$

with $i, j = 1, 2, 3$. Note that the new objects $\hat{e}_i \wedge \hat{e}_j$, $i \neq j$ cannot be further reduced. As relation (2.1) shows, the wedge product is anti-commutative. It is also associative and distributive over addition. The new objects, $\hat{e}_1 \wedge \hat{e}_2$, $\hat{e}_2 \wedge \hat{e}_3$, $\hat{e}_3 \wedge \hat{e}_1$, like the unit vectors $\hat{e}_1, \hat{e}_2, \hat{e}_3$ (and the unit scalar 1) are members of the basis set for 3D GA. They represent oriented unit plane elements, analogues to the basis vectors $\hat{e}_1, \hat{e}_2, \hat{e}_3$ which represent oriented unit line elements.
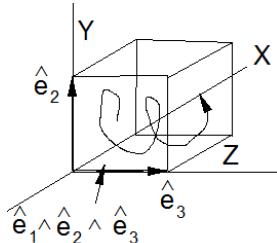
Figure 2.2 The unit area basis elements. The orientation of a plane is denoted in the plane, not perpendicular to the plane.

Additionally, there is a basis object defined by

$$I = \hat{e}_1 \wedge \hat{e}_2 \wedge \hat{e}_3, \tag{2.3}$$

called the pseudo-scalar. It represents the oriented unit volume element.

Figure 2.3 Oriented unit volume basis element: A directional spiral in a unit volume.

The complete basis set for 3D GA is then given by the set

$$\{1,\ \hat{e}_1,\ \hat{e}_2,\ \hat{e}_3,\ \hat{e}_1\wedge\hat{e}_2,\ \hat{e}_2\wedge\hat{e}_3,\ \hat{e}_3\wedge\hat{e}_1,\ \hat{e}_1\wedge\hat{e}_2\wedge\hat{e}_3\}. \tag{2.4}$$

A general element in the 3 dimensional GA can be written as

$$A = s + a\,\hat{e}_1 + b\,\hat{e}_2 + c\,\hat{e}_3 + d\,\hat{e}_1\wedge\hat{e}_2 + f\,\hat{e}_2\wedge\hat{e}_3 + g\,\hat{e}_3\wedge\hat{e}_1 + h\,\hat{e}_1\wedge\hat{e}_2\wedge\hat{e}_3 \ , \tag{2.5}$$

where s, a, b, c, d, f, g, h are real numbers. The s part is called the "scalar" part, the $a\,\hat{e}_1 + b\,\hat{e}_2 + c\,\hat{e}_3$ part is called the "vector" part, the $d\,\hat{e}_1\wedge\hat{e}_2 + f\,\hat{e}_2\wedge\hat{e}_3 + g\,\hat{e}_3\wedge\hat{e}_1$ part is called the "bi-vector" part, and the $h\,\hat{e}_1\wedge\hat{e}_2\wedge\hat{e}_3$ part is called the "tri-vector" part. This nomenclature is derived from the number of unit basis vectors needed to define a single term of the part considered. The object A of Eq. (2.5) is called a multi-vector. This is similar in structure to complex numbers when written in the form $c = real + imaginary$. The complex number c has a real number (scalar) part plus some mathematical object that is not a real number. The whole cannot be further reduced to a real number. In a similar way, objects of GA, for example, $A = 3 + 5\hat{e}_1\wedge\hat{e}_2$, cannot be further reduced to a single real number.

The geometric product, written as a juxtaposition of symbols, of two <u>vectors</u> (but not in general between two multi-vectors) is given by

$$\vec{A}\vec{B} = \vec{A}\bullet\vec{B} + \vec{A}\wedge\vec{B} \ . \tag{2.6}$$

where the "•" is the vector inner product, also called the dot or vector scalar product. Although $\vec{A}$ and $\vec{B}$ are vectors, their geometric product, in general, is a multi-vector. The geometric product is associative, distributive over addition and invertible, but not commutative nor anticommutative. To define the geometric product for arbitrary multi-

vectors requires expanding the definition of the vector dot product for objects which are not simply vectors. The resulting operator is called a "left contraction". Such an extension of the vector dot product is not required in this work. The interested reader is referred to Hestenes [7] and to Dorst, et al. [9] for further details.

Since the dot product between vectors is symmetric $\left(\vec{A}\bullet\vec{B}=\vec{B}\bullet\vec{A}\right)$ and the wedge product is antisymmetric $\left(\vec{A}\wedge\vec{B}=-\vec{B}\wedge\vec{A}\right)$, the following equations are obtained

$$\vec{A}\bullet\vec{B}=\frac{1}{2}\left(\vec{A}\vec{B}+\vec{B}\vec{A}\right),\tag{2.7}$$

$$\vec{A}\wedge\vec{B}=\frac{1}{2}\left(\vec{A}\vec{B}-\vec{B}\vec{A}\right),\tag{2.8}$$

where the juxtaposition of the vectors indicates the geometric product.

Because the basis vectors are orthonormal, so that the dot product between different basis vectors is zero, a shorthand notation is frequently used in the literature for $i\neq j\neq k$

$$\hat{e}_i\wedge\hat{e}_j=\hat{e}_i\,\hat{e}_j=\hat{e}_{ij},\tag{2.9}$$

$$\hat{e}_i\wedge\hat{e}_j\wedge\hat{e}_k=\hat{e}_i\,\hat{e}_j\hat{e}_k=\hat{e}_{ijk}.\tag{2.10}$$

The pseudo-scalar can then be written as

$$I=\hat{e}_1\,\hat{e}_2\hat{e}_3.\tag{2.11}$$

The inverse of the pseudo-scalar is given by

$$I^{-1}=\hat{e}_3\,\hat{e}_2\hat{e}_1.\tag{2.12}$$

Explicitly, this is shown by

$$I\,I^{-1} = \left(\hat{e}_1\,\hat{e}_2\hat{e}_3\right)\left(\hat{e}_3\,\hat{e}_2\hat{e}_1\right) = \hat{e}_1\left(\hat{e}_2\left(\hat{e}_3\hat{e}_3\right)\hat{e}_2\right)\hat{e}_1 = \hat{e}_1\left(\hat{e}_2\,\hat{e}_2\right)\hat{e}_1 = \hat{e}_1\,\hat{e}_1 = 1 \ . \qquad (2.13)$$

Note that for $i \neq j$

$$\hat{e}_i\,\hat{e}_j = -\hat{e}_j\,\hat{e}_i \qquad (2.14)$$

The magnitude of the wedge product of two non-zero vectors is the area of the parallelogram formed by the two vectors. Additionally, for a 3D vector space only, there is a relation between the wedge product and the vector cross product given by

$$\vec{A}\times\vec{B} = \left(\vec{A}\wedge\vec{B}\right)I^{-1} \ , \qquad (2.15)$$

Explicitly, for the nonzero vectors $\vec{A} = A_x\hat{e}_1 + A_y\hat{e}_2 + A_z\hat{e}_3$ and $\vec{B} = B_x\hat{e}_1 + B_y\hat{e}_2 + B_z\hat{e}_3$,

$$\vec{A}\wedge\vec{B} = \left(A_x\hat{e}_1 + A_y\hat{e}_2 + A_z\hat{e}_3\right)\wedge\left(B_x\hat{e}_1 + B_y\hat{e}_2 + B_z\hat{e}_3\right) , \qquad (2.16)$$

$$\begin{aligned}\vec{A}\wedge\vec{B} = A_x\hat{e}_1 \wedge\left(B_x\hat{e}_1 + B_y\hat{e}_2 + B_z\hat{e}_3\right) + A_y\hat{e}_2 \wedge\left(B_x\hat{e}_1 + B_y\hat{e}_2 + B_z\hat{e}_3\right) + \\ A_z\hat{e}_3 \wedge\left(B_x\hat{e}_1 + B_y\hat{e}_2 + B_z\hat{e}_3\right) \end{aligned} \qquad , (2.17)$$

$$\begin{aligned}\vec{A}\wedge\vec{B} = A_xB_x\hat{e}_1 \wedge\hat{e}_1 + A_xB_y\hat{e}_1 \wedge\hat{e}_2 + A_xB_z\hat{e}_1 \wedge\hat{e}_3 + \\ A_yB_x\hat{e}_2 \wedge\hat{e}_1 + A_yB_y\hat{e}_2 \wedge\hat{e}_2 + A_yB_z\hat{e}_2 \wedge\hat{e}_3 + \\ A_zB_x\hat{e}_3 \wedge\hat{e}_1 + A_zB_y\hat{e}_3 \wedge\hat{e}_2 + A_zB_z\hat{e}_3 \wedge\hat{e}_3 \end{aligned} \qquad , (2.18)$$

$$\vec{A}\wedge\vec{B} = A_xB_y\hat{e}_1\hat{e}_2 + A_xB_z\hat{e}_1\hat{e}_3 + A_yB_x\hat{e}_2\hat{e}_1 + A_yB_z\hat{e}_2\hat{e}_3 + A_zB_x\hat{e}_3\hat{e}_1 + A_zB_y\hat{e}_3\hat{e}_2 \ . \qquad (2.19)$$

Multiplying by the inverse pseudo-scalar, gives

$$\left(\vec{A}\wedge\vec{B}\right)I^{-1} = \left(A_xB_y\hat{e}_1\hat{e}_2 + A_xB_z\hat{e}_1\hat{e}_3 + A_yB_x\hat{e}_2\hat{e}_1 + A_yB_z\hat{e}_2\hat{e}_3 + A_zB_x\hat{e}_3\hat{e}_1 + A_zB_y\hat{e}_3\hat{e}_2\right)\hat{e}_3\hat{e}_2\hat{e}_1 , (2.20)$$

$$\begin{aligned}\left(\vec{A}\wedge\vec{B}\right)I^{-1} = A_xB_y\hat{e}_1\hat{e}_2\hat{e}_3\hat{e}_2\hat{e}_1 + A_xB_z\hat{e}_1\hat{e}_3\hat{e}_3\hat{e}_2\hat{e}_1 + A_yB_x\hat{e}_2\hat{e}_1\hat{e}_3\hat{e}_2\hat{e}_1 + A_yB_z\hat{e}_2\hat{e}_3\hat{e}_3\hat{e}_2\hat{e}_1 + \\ A_zB_x\hat{e}_3\hat{e}_1\hat{e}_3\hat{e}_2\hat{e}_1 + A_zB_y\hat{e}_3\hat{e}_2\hat{e}_3\hat{e}_2\hat{e}_1 \end{aligned} \qquad . (2.21)$$

By using Eq. (2.14) this simplifies to

$$\left(\vec{A}\wedge\vec{B}\right)I^{-1}=A_xB_y\hat{e}_3-A_xB_z\hat{e}_2-A_yB_x\hat{e}_3+A_yB_z\hat{e}_1+A_zB_x\hat{e}_2-A_zB_y\hat{e}_1 \ . \tag{2.22}$$

This can be rearranged and written as

$$\left(\vec{A}\wedge\vec{B}\right)I^{-1}=\left(A_yB_z-A_zB_y\right)\hat{e}_1+\left(A_zB_x-A_xB_z\right)\hat{e}_2+\left(A_xB_y-A_yB_x\right)\hat{e}_3 \ , \tag{2.23}$$

which is seen to be the vector cross product of $\vec{A}$ and $\vec{B}$. Therefore, for a 3D vector space only,

$$\left(\vec{A}\wedge\vec{B}\right)I^{-1}=\vec{A}\times\vec{B} \ . \tag{2.24}$$

2.3 Reflection and Rotation Operators

SVP of two 2D vectors results in a third vector, in the plane defined by the two vectors, that is a rotated and scaled version of one of the two vectors involved. Therefore, rotations using 2D GA is here reviewed and illustrated.

The basis set for 2D GA is

$$\{1, \hat{e}_1, \hat{e}_2, \hat{e}_1\wedge\hat{e}_2\}. \tag{2.25}$$

A general multi-vector then has the form

$$V=s+a\hat{e}_1+b\hat{e}_2+c\hat{e}_1\wedge\hat{e}_2, \tag{2.26}$$

where s, a, b, and c are real numbers.

The reflection operation of a vector $\vec{A}$ about a unit vector $\hat{n}$, in the $\hat{e}_1\wedge\hat{e}_2$ plane, is given by the expression

$$\vec{A}'=\hat{n}\,\vec{A}\,\hat{n}. \tag{2.27}$$
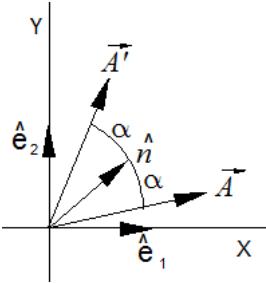
Figure 2.4 Reflection of $\vec{A}$ about the unit vector $\hat{n}$ gives $\vec{A}'$.

As an example, consider the reflection of $\vec{A} = \hat{e}_1$ about the axis line defined by the vector

$\hat{e}_2$. This is written as

$$\vec{A}' = \hat{e}_2 \hat{e}_1 \hat{e}_2. \tag{2.28}$$

Since the geometric product (indicated by the juxtaposition of the vectors) is associative,

this expression can be evaluated in any order. For example

$$\vec{A}' = \left(\hat{e}_2 \hat{e}_1 \right) \hat{e}_2. \tag{2.29}$$

By Eq. (2.9), since the dot product $\hat{e}_2 \cdot \hat{e}_1 = 0$,

$$\vec{A}' = \left(\hat{e}_2 \wedge \hat{e}_1 \right) \hat{e}_2. \tag{2.30}$$

From Eq. (2.1)

$$\vec{A}' = \left(-\hat{e}_1 \wedge \hat{e}_2 \right) \hat{e}_2. \tag{2.31}$$

Using Eq. (2.9) again, in the opposite direction, gives

$$\vec{A}' = \left(-\hat{e}_1 \hat{e}_2 \right) \hat{e}_2. \tag{2.32}$$

By the associativity property, this can be written as

$$\vec{A}' = -\hat{e}_1 \left(\hat{e}_2 \hat{e}_2 \right). \tag{2.33}$$

Finally, using Eq. (2.6) and Eq. (2.2)

$$\vec{A}' = -\hat{e}_1, \tag{2.34}$$

as expected for a reflection of $\hat{e}_1$ about the line defined by $\hat{e}_2$.

Since any rotation can be decomposed into two reflections, any rotation operator acting on a vector $\vec{A}$ can be written as

$$\vec{A}'' = \hat{m}\left(\hat{n}\,\vec{A}\,\hat{n}\right)\hat{m} = \hat{m}\hat{n}\,\vec{A}\,\hat{n}\hat{m} = \left(\hat{m}\hat{n}\right)\vec{A}\left(\hat{n}\hat{m}\right) = R\,\vec{A}\,R^{-1}, \tag{2.35}$$

where $\hat{m}$ and $\hat{n}$ define the two unit vectors about which the reflections are to be performed. R is called a rotor and R^{-1} denotes the inverse of the rotor. That is

$$R\,R^{-1} = \left(\hat{m}\hat{n}\right)\left(\hat{n}\hat{m}\right) = \hat{m}\left(\hat{n}\hat{n}\right)\hat{m} = \hat{m}\hat{m} = 1. \tag{2.36}$$

As Eq. (2.35) shows, rotors are the geometric product of two unit vectors. To better understand rotors, their general form as a multi-vector in a 2 dimensional GA vector space is presented. Let

$$\hat{m} = m_x\,\hat{e}_1 + m_y\,\hat{e}_2, \quad \hat{n} = n_x\,\hat{e}_1 + n_y\,\hat{e}_2 \ , \tag{2.37}$$

where m_x, m_y, n_x, n_y are real numbers with the restriction that $\hat{m}$ and $\hat{n}$ are unit vectors. Then

$$R = \hat{m}\hat{n} = \left(m_x\,\hat{e}_1 + m_y\,\hat{e}_2\right)\left(n_x\,\hat{e}_1 + n_y\,\hat{e}_2\right), \tag{2.38}$$

which can be simplified to

$$R = \left(n_x m_x + n_y\,m_y\right) + \left(n_y\,m_x - n_x\,m_y\right)\hat{e}_1\hat{e}_2 \ . \tag{2.39}$$

Therefore, the rotor can be written as

$$R = \hat{m}\hat{n} = a + bI, \tag{2.40}$$

where $a = \left(n_x m_x + n_y m_y \right)$, $b = \left(n_y m_x - n_x m_y \right)$ and $I = \hat{e}_1 \hat{e}_2 = \hat{e}_1 \wedge \hat{e}_2$ is the pseudo-scalar

in the 2D GA. Note that

$$(I)(I) = \left(\hat{e}_1 \hat{e}_2 \right) \left(\hat{e}_1 \hat{e}_2 \right) = \hat{e}_1 \hat{e}_2 \hat{e}_1 \hat{e}_2 = -\hat{e}_1 \hat{e}_1 \hat{e}_2 \hat{e}_2 = -1. \tag{2.41}$$

Thus the 2D GA includes the complex number algebra.

By analogy with the exponential expansion

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \tag{2.42}$$

and using the geometric product, the following expansion can be written

$$e^{Ix} \equiv \sum_{n=0}^{\infty} \frac{\left(I x \right)^n}{n!} = 1 + I x + \frac{\left(I x \right)^2}{2} + \frac{\left(I x \right)^3}{3!} + \dots . \tag{2.43}$$

Using the property shown in Eq. (2.41) with Eq. (2.43) results in the equations

$$
\begin{aligned}
e^{Ix} &= \left(1 + \frac{I^2 x^2}{2!} + \frac{I^4 x^4}{4!} + \dots \right) + \left(I x + \frac{I^3 x^3}{3!} + \frac{I^5 x^5}{5!} + \dots \right) \\
&= \left(1 - \frac{x^2}{2!} + \dots \right) + I \left(x - \frac{x^3}{3!} + \dots \right) \\
&= \left(\sum_{n=0}^{\infty} (-1)^n \frac{x^{(2n)}}{(2n)!} \right) + I \left(\sum_{n=0}^{\infty} (-1)^n \frac{x^{(2n+1)}}{(2n+1)!} \right).
\end{aligned} \tag{2.44}
$$

Recalling the expansion of the sine and cosine functions leads to

$$e^{Ix} = \cos(x) + I \sin(x). \tag{2.45}$$

This is Euler's equation expressed in 2D GA. One needs to be aware that the pseudo-scalar I is not the same as the complex number $i = \sqrt{-1}$. The pseudo-scalar I is an oriented unit area object. Note that Eq. (2.45) is the same form as the rotor in Eq. (2.40) above. Then let

$$R = e^{I\theta/2}.\tag{2.46}$$

It can be shown that

$$R^{-1} = e^{-I\theta/2},\tag{2.47}$$

so that $RR^{-1} = 1$. A rotor acting on a general 2D vector $\vec{A} = a_x\hat{e}_1 + a_y\hat{e}_2$ is then written as (note the half-angles)

$$R\vec{A}R^{-1} = e^{I\theta/2}\,\vec{A}e^{-I\theta/2}.\tag{2.48}$$

It can be shown that this leads to

$$R\vec{A}R^{-1} = \left(a_x\cos(\theta) + a_y\sin(\theta)\right)\hat{e}_1 + \left(-a_x\sin(\theta) + a_y\cos(\theta)\right)\hat{e}_2,\tag{2.49}$$

and this is recognized as a rotation of the vector $\vec{A}$ in the $I = \hat{e}_1 \wedge \hat{e}_2$ plane by an angular amount θ in the clockwise (x-axis toward $-y$-axis) direction (when $\theta > 0$). Hence the name "rotor" for R.

One important property of 2D GA is that the order of the geometric product of three 2D vectors can be reversed. (This is not necessarily true for vectors of higher dimensional GAs because then the three vectors need not all lay in the same plane.) For example, it can be shown that for 2D vectors

$$\vec{A}\vec{B}\vec{C} = \vec{C}\vec{B}\vec{A}.\tag{2.50}$$

Using this property, it can be shown that rotations in a plane can also be accomplished without having to use the half-angle of the rotation amount. By using Eq. (2.50) it can be shown that

$$R\vec{A} = e^{I\theta/2}\vec{A} = \vec{A}e^{-I\theta/2} = \vec{A}R^{-1},\tag{2.51}$$

and therefore

$$\vec{A}' = R_1\vec{A}R_1^{-1} = \vec{A}R_1^{-1}R_1^{-1} = \vec{A}e^{-I\theta/2}e^{-I\theta/2} = \vec{A}e^{-I\theta} = \vec{A}R_2 .$$ (2.52)

Similarly, it can be shown that

$$\vec{A}' = R_1\,\vec{A}R_1^{-1} = R_1R_1\vec{A} = e^{I\theta/2}e^{I\theta/2}\vec{A} = e^{I\theta}\vec{A} = R_3\vec{A} .$$ (2.53)

Therefore, rotation of a vector in the plane can be accomplished by the application of (the geometric product of) a single rotor on the left or the right of the vector.

Consider a rotor of the form $R = \hat{e}_2\hat{n}$ where $\hat{n}$ is a unit vector in the $I = \hat{e}_1 \wedge \hat{e}_2$ plane.
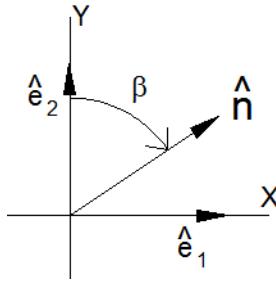
Figure 2.5 An arbitrary unit vector in the $I = \hat{e}_1 \wedge \hat{e}_2$ plane.

Using the angle β as defined in Figure 2.5, we have

$$R = \hat{e}_2\hat{n} = \hat{e}_2\left(n_x\hat{e}_1 + n_y\hat{e}_2\right) = \hat{e}_2\left(\sin(\beta)\hat{e}_1 + \cos(\beta)\hat{e}_2\right)$$
$$= \sin(\beta)\hat{e}_2\hat{e}_1 + \cos(\beta)\hat{e}_2\hat{e}_2 = \cos(\beta) - \sin(\beta)\hat{e}_1\hat{e}_2 ,$$ (2.54)

which gives

$$R = e^{-I\beta} .$$ (2.55)

When used as a single sided rotor on the right of a vector this is a clockwise (x-axis toward $-y$-axis) rotation (assuming β is positive) of the vector in the $I = \hat{e}_1 \wedge \hat{e}_2$ plane by the angular amount β that is the signed angle between $\hat{e}_2$ and $\hat{n}$.

In the next section these properties of GA and rotors are used to develop equations for Shack's vector product.

2.4 Shack's Vector Product

This research makes extensive use of SVP, described as follows. Consider two 2D vectors $\vec{A}$ and $\vec{B}$ in the xy-plane such that vector $\vec{A}$ makes an angle α with respect to the y-axis, and such that vector $\vec{B}$ makes an angle β with respect to the y-axis, both angles being measured in a clockwise direction from the y-axis toward the x-axis. Let the vector $\vec{C}$ be the result of SVP of $\vec{A}$ and $\vec{B}$. Then the vector $\vec{C}$ will have a magnitude equal to the product of the magnitudes of $\vec{A}$ and $\vec{B}$, and will make an angle γ with respect to the y-axis equal to the sum of the angles α and β. This is illustrated in Figure 2.6.
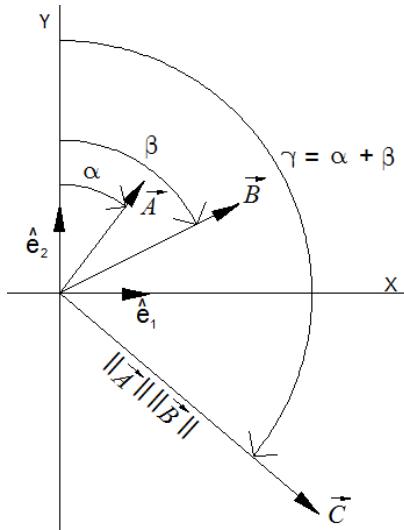
Figure 2.6 Shack's vector product of $\vec{A}$ and $\vec{B}$.

In the Optics literature, SVP is indicated by the juxtaposition of the vectors. This conflicts with the GA geometric product notation. Therefore the symbol "$\star$" will be used to denote SVP in this work. Other notation for SVP will be introduced below as needed.

There are many ways to interpret (and therefore implement) SVP as a procedure (sequence of calculations). Some examples are:

- First rotate the vector $\vec{A}$ by the angular amount β and then multiply the result by the magnitude of vector $\vec{B}$.

- First rotate the vector $\vec{B}$ by the angular amount α and then multiply the result by the magnitude of vector $\vec{A}$.

- First rotate the vector $\hat{e}_2$ by the angular amount α plus the angular amount β and then multiply the result by the magnitude of vector $\vec{A}$ and by the magnitude of vector $\vec{B}$.

- Since any rotation can be decomposed into 2 sequential reflections, first reflect vector $\vec{A}$ about the y-axis (the $\hat{e}_2$ axis) to get a vector $\vec{A}'$. Then reflect $\vec{A}'$ about a unit vector $\hat{n}$ halfway between $\hat{e}_2$ and $\vec{B}$. Finally, multiply the resulting vector by the magnitude of vector $\vec{B}$. This sequence is shown in Figure 2.7 on the following page.

Based on the description of the GA equations for reflections and rotations of 2D vectors provided above, the steps illustrated in Figure 2.7 can be written as the following equation

$$\vec{C} = \vec{A} \star \vec{B} = \left\| \vec{B} \right\| \left(\hat{n} \left(\hat{e}_2 \vec{A} \hat{e}_2 \right) \hat{n} \right), \tag{2.56}$$

where the parentheses have been added for clarity, highlighting the groups of single reflection operations. This is the starting point for all the other formulations of SVP in GA.
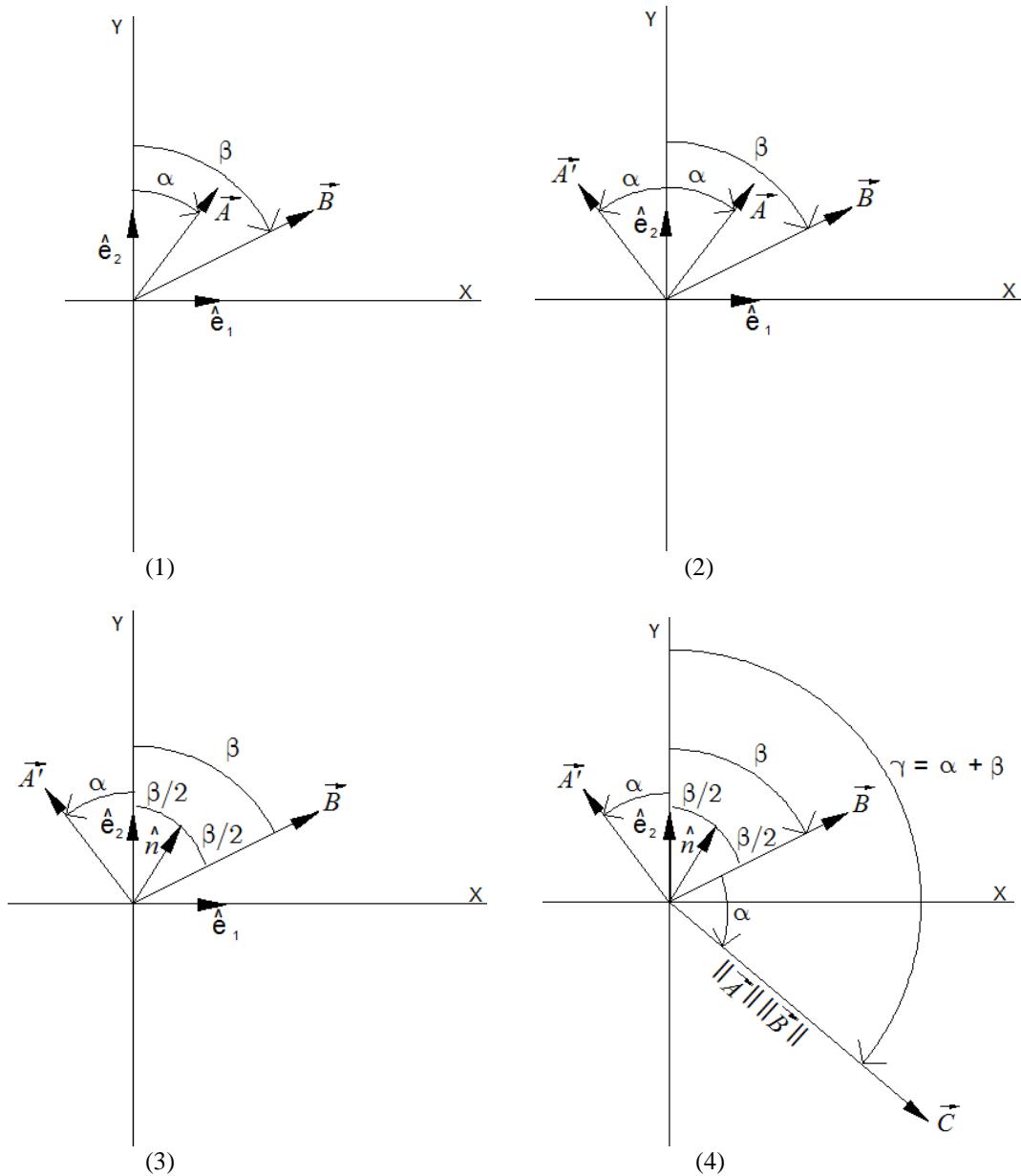
Figure 2.7 **(1) Defining the vectors and their angles. (2)** $\vec{A}$ **is reflected about** $\hat{e}_2$ **to produce** $\vec{A}'$ **. (3) Unit vector** $\hat{n}$ **is defined halfway between** $\hat{e}_2$ **and** $\vec{B}$ **. (4)** $\vec{A}'$ **is reflected about** $\hat{n}$ **and scaled by** $\|\vec{B}\|$ **to produce the resultant vector** $\vec{C}$ **.**

For example, by using a different grouping (recall that the geometric product is associative), SVP can be written as

$$\vec{A} \star \vec{B} = \left\|\vec{B}\right\| (\hat{n}\hat{e}_2) \vec{A} (\hat{e}_2 \hat{n}) = \left\|\vec{B}\right\| R \vec{A} R^{-1},$$ (2.57)

where the rotor R is given by

$$R = \hat{n}\hat{e}_2.$$ (2.58)

Several different expressions and procedural interpretations for SVP are provided in Appendix I. The simplest form developed in this research is

$$\vec{A} \star \vec{B} = \vec{A}\hat{e}_2\vec{B}.$$ (2.59)

Recalling the property shown in Eq. (2.50), that the outer two vectors can be exchanged in the geometric product of three 2D vectors, shows that SVP, Eq. (2.59), is a commutative product,

$$\vec{A} \star \vec{B} = \vec{B} \star \vec{A}.$$ (2.60)

Eq. (2.59) also shows that SVP of two vectors explicitly depends on three vectors, the third vector being a unit vector along the line used as the reference axis. Also, an important property of SVP is that the result is always a 2D vector.

A convention often used in the optical design community is that angles at measured from the positive y-axis toward the positive x-axis. In the optical testing community, the x-axis is often chosen as the reference axis. To accommodate both conventions, the notation and definition of SVP is extended. A subscript on the Shack vector product symbol "$\star$" will be used to indicate which axis is being used as the
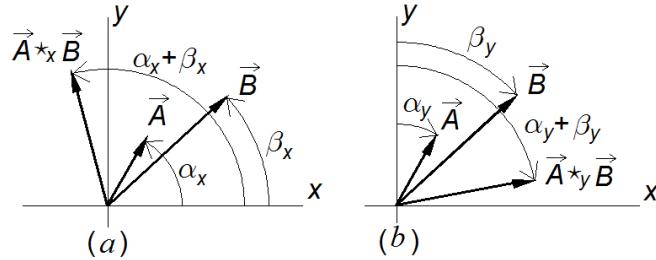
reference axis. Let $\vec{A}=\left(A_x,A_y\right)$ and $\vec{B}=\left(B_x,B_y\right)$ be two arbitrary non-zero 2D vectors.

Then SVP is defined to be

$$\vec{A}\star_x\vec{B}\equiv\vec{A}\hat{e}_1\vec{B}=\left(A_xB_x-A_yB_y,\ A_yB_x+A_xB_y\right)=\vec{C}, \tag{2.61}$$

$$\vec{A}\star_y\vec{B}\equiv\vec{A}\hat{e}_2\vec{B}=\left(A_yB_x+A_xB_y,\ A_yB_y-A_xB_x\right)=\vec{D}, \tag{2.62}$$

where the subscript on the star multiplication symbol indicates which axis is being used as the reference axis. Note that Eq. (2.61) and Eq. (2.62) show that, depending on which axis is being used as the reference axis, SVP of two 2D vectors can produce different results, as Figure 2.8 illustrates.

Figure 2.8 Shack's vector product. (*a*) **graphical representation of** $\vec{A}\star_x\vec{B}$, (*b*) **graphical representation of** $\vec{A}\star_y\vec{B}$.

Also note that Eq. (2.61) and Eq. (2.62) can be used to show that

$$\vec{C}\cdot\vec{D}=0, \tag{2.63}$$

and therefore, for non-zero vectors, $\vec{A}\star_x\vec{B}$ is perpendicular to $\vec{A}\star_y\vec{B}$.

When the two vectors are the same vector, so that $\vec{B}=\vec{A}$, Eq. (2.61) and Eq. (2.62) form the second power or "square" of the vector with respect to SVP. This can be extended to include a vector raised to a positive integer power (or zero). Using curly

braces to indicate powers of a vector with respect to SVP, the following notation is defined

$$\underbrace{\vec{A} \star_x \vec{A} \star_x \cdots \star_x \vec{A}}_{n\ \vec{A}'s} = \underbrace{\vec{A}\hat{e}_1\vec{A}\hat{e}_1\cdots\hat{e}_1\vec{A}}_{n\ \vec{A}'s} \equiv \{\vec{A}\}_x^n = \|\vec{A}\|^n \left(\cos\left(n\alpha_x\right),\ \sin\left(n\alpha_x\right)\right), \qquad (2.64)$$

$$\underbrace{\vec{A} \star_y \vec{A} \star_y \cdots \star_y \vec{A}}_{n\ \vec{A}'s} = \underbrace{\vec{A}\hat{e}_2\vec{A}\hat{e}_2\cdots\hat{e}_2\vec{A}}_{n\ \vec{A}'s} \equiv \{\vec{A}\}_y^n = \|\vec{A}\|^n \left(\sin\left(n\alpha_y\right),\ \cos\left(n\alpha_y\right)\right), \qquad (2.65)$$

where $n \geq 0$ and an integer, $\|\vec{A}\|$ is the magnitude of the vector, and the subscripts x and y indicate which axis is used as the reference axis. To check that these definitions are consistent with Eq. (2.61) and Eq. (2.62), set $n = 2$, and use the trigonometric relations

$$\cos\left(2\theta\right) = \cos^2\left(\theta\right) - \sin^2\left(\theta\right), \qquad (2.66)$$

$$\sin\left(2\theta\right) = 2\cos\left(\theta\right)\sin\left(\theta\right). \qquad (2.67)$$

Then

$$\begin{aligned}
\{\vec{A}\}_x^2 &= \|\vec{A}\|^2 \left(\cos\left(2\alpha_x\right),\ \sin\left(2\alpha_x\right)\right) \\
&= \|\vec{A}\|^2 \left(\cos^2\left(\alpha_x\right) - \sin^2\left(\alpha_x\right),\ 2\cos\left(\alpha_x\right)\sin\left(\alpha_x\right)\right) \\
&= \left(A_x^2 - A_y^2,\ 2A_xA_y\right)
\end{aligned} \qquad (2.68)$$

and this is the same as Eq. (2.61) for the case $\vec{B} = \vec{A}$. Similarly,

$$\begin{aligned}
\{\vec{A}\}_y^2 &= \|\vec{A}\|^2 \left(\sin\left(2\alpha_y\right),\ \cos\left(2\alpha_y\right)\right) \\
&= \|\vec{A}\|^2 \left(2\cos\left(\alpha_y\right)\sin\left(\alpha_y\right),\ \cos^2\left(\alpha_y\right) - \sin^2\left(\alpha_y\right)\right) \\
&= \left(2A_xA_y,\ A_y^2 - A_x^2\right)
\end{aligned} \qquad (2.69)$$

and this is the same as Eq. (2.62) for the case $\vec{B} = \vec{A}$. The definitions in Eq. (2.64) and Eq. (2.65) are therefore consistent with the definitions in Eq. (2.61) and Eq. (2.62).

Further extensions of the definitions of Eq. (2.64) and Eq. (2.65) can be made to define the nth root of a 2D vector with respect to SVP as follows

$$\left\{ \vec{A} \right\}_{x}^{1/n} \equiv \left\| \vec{A} \right\|^{1/n} \left(\cos\left(\alpha_x / n \right), \ \sin\left(\alpha_x / n \right) \right), \tag{2.70}$$

$$\left\{ \vec{A} \right\}_{y}^{1/n} \equiv \left\| \vec{A} \right\|^{1/n} \left(\sin\left(\alpha_y / n \right), \ \cos\left(\alpha_y / n \right) \right), \tag{2.71}$$

where n is a non-zero positive integer.

The cosine of the difference of two angles in terms of SVP has been developed in this research as follows. Consider the trigonometric identity

$$\cos\left(n\alpha_\eta - n\beta_\eta \right) = \cos\left(n\alpha_\eta \right)\cos\left(n\beta_\eta \right) + \sin\left(n\alpha_\eta \right)\sin\left(n\beta_\eta \right), \tag{2.72}$$

where η is either x or y indicating which reference axis is used. This can be written in Cartesian vector component form using the vector dot product as

$$\cos\left(n\alpha_\eta - n\beta_\eta \right) = \left(\cos\left(n\alpha_\eta \right), \ \sin\left(n\alpha_\eta \right) \right) \bullet \left(\cos\left(n\beta_\eta \right), \ \sin\left(n\beta_\eta \right) \right)$$
$$= \left(\sin\left(n\alpha_\eta \right), \ \cos\left(n\alpha_\eta \right) \right) \bullet \left(\sin\left(n\beta_\eta \right), \ \cos\left(n\beta_\eta \right) \right) . \tag{2.73}$$

And, using Eq. (2.64) and Eq. (2.65), this can be written in terms of SVP as

$$\cos\left(n\alpha_\eta - n\beta_\eta \right) = \left\{ \hat{A} \right\}_{\eta}^{n} \bullet \left\{ \hat{B} \right\}_{\eta}^{n} , \tag{2.74}$$

where $\hat{A}$ and $\hat{B}$ are unit vectors having angles α_η and β_η, respectively, with respect to the reference axis η. In writing Eq. (2.74) use of the fact that

$$\left\{ \vec{A} \right\}_{x}^{n} \bullet \left\{ \vec{B} \right\}_{x}^{n} = \left\{ \vec{A} \right\}_{y}^{n} \bullet \left\{ \vec{B} \right\}_{y}^{n} , \tag{2.75}$$

has been employed. Eq. (2.74) is key to writing the wavefront aberration function expansion in terms of SVP.

2.5 Example: The Quadratic Formula for Vectors

As an example of using the above mathematical development, a quadratic formula for vector calculations using SVP is developed. The resulting equation will be used in the chapter on nodal aberration theory (NAT).

For the quadratic equation

$$ax^2 + bx + c = 0, \tag{2.76}$$

the quadratic formula for the (possibly complex) solutions is given by

$$x = \frac{1}{2a}\left[(-b) \pm \sqrt{b^2 - 4ac}\right]. \tag{2.77}$$

An alternative expression exists but will not be considered here [31].

A quadratic equation for 2D vectors using SVP can be written as

$$\vec{a} \star_\eta \left\{\vec{X}\right\}_\eta^2 + \vec{b} \star_\eta \vec{X} + \vec{c} = \vec{0}, \tag{2.78}$$

where $\vec{a}$, $\vec{b}$ and $\vec{c}$ are known constant 2D vectors, and $\vec{0} \equiv (0,0)$ is the 2D zero vector. (Recall that the result of SVP is always a vector, never a scalar.) It will be shown that the vector solutions for the 2D vectors $\vec{X}$ such that Eq. (2.78) holds are given by the formula

$$\vec{X} = \frac{1}{2}\vec{a}^{-1} \star_\eta \left[-\vec{b} \pm \left\{\left\{\vec{b}\right\}_\eta^2 - 4\vec{a} \star_\eta \vec{c}\right\}_\eta^{1/2}\right]. \tag{2.79}$$

In Eq. (2.79) the notation $\vec{a}^{-1}$ is used for the unique vector such that

$$\vec{a} \star_\eta \vec{a}^{-1} = \vec{a}^{-1} \star_\eta \vec{a} = \hat{e}_\eta, \tag{2.80}$$

is the inverse of the non-zero vector $\vec{a}$ with respect to SVP. An explicit expression for $\vec{a}^{-1}$ is given by

$$\vec{a}^{-1} = \begin{cases} \dfrac{1}{\|\vec{a}\|}\left(\cos\left(\theta_x\right), -\sin\left(\theta_x\right)\right) & \eta = x \\[3mm] \dfrac{1}{\|\vec{a}\|}\left(-\sin\left(\theta_y\right), \cos\left(\theta_y\right)\right) & \eta = y \end{cases}, \tag{2.81}$$

where θ_x is the angle that the vector $\vec{a}$ makes with respect to the x-axis, and θ_y is the

angle that the vector $\vec{a}$ makes with respect to the y-axis.

Let

$$\vec{A} = \frac{1}{2}\vec{a}^{-1} \star_\eta \left(-\vec{b}\right), \tag{2.82}$$

and

$$\vec{B} = \frac{1}{2}\vec{a}^{-1} \star_\eta \left\{\left\{\vec{b}\right\}_\eta^2 - 4\vec{a} \star_\eta \vec{c}\right\}_\eta^{1/2}. \tag{2.83}$$

Then the solutions for $\vec{X}$ can be written as

$$\vec{X} = \vec{A} \pm \vec{B}. \tag{2.84}$$

This is demonstrated as follows. Since the geometric product is distributive over

addition,

$$\left\{\vec{X}\right\}_\eta^2 = \left(\vec{A} \pm \vec{B}\right)\hat{e}_\eta\left(\vec{A} \pm \vec{B}\right) = \vec{A}\hat{e}_\eta\vec{A} \pm \vec{B}\hat{e}_\eta\vec{A} \pm \vec{A}\hat{e}_\eta\vec{B} + \vec{B}\hat{e}_\eta\vec{B}. \tag{2.85}$$

Using Eq. (2.50) gives

$$\left\{\vec{X}\right\}_\eta^2 = \vec{A}\hat{e}_\eta\vec{A} \pm 2\vec{A}\hat{e}_\eta\vec{B} + \vec{B}\hat{e}_\eta\vec{B}. \tag{2.86}$$

Putting this into the left hand side of Eq. (2.78) gives

$$\vec{a} \star_\eta \left\{ \vec{X} \right\}_\eta^2 + \vec{b} \star_\eta \vec{X} + \vec{c} = \vec{a} \star_\eta \left(\vec{A}\hat{e}_\eta \vec{A} \pm 2\vec{A}\hat{e}_\eta \vec{B} + \vec{B}\hat{e}_\eta \vec{B} \right) + \vec{b} \star_\eta \vec{X} + \vec{c}$$

$$= \vec{a}\hat{e}_\eta \vec{A}\hat{e}_\eta \vec{A} \pm 2\vec{a}\hat{e}_\eta \vec{A}\hat{e}_\eta \vec{B} + \vec{a}\hat{e}_\eta \vec{B}\hat{e}_\eta \vec{B} + \vec{b} \star_\eta \left(\vec{A} \pm \vec{B} \right) + \vec{c}$$

$$= \vec{a}\hat{e}_\eta \vec{A}\hat{e}_\eta \vec{A} \pm 2\vec{a}\hat{e}_\eta \vec{A}\hat{e}_\eta \vec{B} + \vec{a}\hat{e}_\eta \vec{B}\hat{e}_\eta \vec{B} + \vec{b} \star_\eta \vec{A} \pm \vec{b} \star_\eta \vec{B} + \vec{c} \; . \quad (2.87)$$

Using Eq. (2.82) gives

$$\vec{a} \star_\eta \left\{ \vec{X} \right\}_\eta^2 + \vec{b} \star_\eta \vec{X} + \vec{c} = \vec{a}\hat{e}_\eta \left(\frac{1}{2}\vec{a}^{-1} \star_\eta \left(-\vec{b} \right) \right) \hat{e}_\eta \left(\frac{1}{2}\vec{a}^{-1} \star_\eta \left(-\vec{b} \right) \right)$$

$$\pm 2\vec{a}\hat{e}_\eta \left(\frac{1}{2}\vec{a}^{-1} \star_\eta \left(-\vec{b} \right) \right) \hat{e}_\eta \vec{B}$$

$$+ \vec{a}\hat{e}_\eta \vec{B}\hat{e}_\eta \vec{B} + \vec{b} \star_\eta \left(\frac{1}{2}\vec{a}^{-1} \star_\eta \left(-\vec{b} \right) \right) \pm \vec{b} \star_\eta \vec{B} + \vec{c}. \quad (2.88)$$

Expanding gives

$$\vec{a} \star_\eta \left\{ \vec{X} \right\}_\eta^2 + \vec{b} \star_\eta \vec{X} + \vec{c} = \frac{1}{4}\vec{a}\hat{e}_\eta \vec{a}^{-1}\hat{e}_\eta \vec{b}\hat{e}_\eta \vec{a}^{-1}\hat{e}_\eta \vec{b}$$

$$\mp \vec{a}\hat{e}_\eta \vec{a}^{-1}\hat{e}_\eta \vec{b}\hat{e}_\eta \vec{B}$$

$$+ \vec{a}\hat{e}_\eta \vec{B}\hat{e}_\eta \vec{B} - \frac{1}{2}\vec{b}\hat{e}_\eta \vec{a}^{-1}\hat{e}_\eta \vec{b} \pm \vec{b}\hat{e}_\eta \vec{B} + \vec{c} \; . \quad (2.89)$$

This can be simplified to

$$\vec{a} \star_\eta \left\{ \vec{X} \right\}_\eta^2 + \vec{b} \star_\eta \vec{X} + \vec{c} = \frac{-1}{4}\vec{b}\hat{e}_\eta \vec{a}^{-1}\hat{e}_\eta \vec{b} + \vec{a}\hat{e}_\eta \vec{B}\hat{e}_\eta \vec{B} + \vec{c} \; . \quad (2.90)$$

Now, using Eq. (2.83), $\vec{B}\hat{e}_\eta \vec{B}$ can be written as

$$\vec{B}\hat{e}_\eta \vec{B} = \left\{ \vec{B} \right\}_\eta^2 = \left\{ \frac{1}{2}\vec{a}^{-1} \star_\eta \left\{ \left\{ \vec{b} \right\}_\eta^2 - 4\vec{a} \star_\eta \vec{c} \right\}_\eta^{1/2} \right\}_\eta^2$$

$$= \left[\frac{1}{2}\vec{a}^{-1} \star_\eta \left\{ \left\{ \vec{b} \right\}_\eta^2 - 4\vec{a} \star_\eta \vec{c} \right\}_\eta^{1/2} \right] \star_\eta \left[\frac{1}{2}\vec{a}^{-1} \star_\eta \left\{ \left\{ \vec{b} \right\}_\eta^2 - 4\vec{a} \star_\eta \vec{c} \right\}_\eta^{1/2} \right]$$

$$= \frac{1}{4}\vec{a}^{-1} \star_\eta \vec{a}^{-1} \star_\eta \left(\left\{ \vec{b} \right\}_\eta^2 - 4\vec{a} \star_\eta \vec{c} \right)$$

$$= \frac{1}{4}\vec{a}^{-1}\hat{e}_\eta \vec{a}^{-1}\hat{e}_\eta \left\{ \vec{b} \right\}_\eta^2 - \vec{a}^{-1}\hat{e}_\eta \vec{c} \; . \quad (2.91)$$

Using this result in Eq. (2.90) gives

$$
\begin{aligned}
\vec{a} \star_\eta \left\{\vec{X}\right\}_\eta^2 + \vec{b} \star_\eta \vec{X} + \vec{c} &= \frac{-1}{4}\vec{b}\hat{e}_\eta \vec{a}^{-1}\hat{e}_\eta \vec{b} + \vec{a}\hat{e}_\eta\left(\frac{1}{4}\vec{a}^{-1}\hat{e}_\eta\vec{a}^{-1}\hat{e}_\eta\left\{\vec{b}\right\}_\eta^2 - \vec{a}^{-1}\hat{e}_\eta\vec{c}\right) + \vec{c} \\
&= \frac{-1}{4}\vec{b}\hat{e}_\eta\vec{a}^{-1}\hat{e}_\eta\vec{b} + \frac{1}{4}\vec{a}\hat{e}_\eta\vec{a}^{-1}\hat{e}_\eta\vec{a}^{-1}\hat{e}_\eta\left\{\vec{b}\right\}_\eta^2 - \vec{a}\hat{e}_\eta\vec{a}^{-1}\hat{e}_\eta\vec{c} + \vec{c} \\
&= \frac{-1}{4}\vec{a}^{-1}\hat{e}_\eta\vec{b}\hat{e}_\eta\vec{b} + \frac{1}{4}\vec{a}^{-1}\hat{e}_\eta\vec{b}\hat{e}_\eta\vec{b} - \vec{c} + \vec{c} \\
&= \vec{0}.
\end{aligned}
\tag{2.92}
$$

Therefore,

$$
\vec{X} = \frac{1}{2}\vec{a}^{-1} \star_\eta \left[-\vec{b} \pm \left\{\left\{\vec{b}\right\}_\eta^2 - 4\vec{a} \star_\eta \vec{c}\right\}_\eta^{1/2}\right]
\tag{2.93}
$$

are solutions to the equation

$$
\vec{a} \star_\eta \left\{\vec{X}\right\}_\eta^2 + \vec{b} \star_\eta \vec{X} + \vec{c} = \vec{0}.
\tag{2.94}
$$

For the special case $\vec{a} = \hat{e}_\eta$, Eq. (2.94) can be written as

$$
\left\{\vec{X}\right\}_\eta^2 + \vec{b} \star_\eta \vec{X} + \vec{c} = \vec{0}
\tag{2.95}
$$

with solutions

$$
\vec{X} = \frac{-1}{2}\vec{b} \pm \frac{1}{2}\left\{\left\{\vec{b}\right\}_\eta^2 - 4\vec{c}\right\}_\eta^{1/2}.
\tag{2.96}
$$

Chapter 3 Selected Perspectives on Zernike Polynomials

Zernike polynomials are used extensively in optical design and testing, primarily because they are orthogonal and complete over a unit radius circular disk and they present the balance between multiple orders of H. H. Hopkins' aberrations in the context of minimizing the root mean square (RMS) wavefront error. (A unit radius circular disk may represent a normalized circular optical pupil or a normalized circular optical surface in an optical imaging system.) Zernike polynomials also provide an excellent metric basis for describing and understanding errors in the shape of an optical surface. However, multiple definitions of the Zernike polynomials exist, each having different normalizations and sign conventions, making communication with Zernike polynomials risky due to improper assumptions regarding which convention is being used. In this work, the Zernike polynomials are used to express the pupil and field dependence of the wavefront aberration function expansion and to define the surface shape of a freeform optical element.

In this Chapter, a review of some of the notation conventions used in the literature for specifying Zernike polynomials is presented. Explicit definitions and lists of a few Zernike polynomials using both the zero-to-peak (0-P) and RMS normalizations is given. A discussion and examples of the Fringe ordering compared to other orderings is also presented. Then it is shown how to use the properties of Shack's vector product (SVP), developed in the previous Chapter, to define Zernike vectors that will be used in the definition of the wavefront aberration function expansion in the following Chapter.

Finally, brief comments on the issues of computer implementation and extensions to non-circular pupil shapes are given. Two interesting properties of the Zernike polynomials (orthogonality over a discrete set of points and expressions for the Zernike polynomials as sums of decentered, rotationally symmetric terms) are presented in Appendix III and Appendix IV, respectively.

3.1 Introduction to Zernike Polynomials

The Zernike polynomials were developed by Frits Zernike in 1934 [11, 32]. The original definition of these polynomials used a complex number representation that is still used by many researchers in different branches of optical sciences [11]. In this work, a real number representation commonly used in optical engineering is defined and utilized.

The Zernike polynomials are orthogonal and complete over a unit radius disk [11]. Therefore, they can be used to generate a mathematical expansion of any smooth, continuous function defined over the unit radius disk. These properties of the Zernike polynomials make them ideal for fitting interferometric data as well as surface data of optical elements when dealing with circular shaped pupils and optical elements.

The real circular Zernike polynomials can be written as functions of the form

$$Z_n^{\pm m}(\rho, \varphi) = Z_j(\rho, \varphi) = N_n^m R_n^m(\rho) \begin{cases} \cos(m\varphi) & +m \geq 0 \\ \sin(m\varphi) & -m < 0 \end{cases}, \tag{3.1}$$

where n and m are positive integers such that $0 \leq m \leq n$ and such that $n - m$ is even and m can change in increments or decrements of 2. N_n^m is a normalization constant that may

be different for different values of n and m. $R_n^m(\rho)$ is the radial function factor and is given by [11]

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)!} \rho^{(n-2s)} \; , \tag{3.2}$$

with the radius parameter restricted to $0 \le \rho \le 1$. A useful relation when calculating the norm of the Zernike polynomials is [11]

$$\int_0^1 R_{n_1}^{m_1}(\rho) R_{n_2}^{m_2}(\rho) \rho \, d\rho = \frac{1}{2(n_1+1)} \delta_{n_1 n_2} \delta_{m_1 m_2} \; , \tag{3.3}$$

where

$$\delta_{n_1,n_2} = \begin{cases} 1 & \text{for } n_1 = n_2 \\ 0 & \text{for } n_1 \ne n_2 \end{cases} \; , \tag{3.4}$$

is the Kronecker delta. In some references the m value is allowed to be negative, absorbing the "$\pm$" into m itself. In this case the radial function factor given in (3.2) is defined with the absolute value of m. However, this notation is not universally used in the literature. The azimuthal factors in Eq. (3.1) are written in the form shown to indicate that for a given $m \ne 0$ there is a choice of either a cosine factor or a sine factor in the azimuthal parameter $0 \le \varphi < 2\pi$.

There are two primary normalization conventions in use. These are called the zero-to-peak, 0-P, normalization and the root mean square, RMS, normalization. For the 0-P normalization the normalization constant N_n^m is set to 1. In this case the squared norm of a Zernike polynomial is given by

$$\left|Z_n^{\pm m}\left(\rho,\varphi\right)\right|^2 \equiv \int_0^{2\pi}\int_0^1 Z_n^{\pm m}\left(\rho,\varphi\right)Z_n^{\pm m}\left(\rho,\varphi\right)\rho\,d\rho\,d\varphi = \frac{\left(1+\delta_{0,m}\right)\pi}{2\left(n+1\right)}\ . \tag{3.5}$$

When the wavefront aberration function $W\left(H,\theta,\rho,\varphi\right)$, a function of field parameters H and θ, and pupil parameters ρ and φ, to be defined in the next Chapter, is expanded in terms of the Zernike polynomials over the normalized pupil parameters, the expansion may be written as

$$W = C_1\left(H,\theta\right)Z_0^0\left(\rho,\varphi\right) + C_2\left(H,\theta\right)Z_1^1\left(\rho,\varphi\right) + C_3\left(H,\theta\right)Z_1^{-1}\left(\rho,\varphi\right) + \dots\ , \tag{3.6}$$

where $C_j\left(H,\theta\right)$ are the expansion coefficients. Given an interferogram of the wavefront aberration in the exit pupil of an optical system, the C_j can be approximately determined by a least squares fit of the Zernike polynomials to the interferogram data. The variance of the fit over the unit radius pupil is given by

$$\omega^2 = \int_0^{2\pi}\int_0^1 W^2 P\left(\rho,\phi\right)\rho\,d\rho\,d\phi - \left(\int_0^{2\pi}\int_0^1 W\,P\left(\rho,\phi\right)\rho\,d\rho\,d\phi\right)^2, \tag{3.7}$$

where the probability distribution function for a unit radius circular pupil is

$$P\left(\rho,\phi\right) = \frac{1}{\text{pupil area}} = \frac{1}{\pi}. \tag{3.8}$$

Then, for this 0-P normalization, the variance is calculated to be

$$\omega^2 = \sum_{i=2} \frac{1+\delta_{0,m}}{2\left(n+1\right)} C_i^2. \tag{3.9}$$

Table 3.1 provides a list of 21 Zernike polynomials using the 0-P normalization.

For the RMS normalization convention, employed by Noll [33] and Mahajan, *et al.* [12], the normalization constant is given by

$$N_n^m = \sqrt{\left(2 - \delta_{0,m}\right)\left(n+1\right)}. \tag{3.10}$$

Table 3.1 Examples of 0-P normalized Zernike polynomials up to j=21.

ORDER PARAMETERS			ZERNIKE POLYNOMIAL $Z_n^{\pm m}\left(\rho,\varphi\right) = Z_j\left(\rho,\varphi\right)$
j	n	m	
1	0	0	1
2	1	+1	$\rho\cos\left(\varphi\right)$
3	1	−1	$\rho\sin\left(\varphi\right)$
4	2	+2	$\rho^2\cos\left(2\varphi\right)$
5	2	0	$2\rho^2 - 1$
6	2	−2	$\rho^2\sin\left(2\varphi\right)$
7	3	+3	$\rho^3\cos\left(3\varphi\right)$
8	3	+1	$\left(3\rho^3 - 2\rho\right)\cos\left(\varphi\right)$
9	3	−1	$\left(3\rho^3 - 2\rho\right)\sin\left(\varphi\right)$
10	3	−3	$\rho^3\sin\left(3\varphi\right)$
11	4	+4	$\rho^4\cos\left(4\varphi\right)$
12	4	+2	$\left(4\rho^4 - 3\rho^2\right)\cos\left(2\varphi\right)$
13	4	0	$6\rho^4 - 6\rho^2 + 1$
14	4	−2	$\left(4\rho^4 - 3\rho^2\right)\sin\left(2\varphi\right)$
15	4	−4	$\rho^4\sin\left(4\varphi\right)$
16	5	+5	$\rho^5\cos\left(5\varphi\right)$
17	5	+3	$\left(5\rho^5 - 4\rho^3\right)\cos\left(3\varphi\right)$
18	5	+1	$\left(10\rho^5 - 12\rho^3 + 3\rho\right)\cos\left(\varphi\right)$
19	5	−1	$\left(10\rho^5 - 12\rho^3 + 3\rho\right)\sin\left(\varphi\right)$
20	5	−3	$\left(5\rho^5 - 4\rho^3\right)\sin\left(3\varphi\right)$
21	5	−5	$\rho^5\sin\left(5\varphi\right)$

The squared norm of an RMS normalized Zernike polynomial is

$$\left|Z_n^{\pm m}\left(\rho,\phi\right)\right|^2 \equiv \int\limits_0^{2\pi}\int\limits_0^1 Z_n^{\pm m}\left(\rho,\phi\right) Z_n^{\pm m}\left(\rho,\phi\right) \rho\, d\rho\, d\phi = \pi. \tag{3.11}$$

When the RMS normalized Zernike polynomials are used to expand the wavefront aberration function, as in Eq. (3.6), the variance, Eq. (3.7), of the expansion is calculated to be [12]

$$\omega^2 = \sum_{i=2} C_i^2 \ .$$

(3.12)

Table 3.2, presented on the next page, gives 21 examples of the RMS normalized Zernike polynomials. Note that not only are the normalization constants different, the j-ordering assigned to the Zernike polynomials are different. This set is popular among astronomers as it colocates terms with the same φ-dependence. It was originally introduced by Noll in the context of atmospheric turbulence models [34].

As indicated in Eq. (3.1), there is also the notational convention $Z_j(\rho,\phi)$ used in the literature for specifying the Zernike polynomials. This notation is often employed when there is need for sequentially ordering the Zernike polynomials or a finite subset of the polynomials. One such subset ordering of the Zernike polynomials is called the Fringe Zernike polynomials. Table 3.3 and Table 3.4 provide a list of the Fringe Zernike polynomials. The ISO specification [35] utilizes the ordering label $0 \le j \le 35$ while CODE V®, for example, uses the labeling $1 \le j \le 37$.

Table 3.2 Examples of RMS normalized Zernike polynomials up to j=21.

ORDER PARAMETER			ZERNIKE POLYNOMIAL $Z_n^{\pm m}(\rho,\varphi) = Z_j(\rho,\varphi)$
j	n	m	
1	0	0	1
2	1	+1	$\sqrt{4}\,\rho\cos(\varphi)$
3	1	−1	$\sqrt{4}\,\rho\sin(\varphi)$
4	2	0	$\sqrt{3}\left(2\rho^2-1\right)$
5	2	−2	$\sqrt{6}\,\rho^2\sin(2\varphi)$
6	2	+2	$\sqrt{6}\,\rho^2\cos(2\varphi)$
7	3	−1	$\sqrt{8}\left(3\rho^3-2\rho\right)\sin(\varphi)$
8	3	+1	$\sqrt{8}\left(3\rho^3-2\rho\right)\cos(\varphi)$
9	3	−3	$\sqrt{8}\,\rho^3\sin(3\varphi)$
10	3	+3	$\sqrt{8}\,\rho^3\cos(3\varphi)$
11	4	0	$\sqrt{5}\left(6\rho^4-6\rho^2+1\right)$
12	4	+2	$\sqrt{10}\left(4\rho^4-3\rho^2\right)\cos(2\varphi)$
13	4	−2	$\sqrt{10}\left(4\rho^4-3\rho^2\right)\sin(2\varphi)$
14	4	+4	$\sqrt{10}\,\rho^4\cos(4\varphi)$
15	4	−4	$\sqrt{10}\,\rho^4\sin(4\varphi)$
16	5	+1	$\sqrt{12}\left(10\rho^5-12\rho^3+3\rho\right)\cos(\varphi)$
17	5	−1	$\sqrt{12}\left(10\rho^5-12\rho^3+3\rho\right)\sin(\varphi)$
18	5	+3	$\sqrt{12}\left(5\rho^5-4\rho^3\right)\cos(3\varphi)$
19	5	−3	$\sqrt{12}\left(5\rho^5-4\rho^3\right)\sin(3\varphi)$
20	5	+5	$\sqrt{12}\,\rho^5\cos(5\varphi)$
21	5	−5	$\sqrt{12}\,\rho^5\sin(5\varphi)$

Table 3.3 Fringe Zernike polynomial orderings (0-P normalization).

		CODE V®	ISO	Zernike Polynomial $Z_n^{\pm m}(\rho,\varphi)=Z_j(\rho,\varphi)$
n	*m*	*j*	*j*	
0	0	1	0	1
1	+1	2	1	$\rho\cos(\varphi)$
1	−1	3	2	$\rho\sin(\varphi)$
2	0	4	3	$2\rho^2-1$
2	+2	5	4	$\rho^2\cos(2\varphi)$
2	−2	6	5	$\rho^2\sin(2\varphi)$
3	+1	7	6	$(3\rho^3-2\rho)\cos(\varphi)$
3	−1	8	7	$(3\rho^3-2\rho)\sin(\varphi)$
4	0	9	8	$6\rho^4-6\rho^2+1$
3	+3	10	9	$\rho^3\cos(3\varphi)$
3	−3	11	10	$\rho^3\sin(3\varphi)$
4	+2	12	11	$(4\rho^4-3\rho^2)\cos(2\varphi)$
4	−2	13	12	$(4\rho^4-3\rho^2)\sin(2\varphi)$
5	+1	14	13	$(10\rho^5-12\rho^3+3\rho)\cos(\varphi)$
5	−1	15	14	$(10\rho^5-12\rho^3+3\rho)\sin(\varphi)$

Table 3.4 Continuation of the list of Fringe Zernike polynomials (0-P normalization).

n	m	CODE V ® j	ISO j	Zernike Polynomial $Z_n^{\pm m}(\rho,\varphi)=Z_j(\rho,\varphi)$
6	0	16	15	$20\rho^6-30\rho^4+12\rho^2-1$
4	+4	17	16	$\rho^4\cos(4\varphi)$
4	−4	18	17	$\rho^4\sin(4\varphi)$
5	+3	19	18	$(5\rho^5-4\rho^3)\cos(3\varphi)$
5	−3	20	19	$(5\rho^5-4\rho^3)\sin(3\varphi)$
6	+2	21	20	$(15\rho^6-20\rho^4+6\rho^2)\cos(2\varphi)$
6	−2	22	21	$(15\rho^6-20\rho^4+6\rho^2)\sin(2\varphi)$
7	+1	23	22	$(35\rho^7-60\rho^5+30\rho^3-4\rho)\cos(\varphi)$
7	−1	24	23	$(35\rho^7-60\rho^5+30\rho^3-4\rho)\sin(\varphi)$
8	0	25	24	$70\rho^8-140\rho^6+90\rho^4-20\rho^2+1$
5	+5	26	25	$\rho^5\cos(5\varphi)$
5	−5	27	26	$\rho^5\sin(5\varphi)$
6	+4	28	27	$(6\rho^6-5\rho^4)\cos(4\varphi)$
6	−4	29	28	$(6\rho^6-5\rho^4)\sin(4\varphi)$
7	+3	30	29	$(21\rho^7-30\rho^5+10\rho^3)\cos(3\varphi)$
7	−3	31	30	$(21\rho^7-30\rho^5+10\rho^3)\sin(3\varphi)$
8	+2	32	31	$(56\rho^8-105\rho^6+60\rho^4-10\rho^2)\cos(2\varphi)$
8	−2	33	32	$(56\rho^8-105\rho^6+60\rho^4-10\rho^2)\sin(2\varphi)$
9	+1	34	33	$(126\rho^9-280\rho^7+210\rho^5-60\rho^3+5\rho)\cos(\varphi)$
9	−1	35	34	$(126\rho^9-280\rho^7+210\rho^5-60\rho^3+5\rho)\sin(\varphi)$
10	0	36	35	$252\rho^{10}-630\rho^8+560\rho^6-210\rho^4+30\rho^2-1$
12	0	37	NA	$924\rho^{12}-2772\rho^{10}+3150\rho^8-1680\rho^6+$ $420\rho^4-42\rho^2+1$

In CODE V®, an extension of the Fringe Zernike set is extended through $j=49$.

The Fringe convention is the preferred Zernike polynomial ordering for lens design because it groups terms according to the associated optical wavefront aberration order. When the wavefront aberration function is expanded in terms of Zernike polynomials, the

expansion coefficients are functions of the field parameters (H, θ) [1], each having a

factor of the form H^m, while the Zernike polynomials are functions of the pupil

parameters (ρ, φ), with the pupil dependence of highest exponent value being ρ^n. The

wavefront aberration order for a Zernike expansion term is then given by the sum $m + n$.

As an example, Z_9 of the CODE V® Fringe set has a ρ^4 pupil dependence, whereas

Z_{10}, Z_{11} have a ρ^3 pupil dependence. Z_9 is placed before Z_{10}, Z_{11} in the Fringe

ordering because Z_9 has a H^0 factor in the associated aberration function's field

dependence, whereas Z_{10}, Z_{11} have a H^3 factor in the associated aberration function's

field dependence. In terms of the optical aberration ordering, Z_9 corresponds to a $0 + 4$

$= 4^{\text{th}}$ order aberration whereas Z_{10}, Z_{11} correspond to $3 + 3 = 6^{\text{th}}$ order aberrations.

Therefore, Z_9 is placed before Z_{10}, Z_{11} in the Fringe ordering of the Zernike

polynomials.

The selection of the coordinate axis used for angular reference in optical design is

often selected to be the y-axis. In optical testing the x-axis is often employed as the axis

for angular reference. In Chapter 2, SVP has been defined for both the x- and the y-axis

as reference. It will shortly be shown how to define Zernike vectors in terms of SVP. It

is then necessary to specify which axis is being used as the reference axis for the Zernike

polynomial's azimuthal parameter. Therefore, in this work, a subscript is included on the

angular dependence indicating which reference axis is being used. For example,

$$Z_n^{\pm m}\left(\rho, \varphi_\eta\right) = Z_j\left(\rho, \varphi_\eta\right), \tag{3.13}$$

where $\eta \in \{x, y\}$. When no such axis is explicitly indicated, the *x*-axis will be assumed.

The explicit equation for the radial factor of the Zernike polynomials given in Eq. (3.2) is sufficient for computing low order (low *n* value) Zernike polynomials using standard double precision floating point number representation. But even for moderate values of $n \approx 15$ the factorials can quickly produce very large numbers that, when multiplied by small numbers (the azimuthal factor), can result in numerical inaccuracies in the computation [36]. It is known that the calculation of Zernike polynomials with $n > 25$ can lead to significant numerical errors depending on the way that the radial factor is calculated [37, 38]. The solution is to adapt recurrence relations that do not suffer from such numerical errors even for high order, $n > 50$, Zernike polynomials. A Matlab implementation of the recurrence relation [36, 38] is given in Appendix II. Another approach to the numeric inaccuracy issue is to take advantage of modern computer programming language number representations and computational methods. For example, Matlab® has a symbolic toolbox that includes "variable precision arithmetic" that enables numeric calculations to any specified precision (significant digits below the decimal point). Mathematica® provides for "arbitrary-precision numbers" that provides a similar numeric computational ability to any desired precision. These techniques overcome the numeric accuracy limitations of standard double precision floating point number representation. However, these alternatives tend to be slower in computational speed.

3.2 Defining Zernike Vectors Using SVP

The Zernike polynomials, over a unit radius disk, using the 0-P normalization, Eq. (3.1), can be rewritten as

$$Z_{n\{\eta\}}^{\pm m}\left(\rho,\phi_{\eta}\right) \equiv \mathcal{R}_{n}^{m}\left(\rho\right)\begin{cases} \rho^{m}\cos\left(m\phi_{\eta}\right) & for +m \\ \rho^{m}\sin\left(m\phi_{\eta}\right) & for -m \end{cases}, \tag{3.14}$$

where n and m are positive (or zero) integers, $m \leq n$, $n - m$ is even, and $\mathcal{R}_{n}^{m}\left(\rho\right)$ is given by

$$\mathcal{R}_{n}^{m}(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^{s}(n-s)!}{s!\left(\dfrac{n+m}{2}-s\right)!\left(\dfrac{n-m}{2}-s\right)!}\rho^{n-2s-m}. \tag{3.15}$$

The subscript η is used to indicate which axis is the reference axis. The advantage of pulling out a ρ^{m} factor from the radial component of Eq. (3.1) and including it into the azimuthal factor will be shown below. The "$\pm$" superscript is to be considered a separate parameter and is not to be combined with the integer m. One consequence of this notational convention is that there is now a Zernike polynomial notation for the number zero, $Z_{0\{\eta\}}^{-0}\left(\rho,\varphi_{\eta}\right) \equiv 0$, although zero is not a Zernike polynomial. This is a useful notational extension to be used in the definition of Zernike polynomial vectors that follows.

There are different ways in which Zernike vectors have been defined in the literature [39, 40, 41, 42]. Although the Zernike polynomials are orthogonal to one another, this property is not used in this research for defining the vector basis set for the Zernike vectors. Instead, the $\hat{x}, \hat{y}$ unit vector basis set of the object (image) and entrance

(exit) pupil coordinate planes of the optical system are used. Define a 2D vector consisting of the Zernike polynomials by

$$\vec{Z} = a\, Z_{n_1\{\eta\}}^{\pm m_1}\left(\rho,\phi_\eta\right)\hat{x} + b\, Z_{n_2\{\eta\}}^{\pm m_2}\left(\rho,\phi_\eta\right)\hat{y}, \tag{3.16}$$

where a and b are real constants. Of particular interest to this research is the special subset of Zernike vectors for which $m_1 = m_2 = m$, $n_1 = n_2 = n$, and $a = b = 1$. Denote these Zernike vectors as follows:

$$\vec{Z}_{n\{x\}}^{m}\left(\rho,\phi_x\right) \equiv \left(Z_{n\{x\}}^{+m}\left(\rho,\phi_x\right),\ Z_{n\{x\}}^{-m}\left(\rho,\phi_x\right)\right), \tag{3.17}$$

$$\vec{Z}_{n\{y\}}^{m}\left(\rho,\phi_y\right) \equiv \left(Z_{n\{y\}}^{-m}\left(\rho,\phi_y\right),\ Z_{n\{y\}}^{+m}\left(\rho,\phi_y\right)\right). \tag{3.18}$$

Note that

$$\vec{Z}_{0\{x\}}^{0}\left(\rho,\phi_x\right) \equiv \left(Z_{0\{x\}}^{+0}\left(\rho,\phi_x\right),\ Z_{0\{x\}}^{-0}\left(\rho,\phi_x\right)\right) = \left(1,\ 0\right), \tag{3.19}$$

while

$$\vec{Z}_{0\{y\}}^{0}\left(\rho,\phi_y\right) \equiv \left(Z_{0\{y\}}^{-0}\left(\rho,\phi_y\right),\ Z_{0\{y\}}^{+0}\left(\rho,\phi_y\right)\right) = \left(0,\ 1\right). \tag{3.20}$$

These illustrate the use of the notational convention $Z_{0\{\eta\}}^{-0}\left(\rho,\varphi_\eta\right) \equiv 0$ mentioned above.

Using Eqs. (2.64), (2.65) and (3.14), the Zernike vectors in Eqs. (3.17) and (3.18) can be written in terms of SVP as

$$\vec{Z}_{n\{\eta\}}^{m}\left(\rho,\phi_\eta\right) = \mathcal{R}_n^m(\rho)\{\vec{\rho}\}_\eta^m. \tag{3.21}$$

Table 3.5 provides a list of the Zernike vectors that will be used in the next Chapter to convert the H. H. Hopkins based wavefront aberration function expansion into an expansion in terms of Zernike vectors.

Table 3.5 Examples of 0-P normalized Zernike vectors written in terms of SVP.

Zernike Vector	Zernike Vector in Terms of SVP
$\vec{Z}_{0\{\eta\}}^{0}(\vec{\rho})$	$\{\vec{\rho}\}_{\eta}^{0}$
$\vec{Z}_{1\{\eta\}}^{1}(\vec{\rho})$	$\{\vec{\rho}\}_{\eta}^{1}$
$\vec{Z}_{2\{\eta\}}^{0}(\vec{\rho})$	$(2\rho^{2}-1)\{\vec{\rho}\}_{\eta}^{0}$
$\vec{Z}_{2\{\eta\}}^{2}(\vec{\rho})$	$\{\vec{\rho}\}_{\eta}^{2}$
$\vec{Z}_{3\{\eta\}}^{1}(\vec{\rho})$	$(3\rho^{2}-2)\{\vec{\rho}\}_{\eta}^{1}$
$\vec{Z}_{3\{\eta\}}^{3}(\vec{\rho})$	$\{\vec{\rho}\}_{\eta}^{3}$
$\vec{Z}_{4\{\eta\}}^{0}(\vec{\rho})$	$(6\rho^{4}-6\rho^{2}+1)\{\vec{\rho}\}_{\eta}^{0}$
$\vec{Z}_{4\{\eta\}}^{2}(\vec{\rho})$	$(4\rho^{2}-3)\{\vec{\rho}\}_{\eta}^{2}$
$\vec{Z}_{4\{\eta\}}^{4}(\vec{\rho})$	$\{\vec{\rho}\}_{\eta}^{4}$
$\vec{Z}_{5\{\eta\}}^{1}(\vec{\rho})$	$(10\rho^{4}-12\rho^{2}+3)\{\vec{\rho}\}_{\eta}^{1}$
$\vec{Z}_{5\{\eta\}}^{3}(\vec{\rho})$	$(5\rho^{2}-4)\{\vec{\rho}\}_{\eta}^{3}$
$\vec{Z}_{5\{\eta\}}^{5}(\vec{\rho})$	$\{\vec{\rho}\}_{\eta}^{5}$
$\vec{Z}_{6\{\eta\}}^{0}(\vec{\rho})$	$(20\rho^{6}-30\rho^{4}+12\rho^{2}-1)\{\vec{\rho}\}_{\eta}^{0}$
$\vec{Z}_{6\{\eta\}}^{2}(\vec{\rho})$	$(15\rho^{4}-20\rho^{2}+6)\{\vec{\rho}\}_{\eta}^{2}$
$\vec{Z}_{6\{\eta\}}^{4}(\vec{\rho})$	$(6\rho^{2}-5)\{\vec{\rho}\}_{\eta}^{4}$
$\vec{Z}_{6\{\eta\}}^{6}(\vec{\rho})$	$\{\vec{\rho}\}_{\eta}^{6}$
$\vec{Z}_{8\{\eta\}}^{0}(\vec{\rho})$	$(70\rho^{8}-140\rho^{6}+90\rho^{4}-20\rho^{2}+1)\{\vec{\rho}\}_{\eta}^{0}$
$\vec{Z}_{8\{\eta\}}^{2}(\vec{\rho})$	$(56\rho^{6}-105\rho^{4}+60\rho^{2}-10)\{\vec{\rho}\}_{\eta}^{2}$
$\vec{Z}_{8\{\eta\}}^{4}(\vec{\rho})$	$(28\rho^{4}-42\rho^{2}+15)\{\vec{\rho}\}_{\eta}^{4}$
$\vec{Z}_{8\{\eta\}}^{6}(\vec{\rho})$	$(8\rho^{2}-7)\{\vec{\rho}\}_{\eta}^{6}$
$\vec{Z}_{8\{\eta\}}^{8}(\vec{\rho})$	$\{\vec{\rho}\}_{\eta}^{8}$

3.3 Discrete Orthogonality of Zernike Polynomials

The Zernike polynomials are complete and orthogonal over the unit radius disk. This

means that any continuous function $f(\rho,\varphi)$ defined over the unit radius disk can be

expanded in terms of the Zernike polynomials. The expansion coefficients, for 0-P normalization, are given by

$$C_n^{\pm m} = \frac{2(n+1)}{(1+\delta_{0,m})\pi} \int_0^{2\pi} \int_0^1 f(\rho,\varphi) Z_n^{\pm m}(\rho,\varphi) \rho\, d\rho\, d\varphi \; . \qquad (3.22)$$

However, it is often the case that the continuous function $f(\rho,\varphi)$ is not known in analytic form, but its values are known, or can be determined, at a finite number of points over the disk. In such cases, it is still desirable to know the Zernike expansion coefficients through an upper n and m value to obtain an approximation of the continuous function as a finite series expansion. Such an approximation to the function will be a good representation of the function if the function converges within the n and m upper limits used.

In general, the Zernike polynomials are not orthogonal over a finite, discrete set of points. However, in a 2005 paper [43], Pap and Schipp published a result showing that a finite set of complex number Zernike polynomials are orthogonal over a finite set of discrete points across a unit radius disk. In this way, values for the expansion coefficients can be obtained by special sampling of the function over the discrete set of points.

As mentioned above, the original work presented in [43], as well as in [44, 45], utilized a complex number form of the Zernike polynomials. Appendix III provides a derivation of the discrete orthogonality properties and equations for the real number form of the Zernike polynomials used in this dissertation. The results derived in Appendix III were used in this research as part of a Gaussian quadrature (GQ) method for obtaining the

Zernike expansion coefficients of the wavefront aberration function expansion, to be explained in the following Chapter.

3.4 Zernike Polynomials for Non-Circular Pupils

Optical systems need not have circular shaped pupils. In such cases, it may still be desirable to work with orthogonal polynomials over the non-circular pupil shapes. Zernike polynomials have been used to define sets of orthogonal polynomials over annular, rectangular, elliptical and hexagonal pupil shapes [46, 47]. Each polynomial is then a finite weighted sum of the circular Zernike polynomials. These new polynomials maintain the orthogonality and completeness properties of the circular Zernike polynomials and can therefore be used to fit data from systems having these non-circular pupil shapes. During the research described in this dissertation, a Matlab® program was developed to implement these new polynomials for non-circular pupil shapes. The program, called ZernikeCalc, is available on the MathWork's Matlab® Central File Exchange website [48]. Because the optical models investigated in this research have circular pupils, use of ZernikeCalc was not directly made.

3.5 Zernike Polynomials as a Sum of Decentered Rotationally Symmetric Terms

During the course of this research, it was discovered that the Zernike polynomials can be represented as sums of decentered rotationally symmetric terms. A decentered

rotationally symmetric term is defined here by the sag equation (distance from the *xy*-plane to the surface)

$$z = f(\rho) = a\rho^n \ ,$$ (3.23)

where *a* is a real number, ρ is the radial distance from the rotational symmetry *z*-axis to the surface given by

$$\rho = \sqrt{x^2 + y^2} \ ,$$ (3.24)

and *n* is a positive integer greater than 1. The details are presented in Appendix IV.

Because this research project utilizes Zernike polynomials, decentering of rotationally symmetric optical surfaces, and NAT's method for the mathematical description of the field dependence of the wavefront aberration function for optical imaging systems having decentered surfaces (to be described in Chapter 5), finding that Zernike polynomials can be expressed as a sum of decentered rotationally symmetric terms was an interesting discovery to make. Pursuing this line of research has led to an approach for the application of NAT to optical elements having a freeform surface shape defined by the Zernike polynomials. The mathematical development and examples utilizing NAT with freeform optical surface shapes made possible by expressing the Zernike polynomials as decentered rotationally symmetric terms are presented in Chapter 7.

Chapter 4 SVP, Zernike Vectors and the Wavefront Aberration

Function Expansions

Like the Zernike polynomial definition, there are different conventions and assumptions used in the literature resulting in different definitions of the wavefront aberration function. Each of these satisfies a different need of the researcher employing the definition. Quoting H. H. Hopkins: "Thus one is led to consider a spherical wave-front as free from aberration, and to define the aberration of a wave-front as its departure from any conveniently chosen sphere of reference measured as an optical path length." [3] In this Chapter, the definition of the wavefront aberration function is presented that is used in this research. In presenting the definition, an attempt to explicitly define the terms and assumptions is made. No attempt is made to reconcile the developed definition with any other definition used in the literature. The definition used in this work is that implemented in the optical design software CODE V® and Zemax® that use the real chief ray to define the reference sphere.

Using an extension of H. H. Hopkins' expansion of the wavefront aberration function and his wavefront expansion coefficients as a starting point, it is shown how to convert the power series expansion into a vector expansion using Shack's vector product (SVP). An expansion in terms of Zernike polynomials and in terms of the Zernike vectors, defined in the previous Chapter, is derived. A way to obtain the Hopkins expansion coefficients from the Zernike expansion coefficients is also derived. In this

way, a bridge between the optical design community, using the Hopkins' expansion coefficients, and the optical test and measurement communities, using the Zernike expansion coefficients, can be made.

The method for obtaining the Zernike expansion coefficients used in this research is presented. It utilizes the Gaussian quadrature (GQ) method associated with the discrete Zernike orthogonality property mentioned in the previous Chapter and derived in Appendix III.

Because several different forms of the wavefront aberration function expansion will be developed, a labeling of the expansion coefficients has been used to assist in indicating which expansion is being referred to. The "U" expansion coefficients will be associated with double Zernike expansions of the wavefront aberration function, the "V" expansion coefficients will be associated with the single Zernike expansion, and the "W" expansion coefficients will be associated with expansions involving no Zernike polynomial. It is hoped that this sequential labeling of the expansion coefficients will make the reading of this Chapter easier to follow.

The specific wavefront aberration function expansions to be developed are as follows.

Expansions using W_{klm} coefficients:

1) H. H. Hopkins' Expansion. For rotationally symmetric optical imaging systems. Most often used by optical designers. The field parameter $\vec{H}$ is restricted to be along the y-axis. No Zernike polynomials involved.

2) Shack's Vector Form. For rotationally symmetric optical imaging systems. An extension of H. H. Hopkins' expansion that removes the restriction on the field parameter. No Zernike polynomials involved.

3) Shack Vector Product Form. For rotationally symmetric optical imaging systems. The field vector $\vec{H}$ and pupil vector $\vec{\rho}$ parameters are written in terms of SVP. Field and pupil parameters can be factored by a vector dot product. No Zernike polynomials involved. This form is used to derive the expansions in terms of Zernike polynomials.

Expansion using $V_{k,n}^{m}$ coefficients:

1) An expansion where the pupil vector parameter $\vec{\rho}$ is expanded in terms of Zernike polynomial vectors. For rotationally symmetric optical imaging systems. The $V_{k,n}^{m}$ expansion coefficients can be expressed in terms of the W_{klm} expansion coefficients.
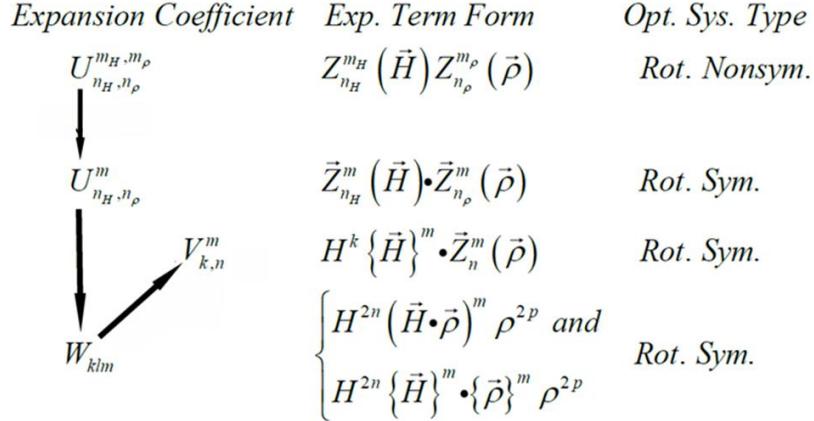
Expansion using U_{n_H,n_ρ}^{m} coefficients:

1) An expansion where the field vector and pupil vector parameters are both expanded in terms of Zernike polynomial vectors. For rotationally symmetric optical imaging systems. The W_{klm} coefficients can be expressed in terms of the U_{n_H,n_ρ}^{m} coefficients.

Expansion using $U_{n_H,n_\rho}^{m_H,m_\rho}$ coefficients:

1) An expansion where the field and pupil parameters are expanded in terms of Zernike polynomials (not Zernike vectors). For rotationally nonsymmetric optical imaging systems. The GQ method uses this form to obtain the $U_{n_H,n_\rho}^{m_H,m_\rho}$ expansion coefficients. The $U_{n_H,n_\rho}^{m_H,m_\rho}$ coefficients reduce to the U_{n_H,n_ρ}^{m} coefficients in the rotationally symmetric optical imaging system case and these U_{n_H,n_ρ}^{m} coefficients are used to calculate the W_{klm} expansion coefficients.

A graphical summary is presented in Figure 4.1.

Expansion Coefficient	Exp. Term Form	Opt. Sys. Type
$U_{n_H,n_\rho}^{m_H,m_\rho}$	$Z_{n_H}^{m_H}(\vec{H})Z_{n_\rho}^{m_\rho}(\vec{\rho})$	Rot. Nonsym.
U_{n_H,n_ρ}^{m}	$\vec{Z}_{n_H}^{m}(\vec{H})\cdot\vec{Z}_{n_\rho}^{m}(\vec{\rho})$	Rot. Sym.
$V_{k,n}^{m}$	$H^{k}\{\vec{H}\}^{m}\cdot\vec{Z}_{n}^{m}(\vec{\rho})$	Rot. Sym.
W_{klm}	$\begin{cases} H^{2n}(\vec{H}\cdot\vec{\rho})^{m}\rho^{2p} \text{ and} \\ H^{2n}\{\vec{H}\}^{m}\cdot\{\vec{\rho}\}^{m}\rho^{2p} \end{cases}$	Rot. Sym.
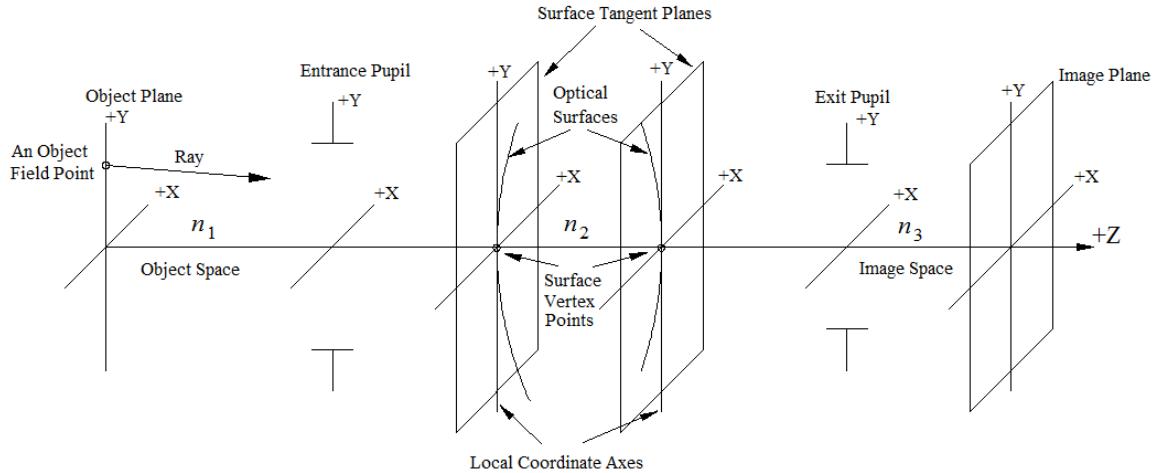
Figure 4.1 Summary of expansion coefficients U, V and W. "*Exp. Term Form*" is an abbreviation for "expansion term form." "*Opt. Sys. Type*" is an abbreviation for "optical system type." "*Rot. Nonsym.*" is an abbreviation for "rotationally nonsymmetric." "*Rot. Sym.*" is an abbreviation for "rotationally symmetric." "•" is the vector dot product. $\{\vec{H}\}^{m}$ and $\{\vec{\rho}\}^{m}$ are SVPs to the mth power.

The details of Figure 4.1 are developed in the following sections.

4.1 The Wavefront Aberration Function Definition

The coordinate systems used in this research are illustrated in Figure 4.2.

Figure 4.2 Optical imaging system's components and local coordinate systems.

The object plane is generally positioned to the left of the other optical system's elements. A ray is emitted from an object point in the object plane and travels to the right toward the entrance pupil. This left to right direction is defined to be the $+z$-axis direction. Each of the elements in the optical model, including the object, entrance pupil, optical surfaces with optical power, exit pupil and image have their own local coordinate systems. These are all centered on the z-axis and form a right handed coordinate system with the $+x$-axis direction considered to be into the page and $+y$-axis direction considered to be "up" or toward the top of the page. A reflective surface does not change the local coordinate system's orientation of any of the surfaces in the optical model. When it is necessary to utilize a global coordinate system from which all the other surfaces' location can be defined, the entrance pupil's coordinate system will be used. For optical imaging models

having the object plane at $z = -\infty$, the entrance pupil will always be to the right of the object plane. For those optical imaging systems having a finite distance between the object surface and the entrance pupil, it may occur that the entrance pupil is to the left of the object plane. Similarly, the exit pupil for an imaging system may occur to the right of the image plane.

The terminology "optical imaging system," "optical system," "optical imaging system model," and "optical model" will all refer to an idealized software model and/or an idealized mathematical model of the real optical imaging system. As part of this idealization away from the real optical imaging system, it will be assumed that diffraction effects can be ignored and that computer ray tracing is sufficient for describing and analyzing the optical system's optical aberrations.

The terminology "paraxial imaging" will refer to ray tracing with rays that are infinitesimally close to the mechanical coordinate axis (MCA) (the symmetry axis for a rotationally symmetric optical imaging system) having infinitesimally small angles with respect to the MCA. "Gaussian imaging" will refer to first order ray tracing [49]. This distinction is not strictly adhered to. Paraxial imaging results in perfect imaging while Gaussian imaging, an extension to finite angles of the infinitesimal angles of paraxial imaging, is an approximation to perfect imaging. "Real ray tracing" will refer to computer ray tracing through an optical model without restrictions to first order idealizations.

The wavefront aberration function has been defined in several different ways in the literature [3, 11, 49]. As an initial definition (meaning without details), it can be

defined to be the optical path difference (OPD) from the ray's wavefront point (labeled *A* in Figure 4.3) when the wavefront is at the system's exit pupil, to the ray's intersection point (labeled *B*) with the image space reference sphere.
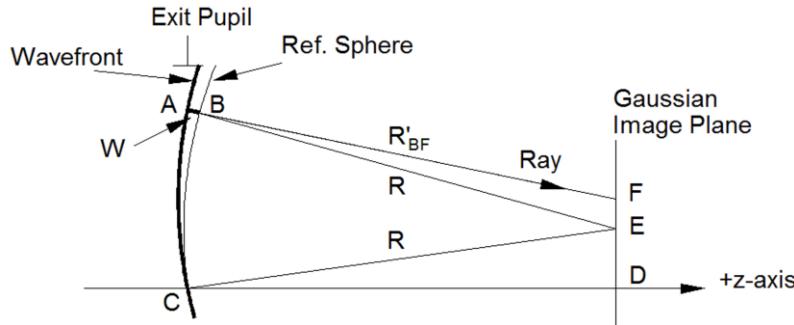
Figure 4.3 Definition of wavefront aberration function *W*.

In Figure 4.3 the point labeled "C" is the intersection point of the optical axis with the image space reference sphere at the center of the exit pupil, the point labeled "D" is the intersection of the optical axis and the Gaussian image plane, the point labeled "E" is the center of the image space reference sphere, the point labeled "F" is the intersection point of the ray being traced with the Gaussian image plane, "R" is the reference sphere's radius, " R'_{BF} " is the ray being traced, and "W" is the OPD for the ray being traced.

Although this initial definition seems simple enough, there are many subtleties that need to be made explicit. For example, there are several choices for the location of the center of the reference sphere as well as choices for how the exit pupil is to be located. Additionally, the parameters of the wavefront aberration function have different meanings depending on the definition used. Therefore, additional details need to be provided.

In this work, the image space reference sphere will be centered at the intersection point of the field point's chief ray and the Gaussian image plane (labeled "E" in Figure 4.3), determined by real ray tracing. The image space reference sphere's surface will be defined to pass through the intersection point of the field point's chief ray and the optical axis ray (OAR) in image space (the z-axis in the case shown in Figure 4.3). This is the center of the real (as opposed to Gaussian) exit pupil. Again, real ray tracing is used to determine this point along the field point's chief ray.

The reason that the professional optical design software packages use this definition for the image space reference sphere is because they are designed to perform far more than just geometrical ray tracing calculations. Specifically, they are designed to perform diffraction related calculations and it is known that diffraction effects are minimal in the real exit pupil.

One consequence of the placement of the image space reference sphere at the intersection of the field point's chief ray with the image plane is that it removes distortion aberrations from the wavefront aberration function even in the case that the final image has significant distortion. This is often acceptable because the distortion aberration is not an image quality reducing aberration and it can be computed independently.

The wavefront aberration function, W, for a monochromatic ray being traced through the optical imaging system model, is a scalar function in four parameters. These four parameters uniquely define the ray that is being traced. Two parameters are associated with the object source point (also called the field point) from which the ray

originates and two parameters specify a point in the optical system through which the ray passes.

There are different ways to specify the two field parameters as well as different conventions for choosing the point in the optical system that the ray passes through. For example, the field parameters, denoted by $\vec{H} = \left(H_x, H_y \right)$, could be the x and y Cartesian coordinates of the object point from which the ray originates, or, when the object is effectively infinitely far away from the rest to the optical system, they could be the angles that a line from the center of the entrance pupil to the object point makes with the z-axis when the line is projected onto the xz- and yz-planes. In this work, the field parameters are taken to be the Cartesian coordinates of the point in the optical system's object plane normalized by the maximum value defining the extent of the object (assumed to be a circle in the object plane). Since these parameters are normalized, the case of the object plane being infinitely far away from the entrance pupil presents no difficulties. An alternative form of the normalized Cartesian coordinates, the polar coordinate form, $\vec{H} = \left(H, \theta_\eta \right)$, where η is either x or y indicating which axis is used as the reference axis, may also be utilized.

The ray intersection point often selected in the optical system for the remaining two (pupil) parameters is the intersection point of the ray with either the (real or Gaussian) entrance pupil plane or the (real or Gaussian) exit pupil plane [11, 49]. Other choices that occur in the literature are the intersection point of the ray being traced with the object space or image space reference spheres [50, 51]. In this work, the Gaussian entrance pupil plane is chosen. The point through which the ray passes in the Gaussian

entrance pupil plane will be denoted by $\vec{\rho} = \left(\rho_x, \, \rho_y \right)$ and these will be called the pupil coordinates of the ray traced. These coordinates are normalized with respect to the pupil's radial extent. Alternatively, a polar form for these normalized pupil coordinates, $\vec{\rho} = \left(\rho, \varphi_\eta \right)$, may be utilized. The wavefront aberration function is then denoted by,

$$W = W\left(\vec{H}, \vec{\rho} \right) = W\left(H_x, H_y, \rho_x, \rho_y \right) = W\left(H, \theta_\eta, \rho, \varphi_\eta \right) . \tag{4.1}$$

The object is idealized and modeled as a finite number of perfect point emitters in the idealized object plane. From each object's field point an expanding spherical wave is imagined. This spherical wavefront, also called the phase front, travels to the Gaussian entrance pupil of the optical imaging system model. The location of the Gaussian entrance pupil plane is determined by paraxial imaging and the pupil will be assumed to have a plane circular shape. Rays are imagined and modeled, in the absence of diffraction, as vectors normal to the wavefront surface. For any object field point there is a unique ray that connects the object field point to the center of the Gaussian entrance pupil. This ray is called the chief ray for that field point, or the field point's chief ray. The sphere centered on the field point and passing through the intersection of the field's chief ray with the center of the Gaussian entrance pupil is called the object space reference sphere for that field point.

A real ray trace is performed through the optical imaging system model from the object plane field point to the Gaussian image plane. The optical path length (OPL) of the ray from the ray's intersection with the object space reference sphere to the image space reference sphere, defined above, is calculated. The optical path difference (OPD)

is then defined to be the OPL of the ray minus the OPL for the field point's chief ray. This means that for any field point's chief ray, the OPD for that chief ray is always zero.

The wavefront aberration function value $W = W\left(\vec{H},\vec{\rho}\right) = W\left(H_x,H_y,\rho_x,\rho_y\right)$ for the ray specified by the normalized parameters $\left(\vec{H},\vec{\rho}\right) = \left(H_x,H_y,\rho_x,\rho_y\right)$ is the OPD value for that ray

$$W\left(H_x,H_y,\rho_x,\rho_y\right) = OPD\left(H_x,H_y,\rho_x,\rho_y\right) . \tag{4.2}$$

There is another aberration used in optical design called the transverse ray aberration. The transverse ray is the vector, in the Gaussian image plane, from the Gaussian image point of the ray to the real ray trace intersection point of the ray with the Gaussian image plane. It is often stated that the relation between the wavefront aberration function and the transverse ray aberration vector function is given by

$$\vec{\epsilon}\left(H_x,H_y,\rho_x,\rho_y\right) = \frac{1}{n'u'}\nabla W\left(H_x,H_y,\rho_x,\rho_y\right) , \tag{4.3}$$

where $\vec{\epsilon}\left(H_x,H_y,\rho_x,\rho_y\right)$ is the transverse ray vector function, n' is the image space index of refraction, u' is the marginal ray's image space angle with respect to the optical axis, $W\left(H_x,H_y,\rho_x,\rho_y\right)$ is the wavefront aberration function and ∇ is the gradient operator with respect to the pupil parameters. This equation is valid for the 4th order wavefront aberration function W but does not hold in general for high order terms of the wavefront expansion [52]. That is, in the derivation of this equation (see for example [49]) the approximations used limit its validity to 3rd order in transverse ray aberrations. It can be shown that Eq. (4.3) leads to inconsistent equations for the 6th order wavefront

aberration expansion coefficients when used with Cox's 5th order equations for the transverse ray aberration function expansion [53]. (See Appendix V for an example.) Also, as pointed out previously, there are several different ways to define the wavefront aberration function. Different meanings of the pupil parameters (defined with respect to entrance or exit pupil planes, or defined with respect to object or image space reference spheres) may not be consistent with the derivation of this equation. Lastly, the selection of which definition of the image space reference sphere to use may result in an incompatibility with the derivation of Eq. (4.3). Therefore, using Eq. (4.3) as the vehicle for obtaining wavefront aberration expansion coefficients, or for work involving higher than 4th order wavefront expansion terms, is perilous.

Because the exact analytic form of the wavefront aberration function for an optical model is not in general known, an expansion of the function in terms of a power series in the ray's four normalized parameters is made. This is written as

$$W\left(H_x, H_y, \rho_x, \rho_y\right) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty} C_{mnpq} H_x^m H_y^n \rho_x^p \rho_y^q \ , \tag{4.4}$$

where C_{mnpq} are the expansion coefficients. It is assumed that this expansion (as well as all other expansions of the wavefront aberration function to be presented) converges. The question of the rate of convergence of this expansion is of particular interest for rotationally non-symmetric optical imaging systems. However, it is an issue not well addressed in the literature.

The z-axis will be used as the MCA. The optical imaging system models to be considered will be restricted to be rotationally symmetric about the MCA in the absence

of any decenters and/or tilts of the model's surfaces. Consequently, the form of the wavefront aberration function consists of only terms that are rotationally invariant to a rotation about the symmetry axis (i.e., MCA). The terms in the expansion Eq. (4.4) are then restricted to be any one or multiples (including powers) of the following four rotationally invariant forms

$$\left(H_x^2 + H_y^2 \right) = \vec{H} \cdot \vec{H} \ , \tag{4.5}$$

$$\left(\rho_x^2 + \rho_y^2 \right) = \vec{\rho} \cdot \vec{\rho} \ , \tag{4.6}$$

$$\left(H_x \rho_x + H_y \rho_y \right) = \vec{H} \cdot \vec{\rho} \ , \tag{4.7}$$

$$\left(H_x \rho_y - H_y \rho_x \right) = \left[\vec{H} \times \vec{\rho} \right]_Z \ , \tag{4.8}$$

where $\left[\cdot \right]_Z$ indicates the z-component of the vector enclosed. Further requiring the wavefront aberration function to have reflection symmetry through a meridional plane (a plane containing the field point and the rotational symmetry z-axis) rules out Eq. (4.8). That is, the wavefront aberration function is not to change sign across the plane containing the field point and the z-axis. The resulting form of the wavefront aberration function's expansion can then be written as

$$W\left(\vec{H}, \vec{\rho} \right) = \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} \sum_{m=0}^{\infty} W_{klm\{S\}} \left(\vec{H} \cdot \vec{H} \right)^n \left(\vec{H} \cdot \vec{\rho} \right)^m \left(\vec{\rho} \cdot \vec{\rho} \right)^p \ , \tag{4.9}$$

where $W_{klm\{S\}}$ are the expansion coefficients with $k = 2n + m, \, l = 2p + m$. The notation "$\{S\}$" is used to indicate that an additional "S" subscript is appended to W_{klm} when the W_{klm} coefficient is associated with the sagittal focal surface. This occurs when, for a W_{klm}, both $n \neq 0$ and $p \neq 0$.

Because the optical system is restricted to be rotationally symmetric, the field point, in optical design, is often selected to be along the positive y-axis. Using the y-axis as the reference axis from which angles are measured, the field parameters may be written as $\vec{H} = \left(H, \theta_y \right)$ with $\theta_y = 0$. Similarly, the pupil parameters may be written as a vector in the plane of the entrance pupil, $\vec{\rho} = \left(\rho, \varphi_y \right)$. Then the wavefront aberration function can be written in the form

$$W\left(\vec{H}, \vec{\rho}\right) = \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} \sum_{m=0}^{\infty} W_{klm\{S\}} H^{2n+m} \rho^{2p+m} \cos^m\left(\varphi_y\right), \tag{4.10}$$

where H is the magnitude of the normalized field parameter vector $\vec{H}$, ρ is the magnitude of the normalized pupil parameter vector $\vec{\rho}$ and φ_y is the angle that the pupil parameter vector makes with respect to the positive y-axis. This is the traditional form presented for the expansion of the wavefront aberration function for rotationally symmetric optical imaging systems.

To explore the full field parameter dependence of the aberration function the restriction that the field point selected be along the y-axis is removed. Then from Eq. (4.9)

$$W\left(\vec{H}, \vec{\rho}\right) = \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} \sum_{m=0}^{\infty} W_{klm\{S\}} H^{2n+m} \rho^{2p+m} \cos^m\left(\theta_\eta - \varphi_\eta\right) \tag{4.11}$$

where $\theta_\eta - \varphi_\eta$ is the angle, in the xy-plane, between the field parameter vector $\vec{H}$ and the pupil parameter $\vec{\rho}$, and the subscript $\eta \in \{x, y\}$ is used to indicate which convention is

used as the reference axis, either the *x*-axis or the *y*-axis, from which the angles are measured. This is a generalization of the form presented in Hopkins [3].

The wavefront aberration order of a term in Eq. (4.11) is given by the sum

$$Order = k + l = 2n + 2p + 2m = 2(n + p + m),$$
<div align="right">(4.12)</div>

and is therefore always even.

For clarity, it is stated again that the field parameters used in this work are the normalized Cartesian coordinates of the object point of the ray under consideration.

4.2 Wavefront Aberration Function in Terms of SVP

It is most desirable to write the wavefront aberration function's terms in such a way that the two field parameters are factored (separated except by multiplication) from the two pupil parameters. As the function is written in Eq. (4.11), the factor $\cos^m\left(\theta_\eta - \varphi_\eta\right)$ appears to prevent this. However, it has been found in this research that this separation of field and pupil parameters can be accomplished by using the trigonometry relation

$$\cos^m\left(\theta_\eta - \varphi_\eta\right) = \frac{1}{2^m}\begin{pmatrix} m \\ \dfrac{m}{2} \end{pmatrix}\delta_{E(m)} + \frac{1}{2^{m-1}}\sum_{t=0}^{\frac{1}{2}\left(m-1-\delta_{E(m)}\right)}\begin{pmatrix} m \\ t \end{pmatrix}\cos\left((m-2t)\left(\theta_\eta - \varphi_\eta\right)\right),$$
<div align="right">(4.13)</div>

where

$$\delta_{E(m)} \equiv \begin{cases} 1 & \textit{for m even} \\ 0 & \textit{otherwise} \end{cases}.$$
<div align="right">(4.14)</div>

Writing

$$H^m = H^{m-2t}H^{2t},$$
<div align="right">(4.15)</div>

$$\rho^m = \rho^{m-2t}\rho^{2t}, \qquad (4.16)$$

the following relation is obtained,

$$H^m\rho^m\cos^m\left(\theta_\eta-\varphi_\eta\right)=\frac{1}{2^m}\begin{pmatrix}m\\ \frac{m}{2}\end{pmatrix}H^m\rho^m\delta_{E(m)}$$

$$+\frac{1}{2^{m-1}}\sum_{t=0}^{\frac{1}{2}\left(m-1-\delta_{E(m)}\right)}\begin{pmatrix}m\\ t\end{pmatrix}H^{2t}\rho^{2t}H^{m-2t}\rho^{m-2t}\cos\left(\left(m-2t\right)\theta_\eta-\left(m-2t\right)\varphi_\eta\right). \qquad (4.17)$$

Recalling Eq. (2.74), repeated here for convenience,

$$\cos\left(n\alpha_\eta-n\beta_\eta\right)=\left\{\hat{A}\right\}_\eta^n\bullet\left\{\hat{B}\right\}_\eta^n, \qquad (4.18)$$

the cosine terms in Eq. (4.17) can be written as

$$H^{m-2t}\rho^{m-2t}\cos\left(\left(m-2t\right)\theta_\eta-\left(m-2t\right)\varphi_\eta\right)=\left\{\vec{H}\right\}_\eta^{m-2t}\bullet\left\{\vec{\rho}\right\}_\eta^{m-2t}. \qquad (4.19)$$

Then

$$H^m\rho^m\cos^m\left(\theta_\eta-\varphi_\eta\right)=\frac{1}{2^m}\begin{pmatrix}m\\ \frac{m}{2}\end{pmatrix}H^m\left\{\vec{H}\right\}_\eta^0\bullet\left\{\vec{\rho}\right\}_\eta^0\rho^m\delta_{E(m)}$$

$$+\frac{1}{2^{m-1}}\sum_{t=0}^{\frac{1}{2}\left(m-1-\delta_{E(m)}\right)}\begin{pmatrix}m\\ t\end{pmatrix}H^{2t}\rho^{2t}\left\{\vec{H}\right\}_\eta^{m-2t}\bullet\left\{\vec{\rho}\right\}_\eta^{m-2t}, \qquad (4.20)$$

where the relation $\left\{\vec{H}\right\}_\eta^0\bullet\left\{\vec{\rho}\right\}_\eta^0=1$ has been used. This result can be used to write Eq.

(4.11) as

$$W\left(\vec{H},\vec{\rho}\right)=\sum_{n=0}^\infty\sum_{p=0}^\infty\sum_{m=0}^\infty\left[W_{klm}\frac{1}{2^m}\begin{pmatrix}m\\ \frac{m}{2}\end{pmatrix}H^{2n+m}\left\{\vec{H}\right\}_\eta^0\bullet\left\{\vec{\rho}\right\}_\eta^0\rho^{2p+m}\delta_{E(m)}\right.$$

$$\left.+W_{klm\{S\}}\frac{1}{2^{m-1}}\sum_{t=0}^{\frac{1}{2}\left(m-1-\delta_{E(m)}\right)}\begin{pmatrix}m\\ t\end{pmatrix}H^{2(t+n)}\left\{\vec{H}\right\}_\eta^{m-2t}\bullet\left\{\vec{\rho}\right\}_\eta^{m-2t}\rho^{2(t+p)}\right], \qquad (4.21)$$

where the field and pupil parameters have been completely separated by a vector dot product for all orders of the expansion.

Keeping only through 8th order terms, the wavefront aberration function expansion can be written as follows.

$$
\begin{aligned}
W = {}& W_{000}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0} + W_{200}H^{2}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0} + W_{020}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{2} + W_{111}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1} + \\
& W_{400}H^{4}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0} + W_{040}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{4} + W_{131}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{2} + \frac{1}{2}W_{222}H^{2}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{2} + \\
& \frac{1}{2}W_{222}\left\{\vec{H}\right\}_{\eta}^{2}\bullet\left\{\vec{\rho}\right\}_{\eta}^{2} + W_{220S}H^{2}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{2} + W_{311}H^{2}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1} + \\
& W_{600}H^{6}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0} + W_{060}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{6} + W_{151}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{4} + \frac{1}{2}W_{242}\left\{\vec{H}\right\}_{\eta}^{2}\bullet\left\{\vec{\rho}\right\}_{\eta}^{2}\rho^{2} + \\
& \frac{1}{2}W_{242}H^{2}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{4} + W_{240S}H^{2}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{4} + W_{331}H^{2}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{2} + \\
& \frac{1}{4}W_{333}\left\{\vec{H}\right\}_{\eta}^{3}\bullet\left\{\vec{\rho}\right\}_{\eta}^{3} + \frac{3}{4}W_{333}H^{2}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{2} + \frac{1}{2}W_{422}H^{2}\left\{\vec{H}\right\}_{\eta}^{2}\bullet\left\{\vec{\rho}\right\}_{\eta}^{2} + \\
& \frac{1}{2}W_{422}H^{4}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{2} + W_{420S}H^{4}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{2} + W_{511}H^{4}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1} + \\
& W_{800}H^{8}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0} + W_{080}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{8} + W_{171}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{6} + \frac{1}{2}W_{262}\left\{\vec{H}\right\}_{\eta}^{2}\bullet\left\{\vec{\rho}\right\}_{\eta}^{2}\rho^{4} + \\
& \frac{1}{2}W_{262}H^{2}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{6} + W_{260S}H^{2}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{6} + W_{351S}H^{2}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{4} + \\
& \frac{1}{4}W_{353}\left\{\vec{H}\right\}_{\eta}^{3}\bullet\left\{\vec{\rho}\right\}_{\eta}^{3}\rho^{2} + \frac{3}{4}W_{353}H^{2}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{4} + W_{440S}H^{4}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{4} + \\
& \frac{1}{2}W_{442S}H^{2}\left\{\vec{H}\right\}_{\eta}^{2}\bullet\left\{\vec{\rho}\right\}_{\eta}^{2}\rho^{2} + \frac{1}{2}W_{442S}H^{4}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{4} + \frac{1}{8}W_{444}\left\{\vec{H}\right\}_{\eta}^{4}\bullet\left\{\vec{\rho}\right\}_{\eta}^{4} + \\
& \frac{1}{2}W_{444}H^{2}\left\{\vec{H}\right\}_{\eta}^{2}\bullet\left\{\vec{\rho}\right\}_{\eta}^{2}\rho^{2} + \frac{3}{8}W_{444}H^{4}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{4} + W_{531S}H^{4}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{2} + \\
& \frac{1}{4}W_{533}H^{2}\left\{\vec{H}\right\}_{\eta}^{3}\bullet\left\{\vec{\rho}\right\}_{\eta}^{3} + \frac{3}{4}W_{533}H^{4}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}\rho^{2} + W_{620S}H^{6}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{2} + \\
& \frac{1}{2}W_{622}H^{4}\left\{\vec{H}\right\}_{\eta}^{2}\bullet\left\{\vec{\rho}\right\}_{\eta}^{2} + \frac{1}{2}W_{622}H^{6}\left\{\vec{H}\right\}_{\eta}^{0}\bullet\left\{\vec{\rho}\right\}_{\eta}^{0}\rho^{2} + W_{711}H^{6}\left\{\vec{H}\right\}_{\eta}^{1}\bullet\left\{\vec{\rho}\right\}_{\eta}^{1}.
\end{aligned}
\tag{4.22}
$$

Collecting like terms in $\left\{\vec{H}\right\}_{\eta}^{u}\bullet\left\{\vec{\rho}\right\}_{\eta}^{u}\rho^{v}$ results in

$$W = \left(W_{000} + W_{200}H^2 + W_{400}H^4 + W_{600}H^6 + W_{800}H^8 \right) \left\{ \vec{H} \right\}_\eta^0 \cdot \left\{ \vec{\rho} \right\}_\eta^0 +$$

$$\left(W_{111} + W_{311}H^2 + W_{511}H^4 + W_{711}H^6 \right) \left\{ \vec{H} \right\}_\eta^1 \cdot \left\{ \vec{\rho} \right\}_\eta^1 +$$

$$\left(W_{020} + \frac{1}{2}W_{222}H^2 + W_{220S}H^2 + \frac{1}{2}W_{422}H^4 + W_{420S}H^4 + W_{620S}H^6 + \frac{1}{2}W_{622}H^6 \right) \left\{ \vec{H} \right\}_\eta^0 \cdot \left\{ \vec{\rho} \right\}_\eta^0 \rho^2 +$$

$$\left(\frac{1}{2}W_{222} + \frac{1}{2}W_{422}H^2 + \frac{1}{2}W_{622}H^4 \right) \left\{ \vec{H} \right\}_\eta^2 \cdot \left\{ \vec{\rho} \right\}_\eta^2 +$$

$$\left(W_{131} + W_{331S}H^2 + \frac{3}{4}W_{333}H^2 + W_{531S}H^4 + \frac{3}{4}W_{533}H^4 \right) \left\{ \vec{H} \right\}_\eta^1 \cdot \left\{ \vec{\rho} \right\}_\eta^1 \rho^2 +$$

$$\left(\frac{1}{4}W_{333} + \frac{1}{4}W_{533}H^2 \right) \left\{ \vec{H} \right\}_\eta^3 \cdot \left\{ \vec{\rho} \right\}_\eta^3 +$$

$$\left(W_{040} + \frac{1}{2}W_{242}H^2 + W_{240S}H^2 + W_{440S}H^4 + \frac{1}{2}W_{442S}H^4 + \frac{3}{8}W_{444}H^4 \right) \left\{ \vec{H} \right\}_\eta^0 \cdot \left\{ \vec{\rho} \right\}_\eta^0 \rho^4 +$$

$$\left(\frac{1}{2}W_{242} + \frac{1}{2}W_{442S}H^2 + \frac{1}{2}W_{444}H^2 \right) \left\{ \vec{H} \right\}_\eta^2 \cdot \left\{ \vec{\rho} \right\}_\eta^2 \rho^2 + \frac{1}{8}W_{444} \left\{ \vec{H} \right\}_\eta^4 \cdot \left\{ \vec{\rho} \right\}_\eta^4 + \frac{1}{4}W_{353} \left\{ \vec{H} \right\}_\eta^3 \cdot \left\{ \vec{\rho} \right\}_\eta^3 \rho^2 +$$

$$\left(W_{151} + W_{351S}H^2 + \frac{3}{4}W_{353}H^2 \right) \left\{ \vec{H} \right\}_\eta^1 \cdot \left\{ \vec{\rho} \right\}_\eta^1 \rho^4 + \left(W_{060} + W_{260}H^2 + \frac{1}{2}W_{262}H^2 \right) \left\{ \vec{H} \right\}_\eta^0 \cdot \left\{ \vec{\rho} \right\}_\eta^0 \rho^6 +$$

$$\frac{1}{2}W_{262} \left\{ \vec{H} \right\}_\eta^2 \cdot \left\{ \vec{\rho} \right\}_\eta^2 \rho^4 + W_{171} \left\{ \vec{H} \right\}_\eta^1 \cdot \left\{ \vec{\rho} \right\}_\eta^1 \rho^6 + W_{080} \left\{ \vec{H} \right\}_\eta^0 \cdot \left\{ \vec{\rho} \right\}_\eta^0 \rho^8 . \tag{4.23}$$

By using the definitions [54]

$$W_{220M} \equiv W_{220S} + \frac{1}{2}W_{222} , \tag{4.24}$$

$$W_{240M} \equiv W_{240S} + \frac{1}{2}W_{242} , \tag{4.25}$$

$$W_{260M} \equiv W_{260S} + \frac{1}{2}W_{262} , \tag{4.26}$$

$$W_{331M} \equiv W_{331S} + \frac{3}{4}W_{333} , \tag{4.27}$$

$$W_{351M} \equiv W_{351S} + \frac{3}{4}W_{353} , \tag{4.28}$$

$$W_{420M} \equiv W_{420S} + \frac{1}{2}W_{422} , \tag{4.29}$$

$$W_{440M} \equiv W_{440S} + \frac{1}{2}W_{442S} + \frac{3}{8}W_{444} , \qquad (4.30)$$

$$W_{442M} \equiv W_{442S} + W_{444} , \qquad (4.31)$$

$$W_{531M} \equiv W_{531S} + \frac{3}{4}W_{533} , \qquad (4.32)$$

$$W_{620M} \equiv W_{620S} + \frac{1}{2}W_{622} , \qquad (4.33)$$

where the subscript "M" is used to indicate quantities associated with the medial focal surface, the wavefront imaging aberration function through 8th order can be written as

$$
\begin{aligned}
W = & \left(W_{000} + W_{200}H^2 + W_{400}H^4 + W_{600}H^6 + W_{800}H^8\right)\{\vec{H}\}_\eta^0 \bullet \{\vec{\rho}\}_\eta^0 + \\
& \left(W_{111} + W_{311}H^2 + W_{511}H^4 + W_{711}H^6\right)\{\vec{H}\}_\eta^1 \bullet \{\vec{\rho}\}_\eta^1 + \\
& \left(W_{020} + W_{220M}H^2 + W_{420M}H^4 + W_{620M}H^6\right)\{\vec{H}\}_\eta^0 \bullet \{\vec{\rho}\}_\eta^0 \rho^2 + \\
& \left(\frac{1}{2}W_{222} + \frac{1}{2}W_{422}H^2 + \frac{1}{2}W_{622}H^4\right)\{\vec{H}\}_\eta^2 \bullet \{\vec{\rho}\}_\eta^2 + \left(W_{131} + W_{331M}H^2 + W_{531M}H^4\right)\{\vec{H}\}_\eta^1 \bullet \{\vec{\rho}\}_\eta^1 \rho^2 + \\
& \left(\frac{1}{4}W_{333} + \frac{1}{4}W_{533}H^2\right)\{\vec{H}\}_\eta^3 \bullet \{\vec{\rho}\}_\eta^3 + \left(W_{040} + W_{240M}H^2 + W_{440M}H^4\right)\{\vec{H}\}_\eta^0 \bullet \{\vec{\rho}\}_\eta^0 \rho^4 + \\
& \left(\frac{1}{2}W_{242} + \frac{1}{2}W_{442M}H^2\right)\{\vec{H}\}_\eta^2 \bullet \{\vec{\rho}\}_\eta^2 \rho^2 + \left(\frac{1}{8}W_{444}\right)\{\vec{H}\}_\eta^4 \bullet \{\vec{\rho}\}_\eta^4 + \left(W_{151} + W_{351M}H^2\right)\{\vec{H}\}_\eta^1 \bullet \{\vec{\rho}\}_\eta^1 \rho^4 + \\
& \left(\frac{1}{4}W_{353}\right)\{\vec{H}\}_\eta^3 \bullet \{\vec{\rho}\}_\eta^3 \rho^2 + \left(W_{060} + W_{260M}H^2\right)\{\vec{H}\}_\eta^0 \bullet \{\vec{\rho}\}_\eta^0 \rho^6 + \\
& \left(\frac{1}{2}W_{262}\right)\{\vec{H}\}_\eta^2 \bullet \{\vec{\rho}\}_\eta^2 \rho^4 + \left(W_{171}\right)\{\vec{H}\}_\eta^1 \bullet \{\vec{\rho}\}_\eta^1 \rho^6 + \left(W_{080}\right)\{\vec{H}\}_\eta^0 \bullet \{\vec{\rho}\}_\eta^0 \rho^8 .
\end{aligned} \qquad (4.34)
$$

A general pattern of the field dependent vector functions is that they factor into a scalar part, shown in parentheses, containing even powers of H, and a Shack vector product part, shown as curly braces. For example

$$W = \cdots + \overbrace{\underbrace{\left(W_{131} + W_{331M}H^2 + W_{531M}H^4\right)}_{Scalar\ Part}}^{Field\ Dependent\ Part} \underbrace{\{\vec{H}\}_\eta^1 \bullet \{\vec{\rho}\}_\eta^1}_{Vector\ Part} \rho^2 + \cdots . \qquad (4.35)$$

Of primary interest in this work are the characteristics of the field dependent parts of the wavefront aberration function expansion terms, called the "field functions", particularly when the rotationally symmetric optical system's elements are perturbed, breaking the rotational symmetry restriction. This is the subject of NAT, briefly reviewed in the next Chapter.

From the notational patterns found in Eq. (4.34), a summation equation for the wavefront aberration function expansion using the medial focus surface W_{klmM} coefficients has been developed in this research. The wavefront aberration function expansion can be written as

$$W\left(\vec{H},\vec{\rho}\right)=\sum_{p=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\ \left(\frac{1}{2}\right)^{m-1+\delta_{0,m}}W_{klm\{M\}}\ H^{2n}\left\{\vec{H}\right\}_{\eta}^{m}\bullet\left\{\vec{\rho}\right\}_{\eta}^{m}\rho^{2p}\ ,\qquad(4.36)$$

where

$$\delta_{0,m}=\begin{cases}1 & m=0\\0 & m\neq0\end{cases},\qquad(4.37)$$

and $k=2n+m,\ l=2p+m$. The "$\{M\}$" subscript indicates inclusion of the "M" subscript when $p\neq0$ and $n\neq0$ for a given W_{klm} coefficient. The wavefront aberration order of a term is, as before, given by

$$Order=2\left(n+p+m\right)\ .\qquad(4.38)$$

In terms of the $W_{klm\{S\}}$ coefficients used in the original expansion, Eq. (4.11), the following relation has been derived

$$W_{klmM}=2^{m-1+\delta_{0,m}}\sum_{\substack{u=m\\u+2}}^{\lfloor k,l\rfloor}K_{u,m}\ W_{klu\{S\}}\ ,\qquad(4.39)$$

where

$$
K_{u,m} = \begin{cases} \dfrac{1}{2^u} \begin{pmatrix} u \\ \dfrac{u}{2} \end{pmatrix} & , \; m = 0 \text{ and } u \text{ even} \\[2em] \dfrac{1}{2^{u-1}} \begin{pmatrix} u \\ \dfrac{u-m}{2} \end{pmatrix} & , \; m > 0 \end{cases} , \qquad (4.40)
$$

and

$$
\lfloor k, l \rfloor = \begin{cases} k & \text{for } k \le l \\ l & \text{for } l \le k \end{cases} . \qquad (4.41)
$$

By defining

$$
W'_{klm\{M\}} \equiv \left(\frac{1}{2} \right)^{m-1+\delta_{0,m}} W_{klm\{M\}} , \qquad (4.42)
$$

the wavefront aberration function expansion can be written as

$$
W\left(\vec{H},\vec{\rho}\right) = \sum_{p=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} W'_{klm\{M\}} \left(\vec{H}\cdot\vec{H}\right)^n \left\{\vec{H}\right\}_\eta^m \cdot \left\{\vec{\rho}\right\}_\eta^m \left(\vec{\rho}\cdot\vec{\rho}\right)^p . \qquad (4.43)
$$

Thus by using the medial focal surface based coefficients, $W'_{klm\{M\}}$, and SVP, a simple

equation (compared to Eq. (4.21)) for the aberration function expansion with the field and

pupil parameters factored has been achieved.

It is interesting to note the similarities and differences of Eq. (4.43) and Shack's

vector form of the wavefront aberration function expansion, Eq. (4.9), repeated here for

ease of comparison.

$$
W\left(\vec{H},\vec{\rho}\right) = \sum_{p=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} W_{klm\{S\}} \left(\vec{H}\cdot\vec{H}\right)^n \left(\vec{H}\cdot\vec{\rho}\right)^m \left(\vec{\rho}\cdot\vec{\rho}\right)^p . \qquad (4.44)
$$

4.3 Wavefront Aberration Function Expansion in Terms of Zernike Vectors

There are several different ways that the pupil dependence of the wavefront aberration function expansion terms, through 8th order, Eq. (4.34), can be converted into Zernike vector functions of the pupil parameters. One way is to simply use algebra. Specifically, the entries in Table 3.5, listing the Zernike vectors in terms of SVP, can be used to solve for the pupil dependencies in terms of Zernike vectors. For example, the pupil dependence $\{\vec{\rho}\}_\eta^1 \rho^2$ is determined by first noticing that (from Table 3.5)

$$\vec{Z}^1_{3\{\eta\}}(\vec{\rho}) = 3\{\vec{\rho}\}_\eta^1 \rho^2 - 2\{\vec{\rho}\}_\eta^1 \quad . \tag{4.45}$$

Then

$$\frac{1}{3}\vec{Z}^1_{3\{\eta\}}(\vec{\rho}) = \{\vec{\rho}\}_\eta^1 \rho^2 - \frac{2}{3}\{\vec{\rho}\}_\eta^1 \, . \tag{4.46}$$

And since (from Table 3.5)

$$\vec{Z}^1_{1\{\eta\}}(\vec{\rho}) = \{\vec{\rho}\}_\eta^1 \quad , \tag{4.47}$$

the following is obtained

$$\{\vec{\rho}\}_\eta^1 \rho^2 = \frac{1}{3}\vec{Z}^1_{3\{\eta\}}(\vec{\rho}) + \frac{2}{3}\vec{Z}^1_{1\{\eta\}}(\vec{\rho}) \, . \tag{4.48}$$

Proceeding in a similar way, a tabulation of all the pupil dependencies of Eq. (4.34) in terms of the Zernike vectors is obtained.

Table 4.1 Pupil dependence in terms of Zernike vectors.

Pupil Dependence	Equivalent Zernike Vector Combination
$\{\vec{\rho}\}_\eta^0$	$\vec{Z}_{0\{\eta\}}^0(\vec{\rho})$
$\{\vec{\rho}\}_\eta^0 \rho^2$	$\frac{1}{2}\vec{Z}_{2\{\eta\}}^0(\vec{\rho})+\frac{1}{2}\vec{Z}_{0\{\eta\}}^0(\vec{\rho})$
$\{\vec{\rho}\}_\eta^0 \rho^4$	$\frac{1}{6}\vec{Z}_{4\{\eta\}}^0(\vec{\rho})+\frac{1}{2}\vec{Z}_{2\{\eta\}}^0(\vec{\rho})+\frac{1}{3}\vec{Z}_{0\{\eta\}}^0(\vec{\rho})$
$\{\vec{\rho}\}_\eta^0 \rho^6$	$\frac{1}{20}\vec{Z}_{6\{\eta\}}^0(\vec{\rho})+\frac{1}{4}\vec{Z}_{4\{\eta\}}^0(\vec{\rho})+\frac{9}{20}\vec{Z}_{2\{\eta\}}^0(\vec{\rho})+\frac{1}{4}\vec{Z}_{0\{\eta\}}^0(\vec{\rho})$
$\{\vec{\rho}\}_\eta^0 \rho^8$	$\frac{1}{70}\vec{Z}_{8\{\eta\}}^0(\vec{\rho})+\frac{1}{10}\vec{Z}_{6\{\eta\}}^0(\vec{\rho})+\frac{2}{7}\vec{Z}_{4\{\eta\}}^0(\vec{\rho})+\frac{2}{5}\vec{Z}_{2\{\eta\}}^0(\vec{\rho})+\frac{1}{5}\vec{Z}_{0\{\eta\}}^0(\vec{\rho})$
$\{\vec{\rho}\}_\eta^1$	$\vec{Z}_{1\{\eta\}}^1(\vec{\rho})$
$\{\vec{\rho}\}_\eta^1 \rho^2$	$\frac{1}{3}\vec{Z}_{3\{\eta\}}^1(\vec{\rho})+\frac{2}{3}\vec{Z}_{1\{\eta\}}^1(\vec{\rho})$
$\{\vec{\rho}\}_\eta^1 \rho^4$	$\frac{1}{10}\vec{Z}_{5\{\eta\}}^1(\vec{\rho})+\frac{2}{5}\vec{Z}_{3\{\eta\}}^1(\vec{\rho})+\frac{1}{2}\vec{Z}_{1\{\eta\}}^1(\vec{\rho})$
$\{\vec{\rho}\}_\eta^1 \rho^6$	$\frac{1}{35}\vec{Z}_{7\{\eta\}}^1(\vec{\rho})+\frac{6}{35}\vec{Z}_{5\{\eta\}}^1(\vec{\rho})+\frac{2}{5}\vec{Z}_{3\{\eta\}}^1(\vec{\rho})+\frac{2}{5}\vec{Z}_{1\{\eta\}}^1(\vec{\rho})$
$\{\vec{\rho}\}_\eta^2$	$\vec{Z}_{2\{\eta\}}^2(\vec{\rho})$
$\{\vec{\rho}\}_\eta^2 \rho^2$	$\frac{1}{4}\vec{Z}_{4\{\eta\}}^2(\vec{\rho})+\frac{3}{4}\vec{Z}_{2\{\eta\}}^2(\vec{\rho})$
$\{\vec{\rho}\}_\eta^2 \rho^4$	$\frac{1}{15}\vec{Z}_{6\{\eta\}}^2(\vec{\rho})+\frac{1}{3}\vec{Z}_{4\{\eta\}}^2(\vec{\rho})+\frac{3}{5}\vec{Z}_{2\{\eta\}}^2(\vec{\rho})$
$\{\vec{\rho}\}_\eta^2 \rho^6$	$\frac{1}{56}\vec{Z}_{8\{\eta\}}^2(\vec{\rho})+\frac{1}{8}\vec{Z}_{6\{\eta\}}^2(\vec{\rho})+\frac{5}{14}\vec{Z}_{4\{\eta\}}^2(\vec{\rho})+\frac{1}{2}\vec{Z}_{2\{\eta\}}^2(\vec{\rho})$
$\{\vec{\rho}\}_\eta^3$	$\vec{Z}_{3\{\eta\}}^3(\vec{\rho})$
$\{\vec{\rho}\}_\eta^3 \rho^2$	$\frac{1}{5}\vec{Z}_{5\{\eta\}}^3(\vec{\rho})+\frac{4}{5}\vec{Z}_{3\{\eta\}}^3(\vec{\rho})$
$\{\vec{\rho}\}_\eta^4$	$\vec{Z}_{4\{\eta\}}^4(\vec{\rho})$

The pupil dependence of the terms in the wavefront aberration function expansion of Eq. (4.34) can now be written in terms of the pupil Zernike vectors. Substituting the Zernike vector equivalent pupil dependence of Table 4.1 for the pupil dependencies in

Eq. (4.34) and then collecting terms of like Zernike pupil dependence results in the wavefront aberration function as shown in Eq. (4.49).

$$
\begin{aligned}
W = &\left(\left(W_{000} + \frac{1}{2}W_{020} + \frac{1}{3}W_{040} + \frac{1}{4}W_{060} + \frac{1}{5}W_{080} \right) + \left(W_{200} + \frac{1}{2}W_{220M} + \frac{1}{3}W_{240M} + \frac{1}{4}W_{260M} \right)H^2 + \right. \\
&\left. \left(W_{400} + \frac{1}{2}W_{420M} + \frac{1}{3}W_{440M} \right)H^4 + \left(W_{600} + \frac{1}{2}W_{620M} \right)H^6 + \left(W_{800} \right)H^8 \right)\left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{0\{\eta\}}^0 \left(\vec{\rho} \right) + \\
&\left(\left(W_{111} + \frac{2}{3}W_{131} + \frac{1}{2}W_{151} + \frac{2}{5}W_{171} \right) + \left(W_{311} + \frac{2}{3}W_{331M} + \frac{1}{2}W_{351M} \right)H^2 + \right. \\
&\left. \left(W_{511} + \frac{2}{3}W_{531M} \right)H^4 + \left(W_{711} \right)H^6 \right)\left\{ \vec{H} \right\}_\eta^1 \bullet \vec{Z}_{1\{\eta\}}^1 \left(\vec{\rho} \right) + \\
&\left(\left(\frac{1}{2}W_{020} + \frac{1}{2}W_{040} + \frac{9}{20}W_{060} + \frac{2}{5}W_{080} \right) + \left(\frac{1}{2}W_{220M} + \frac{1}{2}W_{240M} + \frac{9}{20}W_{260M} \right)H^2 + \right. \\
&\left. \left(\frac{1}{2}W_{420M} + \frac{1}{2}W_{440M} \right)H^4 + \frac{1}{2}W_{620M}H^6 \right)\left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{2\{\eta\}}^0 \left(\vec{\rho} \right) + \\
&\left(\left(\frac{1}{2}W_{222} + \frac{3}{8}W_{242} + \frac{3}{10}W_{262} \right) + \left(\frac{1}{2}W_{422} + \frac{3}{8}W_{442M} \right)H^2 + \frac{1}{2}W_{622}H^4 \right)\left\{ \vec{H} \right\}_\eta^2 \bullet \vec{Z}_{2\{\eta\}}^2 \left(\vec{\rho} \right) + \\
&\left(\left(\frac{1}{3}W_{131} + \frac{2}{5}W_{151} + \frac{2}{5}W_{171} \right) + \left(\frac{1}{3}W_{331M} + \frac{2}{5}W_{351M} \right)H^2 + \frac{1}{3}W_{531M}H^4 \right)\left\{ \vec{H} \right\}_\eta^1 \bullet \vec{Z}_{3\{\eta\}}^1 \left(\vec{\rho} \right) + \\
&\left(\left(\frac{1}{4}W_{333} + \frac{1}{5}W_{353} \right) + \frac{1}{4}W_{533}H^2 \right)\left\{ \vec{H} \right\}_\eta^3 \bullet \vec{Z}_{3\{\eta\}}^3 \left(\vec{\rho} \right) + \\
&\left(\left(\frac{1}{6}W_{040} + \frac{1}{4}W_{060} + \frac{2}{7}W_{080} \right) + \left(\frac{1}{6}W_{240M} + \frac{1}{4}W_{260M} \right)H^2 + \frac{1}{6}W_{440M}H^4 \right)\left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{4\{\eta\}}^0 \left(\vec{\rho} \right) + \\
&\left(\left(\frac{1}{8}W_{242} + \frac{1}{6}W_{262} \right) + \frac{1}{8}W_{442M}H^2 \right)\left\{ \vec{H} \right\}_\eta^2 \bullet \vec{Z}_{4\{\eta\}}^2 \left(\vec{\rho} \right) + \left(\frac{1}{8}W_{444} \right)\left\{ \vec{H} \right\}_\eta^4 \bullet \vec{Z}_{4\{\eta\}}^4 \left(\vec{\rho} \right) + \\
&\left(\frac{1}{20}W_{353} \right)\left\{ \vec{H} \right\}_\eta^3 \bullet \vec{Z}_{5\{\eta\}}^3 \left(\vec{\rho} \right) + \left(\left(\frac{1}{10}W_{151} + \frac{6}{35}W_{171} \right) + \frac{1}{10}W_{351M}H^2 \right)\left\{ \vec{H} \right\}_\eta^1 \bullet \vec{Z}_{5\{\eta\}}^1 \left(\vec{\rho} \right) + \\
&\left(\left(\frac{1}{20}W_{060} + \frac{1}{10}W_{080} \right) + \frac{1}{20}W_{260M}H^2 \right)\left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{6\{\eta\}}^0 \left(\vec{\rho} \right) + \\
&\left(\frac{1}{30}W_{262} \right)\left\{ \vec{H} \right\}_\eta^2 \bullet \vec{Z}_{6\{\eta\}}^2 \left(\vec{\rho} \right) + \left(\frac{1}{35}W_{171} \right)\left\{ \vec{H} \right\}_\eta^1 \bullet \vec{Z}_{7\{\eta\}}^1 \left(\vec{\rho} \right) + \left(\frac{1}{70}W_{080} \right)\left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{8\{\eta\}}^0 \left(\vec{\rho} \right).
\end{aligned}
\tag{4.49}
$$

Using a single symbol $V_{k,n}^m$ for the expansion coefficient of each term gives,

$$W = \left(V_{0,0}^0 + V_{2,0}^0 H^2 + V_{4,0}^0 H^4 + V_{6,0}^0 H^6 + V_{8,0}^0 H^8 \right) \left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{0\{\eta\}}^0 \left(\vec{\rho} \right) +$$

$$\left(V_{0,1}^1 + V_{2,1}^1 H^2 + V_{4,1}^1 H^4 + V_{6,1}^1 H^6 \right) \left\{ \vec{H} \right\}_\eta^1 \bullet \vec{Z}_{1\{\eta\}}^1 \left(\vec{\rho} \right) +$$

$$\left(V_{0,2}^0 + V_{2,2}^0 H^2 + V_{4,2}^0 H^4 + V_{6,2}^0 H^6 \right) \left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{2\{\eta\}}^0 \left(\vec{\rho} \right) +$$

$$\left(V_{0,2}^2 + V_{2,2}^2 H^2 + V_{4,2}^2 H^4 \right) \left\{ \vec{H} \right\}_\eta^2 \bullet \vec{Z}_{2\{\eta\}}^2 \left(\vec{\rho} \right) + \left(V_{0,3}^1 + V_{2,3}^1 H^2 + V_{4,3}^1 H^4 \right) \left\{ \vec{H} \right\}_\eta^1 \bullet \vec{Z}_{3\{\eta\}}^1 \left(\vec{\rho} \right) +$$

$$\left(V_{0,3}^3 + V_{2,3}^3 H^2 \right) \left\{ \vec{H} \right\}_\eta^3 \bullet \vec{Z}_{3\{\eta\}}^3 \left(\vec{\rho} \right) + \left(V_{0,4}^0 + V_{2,4}^0 H^2 + V_{4,4}^0 H^4 \right) \left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{4\{\eta\}}^0 \left(\vec{\rho} \right) +$$

$$\left(V_{0,4}^2 + V_{2,4}^2 H^2 \right) \left\{ \vec{H} \right\}_\eta^2 \bullet \vec{Z}_{4\{\eta\}}^2 \left(\vec{\rho} \right) + \left(V_{0,4}^4 \right) \left\{ \vec{H} \right\}_\eta^4 \bullet \vec{Z}_{4\{\eta\}}^4 \left(\vec{\rho} \right) + \left(V_{0,5}^3 \right) \left\{ \vec{H} \right\}_\eta^3 \bullet \vec{Z}_{5\{\eta\}}^3 \left(\vec{\rho} \right) +$$

$$\left(V_{0,5}^1 + V_{2,5}^1 H^2 \right) \left\{ \vec{H} \right\}_\eta^1 \bullet \vec{Z}_{5\{\eta\}}^1 \left(\vec{\rho} \right) + \left(V_{0,6}^0 + V_{2,6}^0 H^2 \right) \left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{6\{\eta\}}^0 \left(\vec{\rho} \right) +$$

$$\left(V_{0,6}^2 \right) \left\{ \vec{H} \right\}_\eta^2 \bullet \vec{Z}_{6\{\eta\}}^2 \left(\vec{\rho} \right) + \left(V_{0,7}^1 \right) \left\{ \vec{H} \right\}_\eta^1 \bullet \vec{Z}_{7\{\eta\}}^1 \left(\vec{\rho} \right) + \left(V_{0,8}^0 \right) \left\{ \vec{H} \right\}_\eta^0 \bullet \vec{Z}_{8\{\eta\}}^0 \left(\vec{\rho} \right) . \tag{4.50}$$

The general form for rotationally symmetric optical imaging models is given by

$$W \left(\vec{H}, \vec{\rho} \right) = \sum_{n=0}^{\infty} \sum_{\substack{m=n \\ m-2}}^{0} \sum_{\substack{k=0 \\ k+2}}^{\infty} V_{k,n}^m H^k \left\{ \vec{H} \right\}_\eta^m \bullet \vec{Z}_{n\{\eta\}}^m \left(\vec{\rho} \right) \tag{4.51}$$

The terms most often of interest are the Zernike astigmatism $\left(\vec{Z}_{2\{\eta\}}^2 \left(\vec{\rho} \right) \right)$, Zernike coma $\left(\vec{Z}_{3\{\eta\}}^1 \left(\vec{\rho} \right) \right)$, and Zernike spherical $\left(\vec{Z}_{4\{\eta\}}^0 \left(\vec{\rho} \right) \right)$ terms. In Chapter 6, the examples provided will utilize this through 8th order expansion of the wavefront aberration function. Explicit expressions for the $V_{k,n}^m$ coefficients in terms of the $W_{klm\{M\}}$ coefficients for the through 12th order expansion are presented in Table 4.2 on the following page.

It is also desirable to write the wavefront aberration function's field dependence in terms of Zernike vectors. To write the field dependence in Eq. (4.49) in terms of the Zernike vectors, the $\vec{\rho}$ and ρ of Table 4.1 are replaced with $\vec{H}$ and H, respectively. The field functions in Eq. (4.49) can then be replaced with field dependent Zernike

vectors. The idea of expanding both the pupil and the field dependence in terms of Zernike polynomials for rotationally symmetric optical imaging systems has been previously suggested in the literature [55, 56, 57, 58]. However, these other double Zernike expansions have not been in terms of the Zernike vectors nor SVP, a result of this research.

Table 4.2 $V_{k,n}^m$ expansion coefficients.

$V_{k,n}^m$	Equivalent $W_{klm\{M\}}$ Expression
$V_{0,2}^2$	$\frac{1}{2}W_{222} + \frac{3}{8}W_{242} + \frac{3}{10}W_{262} + \frac{1}{4}W_{282} + \frac{3}{14}W_{2,10,2}$
$V_{2,2}^2$	$\frac{1}{2}W_{422} + \frac{3}{8}W_{442M} + \frac{3}{10}W_{462M} + \frac{1}{4}W_{482M}$
$V_{4,2}^2$	$\frac{1}{2}W_{622} + \frac{3}{8}W_{642M} + \frac{3}{10}W_{662M}$
$V_{6,2}^2$	$\frac{1}{2}W_{822} + \frac{3}{8}W_{842M}$
$V_{8,2}^2$	$\frac{1}{2}W_{10,2,2}$
$V_{0,3}^1$	$\frac{1}{3}W_{131} + \frac{2}{5}W_{151} + \frac{2}{5}W_{171} + \frac{8}{21}W_{191} + \frac{5}{14}W_{1,11,1}$
$V_{2,3}^1$	$\frac{1}{3}W_{331M} + \frac{2}{5}W_{351M} + \frac{2}{5}W_{371M} + \frac{8}{21}W_{391M}$
$V_{4,3}^1$	$\frac{1}{3}W_{531M} + \frac{2}{5}W_{551M} + \frac{2}{5}W_{571M}$
$V_{6,3}^1$	$\frac{1}{3}W_{731M} + \frac{2}{5}W_{751M}$
$V_{8,3}^1$	$\frac{1}{3}W_{931M}$
$V_{0,4}^0$	$\frac{1}{6}W_{040} + \frac{1}{4}W_{060} + \frac{2}{7}W_{080} + \frac{25}{84}W_{0,10,0} + \frac{25}{84}W_{0,12,0}$
$V_{2,4}^0$	$\frac{1}{6}W_{240M} + \frac{1}{4}W_{260M} + \frac{2}{7}W_{280M} + \frac{25}{84}W_{2,10,0M}$
$V_{4,4}^0$	$\frac{1}{6}W_{440M} + \frac{1}{4}W_{460M} + \frac{2}{7}W_{480M}$
$V_{6,4}^0$	$\frac{1}{6}W_{640M} + \frac{1}{4}W_{660M}$
$V_{8,4}^0$	$\frac{1}{6}W_{840M}$

For a rotationally symmetric optical imaging model, the double Zernike vector expansion of the wavefront aberration function is given by

$$W\left(\vec{H},\vec{\rho}\right)=\sum_{n_\rho=0}^{\infty}\sum_{\substack{m=n_\rho\\m-2}}^{0}\sum_{n_H=0}^{\infty} U_{n_H,n_\rho}^{m}\,\vec{Z}_{n_H\{\eta\}}^{m}\left(\vec{H}\right)\bullet\vec{Z}_{n_\rho\{\eta\}}^{m}\left(\vec{\rho}\right) \tag{4.52}$$

where the "$m-2$" below the summation is used to indicate decrements of m by 2 rather than increments of 1.

The preceding development has been for rotationally symmetric optical systems. For the case of rotationally nonsymmetric optical systems, an expansion in field and pupil parameters in terms of the Zernike polynomials, in field and pupil, is given by

$$W\left(\vec{H},\vec{\rho}\right)=\sum_{n_\rho=0}^{\infty}\sum_{\substack{m_\rho=-n_\rho\\m_\rho+2}}^{n_\rho}\sum_{n_H=0}^{\infty}\sum_{\substack{m_H=-n_H\\m_H+2}}^{n_H} U_{n_H,n_\rho}^{m_H,m_\rho}Z_{n_H\{\eta\}}^{m_H}\left(\vec{H}\right)Z_{n_\rho\{\eta\}}^{m_\rho}\left(\vec{\rho}\right)\,, \tag{4.53}$$

where "$m_\rho+2$" and "m_H+2" indicate addition in increments of 2. Note that these are Zernike polynomials, not the Zernike vectors previously used for the case of rotationally symmetric optical systems. Recall that the Zernike polynomials are complete over a unit radius disk (normalized object and pupil) and thus any function defined over the unit radius object disk and unit radius pupil disk may be so expanded [56, 59, 60]. This, Eq. (4.53), will be called the double Zernike expansion of the wavefront aberration function. It is this form of the wavefront aberration function expansion that is used together with the GQ method to obtain the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$ from which, for rotationally

symmetric optical imaging systems, the $W_{klm\{M\}}$ expansion coefficients of Eq. (4.49) can be obtained, as will be shown below.

4.4 Method for Obtaining Expansion Coefficients

The $U_{n_H,n_\rho}^{m_H,m_\rho}$ expansion coefficients of the general double Zernike expansion of the wavefront aberration function, Eq. (4.53), may be calculated in the following way. An optical design, ray tracing software package is used to implement the optical imaging system models and to perform the real ray tracing. The OPD value for each ray traced (real ray tracing) is calculated. A data file containing five numerical values for each ray traced is obtained. These five values are the 2 field and 2 pupil parameter values identifying the ray traced, and the calculated OPD value for that ray. Using the discrete Zernike orthogonality properties described in Appendix III, the GQ method is used twice, once in pupil and once in field parameter values, to obtain the $U_{n_H,n_\rho}^{m_H,m_\rho}$ expansion coefficients.

For the through 12th order expansion of the wavefront aberration function, the maximum radial, n, Zernike index is set to $n_{\text{max}} = 12$. Then the number of radial (field and pupil) values needed to perform the GQ procedure, as calculated by Eq. (III.61), is

$$N_{radial} = \left\lfloor \frac{n_{\text{max}}}{2} \right\rfloor + 1 = \left\lfloor \frac{12}{2} \right\rfloor + 1 = 7 \ , \tag{4.54}$$

while the number of angular values required, as given by Eq. (III.65), is

$$N_{angle} = 2n_{\text{max}} + 1 = 2 \times 12 + 1 = 25 \ . \tag{4.55}$$

There are then $N_{radial} \times N_{angle} = 7 \times 25 = 175$ field and pupil points required for one application of the GQ method. This means that $175 \times 175 + 175 = 30,800$ rays need to be traced, and their OPD values calculated. The "+175" is to account for the chief ray, which is not one of the rays needed by the GQ method and so must be traced in addition to the GQ rays. Depending on the implementation of the real ray tracing program, this can take a long time. Fortunately, the available commercial optical design software packages today are very fast for tracing rays.

Note that the number of rays to be traced, 30800, is for a single optical model to determine the optical model's total $W_{klm\{M\}}$ values. As will be detailed in the next section, to obtain the required per surface, sphere/asphere, and intrinsic/extrinsic expansion coefficient values, $W_{klm\{M\}}$, requires at most 4 separate optical models *per surface* of the original optical model. So, depending on the number of surfaces in the original optical model, several times this number of rays may need to be traced.

With the double Zernike expansion coefficients U_{n_H,n_ρ}^{m} of Eq. (4.52) for a rotationally symmetric optical imaging model known (the $U_{n_H,n_\rho}^{m_H,m_\rho}$ coefficients reduce the to U_{n_H,n_ρ}^{m} expansion coefficients for the rotationally symmetric case), a set of equations can be derived for the calculation of the $W_{klm\{M\}}$ expansion coefficients. The derivation of these equations has been accomplished for the through 12th order expansion of the wavefront aberration function. The method of derivation is here described.

A Matlab® program, using the add-on symbolic toolbox for symbolic algebraic manipulation, was written to implement Eq. (4.43), providing the wavefront aberration function expansion in terms of the $W_{klm\{M\}}$ expansion coefficients and SVP in field and pupil parameters. Another Matlab® function was written to calculate the conversion of the pupil and field dependencies occurring in the wavefront aberration function expansion into the equivalent real number Zernike polynomials. A portion of these conversions has been presented in Table 4.1. Then, in software, collecting like terms in the resulting symbolic expression of the $W\left(\vec{H},\vec{\rho}\right)$ expansion provides the Zernike-in-field, Zernike-in-pupil expansion of the wavefront aberration function in terms of the $W_{klm\{M\}}$ expansion coefficients. Comparison with the double Zernike expression of Eq. (4.52) provides equations for the $U^m_{n_H,n_\rho}$ expansion coefficients in terms of the $W_{klm\{M\}}$ expansion coefficients. A sample of these equations is presented here.

$$U^2_{10,2} = \left(\frac{1}{420}W_{1022}\right). \tag{4.56}$$

$$U^2_{8,2} = \left(\frac{1}{112}W_{822} + \frac{3}{448}W_{842M} + \frac{3}{140}W_{1022}\right). \tag{4.57}$$

$$U^2_{6,2} = \left(\frac{1}{30}W_{622} + \frac{1}{40}W_{642M} + \frac{1}{50}W_{662M} + \frac{1}{16}W_{822} + \frac{3}{64}W_{842M} + \frac{1}{12}W_{1022}\right). \tag{4.58}$$

$$U^2_{4,2} = \left(\frac{1}{8}W_{422M} + \frac{3}{32}W_{442M} + \frac{3}{40}W_{462M} + \frac{1}{16}W_{482M} + \frac{1}{6}W_{622} + \frac{1}{8}W_{642M}\right.$$
$$\left. + \frac{1}{10}W_{662M} + \frac{5}{28}W_{822} + \frac{15}{112}W_{842M} + \frac{5}{28}W_{1022}\right). \tag{4.59}$$

$$U_{2,2}^2 = \left(\frac{1}{2} W_{222} + \frac{3}{8} W_{242} + \frac{3}{10} W_{262} + \frac{1}{4} W_{282} + \frac{3}{14} W_{2102} + \frac{3}{8} W_{422} + \frac{9}{32} W_{442M} + \frac{9}{40} W_{462M} \right.$$
$$\left. + \frac{3}{16} W_{482M} + \frac{3}{10} W_{622} + \frac{9}{40} W_{642M} + \frac{9}{50} W_{662M} + \frac{1}{4} W_{822M} + \frac{3}{16} W_{842M} + \frac{3}{14} W_{1022M} \right). \quad (4.60)$$

$$U_{9,3}^1 = \left(\frac{1}{378} W_{931M} \right). \quad (4.61)$$

$$U_{7,3}^1 = \left(\frac{1}{105} W_{731M} + \frac{2}{175} W_{751M} + \frac{4}{189} W_{931M} \right). \quad (4.62)$$

$$U_{5,3}^1 = \left(\frac{1}{30} W_{531M} + \frac{1}{25} W_{551M} + \frac{1}{25} W_{571M} + \frac{2}{35} W_{731M} + \frac{12}{175} W_{751M} + \frac{1}{14} W_{931M} \right). \quad (4.63)$$

$$U_{3,3}^1 = \left(\frac{1}{9} W_{331M} + \frac{2}{15} W_{351M} + \frac{2}{15} W_{371M} + \frac{8}{63} W_{391M} + \frac{2}{15} W_{531M} + \frac{4}{25} W_{551M} \right.$$
$$\left. + \frac{4}{25} W_{571M} + \frac{2}{15} W_{731M} + \frac{4}{25} W_{751M} + \frac{8}{63} W_{931M} \right). \quad (4.64)$$

$$U_{1,3}^1 = \left(\frac{1}{3} W_{131} + \frac{2}{5} W_{151} + \frac{2}{5} W_{171} + \frac{8}{21} W_{191} + \frac{5}{14} W_{1,11,1} + \frac{2}{9} W_{331M} + \frac{4}{15} W_{351M} + \frac{4}{15} W_{371M} \right.$$
$$\left. + \frac{16}{63} W_{391M} + \frac{1}{6} W_{531M} + \frac{1}{5} W_{551M} + \frac{1}{5} W_{571M} + \frac{2}{15} W_{731M} + \frac{4}{25} W_{751M} + \frac{1}{9} W_{931M} \right). \quad (4.65)$$

The complete set of such equations through 12$^{\text{th}}$ order (not shown) are then inverted to provide the $W_{klm\{M\}}$ coefficients in terms of the U_{n_H,n_ρ}^m coefficients. Some of the resulting equations are provided here.

$$W_{040} = 6 \left(U_{0,4}^0 - 5U_{0,6}^0 + 15U_{0,8}^0 - 35U_{0,10}^0 + 70U_{0,12}^0 - U_{2,4}^0 + 5U_{2,6}^0 \right.$$
$$\left. - 15U_{2,8}^0 + 35U_{2,10}^0 + U_{4,4}^0 - 5U_{4,6}^0 + 15U_{4,8}^0 - U_{6,4}^0 + 5U_{6,6}^0 + U_{8,4}^0 \right). \quad (4.66)$$

$$W_{131} = 3\left(U_{1,3}^1 - 4U_{1,5}^1 + 10U_{1,7}^1 - 20U_{1,9}^1 + 35U_{1,11}^1 - 2U_{3,3}^1 + 8U_{3,5}^1 - 20U_{3,7}^1 + 40U_{3,9}^1 \right.$$
$$\left. + 3U_{5,3}^1 - 12U_{5,5}^1 + 30U_{5,7}^1 - 4U_{7,3}^1 + 16U_{7,5}^1 + 5U_{9,3}^1 \right).$$

(4.67)

$$W_{220M} = 4\left(U_{2,2}^0 - 3U_{2,4}^0 + 6U_{2,6}^0 - 10U_{2,8}^0 + 15U_{2,10}^0 - 3U_{4,2}^0 + 9U_{4,4}^0 - 18U_{4,6}^0 + 30U_{4,8}^0 \right.$$
$$\left. + 6U_{6,2}^0 - 18U_{6,4}^0 + 36U_{6,6}^0 - 10U_{8,2}^0 + 30U_{8,4}^0 + 15U_{10,2}^0 \right)$$

(4.68)

$$W_{222} = 2\left(U_{2,2}^2 - 3U_{2,4}^2 + 6U_{2,6}^2 - 10U_{2,8}^2 + 15U_{2,10}^2 - 3U_{4,2}^2 + 9U_{4,4}^2 - 18U_{4,6}^2 + 30U_{4,8}^2 \right.$$
$$\left. + 6U_{6,2}^2 - 18U_{6,4}^2 + 36U_{6,6}^2 - 10U_{8,2}^2 + 30U_{8,4}^2 + 15U_{10,2}^2 \right).$$

(4.69)

$$W_{311} = 3\left(U_{3,1}^1 - 2U_{3,3}^1 + 3U_{3,5}^1 - 4U_{3,7}^1 + 5U_{3,9}^1 - 4U_{5,1}^1 + 8U_{5,3}^1 - 12U_{5,5}^1 + 16U_{5,7}^1 \right.$$
$$\left. + 10U_{7,1}^1 - 20U_{7,3}^1 + 30U_{7,5}^1 - 20U_{9,1}^1 + 40U_{9,3}^1 + 35U_{11,1}^1 \right).$$

(4.70)

$$W_{060} = 20\left(U_{0,6}^0 - 7U_{0,8}^0 + 28U_{0,10}^0 - 84U_{0,12}^0 - U_{2,6}^0 + 7U_{2,8}^0 - 28U_{2,10}^0 \right.$$
$$\left. + U_{4,6}^0 - 7U_{4,8}^0 - U_{6,6}^0 \right).$$

(4.71)

$$W_{151} = 10\left(U_{1,5}^1 - 6U_{1,7}^1 + 21U_{1,9}^1 - 56U_{1,11}^1 - 2U_{3,5}^1 + 12U_{3,7}^1 - 42U_{3,9}^1 \right.$$
$$\left. + 3U_{5,5}^1 - 18U_{5,7}^1 - 4U_{7,5}^1 \right).$$

(4.72)

$$W_{242} = 8\left(U_{2,4}^2 - 5U_{2,6}^2 + 15U_{2,8}^2 - 35U_{2,10}^2 - 3U_{4,4}^2 + 15U_{4,6}^2 - 45U_{4,8}^2 \right.$$
$$\left. + 6U_{6,4}^2 - 30U_{6,6}^2 - 10U_{8,4}^2 \right).$$

(4.73)

$$W_{240M} = 12\left(U_{2,4}^0 - 5U_{2,6}^0 + 15U_{2,8}^0 - 35U_{2,10}^0 - 3U_{4,4}^0 + 15U_{4,6}^0 - 45U_{4,8}^0 \right.$$
$$\left. + 6U_{6,4}^0 - 30U_{6,6}^0 - 10U_{8,4}^0 \right)$$

(4.74)

$$W_{331M} = 9\left(U_{3,3}^1 - 4U_{3,5}^1 + 10U_{3,7}^1 - 20U_{3,9}^1 - 4U_{5,3}^1 + 16U_{5,5}^1 - 40U_{5,7}^1 \right.$$
$$\left. + 10U_{7,3}^1 - 40U_{7,5}^1 - 20U_{9,3}^1 \right).$$

(4.75)

$$W_{333} = 4\left(U_{3,3}^3 - 4U_{3,5}^3 + 10U_{3,7}^3 - 20U_{3,9}^3 - 4U_{5,3}^3 + 16U_{5,5}^3 - 40U_{5,7}^3\right.$$
$$\left. + 10U_{7,3}^3 - 40U_{7,5}^3 - 20U_{9,3}^3\right). \tag{4.76}$$

$$W_{420M} = 12\left(U_{4,2}^0 - 3U_{4,4}^0 + 6U_{4,6}^0 - 10U_{4,8}^0 - 5U_{6,2}^0 + 15U_{6,4}^0 - 30U_{6,6}^0\right.$$
$$\left. + 15U_{8,2}^0 - 45U_{8,4}^0 - 35U_{10,2}^0\right). \tag{4.77}$$

$$W_{422} = 8\left(U_{4,2}^2 - 3U_{4,4}^2 + 6U_{4,6}^2 - 10U_{4,8}^2 - 5U_{6,2}^2 + 15U_{6,4}^2 - 30U_{6,6}^2\right.$$
$$\left. + 15U_{8,2}^2 - 45U_{8,4}^2 - 35U_{10,2}^2\right). \tag{4.78}$$

$$W_{511} = 10\left(U_{5,1}^1 - 2U_{5,3}^1 + 3U_{5,5}^1 - 4U_{5,7}^1\right.$$
$$\left. - 6U_{7,1}^1 + 12U_{7,3}^1 - 18U_{7,5}^1 + 21U_{9,1}^1 - 42U_{9,3}^1 - 56U_{11,1}^1\right). \tag{4.79}$$

$$W_{080} = 70\left(U_{0,8}^0 - U_{2,8}^0 + U_{4,8}^0 - 9U_{0,10}^0 + 9U_{2,10}^0 + 45U_{0,12}^0\right). \tag{4.80}$$

Thus a bridge between the W_{klm} expansion coefficients often used by the optical design community and the U_{n_H,n_ρ}^m expansion coefficients of the double Zernike expansion has been established at the total system level. In the following section, the per system $W_{klm;j}$ expansion coefficients are obtained.

For a rotationally symmetric optical imaging model, some of the expansion coefficients, $U_{n_H,n_\rho}^{m_H,m_\rho}$, of the general expansion Eq. (4.53) should be zero. The deviation from zero of such coefficients provides some indication of the accuracy of the method used to obtain all of the U_{n_H,n_ρ}^m coefficient values.

Since, for rotationally symmetric systems, the $U_{n_H,n_\rho}^{m_H,m_\rho}$ occur in pairs, $\left(U_{n_H,n_\rho}^{+m,+m},U_{n_H,n_\rho}^{-m,-m}\right)$, it should be the case that $U_{n_H,n_\rho}^{+m,+m}=U_{n_H,n_\rho}^{-m,-m}\left(\equiv U_{n_H,n_\rho}^{m}\right)$. Then the difference in the pairs' values is another indication of the accuracy of these coefficients. This method for estimating the computational error of the U_{n_H,n_ρ}^{m} values is used in Chapter 6.

Using error propagation techniques, and the estimate of the U_{n_H,n_ρ}^{m} error, an upper bound for the estimate of the computational errors in the final $W_{klm\{M\}}$ expansion coefficients can be calculated. Examples will be provided in Chapter 6. Estimations for the errors in the calculated $W_{klm\{M\}}$ expansion coefficients appears to be lacking in the literature. The technique used in this research provides a method to obtain computational error estimates for the values calculated. These errors occur due to limitations of numerical representation and error propagation during calculations.

4.5 Per Surface, Sphere/Asphere, and Intrinsic/Extrinsic Contributions

In optical design, and for a large part of the present work, specifically for NAT, it is desirable to know the contribution that each optical element in the optical model, surface by surface, makes to the total system's aberration function. An assumption is made, for rotationally symmetric optical models, that the contribution each optical element (surface with optical power) makes to the total system's aberration function can be summed to

produce the total system's aberration function. Therefore a sum over all surfaces is introduced. For example, Eq. (4.9) is now written as

$$W\left(\vec{H},\vec{\rho}\right) = \sum_{j=1}^{\# \, Surfaces} \sum_{n=0}^{\infty}\sum_{p=0}^{\infty}\sum_{m=0}^{\infty} W_{klm\{S\};j}\left(\vec{H}\cdot\vec{H}\right)^{n}\left(\vec{H}\cdot\vec{\rho}\right)^{m}\left(\vec{\rho}\cdot\vec{\rho}\right)^{p} \, , \qquad (4.81)$$

where the expansion coefficients $W_{klm\{S\};j}$ now carry the subscript j to indicate the jth surface's contribution.

When the optical surface j has a rotationally symmetric non-spherical shape, the surface is considered to be composed of a sphere base shape plus an "aspheric cap" shape that is added to the sphere base shape. It is then desirable to know the contribution to the surface's total aberration contribution provided by the sphere base shape and separately the contribution attributable to the aspheric cap shape. In this terminology, an asphere is composed of a sphere base shape plus an "aspheric cap" given by

$$aspheric \; cap = asphere - sphere \; . \qquad (4.82)$$

A conic is one type of asphere.

A spherical wavefront is a wavefront without aberrations. When a spherical wavefront interacts with (is refracted or reflected by) an optical element, the resulting wavefront will in general have a nonspherical shape. The difference between the spherical shape of a nonaberrated incident wavefront and the resulting nonspherical shape of the wavefront due to aberrations is called the intrinsic aberration contribution.

It is not always the case that the incident wavefront onto an optical surface is free of aberrations. Then the resulting wavefront, after refracting/reflecting from the optical surface, will have components due to the aberration free spherical incident wavefront and
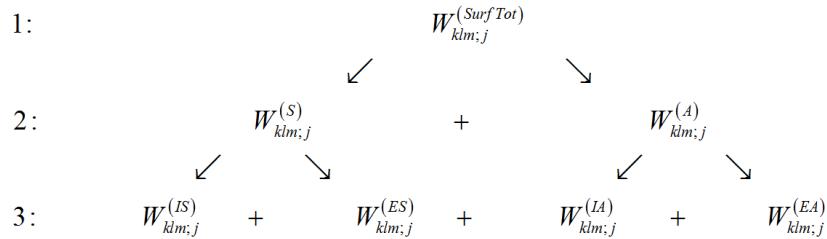
a component due to the aberrations present in the incident wavefront. The contribution due to the aberrations in the incident wavefront is called the extrinsic aberration contributions. The total aberration contribution of a surface is the sum of the intrinsic and the extrinsic aberration contributions.

Because the surface is divided into a sphere base shape plus an aspheric cap shape, there are four contributions to the surface's total aberration contribution. These are denoted by $W_{klm;j}^{(IS)}$, $W_{klm;j}^{(IA)}$, $W_{klm;j}^{(ES)}$, and $W_{klm;j}^{(EA)}$ where "IS" denotes the intrinsic contribution from surface j's sphere base shape, "IA" denotes the intrinsic contribution from surface j's aspheric cap shape, "ES" denotes the extrinsic contribution from surface j's sphere base shape, and "EA" denotes the extrinsic contribution from surface j's aspheric cap shape. Then

$$W_{klm;j} \equiv W_{klm;j}^{(Surf\ Tot)} = W_{klm;j}^{(IS)} + W_{klm;j}^{(IA)} + W_{klm;j}^{(ES)} + W_{klm;j}^{(EA)} \tag{4.83}$$

where $W_{klm;j}^{(Surf\ Tot)}$ is the total expansion coefficient aberration contribution for surface j.

Figure 4.4 provides an illustration of one example of the hierarchy of contributions to the surface's total aberration coefficient value.

Figure 4.4 Hierarchy of total surface's aberration coefficient. The numbers at the left indicates the level of separation in the hierarchy tree. "S" = sphere, "A" = aspheric cap, "IS" = intrinsic sphere, "ES" = extrinsic sphere, "IA" = intrinsic aspheric cap, "EA" = extrinsic aspheric cap.

There are other ways that this hierarchy can be arranged. For example, at the second level, the $W_{klm;j}^{(Surf\ Tot)}$ could be separated into the total intrinsic contribution $W_{klm;j}^{(I)}$ and the total extrinsic contribution $W_{klm;j}^{(E)}$ rather than the sphere and aspheric cap contributions.

Assuming that the total optical model's W_{klm} values can be obtained, as outlined in the preceding section or by other means, the following method has been defined and implemented in this work to obtain the per surface, sphere/asphere cap, and intrinsic/extrinsic $W_{klm;j}^{(Type)}$ contributions. These expansion coefficients are utilized in NAT to be described in the next Chapter for the analysis of the wavefront aberration function's field dependence.

Assume that the optical model consists of *N* optical surfaces that have non-zero optical power. For completeness of the method to be described, assume that all the surfaces are aspheres. For any surface $j \leq N$ in the optical system, it is assumed that all the surfaces preceding surface *j* form an optical imaging subsystem, denoted by $S_{j'<j}$, of the total optical system S_{Total}. Each such subsystem will have its own object, entrance, exit, and Gaussian image planes. It is further assumed that each individual surface *j*, separated from the rest of the original optical system, will form its own optical imaging subsystem, denoted by S_j. These subsystems are defined and modeled such that the Gaussian entrance pupil for S_j is the Gaussian exit pupil of $S_{j'<j}$ and the Gaussian

object plane for S_j is the Gaussian image plane of $S_{j'<j}$. Then the total per surface

expansion coefficients for surface j are defined to be

$$W_{klm;j}^{(Total)} \equiv W_{klm;j'\leq j}^{(Total)} - W_{klm;j'<j}^{(Total)} \ . \tag{4.84}$$

To obtain the four wavefront expansion coefficient types listed above, changes to

both the shape of the wavefront that is incident onto the jth surface as well as to the shape

of the jth surface are needed. Additional notation to keep these changes clear will now be

defined. The incident wavefront onto the jth surface will be either the real wavefront

(RW) shape from subsystem $S_{j'<j}$ or, by constructing a new optical model consisting of

just surface j, an aberration free spherical wavefront (SW). The jth surface may have

either the real surface (RS) shape of the original surface or it is replaced in the optical

model with just the base sphere shape (SS) component of the jth surface's shape. Then

Eq. (4.84) can be written as

$$W_{klm;j}^{(RWRS)} = W_{klm;j'\leq j}^{(Total)} - W_{klm;j'<j}^{(Total)} \ , \tag{4.85}$$

denoting the aberrations attributable to surface j as a result of the real wavefront incident

onto surface j using surface j's real surface shape.

By replacing the jth optical surface by its spherical base shape in the $S_{j'\leq j}$ model

$$W_{klm;j}^{(RWSS)} = W_{klm;j'\leq j}^{(RWSS)} - W_{klm;j'<j}^{(Total)} \ , \tag{4.86}$$

is obtained.

Replacing the incident wavefront onto the jth surface by a spherical wavefront

originating from the local, surface j's, object plane, the coefficient values $W_{klm;j}^{(SWRS)}$ are

obtained. This is accomplished by constructing a separate optical model, S_j , consisting

of only one surface with optical power, namely surface j.

Finally, replacing the jth surface shape in the S_j optical model by the sphere base

shape of the jth surface, $W_{klm;j}^{(SWSS)}$ is obtained. The four per surface expansion coefficient

types can now be calculated as follows.

$$W_{klm;j}^{(IS)} = W_{klm;j}^{(SWSS)} \; , \tag{4.87}$$

$$W_{klm;j}^{(IA)} = W_{klm;j}^{(SWRS)} - W_{klm;j}^{(SWSS)} \; , \tag{4.88}$$

$$W_{klm;j}^{(ES)} = W_{klm;j}^{(RWSS)} - W_{klm;j}^{(SWSS)} \; , \tag{4.89}$$

$$W_{klm;j}^{(EA)} = W_{klm;j}^{(RWRS)} - W_{klm;j}^{(IS)} - W_{klm;j}^{(IA)} - W_{klm;j}^{(ES)} \; . \tag{4.90}$$
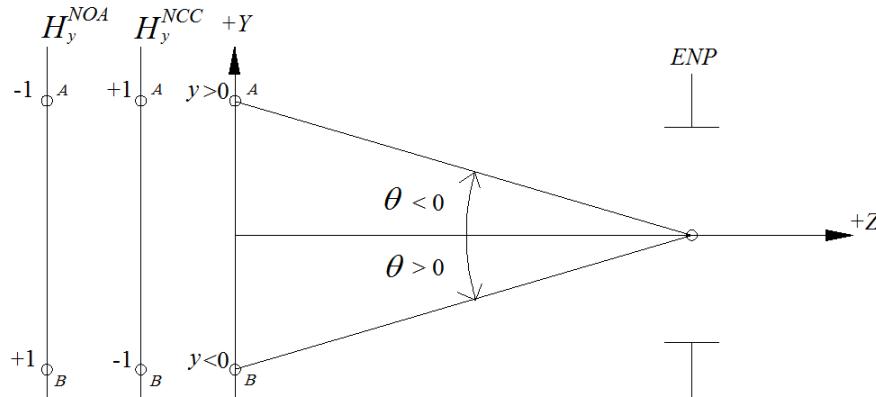
Note that by the construction of these equations, it is always the case that

$$W_{klm;j}^{(Total)} = W_{klm;j}^{(IS)} + W_{klm;j}^{(IA)} + W_{klm;j}^{(ES)} + W_{klm;j}^{(EA)} \; . \tag{4.91}$$

This is unfortunate because it does not provide for a way to independently validate the

separation of the total surface contribution into the four component values. However, as

mentioned in other work [50] the 4th order extrinsic coefficients $W_{klm;j}^{(ES)}$ and $W_{klm;j}^{(EA)}$ are

always zero. This has been a valuable debugging condition during the implementation of

this method for calculating the $W_{klm;j}^{(Type)}$ coefficients. Using Eq. (4.87) through Eq. (4.90)

for the 4th order extrinsic coefficients should always produce a number close to zero.

When this does not occur an error in implementation can be assumed. It also provides

some indication of how accurate the calculated $W_{klm;j}$ values are. The inaccuracies arise

from the limitations of the number representation used for the calculations, and from error propagation associated with the implemented equations. The 4th order extrinsic coefficients, being theoretically zero, are not programmed into the generation of any full field display (FFD) plots to be presented in Chapters 5 through 7.

Although the above method for obtaining the $W_{klm;j}^{(Type)}$ expansion coefficients appears to be straight forward, there is a subtle point that needs to be made explicit. Recall that the wavefront aberration function expansion is defined in this dissertation using the normalized Cartesian coordinate object field parameters $\vec{H}$. Not all researchers may choose to use this parameterization of the wavefront aberration function. It is possible that some researchers may choose to use a normalized object angle parameterization. An illustration comparing these choices is given in Figure 4.5.

Figure 4.5 Field parameterization choice comparison. Points labeled "A" are all the same point in the object plane along the +y-axis. Points labeled "B" are all the same points in the object plane along the –y-axis. NCC = normalized Cartesian coordinates. NOA = normalized object angle. The object point's angle, along with its sign convention, is shown as θ. ENP = entrance pupil. +Z is the direction of positive ray translation.

In Figure 4.5, the normalized field component along the y-axis, H_y, may be either the normalized Cartesian coordinate H_y^{NCC} or the normalized object angle H_y^{NOA}. The object angle to the point (either point "A" or to point "B") is given by θ. The sign convention for θ is indicated in Figure 4.5.

From Figure 4.5 it is seen that the following relation holds between the two different normalized field choices,

$$\vec{H}^{NCC} = -\vec{H}^{NOA} \ . \tag{4.92}$$

Why this matters is because it changes the sign of all the W_{klm} values with k odd. k is the power of the field parameter H. As an example, consider the 4$^{\text{th}}$ order coma term

$$W_{coma}\left(\vec{H}, \vec{\rho}\right) = W_{131}\vec{H} \bullet \vec{\rho}\,\rho^2 \ . \tag{4.93}$$

Explicitly stating that the field parameters are the normalized Cartesian coordinates, this can be written as

$$W_{coma}\left(\vec{H}^{NCC}, \vec{\rho}\right) = W_{131}^{NCC}\vec{H}^{NCC} \bullet \vec{\rho}\,\rho^2 \ . \tag{4.94}$$

Explicitly stating that the field parameters are the normalized object angles, the coma term is written as

$$W_{coma}\left(\vec{H}^{NOA}, \vec{\rho}\right) = W_{131}^{NOA}\vec{H}^{NOA} \bullet \vec{\rho}\,\rho^2 \tag{4.95}$$

Since the amount of the departure of the wavefront from the reference sphere, the amount of wavefront aberration, is independent of the choice of field normalization, the relation

$$W_{coma}\left(\vec{H}^{NCC}, \vec{\rho}\right) = W_{coma}\left(\vec{H}^{NOA}, \vec{\rho}\right) \ , \tag{4.96}$$

holds. This implies that

$$W_{131}^{NCC} \vec{H}^{NCC} = W_{131}^{NOA} \vec{H}^{NOA} \ . \tag{4.97}$$

Using Eq. (4.92) in Eq. (4.97), the following equations are obtained

$$W_{131}^{NCC} \vec{H}^{NCC} = W_{131}^{NOA} \left(-\vec{H}^{NCC} \right) \ , \tag{4.98}$$

$$W_{131}^{NCC} = -W_{131}^{NOA} \ . \tag{4.99}$$

Therefore, the sign of the odd k W_{klm} expansion coefficients is different depending on the normalized field parameter chosen.

This result may be the cause of a sign issue in the examples of Chapter 6. In that Chapter, the 4th order W_{klm} values calculated using the GQ method are compared to the values reported by the CODE V® macros fifthdef and FORDER, which also calculate the 4th order W_{klm} values. It is there seen that the W_{131} values of fifthdef/FORDER have opposite signs to that of the GQ calculated values. However, it is not known exactly what field parameterization the equations implemented in fifthdef/FORDER are for.

The important point of this discovery of this research is that the W_{klm} values are not all independent of the choice of the normalized field parameterization used, and it is therefore critical to report which field normalization is being used when calculating and reporting W_{klm} expansion coefficient values.

Chapter 5 Sigma Offset Vectors, Full Field Displays, and Nodal Aberration Theory's Nodal Splits

Nodal aberration theory (NAT) was originally developed by Kevin P. Thompson [5, 54, 61, 62, 63, 64], based on early insight into binodal astigmatism from Roland V. Shack. It builds on the work of Buchroeder, also working with Shack, who determined that when an optical surface of a rotationally symmetric optical imaging system is tilted, the center of the field dependence of the wavefront aberration function in the image plane shifts. [65] The quantification of this shift is denoted by a per surface 2D sigma "offset" vector, $\vec{\sigma}_j$. If the optical surface is composed of a sphere base shape plus an aspheric cap shape (which includes the conic shapes), then each of these two component shapes of the overall surface shape contributes a different shift to the total field dependence. Therefore, for a single optical surface, there is a sigma offset vector for the sphere base shape and a different sigma offset vector for the additional aspheric cap shape.
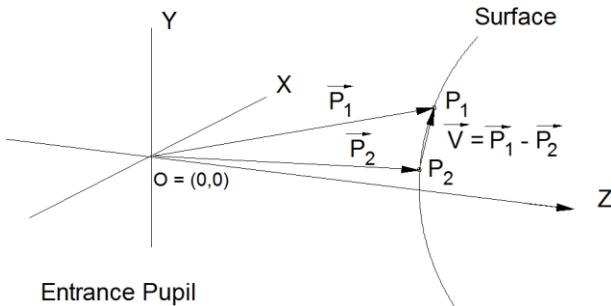
In NAT, a nodal point (a.k.a. node point) for an aberration term in the wavefront aberration function expansion is the object field point $\vec{H}$ that makes the field dependent vector factor (the field function) of the wavefront aberration expansion term zero. Then the contribution of that aberration term to the total wavefront aberration function at that object field point will be zero.

The vector valued field dependence of a scalar aberration term of the wavefront aberration function expansion, either as an expansion in terms of Shack's vector product

(SVP), as in Eq. (4.43), or as an expansion in Zernike vectors in pupil coordinates, as in Eq. (4.51), can be graphically displayed as 2-dimensional plots called full field displays (FFDs). For a rotationally symmetric optical imaging system, the node points for a field dependent term in the wavefront aberration function expansion is located at the coordinate origin $\vec{H} = (0,0)$. When one or more of the imaging system's optical surfaces is perturbed (tilted and/or decentered laterally from the mechanical coordinate axis (MCA)) then the node points for a field dependent term in the aberration function expansion may shift away from the coordinate origin, and may also split into multiple nodes depending on the characteristics of the field dependence of the expansion term. NAT is the study of these characteristics of the field dependence of the wavefront aberration function expansion terms.

In this Chapter, a summary of the concepts and equations used in NAT is presented. A review of the definitions and equations for the 2D field sigma offset vectors is given and it is shown how these vectors are introduced into the wavefront aberration function expansion terms. Clarification of previously published work is made. The development of equations for the sigma offset vectors, correcting for sign issues that have been identified in this research, is presented.
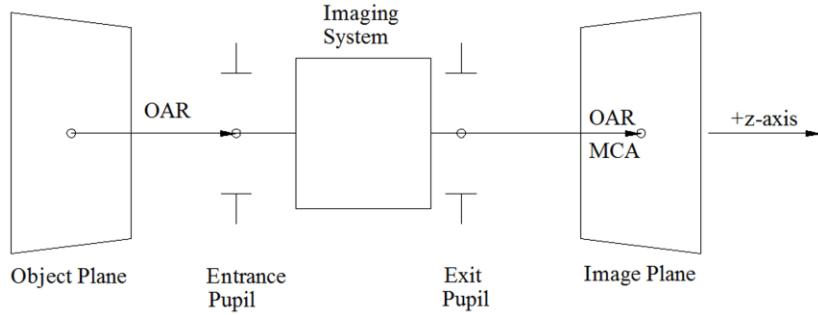
A global coordinate system will be used to locate surfaces and points of intersection, e.g. the intersection point of the optical axis ray (OAR) and a surface or the aberration field axis (AFA) and a surface. A vector may then be formed between two points as the difference of the two vectors to the points with respect to the global coordinate system. This is illustrated in Figure 5.1.

Figure 5.1 The vector $\vec{V}$ is obtained as the subtraction of the two vectors $\vec{P}_1$ and $\vec{P}_2$ associated with the surface points P_1 and P_2, respectively, and referenced with respect to the global coordinate system. As an example, P_1 may be the intersection point of the OAR with the surface, while P_2 may be the intersection point of the AFA with the surface.
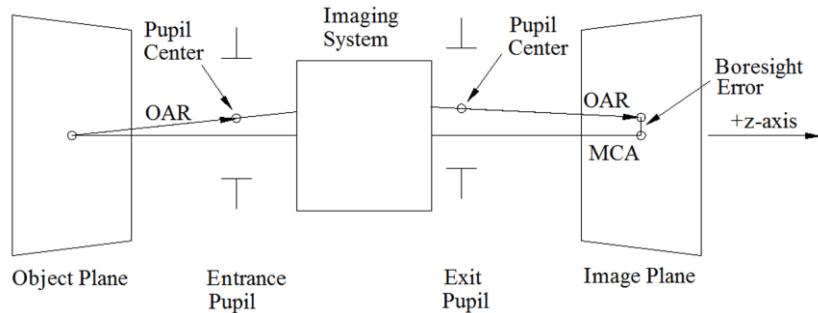
5.1 Calculating 2D Sigma Offset Vectors

Consider a rotationally symmetric optical imaging system. The axis of symmetry will be defined to be along the *z*-axis and it will be called the mechanical coordinate axis (MCA). Rays travel from the object plane, through the imaging system, to the Gaussian image plane. The positive direction of a ray will be selected to be in the sense of the positive *z*-axis. The ray from the center of the object (the intersection of the negative *z*-axis with the object plane) through the center of the entrance pupil and on through the imaging system, to the Gaussian image plane, is called the optical axis ray (OAR). As shown in Figure 5.2, this ray is coincident with the MCA for this case of a rotationally symmetric optical imaging system.

Figure 5.2 OAR and MCA coincide for rotationally symmetric imaging system.

But when one or more of the optical system's surfaces is perturbed by tilts and/or decenters, the OAR will not coincide with the MCA. In all cases, the intersection of the OAR with the Gaussian image plane is defined to be the coordinate origin in the image plane. Therefore, the coordinate origin in the image plane need not coincide with the intersection of the MCA with the image plane. The difference between the OAR and the MCA intersection points with the image plane is called the boresight error of the imaging system. See Figure 5.3.
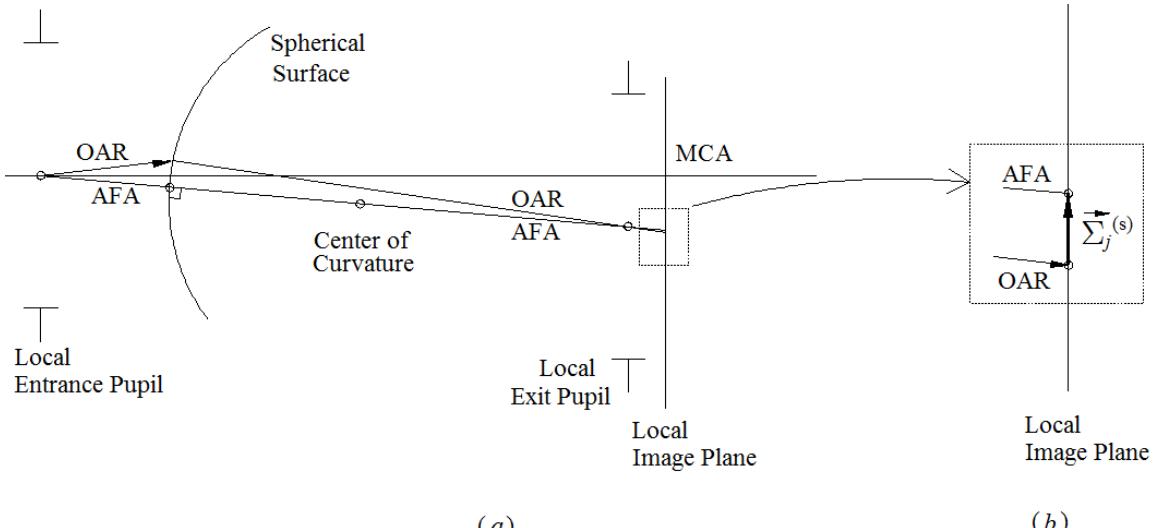
Figure 5.3   For a rotationally symmetric imaging system with one or more tilted and/or decentered optical surfaces, the OAR and MCA need not coincide, giving rise to the boresight error vector.

When one or more of the optical imaging system's elements (optical surfaces with optical power) are perturbed (decentered and/or tilted) the center of the field dependence of the wavefront aberration function changes. To account for this re-centering of the field dependence, per surface, normalized, sigma offset vectors are introduced into the field dependence of the wavefront aberration function. There is one sigma offset vector associated with the spherical base shape of the surface, $\vec{\sigma}_j^{(S)}$, and one sigma offset vector associated with the aspheric cap shape of the surface, $\vec{\sigma}_j^{(A)}$. For both the per surface normalized sphere sigma offset vector $\vec{\sigma}_j^{(S)}$ and the per surface normalized asphere sigma offset vector $\vec{\sigma}_j^{(A)}$ there is a choice to define them either in the normalized local image plane or the normalized local object plane. In this work, these vectors will be defined in the normalized local image plane. Once these normalized sigma vector values have been obtained with respect to the normalized local image plane of an optical surface *j*, they need to be propagated to the system's normalized object plane so that they may be used together with (combined with) the normalized field parameters $\vec{H}$, defined in this work in the system's normalized object plane. This may involve a "sign flip" for the normalized local sigma offset vector values depending on the orientation of the local image plane with respect to the system's object plane. The details are presented below.

Thompson, et al., [66] proposed an equation for the sphere sigma offset vector to facilitate its calculation irrespective of how the local coordinate system is defined in optical design software. However, this research has discovered that while the proposed equation provides the correct magnitude for the sphere sigma vector components there is

an overall sign issue with the proposed equation for some optical systems. Likewise, there is a sign issue for the reported [66] equation for the asphere sigma offset vector for some optical systems. For this reason, tracking the necessary sign flips is made explicit in the following development.

For the sphere base shape of a refracting/reflecting optical surface j, the center of the field dependence for an aberration term in the wavefront aberration function expansion is located in the normalized local image (LI) plane by the 2D normalized sphere sigma offset vector $\vec{\sigma}_j^{(s)LI}$. It is defined to be the 2D vector $\vec{\Sigma}_j^{(s)}$ in surface j's local image plane divided by the absolute value of the chief ray height in the local image plane for the unperturbed optical model. The vector $\vec{\Sigma}_j^{(s)}$ is defined in the local image plane from the intersection point of the OAR to the intersection point of the AFA. The AFA passes through the center of the local entrance (and exit) pupil and through the center of curvature of the surface. This is illustrated in Figure 5.4. Note that the AFA considered as a ray has an incident and refraction/reflection angle of zero at surface j because it is perpendicular to the surface at the intersection point by construction.

Figure 5.4 Definition of the vector $\vec{\Sigma}_j^{(s)}$ for a sphere base shape surface. (*a*) The general setup. (*b*) Magnified view showing the vector definition. "AFA" is a line through the center of the entrance pupil and through the center of curvature of the sphere base shape that intersects with the local image plane.

Based on the definition of the $\vec{\Sigma}_j^{(s)}$ vector shown in Figure 5.4 the following definition can be written for the local image space, normalized sigma offset vector

$$\vec{\sigma}_j^{(s)LI} \equiv \left[\frac{\vec{\Sigma}_j^{(s)}}{\left| HCY_j^{LI} \right|} \right]_{(x,y)} = \left[\frac{\overline{AFA_j^{LI}} - \overline{OAR_j^{LI}}}{\left| HCY_j^{LI} \right|} \right]_{(x,y)} , \tag{5.1}$$

where superscript "*LI* " stands for "local image" plane of surface j, $\overline{AFA_j^{LI}}$ is the vector to the intersection point of the AFA with the local image plane of surface j, $\overline{OAR_j^{LI}}$ is the vector to the intersection of the OAR with the local image plane of surface j, and $\left| HCY_j^{LI} \right|$ is the absolute value of the height of the maximal chief ray in the local image plane of surface j for the unperturbed optical imaging system. The $\left[. \right]_{(x,y)}$ notation is used to explicitly indicate that the resulting vector is a 2D vector in the *xy*-plane.

For the aspheric cap component of the surface's overall shape, unlike for the base sphere, there is a unique surface vertex point, which is labeled "V" in Figure 5.5, about which the aspheric cap shape is rotationally symmetric. The center of the field dependence for an aberration term in the wavefront aberration function expansion associated with the aspheric cap of a surface j is located in the local normalized image plane by the 2D local image space normalized asphere sigma offset vector $\vec{\sigma}_j^{(A)LI}$. However, to obtain this asphere sigma vector, a vector $\vec{\Sigma}_j^{(A)(Surf)}$ is first defined at the local surface j. Once the $\vec{\Sigma}_j^{(A)(Surf)}$ vector value at surface j is obtained it is projected to the local image plane of surface j and normalized by the absolute value of the maximal chief ray height at the local image plane of the unperturbed system. The $\vec{\Sigma}_j^{(A)(Surf)}$ vector defined at surface j is the vector from the OAR's intersection point at surface j to the aspheric cap's vertex point "V". See Figure 5.5. As an equation, this can be written as

$$\vec{\Sigma}_j^{(A)(surf)} = \left[\vec{V}_j - \overrightarrow{OAR}_j \right]_{(x,y)} ,$$
(5.2)

where $\vec{V}_j$ is a vector to surface j's vertex point, and $\overrightarrow{OAR}_j$ is a vector to the intersection point of the OAR with surface j. The $\left[. \right]_{(x,y)}$ notation is used to explicitly indicate that the resulting vector is a 2D vector in the xy vertex plane of the surface. The surface's vertex plane is the plane containing the surface's vertex point and is perpendicular to the MCA (the z-axis) regardless of the tilts and decenters applied to the surface.
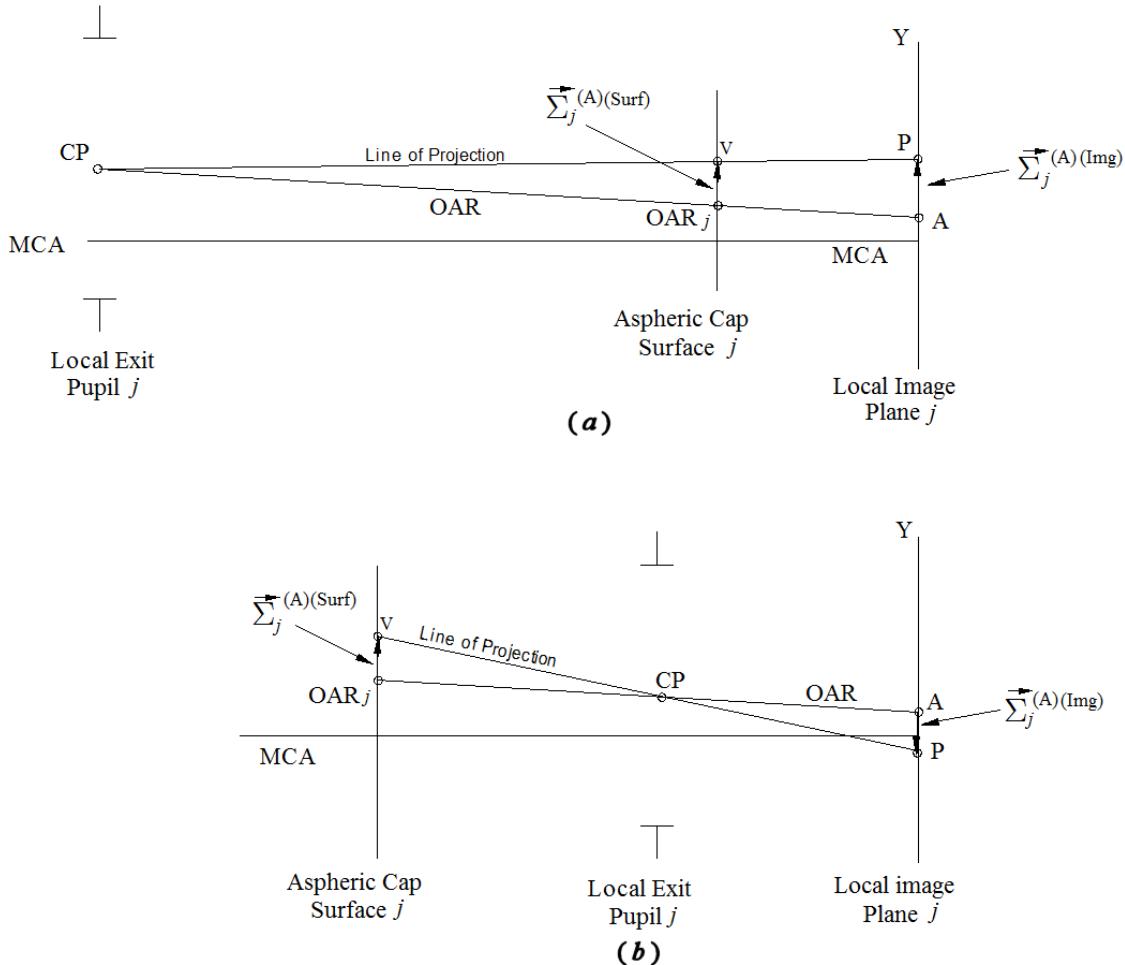
Figure 5.5 The $\vec{\Sigma}_j^{(A)(Surf)}$ vector is defined to be the vector from the OAR at surface *j* to the aspheric cap's vertex point "V". It is to be projected to the local image plane. (*a*) For the case that the local exit pupil occurs to the left of the local image plane and the surface *j*. (Based on Fig. 2.8 of [54].) (*b*) For the case that the local exit pupil occurs to the left of the local image plane but to the right of surface *j*. Notice the sign flip for case (*b*) in going from surface *j* to the local image plane. CP = center of the exit pupil.

By defining the "Line of Projection" to be the line connecting the center of the local exit pupil "CP" and the asphere's vertex point "V", and extending this line to the local image plane, similar triangles are formed. In this way, the $\vec{\Sigma}_j^{(A)(surf)}$ vector is projected to the local image plane. Note that, as shown in Figure 5.5(*b*), and explicitly identified in this research, there may be a sign flip associated with this projection. To

account for this sign flip, the sign of the chief ray height at the local image plane is multiplied by the sign of the chief ray height at surface j, both for the unperturbed optical model. This is written as

$$\vec{\Sigma}_j^{(A)(img)} = \text{sgn}\left(HCY_j^{img}\right)\text{sgn}\left(HCY_j^{surf}\right)\frac{\left|HCY_j^{img}\right|}{\left|HCY_j^{surf}\right|}\vec{\Sigma}_j^{(A)(surf)} \ , \tag{5.3}$$

where $\text{sgn}\left(HCY_j^{img}\right)$ is the sign of the chief ray at the unperturbed local image plane, $\text{sgn}\left(HCY_j^{surf}\right)$ is the sign of the chief ray at the unperturbed surface j, $\left|HCY_j^{img}\right|$ is the absolute value of the chief ray height in the local image plane of surface j for the unperturbed model, and $\left|HCY_j^{surf}\right|$ is the absolute value of the chief ray height at surface j also for the unperturbed model.

The normalized asphere sigma offset vector associated with surface j, in the local normalized image plane, is then given by

$$\vec{\sigma}_j^{(A)LI} = \text{sgn}\left(HCY_j^{img}\right)\text{sgn}\left(HCY_j^{surf}\right)\frac{1}{\left|HCY_j^{surf}\right|}\vec{\Sigma}_j^{(A)(surf)} \ . \tag{5.4}$$
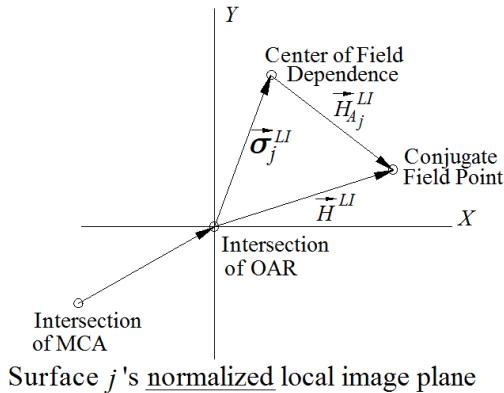
Simplifying gives

$$\vec{\sigma}_j^{(A)LI} = \text{sgn}\left(HCY_j^{img}\right)\frac{\left[\vec{V}_j - \overrightarrow{OAR}_j\right]_{(x,y)}}{HCY_j^{surf}} \ , \tag{5.5}$$

where the absolute value of the height of the chief ray at surface j, HCY_j^{surf}, in the unperturbed system has been removed.

Note that since neither $\vec{V}_j$ nor $\overrightarrow{OAR}_j$ at surface j depend on the tilt of the aspheric

cap (about the surface's vertex $\vec{V}_j$), the $\vec{\sigma}_j^{(A)LI}$ offset vectors are independent of surface

j's tilt [54].

These per surface normalized sigma offset vectors (sphere $\vec{\sigma}_j^{(S)LI}$ and asphere

$\vec{\sigma}_j^{(A)LI}$) are used to identify the center of the field dependence in the optical system's

normalized *local* Gaussian image plane for the aberration terms in the wavefront

aberration function expansion. This is illustrated in Figure 5.6.

Figure 5.6 A normalized sigma offset vector (sphere or aspheric cap) signifies a shift in the center of the field dependence of an aberration term in the normalized <u>local</u> image (LI) plane. (Based on Fig. 9 of [66].)

The "*LI*" superscripts in Figure 5.6 are used to emphasize that the parameters

shown are with respect to the <u>local</u> image space of surface j. Recall that the wavefront

aberration function has been defined in this work to be a function of the total system's

object space normalized Cartesian field parameter $\vec{H}$. Because the normalized object

field parameter $\vec{H}$ and the image conjugate normalized field parameter $\vec{H}^{LI}$, defined in

the normalized local image plane, are the same up to sign, a sign flip may be necessary

when introducing the local image space defined sigma offset vectors into the object space defined wavefront aberration function. The identification of this sign flip is made explicit in this research and is shown in the following derivation.

Consider the wavefront aberration function expansion written in the following way

$$W\left(\vec{H},\vec{\rho}\right)=W\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right)=W_{j=1}\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right)+W_{j=2}\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right)+\cdots+W_{j=N}\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right),(5.6)$$

where the field and pupil parameters are normalized parameters, and the j subscript is indicating the surface number. The $W_{j}\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right)$ are the aberration contributions to the total wavefront aberration function attributable to surface j. There are two cases to consider. The first case is that the optical subsystem consisting of all surfaces up to and including surface j is an image inverting optical subsystem. The second case is that the optical subsystem consisting of all surfaces up to and including surface j is not an image inverting optical subsystem.

CASE I:

For an arbitrary but fixed $j \leq N$, and assuming that subsystem $S_{j' \leq j}$ is an *image inverting* optical system defined by the relations, for normalized parameters,

$$\vec{H}^{Obj} = -\vec{H}^{LI} ,\tag{5.7}$$

$$\vec{\rho}^{Obj} = -\vec{\rho}^{LI} ,\tag{5.8}$$

then

$$W_{j}\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right)=W_{j}\left(-\vec{H}^{LI},-\vec{\rho}^{LI}\right) .\tag{5.9}$$

Following the NAT prescription to re-center the field dependence for an optical imaging system having decentered or tilted optical elements, the substitution $\vec{H}^{LI} \to \vec{H}^{LI}_{Aj}$ [54] is made, giving

$$W_j\left(-\vec{H}^{LI}, -\vec{\rho}^{LI}\right) \to W_j\left(-\vec{H}^{LI}_{Aj}, -\vec{\rho}^{LI}\right) . \tag{5.10}$$

Using the vector relation, illustrated in Figure 5.6,

$$\vec{H}^{LI}_{Aj} = \vec{H}^{LI} - \vec{\sigma}^{LI}_j , \tag{5.11}$$

gives

$$W_j\left(-\vec{H}^{LI}_{Aj}, -\vec{\rho}^{LI}\right) = W_j\left(-\left[\vec{H}^{LI} - \vec{\sigma}^{LI}_j\right], -\vec{\rho}^{LI}\right) = W_j\left(-\vec{H}^{LI} + \vec{\sigma}^{LI}_j, -\vec{\rho}^{LI}\right) . \tag{5.12}$$

Again using Eq. (5.7) and Eq. (5.8) for an image inverting optical subsystem,

$$W_j\left(-\vec{H}^{LI} + \vec{\sigma}^{LI}_j, -\vec{\rho}^{LI}\right) = W_j\left(\vec{H}^{Obj} + \vec{\sigma}^{LI}_j, \vec{\rho}^{Obj}\right) . \tag{5.13}$$

This sequence of relations has shown, for the case that the subsystem $S_{j' \le j}$ is image inverting, that the NAT prescription for the re-centering of the field dependence for rotationally nonsymmetric optical imaging systems is given by

$$W_j\left(\vec{H}^{Obj}, \vec{\rho}^{Obj}\right) \to W_j\left(\vec{H}^{Obj} + \vec{\sigma}^{LI}_j, \vec{\rho}^{Obj}\right) . \tag{5.14}$$

CASE II:

For the case that the subsystem $S_{j' \le j}$ is *not an image inverting* optical system,

$$\vec{H}^{Obj} = \vec{H}^{LI} , \tag{5.15}$$

$$\vec{\rho}^{Obj} = \vec{\rho}^{LI} . \tag{5.16}$$

Then the following relation holds,

$$W_j\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right)=W_j\left(\vec{H}^{LI},\vec{\rho}^{LI}\right) \; . \tag{5.17}$$

Following the NAT prescription to re-center the field dependence for an optical imaging system having decentered or tilted optical elements, the substitution $\vec{H}^{LI} \rightarrow \vec{H}^{LI}_{Aj}$ is made, giving

$$W_j\left(\vec{H}^{LI},\vec{\rho}^{LI}\right)\rightarrow W_j\left(\vec{H}^{LI}_{Aj},\vec{\rho}^{LI}\right) \; . \tag{5.18}$$

Using the vector relation, illustrated in Figure 5.6,

$$\vec{H}^{LI}_{Aj} = \vec{H}^{LI} - \vec{\sigma}^{LI}_{j} \; , \tag{5.19}$$

gives

$$W_j\left(\vec{H}^{LI}_{Aj},\vec{\rho}^{LI}\right)=W_j\left(\left[\vec{H}^{LI}-\vec{\sigma}^{LI}_{j}\right],\vec{\rho}^{LI}\right)=W_j\left(\vec{H}^{LI}-\vec{\sigma}^{LI}_{j},\vec{\rho}^{LI}\right) \; . \tag{5.20}$$

Again using Eq. (5.15) and Eq. (5.16) for a non-inverting imaging optical subsystem,

$$W_j\left(\vec{H}^{LI}-\vec{\sigma}^{LI}_{j},\vec{\rho}^{LI}\right)=W_j\left(\vec{H}^{Obj}-\vec{\sigma}^{LI}_{j},\vec{\rho}^{Obj}\right) \; . \tag{5.21}$$

This sequence of relations has shown, for the case that the subsystem $S_{j'\leq j}$ is _not_ image inverting, that the NAT prescription for the re-centering of the field dependence for rotationally nonsymmetric optical imaging systems is given by

$$W_j\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right)\rightarrow W_j\left(\vec{H}^{Obj}-\vec{\sigma}^{LI}_{j},\vec{\rho}^{Obj}\right) \; . \tag{5.22}$$

Then, from these two cases, the re-centering of the field dependence in the wavefront aberration function is written as

$$W_j\left(\vec{H}^{Obj},\vec{\rho}^{Obj}\right)\rightarrow W_j\left(\vec{H}^{Obj}\pm\vec{\sigma}^{LI}_{j},\vec{\rho}^{Obj}\right) \tag{5.23}$$

One way to avoid this use of the "$\pm$" symbol, as well as the superscripts "*Obj*" and "*LI*", while retaining the form previously published in the literature [5] for re-centering the field dependence, is to change the equations for the sigma offset vectors to incorporate the potential orientation sign flip of the local image plane with respect to the total system's object plane. That is, define

$$\vec{\sigma}_j^{(S)} = \pm\vec{\sigma}_j^{(S)LI} \ , \tag{5.24}$$

and

$$\vec{\sigma}_j^{(A)} = \pm\vec{\sigma}_j^{(A)LI} \ , \tag{5.25}$$

where the "+" is used when the optical subsystem $S_{j' \le j}$ is not image inverting, and "$-$" is used when it is an image inverting subsystem.

By using HCY_{Tot}^{Obj} for the (maximal) chief ray height at the total system's object plane (which may be infinite), and HCY_j^{Img} for the chief ray height at surface *j*'s <u>local</u> image plane, the sphere sigma equation can be written as

$$\vec{\sigma}_j^{(S)} = \operatorname{sgn}\left(HCY_{Tot}^{Obj}\right)\operatorname{sgn}\left(HCY_j^{Img}\right)\vec{\sigma}_j^{(S)LI} \ . \tag{5.26}$$

And for the asphere sigma offset vector, it may be written as

$$\vec{\sigma}_j^{(A)} = \operatorname{sgn}\left(HCY_{Tot}^{Obj}\right)\operatorname{sgn}\left(HCY_j^{Img}\right)\vec{\sigma}_j^{(A)LI} \ . \tag{5.27}$$

Using Eq. (5.1) and Eq. (5.5) these may be written as

$$\vec{\sigma}_j^{(S)} = \operatorname{sgn}\left(HCY_{Tot}^{Obj}\right)\operatorname{sgn}\left(HCY_j^{LI}\right)\frac{\left[\overrightarrow{AFA_j^{LI}} - \overrightarrow{OAR_j^{LI}}\right]_{(x,y)}}{\left|HCY_j^{LI}\right|} \ , \tag{5.28}$$

and

$$\vec{\sigma}_j^{(A)} = \mathrm{sgn}\left(HCY_{Tot}^{Obj}\right)\mathrm{sgn}\left(HCY_j^{surf}\right)\frac{\left[\vec{V}_j - \overrightarrow{OAR}_j\right]_{(x,y)}}{\left|HCY_j^{surf}\right|} \ . \tag{5.29}$$

Note that it is customary to use the absolute value of a quantity that is to be used as a normalization factor so that the direction of the vector to be normalized is not inadvertently changed. Hence the absolute value is not removed from Eq. (5.28) and Eq. (5.29). This also helps to keep track of the source of the sign flips.

Explicitly providing the sign flips in equations Eq. (5.28) and Eq. (5.29) is new to this research, although their concept (the definitions of the sigma offset vectors) is not. These equations are to be assumed when referring to the sigma offset vectors $\vec{\sigma}_j$ throughout the rest of this dissertation unless explicitly stated otherwise. They are written in this form, using the sgn(.) notation, to keep track of the source of sign flips and to aid in understanding how they may be implemented in a computer program.

Note that these equations are applicable for the case that the wavefront aberration function's normalized field parameters $\vec{H}$ are the object space normalized Cartesian coordinate field parameters, as used throughout this dissertation. Some researcher's may prefer to use normalized image space field parameters when defining the wavefront aberration function's field parameters. Let HCY_{Tot}^{Img} be the chief ray height at the total system's image plane. The sigma offset vector equations, when using the total system's image plane to define the field parameters of the wavefront aberration function are given by

$$\vec{\sigma}_j^{(S)} = \mathrm{sgn}\left(HCY_{Tot}^{Img}\right)\mathrm{sgn}\left(HCY_j^{Img}\right)\vec{\sigma}_j^{(S)LI} \ , \tag{5.30}$$

and

$$\vec{\sigma}_j^{(A)} = \text{sgn}\left(HCY_{Tot}^{Img}\right)\text{sgn}\left(HCY_j^{Img}\right)\vec{\sigma}_j^{(A)LI} \ . \qquad (5.31)$$

Whether the normalized field parameters $\vec{H}$ are defined using the system's object plane or the system's image plane, the form of the equations for re-centering of the wavefront aberration function is generally given as [5]

$$W_j\left(\vec{H}, \vec{\rho}\right) \rightarrow W_j\left(\vec{H} - \vec{\sigma}_j, \vec{\rho}\right), \qquad (5.32)$$

where it is now to be understood that the sigma offset vectors incorporates the possible sign flips as given in Eq. (5.26) and Eq. (5.27) or in Eq. (5.30) and Eq. (5.31) depending on the choice of the wavefront aberration function's field parameterization.

As discussed in the previous Chapter, there are four types of $W_{klm;j}^{(Type)}$ expansion coefficients (*Type* = *IS*, *IA*, *ES*, *EA*) and, at this time, only two types of sigma offset vectors (sphere and aspheric cap). When re-centering the field dependence of the wavefront aberration function, the *IS*, *ES* and *EA* are to be associated with the sphere sigma offset vectors, and the *IA* is to be associated with the aspheric cap sigma vectors. As an example, consider the through 6th order field dependence for the Zernike astigmatism term $\vec{Z}_2^2\left(\vec{\rho}\right)$ for a rotationally symmetric optical imaging system

$$\vec{F} = \sum_{j}^{\substack{\# \, Surfaces \\ \& \, Types}} \left(\frac{1}{2}W_{222;j} + \frac{3}{8}W_{242;j} + \frac{1}{2}W_{422;j}H^2\right)\left\{\vec{H}\right\}_\eta^2 \ . \qquad (5.33)$$

Writing the expansion coefficient types explicitly gives

$$\vec{F} = \sum_{j}^{\# Surfaces} \left(\frac{1}{2}W_{222;j}^{IS} + \frac{3}{8}W_{242;j}^{IS} + \frac{1}{2}W_{422;j}^{IS}H^2\right)\left\{\vec{H}\right\}_{\eta}^2 +$$

$$\left(\frac{1}{2}W_{222;j}^{IA} + \frac{3}{8}W_{242;j}^{IA} + \frac{1}{2}W_{422;j}^{IA}H^2\right)\left\{\vec{H}\right\}_{\eta}^2 +$$

$$\left(\frac{1}{2}W_{222;j}^{ES} + \frac{3}{8}W_{242;j}^{ES} + \frac{1}{2}W_{422;j}^{ES}H^2\right)\left\{\vec{H}\right\}_{\eta}^2 +$$

$$\left(\frac{1}{2}W_{222;j}^{EA} + \frac{3}{8}W_{242;j}^{EA} + \frac{1}{2}W_{422;j}^{EA}H^2\right)\left\{\vec{H}\right\}_{\eta}^2 \quad . \tag{5.34}$$

For the case of a rotationally nonsymmetric optical system, the re-centering of the field dependence is accomplished by the NAT prescription described above. This gives

$$\vec{F} = \sum_{j}^{\# Surfaces} \left(\frac{1}{2}W_{222;j}^{IS} + \frac{3}{8}W_{242;j}^{IS} + \frac{1}{2}W_{422;j}^{IS}\left(\vec{H} - \vec{\sigma}_j^{(S)}\right)\bullet\left(\vec{H} - \vec{\sigma}_j^{(S)}\right)\right)\left\{\vec{H} - \vec{\sigma}_j^{(S)}\right\}_{\eta}^2 +$$

$$\left(\frac{1}{2}W_{222;j}^{IA} + \frac{3}{8}W_{242;j}^{IA} + \frac{1}{2}W_{422;j}^{IA}\left(\vec{H} - \vec{\sigma}_j^{(A)}\right)\bullet\left(\vec{H} - \vec{\sigma}_j^{(A)}\right)\right)\left\{\vec{H} - \vec{\sigma}_j^{(A)}\right\}_{\eta}^2 +$$

$$\left(\frac{1}{2}W_{222;j}^{ES} + \frac{3}{8}W_{242;j}^{ES} + \frac{1}{2}W_{422;j}^{ES}\left(\vec{H} - \vec{\sigma}_j^{(S)}\right)\bullet\left(\vec{H} - \vec{\sigma}_j^{(S)}\right)\right)\left\{\vec{H} - \vec{\sigma}_j^{(S)}\right\}_{\eta}^2 +$$

$$\left(\frac{1}{2}W_{222;j}^{EA} + \frac{3}{8}W_{242;j}^{EA} + \frac{1}{2}W_{422;j}^{EA}\left(\vec{H} - \vec{\sigma}_j^{(S)}\right)\bullet\left(\vec{H} - \vec{\sigma}_j^{(S)}\right)\right)\left\{\vec{H} - \vec{\sigma}_j^{(S)}\right\}_{\eta}^2 \quad , \tag{5.35}$$

where the H^2 have first been replaced by their vector equivalent $\vec{H}\bullet\vec{H}$, and where $\vec{\sigma}_j^{(S)}$ is the sigma offset vector associated with the sphere base shape of surface j, Eq. (5.26), and $\vec{\sigma}_j^{(A)}$ is the sigma offset vector associated with the aspheric cap of surface j, Eq. (5.27).

Note that once the OAR is deflected from the MCA, for example, by surface j, all other optical surfaces after surface j may have sigma offset vectors associated with them even though these surfaces may not be tilted nor decentered. For this reason, the sigma
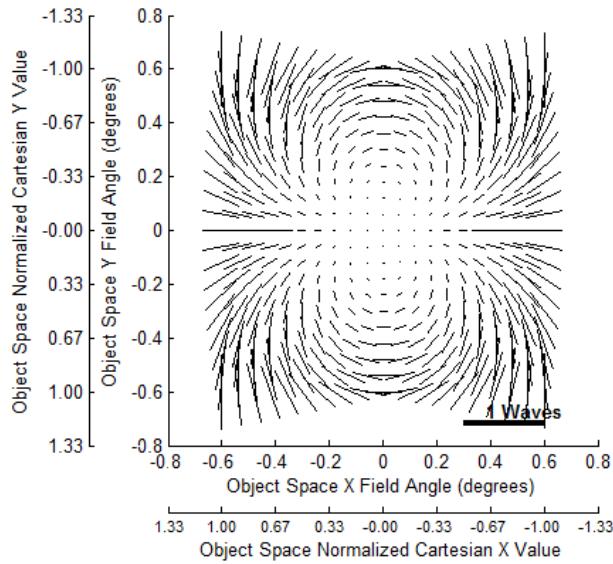
offset vectors are not necessarily directly associated with the tilt or decenter parameters of a surface in the optical model.

It is expected that the magnitude of the normalized sigma offset vectors, for small perturbations of an optical surface, will be less than or approximately one. For larger values, it is not clear that the use of the sigma vectors in the wavefront aberration function expansion doesn't significantly alter the convergence of the function's expansion. In one example to be presented in Chapter 6, the magnitude of a normalized sigma vector is greater than 20. This, together with the limited accuracy of the W_{klm} expansion coefficients, and the truncation order of the expansion used, may lead to inaccuracy in the resulting field dependence for the optical model considered. The range of acceptability for the magnitude of the sigma offset values has not been formally addressed here. However, it seems clear that sigma values much greater than one may present issues in the computational accuracy of the field functions.

5.2 Review of Full Field Displays (FFDs)

FFDs are a graphical way of illustrating the vector field dependence (the field functions) of the terms in the wavefront aberration function expansion in the field parameter space (the $\vec{H} = \left(H_x, H_y \right)$ space). (Note that this is not the image plane of the system. Although the normalized field parameters can be mapped to the image plane, the wavefront aberration function, and hence the field functions, is not defined in the image plane.) The FFDs are the primary method for quantitatively and qualitatively validating the mathematical development in this research. They are the primary means for assessing

the field dependence of the wavefront aberration function for optical imaging systems.

An example of a FFD is shown in Figure 5.7. This FFD was created from real ray trace

data for a 2 mirror Ritchey-Chrétien telescope model. The half-field of view (HFOV) for

the optical model is 0.6 degrees. Other details of the optical model will be presented in

the next Chapter.

Figure 5.7 Example of a FFD for the Zernike astigmatism term $\vec{Z}_2^2(\vec{\rho})$.

Figure 5.7 shows a plot of the field dependence of the wavefront aberration

function expansion for the Zernike astigmatism terms $\vec{Z}_2^2(\vec{\rho}) = \left(Z_2^{+2}(\vec{\rho}), Z_2^{-2}(\vec{\rho}) \right)$

through 12th order. Specifically, the double Zernike expansion of the wavefront

aberration function, Eq. (4.53), was used together with the GQ method described in

Appendix III to obtain the expansion coefficients $U_{n_H, n_\rho}^{m_H, m_\rho}$. The input to the GQ method

was the OPD real ray trace data obtained from CODE V® for the optical model. With

the   expansion   coefficients   determined,   the   field   dependent   vector   function

$\vec{F}_2^2 \left(H_x, H_y \right) = \left(F_2^{+2}, F_2^{-2} \right)$ associated with the $\vec{Z}_2^2 (\vec{\rho})$ terms were then known as

functions of the normalized Cartesian object field parameters $\vec{H} = \left(H_x, H_y \right)$. From Eq.

(4.53)

$$F_2^{+2} = \sum_{n_H = 0}^{12} \sum_{\substack{m_H = -n_H \\ m_H + 2}}^{n_H} U_{n_H, 2}^{m_H, +2} Z_{n_H}^{m_H} \left(\vec{H} \right) , \qquad (5.36)$$

$$F_2^{-2} = \sum_{n_H = 0}^{12} \sum_{\substack{m_H = -n_H \\ m_H + 2}}^{n_H} U_{n_H, 2}^{m_H, -2} Z_{n_H}^{m_H} \left(\vec{H} \right) \qquad (5.37)$$
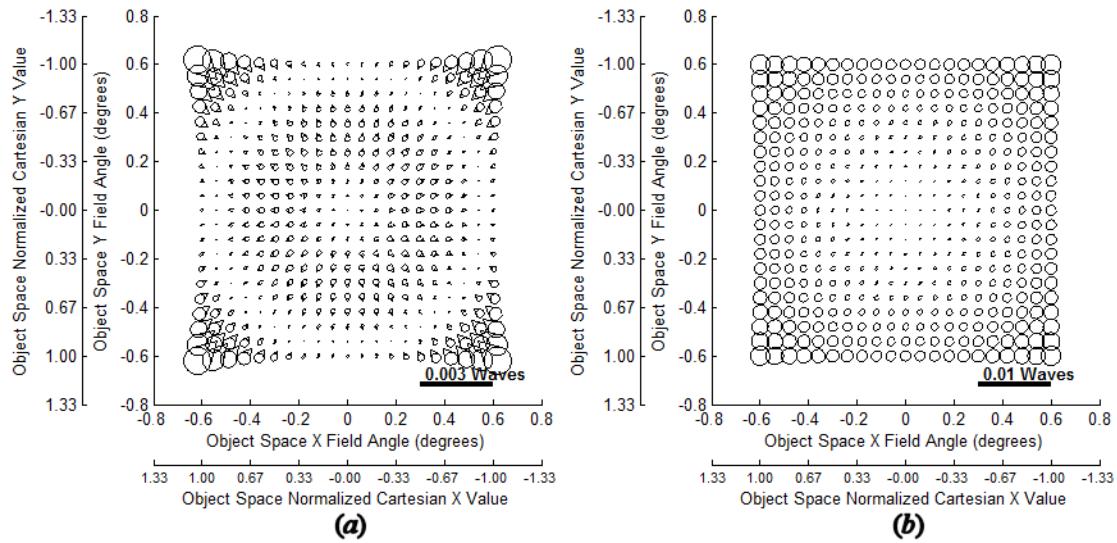
A grid of 21×21 normalized object field points was used together with the field

dependent vector function $\vec{F}_2^2$ to calculate the vector value at each of the normalized

field grid points. A symbol was plotted at each of the grid points to represent the vector

values $\vec{F}_2^2$ of the field dependence of the $\vec{Z}_2^2 (\vec{\rho})$ aberration function's expansion terms.

In Figure 5.7, the symbol used is a line segment. The length of the line segment

corresponds to the magnitude of the vector's magnitude, $\left\| \vec{F}_2^2 \right\|$. The orientation of the

line segment likewise corresponds to the orientation of the $\vec{F}_2^2$ vector's orientation in the

field parameter's space $\left(H_x, H_y \right)$ coordinate plane. It is customary to ignore the "head"

and "tail" aspect of the symbol used to represent the Zernike astigmatism aberration $\vec{F}_2^2$

vector, hence the line segment is used as the symbol plotted. A scale bar is provided in

the lower right hand corner of the plot, in units of waves.

The FFD of Figure 5.7 is presented with two axis scales for both the *x*- and *y*-

axes. In this work the field parameters to the wavefront aberration function are the

normalized Cartesian x and y values $\left(H_x, H_y\right)$ identifying the object source point for the ray being traced through the optical system. In optical design it is more common to use the associated object angles to identify the object source point, particularly when the object is effectively at (minus) infinity along the z-axis. However, since the relation between the object angles and the Cartesian coordinate system of the object plane changes sign depending on whether or not the entrance pupil occurs to the left or to the right of the object plane, both parameterizations are here shown to avoid any confusion.

In addition to the Zernike astigmatism FFD shown above, two other FFDs are often produced to display the field dependence of the Zernike coma $\vec{Z}_3^1(\vec{\rho}) = \left(Z_3^{+1}(\vec{\rho}), Z_3^{-1}(\vec{\rho})\right)$ and the Zernike spherical $\vec{Z}_4^0(\vec{\rho}) = \left(Z_4^{+0}(\vec{\rho}), 0\right)$ aberration expansion terms. The following FFDs provide examples for these FFD types using the same telescope model as for the Zernike astigmatism plot above.

Figure 5.8 Examples of (a) Zernike coma and (b) Zernike spherical FFDs.

As shown in Figure 5.8, the Zernike coma plots use cone-like symbols $\bigtriangledown$ to represent the $\vec{F}_3^1$ field dependent vector function values with the tip of the cone pointing in the vector direction, while the Zernike spherical plots use a circle symbols to represent the $\vec{F}_4^0$ field dependent vector function values. Other symbols are used in the FFDs associated with other types of Zernike aberration terms of the wavefront aberration function expansion.

As can be seen in the Zernike coma plot, there is a center nodal point (zero value) at $(0, 0)$ and a set of field points forming a ring of zeros. In the terminology of NAT, this ring of zeros is not considered to be "nodal points". It is simply a ring of zeros. This ring of zeros in the field dependence of the Zernike coma term $\vec{Z}_3^1(\vec{\rho})$ results from the associated field function (shown here through 6$^{\text{th}}$ order)

$$\vec{F}_3^1 = \left(\frac{1}{3}W_{131} + \frac{2}{5}W_{151}\right)\{\vec{H}\}^1 + \left(\frac{1}{3}W_{331M}H^2\right)\{\vec{H}\}^1 , \tag{5.38}$$
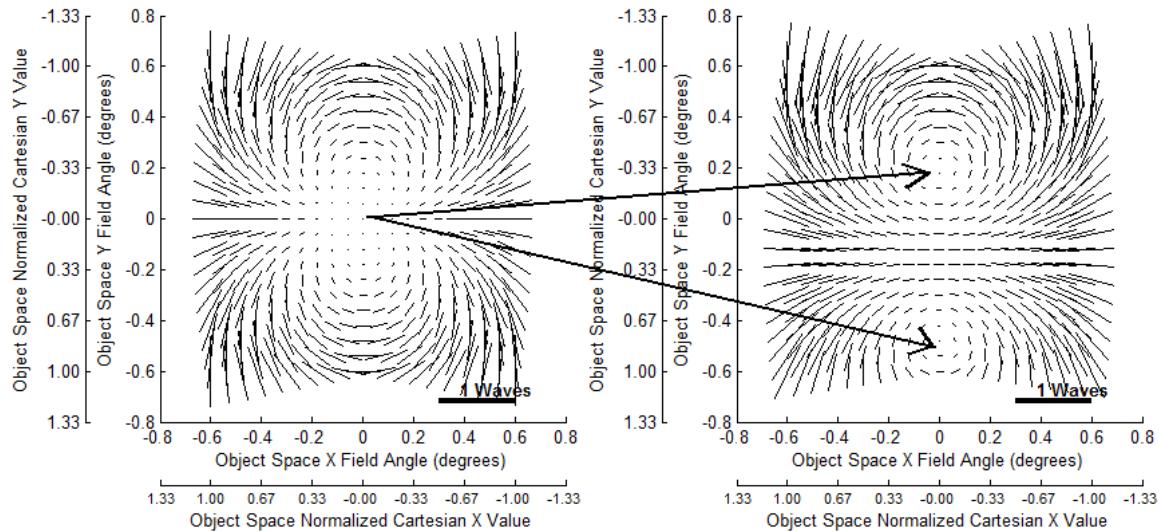
having a field linear part and a field cubic part. Where the two contributing parts cancel defines the ring of zeros.

5.3 An Example of Nodal Point Splits and Nodal Locations

When an optical surface in a rotationally symmetric optical imaging system is perturbed (decentered and/or tilted) the center node point in a FFD may split into multiple nodes. Thompson [5, 54, 61] originally developed analytic equations for the resulting nodal point locations using the Shack vector expansion of the wavefront aberration function, see Eq. (4.9), through 4$^{\text{th}}$ order using a complex number formalism for SVP. Later, in a

series of papers [62, 63, 64], Thompson extended this to include through 6th order aberration terms. For illustrative purposes only, the binodal split of the Zernike astigmatism term for 4th order is presented here using the Zernike expansion of the wavefront aberration function and the GA definition for SVP as developed in this research.

Continuing with the telescope model of the previous section, the primary mirror of the model is decenter in the +y-axis direction by 3 *mm* and the Zernike astigmatism FFD is again generated. The single nodal point in the Zernike astigmatism FFD, originally located at the center $\vec{H} = (0,0)$ of the plot is now seen to split into two nodal points.

Figure 5.9   Example of a nodal point split of Zernike astigmatism.   Left: Rotationally symmetric optical imaging system. Right: Primary mirror decentered by y = +3mm.

This behavior can be calculated by considering the Zernike astigmatism vector field function $\vec{F}_2^2$. Assuming that the Zernike astigmatism is dominated by 4th order

aberrations, the field dependence of $\vec{Z}_2^2(\vec{\rho})$ for the rotationally symmetric model (see Eq. (4.49)) can be approximated to be

$$\vec{F}_2^2 = \sum_{j}^{\substack{\#\,Surfaces \\ \&\,Types}} \frac{1}{2} W_{222;j} \left\{\vec{H}\right\}_{\eta}^{2} \; . \tag{5.39}$$

For the rotationally nonsymmetric case, this vector function becomes, by NAT recentering of the field parameter

$$\vec{F}_2^2 = \sum_{j}^{\substack{\#\,Surfaces \\ \&\,Types}} \frac{1}{2} W_{222;j} \left\{\vec{H} - \vec{\sigma}_j^{(Type)}\right\}_{\eta}^{2} \; . \tag{5.40}$$

Setting this equal to the zero vector, and recalling the algebraic properties of SVP, the equation for the nodal points can be written as

$$\sum_{j}^{\substack{\#\,Surfaces \\ \&\,Types}} \frac{1}{2} W_{222;j} \left[\left\{\vec{H}\right\}_{\eta}^{2} - 2\vec{H} \star_{\eta} \vec{\sigma}_j^{(Type)} + \left\{\vec{\sigma}_j^{(Type)}\right\}_{\eta}^{2} \right] = \vec{0} \; . \tag{5.41}$$

Distributing the summation gives

$$\sum_{j}^{\substack{\#\,Surfaces \\ \&\,Types}} \frac{1}{2} W_{222;j} \left\{\vec{H}\right\}_{\eta}^{2} + \sum_{j}^{\substack{\#\,Surfaces \\ \&\,Types}} \frac{1}{2} W_{222;j} \left(-2\vec{H} \star_{\eta} \vec{\sigma}_j^{(Type)}\right) +$$
$$\sum_{j}^{\substack{\#\,Surfaces \\ \&\,Types}} \frac{1}{2} W_{222;j} \left\{\vec{\sigma}_j^{(Type)}\right\}_{\eta}^{2} = \vec{0} \; . \tag{5.42}$$

And this can be written as

$$\left\{\vec{H}\right\}_{\eta}^{2} + \frac{\vec{B}}{A} \star_{\eta} \vec{H} + \frac{\vec{C}}{A} = \vec{0} \; , \tag{5.43}$$

where

$$A = \sum_{j}^{\substack{\text{\# Surfaces} \\ \text{\& Types}}} \frac{1}{2} W_{222;j} \quad \text{, a scalar,} \tag{5.44}$$

$$\vec{B} = \sum_{j}^{\substack{\text{\# Surfaces} \\ \text{\& Types}}} -W_{222;j} \vec{\sigma}_{j}^{(Type)} \quad \text{, a vector,} \tag{5.45}$$

$$\vec{C} = \sum_{j}^{\substack{\text{\# Surfaces} \\ \text{\& Types}}} \frac{1}{2} W_{222;j} \left\{ \vec{\sigma}_{j}^{(Type)} \right\}_{\eta}^{2} \quad \text{, a vector.} \tag{5.46}$$

Using Eq. (2.96), the vector form of the quadratic equation using SVP, the two nodal point locations $\vec{H}_{1,2}$ are given by
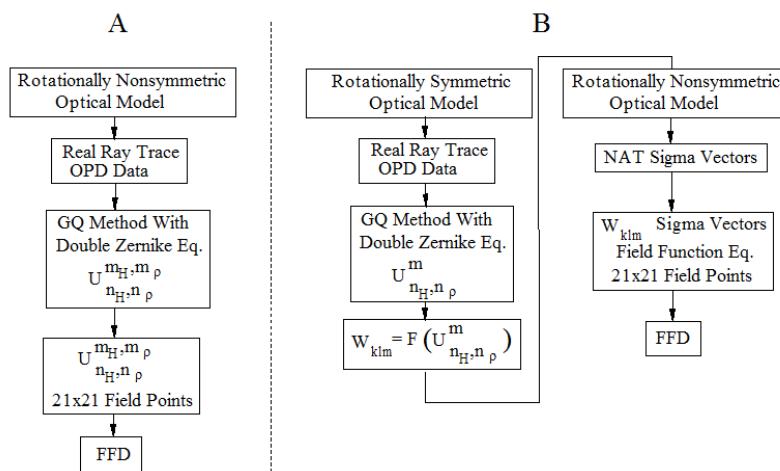
$$\vec{H}_{1,2} = \frac{-1}{2} \frac{\vec{B}}{A} \pm \frac{1}{2} \left\{ \left\{ \frac{\vec{B}}{A} \right\}_{\eta}^{2} - 4 \frac{\vec{C}}{A} \right\}_{\eta}^{1/2} . \tag{5.47}$$

As Thompson has shown, far more complicated nodal patterns and equations for the node locations can emerge when higher order aberration terms are considered [62,63].

Chapter 6 Application of the Theoretical Development

This chapter provides several examples of the application of the material developed in the previous Chapters to idealized optical imaging system models as a means for validating the mathematical development. These examples will consist of mirror based telescope systems. Not all models used as examples will be realizable as actual telescopes. The goal here is to qualitatively and quantitatively check the mathematical development previously presented rather than the performance evaluation of any particular telescope system.

The primary method for qualitatively checking the mathematical development is the visual comparison of the full field displays (FFDs) calculated using the vector field function of terms from the wavefront aberration function's expansions. Two different expansions, and thus two different processes, are used. These processes are graphically presented in Figure 6.1.

Figure 6.1   Outline of the processes for generating FFDs for rotationally nonsymmetric optical imaging systems. Process "A" uses the double Zernike expansion of the wavefront aberration function. Process "B" uses the W_{klm} values, NAT sigma offset vectors, and the field functions from the single Zernike expansion.

The first expansion of the wavefront aberration function used is that of Eq. (4.53) utilizing the double Zernike form, through 12th order in pupil and field, having the $U_{n_H,n_\rho}^{m_H,m_\rho}$ as expansion coefficients. This expansion is valid for rotationally symmetric and rotationally nonsymmetric optical models. The expansion coefficients are obtained from using real ray trace data (OPD values for the rays traced) from CODE V® and the GQ method described in Appendix III. With the expansion coefficients known, and using the double Zernike expansion, the desired FFDs are produced. Typically the FFDs of the vector field functions of the Zernike astigmatism $\vec{Z}_2^2(\vec{\rho})$, the Zernike coma $\vec{Z}_3^1(\vec{\rho})$, and the Zernike spherical $\vec{Z}_4^0(\vec{\rho})$ terms are used.

The second expansion of the wavefront aberration function used is the single Zernike (in pupil parameters) expansion given by Eq. (4.51) valid for rotationally symmetric optical imaging systems only. This expansion is in terms of the W_{klm} expansion coefficients. For a rotationally symmetric optical imaging model, the $U_{n_H,n_\rho}^{m_H,m_\rho}$ expansion coefficients are the U_{n_H,n_ρ}^m expansion coefficients. The W_{klm} expansion coefficients are then obtainable by using Eq. (4.66) through Eq. (4.80) that are in terms of the U_{n_H,n_ρ}^m, as well as additional similar equations for higher order W_{klm} when needed. The per surface, sphere/asphere, intrinsic/extrinsic W_{klm} values are then obtained as detailed in Chapter 4. After obtaining the W_{klm} coefficients, the optical model is perturbed (some of the optical elements are decentered and/or tilted) and the sigma offset vectors of NAT are calculated. Using the vector field functions of the expansion, Eq.

(4.51), the W_{klm} coefficients, and the sigma offset vectors, the FFDs are again produced. These FFDs are visually compared to the FFDs produced using the double Zernike expansion for the same perturbed optical imaging model.

The quantitative check of the mathematical development is performed by calculating the difference between the two FFD's plot data generated as described above. Plots of the differences for each FFD, as well as the numeric value of the mean, standard deviation, and signed maximum difference of the FFD difference data, are provided.

6.1 Early Development, Early Results

During the early stages of the research described in this dissertation, the wavefront aberration function expansion in terms of Zernike polynomials in the normalized pupil parameters, $\vec{\rho}$, for rotationally symmetric optical imaging systems was developed. This resulted in the publication of the article [1] detailing the results obtained. The material presented in that paper occurred before the development of the Zernike vector concept as presented in Chapter 3 as well as many other concepts developed in Chapters 2 through 5. This also occurred before the quantitative analysis of the FFD comparison was developed. For this reason, only qualitative comparisons are presented in this section as they are presented in the original article [1]. Quantitative analyses of FFD comparisons are presented in section 2 of this Chapter for the optical models presented there.

At the time of the writing of article [1], the wavefront aberration function expansion was written in the form

$$W\left(\vec{H},\vec{\rho}\right)=\cdots+g\left(H^{2}\right)H^{m}\cos\left(m\theta\right)Z_{n}^{+m}\left(\vec{\rho}\right)+g\left(H^{2}\right)H^{m}\sin\left(m\theta\right)Z_{n}^{-m}\left(\vec{\rho}\right)+\cdots\ ,\quad (6.1)$$
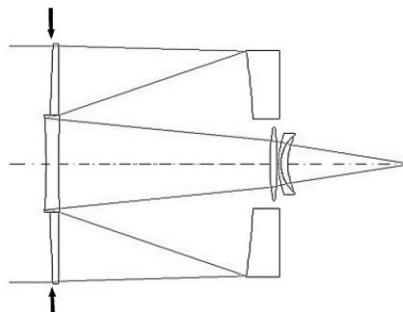
where H is the magnitude of the normalized field parameter vector, θ is the angle that the field vector makes with the $+x$-axis, $\vec{\rho}$ is the normalized pupil vector, and

$$g\left(H^{2}\right)=\sum_{q=0}C_{q}\left(H^{2}\right)^{q}\ ,\qquad (6.2)$$

where C_q are expansion coefficients.

The qualitative check of the expansion, through 6th order plus 8th order spherical, was performed by utilizing CODE V® version 10.4's FFD capabilities and three optical models. A modified version of the fifthdef.seq and FORDER.seq CODE V® macros (modified and provided by Kevin Thompson) for the calculation of the per surface wavefront aberration expansion coefficients $W_{\text{klm;j}}$ through 6th order plus 8th order spherical, but not including the separation of the extrinsic sphere and extrinsic asphere coefficients, was utilized. The results are described in this section.

The first model presented in the paper [1] is that of a Baker telescope model. The layout of the model is presented in Figure 6.2.

Figure 6.2 Baker model layout. The arrows indicate the location of the aperture stop. Originally published in [1].

The model has an entrance pupil diameter of 468.75 *mm* and a half-field of view (HFOV) of 1.0°. The surface parameters for the model are presented in Table 6.1.

Table 6.1 Surface definitions for the Baker telescope model. Table generated by CODE V®.

Surface #	Surface Name	Surface Type	Y Radius	Thickness	Glass	Refract Mode	Y Semi-Aperture
1		Sphere	Infinity	6.4516		Refract	238.7600
Stop		Sphere	4177.1900	18.6090	517645_C	Refract	234.3750
3		Sphere	Infinity	392.8648		Refract	234.3985
4		Sphere	-1515.0917	-392.8648		Reflect	230.0712
5		Sphere	4594.9095	-25.5000	517645_C	Refract	107.6064
6		Sphere	-1378.9582	25.5018	517645_C	Reflect	101.4760
7		Sphere	4594.9095	423.8800		Refract	99.9264
8		Sphere	403.6685	13.8770	SSKN8_SC	Refract	71.4907
9		Sphere	-796.0580	3.4500		Refract	71.0183
10		Sphere	175.2720	4.8250	LLF6HT_S	Refract	66.3261
11		Sphere	73.9082	13.7850	NBK7_SCH	Refract	59.7542
12		Sphere	117.3800	178.1550		Refract	59.8231
13		Sphere	Infinity	69.9987		Refract	36.4116
Image		Sphere	Infinity	-0.0358		Refract	26.3121

The FFD comparison is presented in Figure 6.3. It shows that the equations developed and presented in the paper [1] qualitatively reproduce the FFDs produced by CODE V® (version 10.4) for this model.

Initially, the Zernike astigmatism FFD, the top row column (*b*) of Figure 6.3, calculated by the equations, did not match the CODE V generated FFD shown in column (*a*). As a result of this investigation it was found that CODE V® (version 10.4) divides the azimuthal angular dependence of the field parameters of Zernike astigmatism by ½ in order to present the results with respect to the image plane rather than with respect to the exit pupil where the wavefront aberration function is defined. The other models to be presented in this section of this Chapter similarly use the reduced angular value to match CODE V® (v10.4). All other Chapters, and all other sections in this Chapter, restore the factor of 2 for the azimuthal angle in the field dependence of the Zernike astigmatism FFDs to be consistent with the definition of the wavefront aberration function.
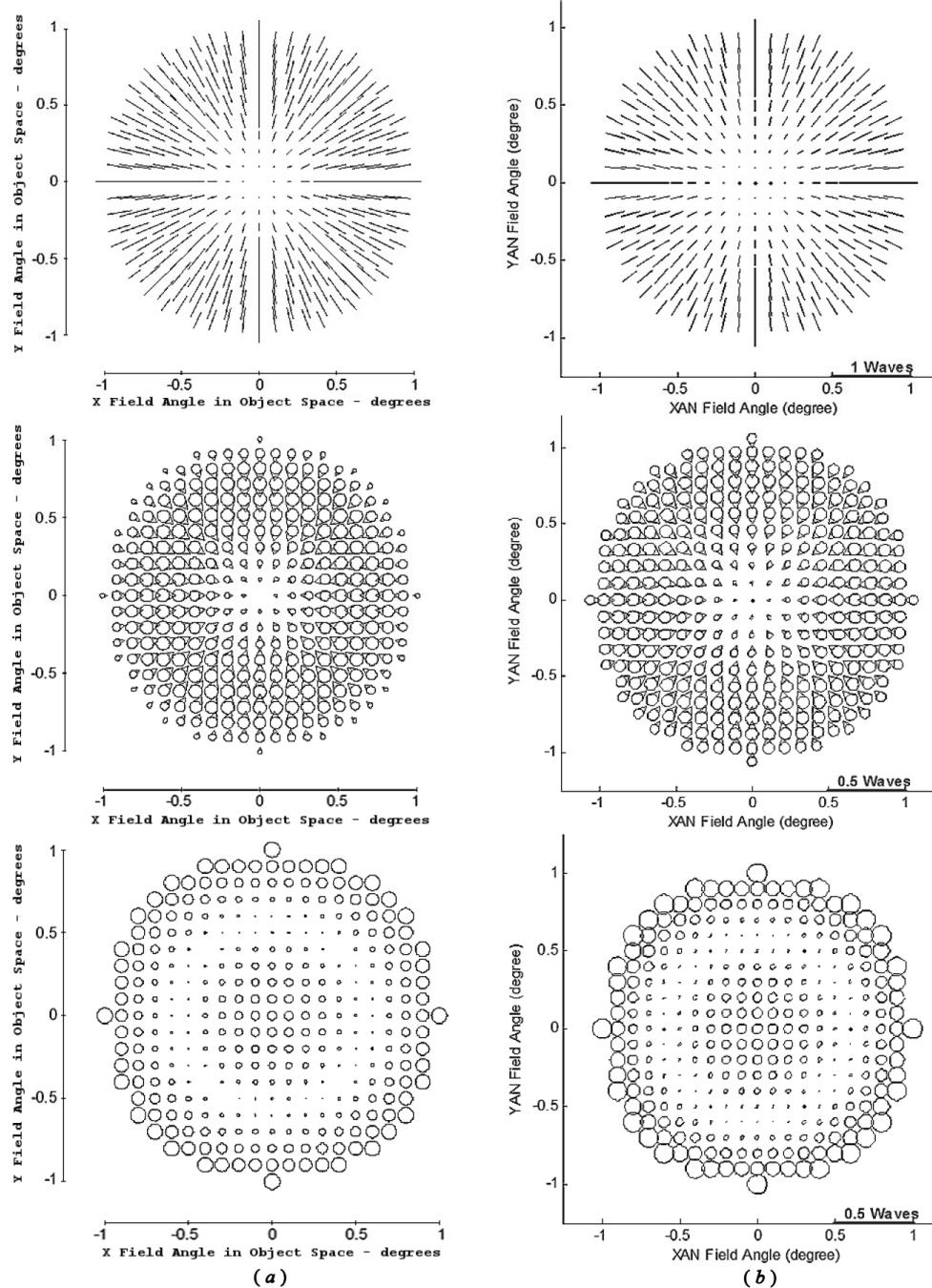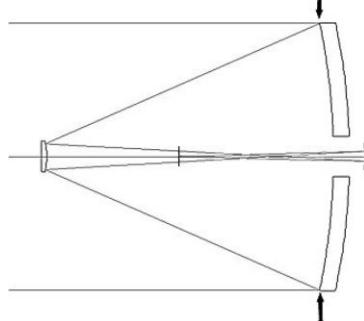
Figure 6.3 (*a*) **The left plots are from CODE V (V10.4), while** (*b*) **the right plots are based on analytic calculations using Eq. (12) of paper [1] (wavefront expanded through 6th order plus 8th order spherical). Top row: Z5 + Z6 (Astigmatism). Middle row: Z7 + Z8 (Coma). Bottom row: Z9 (Spherical). Originally published in [1].**

The second model presented in paper [1] is that of an imaging system based on the James Webb telescope. The layout is presented in Figure 6.4.

Figure 6.4 Model based on the James Web telescope. Arrows indicate the location of the aperture stop. Based on a Figure originally published in [1].
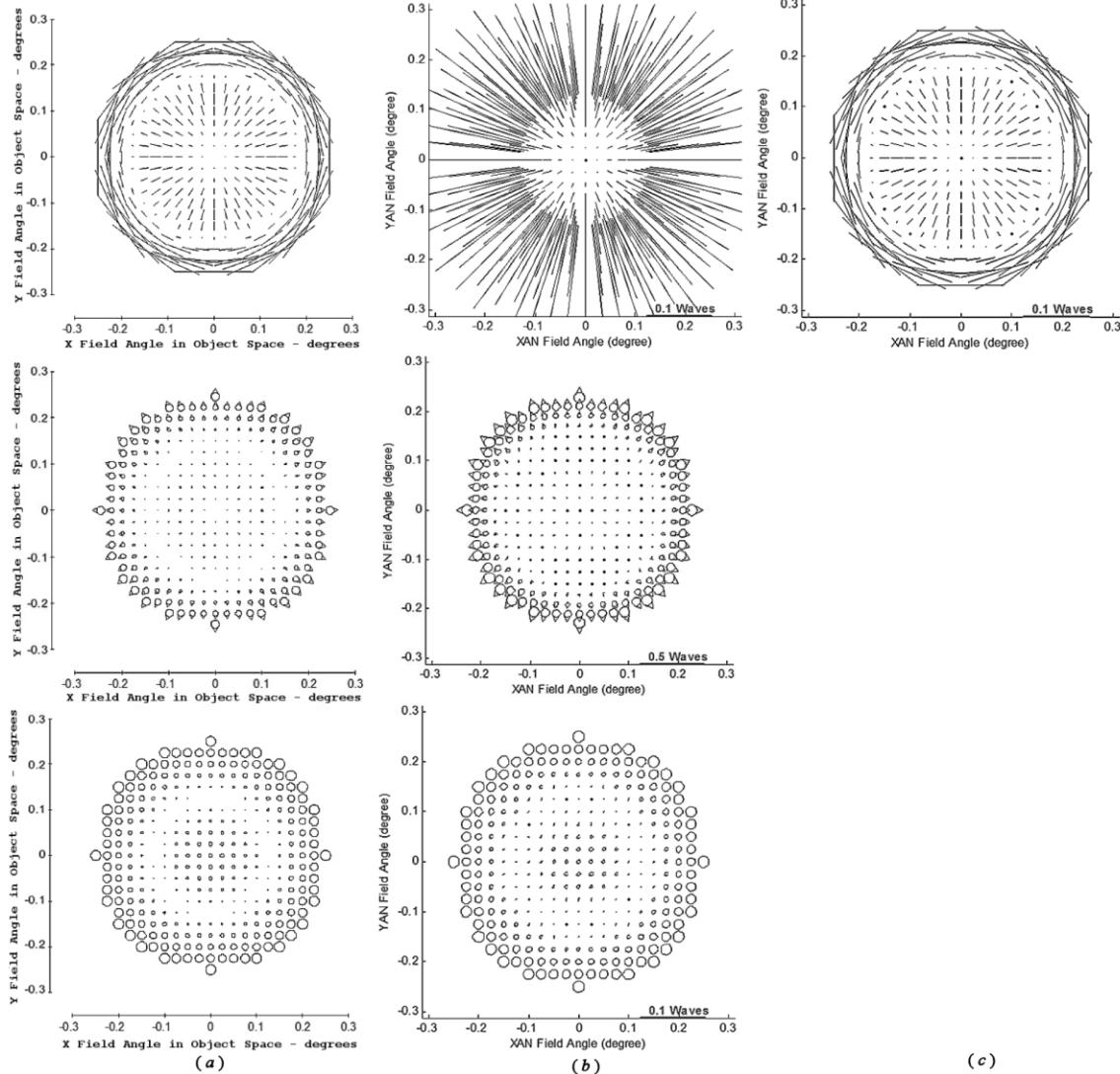
The entrance pupil is defined to have a diameter of 6603.5 *mm* and the HFOV is defined to be 0.25°. The surface definitions are presented in Table 6.2. The conic constant for the primary, secondary, and tertiary mirrors are − 0.9967, −1.6598, and –0.6595, respectively.

Table 6.2 Surface definitions for the James Webb-like model. Table generated by CODE V®.

Surface #	Surface Name	Surface Type	Y Radius	Thickness	Glass	Refract Mode	Y Semi-Aperture
Object		Sphere	Infinity	Infinity		Refract	0
1		Sphere	Infinity	7294.0420		Refract	3333.5765
Stop	PRIMARY	Conic	−15879.7200	−7169.0420		Reflect	3301.7500
3	SECONDAR	Conic	−1778.9130	7965.3130		Reflect	354.8739
4	TERTIARY	Conic	−3016.2270	−4861.5912		Reflect	468.5729
Image		Sphere	3017.5600	0.1170		Refract	570.6737

The FFDs generated by CODE V® and by the analytic equations presented in [1] are reproduced in Figure 6.5. Because the Zernike astigmatism FFD obtained by using the through 6th order wavefront aberration function expansion did not reproduce the full field plot generated by CODE V® (version 10.4), it was hypothesized that the Zernike

astigmatism for this model had a significant amount of higher than 6th order astigmatism

contributions.

Figure 6.5 Qualitative comparison of CODE V (version 10.4) FFDs to the FFDs produced by the equations presented in [1] for the wavefront aberration function expansion through 6th order plus 8th order spherical (*a*) The left plots are CODE V generated FFDs (field range ± 0.25°), while (*b*) the center displays are based on analytic calculations using Eq. (12) of paper [1] (wavefront aberration function expanded through 6th order) with a field range ± 0.25°. The right display (*c*) shows the result for an equation for Zernike astigmatism expanded through 8th order (Eq. (16) of paper [1]) and a least squares fit to CODE V data providing a far better qualitative match to the CODE V® results. Top row: Z5 + Z6 (Astigmatism). Middle row: Z7 + Z8 (Coma). Bottom row: Z9 (Spherical). Originally published in [1].
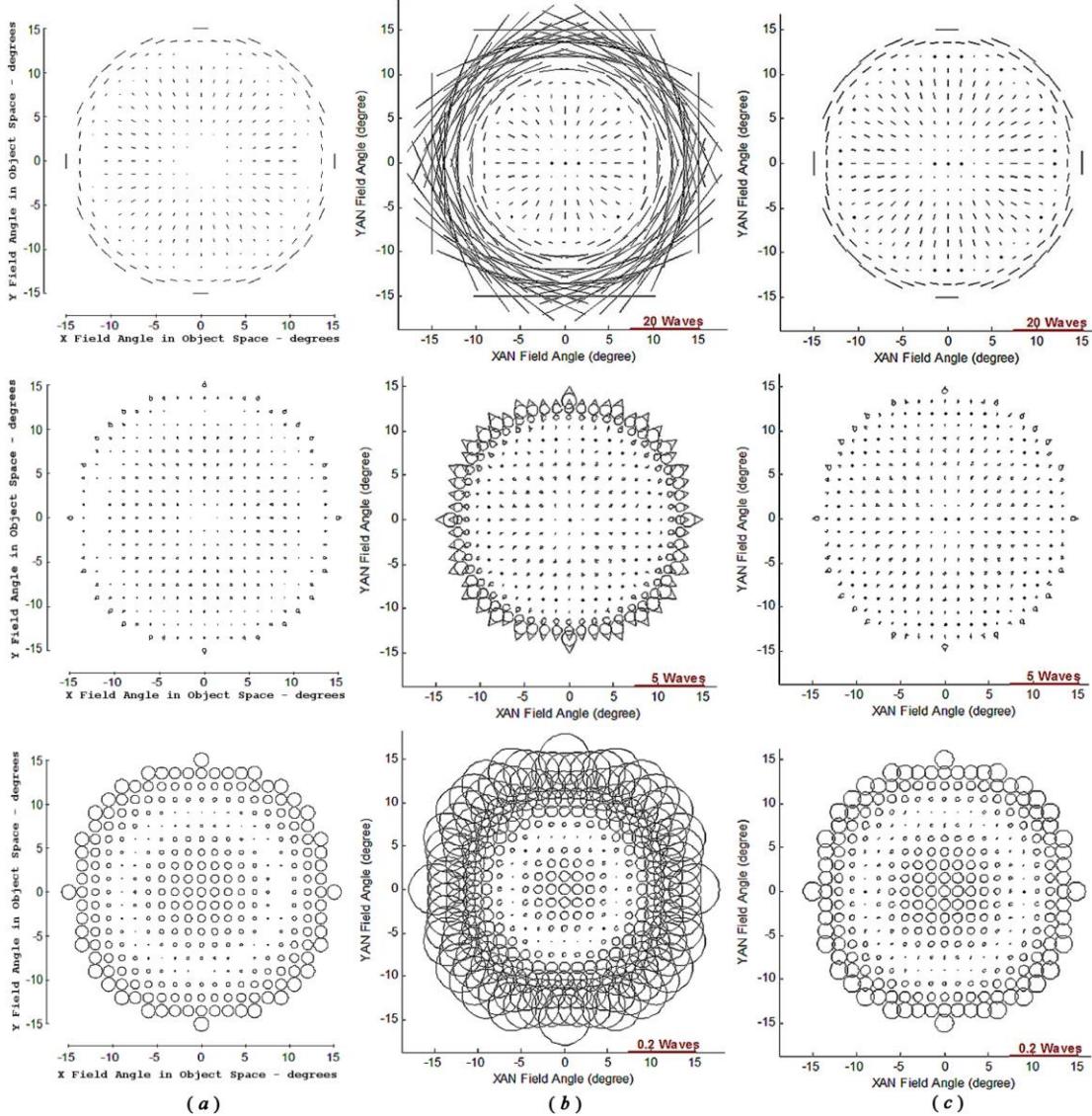
To test this hypothesis, and because at the time of this part of the research higher than 6th order W_{klm} values were not available, a Matlab® function was written to perform a least squares fit to some of the Zernike astigmatism data used by CODE V® to generate the CODE V® Zernike astigmatism FFD. The Zernike astigmatism (Fringe Zernike Z_5, Z_6) terms of the wavefront aberration function was expanded through 8th order and, for the field angle $\theta = 0$, they take the form

$$W = \cdots + \left(w_{5,2} H^2 + w_{5,4} H^4 + w_{5,6} H^6 \right) Z_5 \left(\vec{\rho} \right) + \cdots \tag{6.3}$$

where $w_{r,s}$ are the expansion coefficients to be determined by least squares fit to the CODE V® full field data. These coefficients are related to the W_{klm} coefficients, but do not provide the W_{klm} coefficients. They are undetermined sums of the W_{klm} coefficients. "Undetermined" because it is not known exactly which orders contribute to their values.

With the coefficient values obtained from the least squares fit, the Zernike astigmatism FFD of Figure 6.5(*c*) was generated. As can be seen, there is now a very good match to the CODE V® generated Zernike astigmatism FFD validating the assumption that higher order contributions contribute significantly to Zernike astigmatism for this model and that the equations can reproduce the CODE V® results.

The third model presented in [1] is that of a 3-mirror proprietary telescope model. Therefore, details of the model are not given. The HFOV for the model is 15°. The FFD comparison between the CODE V® and analytic equations of [1] are presented in Figure 6.6.

Figure 6.6 Qualitative comparison of CODE V® FFDs for a proprietary telescope model. (*a*) The left displays are generated by CODE V® (version 10.4), while (*b*) the center displays are based on analytic calculations using Eq. (12) of [1] (wavefront expanded through 6th order). The right plots (*c*) show the results for the wavefront aberration function expanded through 8th order (Eq. (16) of [1]) and a least squares fit to CODE V's display data. Top row: Z5 + Z6 (Astigmatism). Middle row: Z7 + Z8 (Coma). Bottom row: Z9 (Spherical). The field range is ± 15°. Originally published in [1].

It is seen that the analytic equations for the expansion of the wavefront aberration function through 6th order do not reproduce the FFDs produced by CODE V®. The wavefront aberration function was then expanded through 8th order for Zernike coma

(Fringe Z_7, Z_8) and spherical (Fringe Z_9) as well as for astigmatism. For the case $\theta = 0°$ these terms then have the form

$$W = \cdots + \left(w_{5,2}H^2 + w_{5,4}H^4 + w_{5,6}H^6 \right) Z_5\left(\vec{\rho}\right) + \left(w_{7,1}H + w_{7,3}H^3 + w_{7,5}H^5 \right) Z_7\left(\vec{\rho}\right) +$$
$$+ \left(w_{9,0} + w_{9,2}H^2 + w_{9,4}H^4 \right) Z_9\left(\vec{\rho}\right) + \cdots \tag{6.4}$$

Using this expansion, a least squares fit, and the CODE V® data for the FFDs, as with the previous model, the $w_{r,s}$ expansion coefficients were obtained. Using the obtained values and the equations in [1] the FFDs shown in column "*c*" of Figure 6.6 were obtained. These plots show good qualitative agreement with the CODE V® generated FFDs.
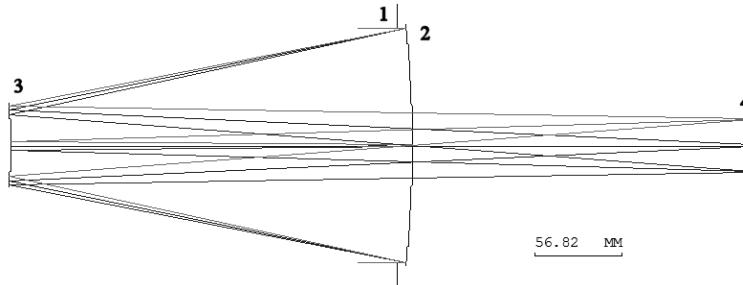
The results indicated the need, for some optical models, to include the 8th order wavefront aberration expansion coefficients. These are not available using the fifthdef and FORDER CODE V® macros. Therefore, an alternative method for the calculation of the wavefront aberration expansion coefficients, $W_{klm;j}$, that would not be limited to through 6th order was sought. Additionally, it was discovered that the equations used in the CODE V® macros for some of the 6th order expansion coefficient calculations could not be reproduced, and that no published detailed documentation of the derivation could be located [67, 68]. (See Appendix V for an example of the issues found when attempting one approach to derive the 6th order expansion coefficients.) Finally, it is not clear what definition (what reference sphere location, what exit pupil (real or Gaussian) location) is used for the $W_{klm;j}$ coefficients calculated by the fifthdef.seq/FORDER.seq CODE V® macros. The development of the GQ method for the calculation of the $W_{klm;j}$ expansion coefficients as described in this dissertation overcomes these limitations.

In the remainder of this Chapter, the FFDs from CODE V® version 10.5 are used as part of the qualitative comparisons of the FFDs. As pointed out previously, the CODE V® version 10.4 Zernike astigmatism FFD has a factor of 2 difference in its azimuthal angle parameter compared to the equations developed in this dissertation. CODE V® version 10.5 restored this factor of 2.

The mathematical and other developments detailed in the previous Chapters, will now be used in the remaining examples of this Chapter. The examples are used to quantitatively and qualitatively check the mathematical development since the publication of [1].

6.2 Two Mirror Telescope Model

The first example is a two mirror, rotationally symmetric telescope system. The layout of the model is presented in Figure 6.7, generated by CODE V®, modified to include the surface numbers and scale bar. The model started as a Ritchey-Chrétien telescope design but the surface shapes were then modified. That is, the surface shapes have purposely not been optimized for minimal optical aberrations so that the sphere and coma FFDs will show non-negligible aberrations. The purpose here is not to design optical imaging systems but to check the development detailed in previous Chapters and to illustrate the information available to a designer by performing similar field analysis during the design process.

Figure 6.7 Two mirror telescope. Surface 1 is the aperture stop and entrance pupil. Surfaces 2 and 3 are the primary and secondary mirrors, respectively. Surface 4 is the image plane.

This model is defined to have an entrance pupil diameter of 150 *mm*, and uses a wavelength of $632.8 \, nm$. The HFOV is 0.6 degrees. The primary mirror has a conic constant of -1.1 and the secondary mirror has a conic constant of -2.5. Additional surface and layout data are provided in Table 6.3 from CODE V®.

Table 6.3 Surface data for optical model.

Surface #	Surface Name	Surface Type	Y Radius	Thickness	Refract Mode	Y Semi-Aperture
Object		Sphere	Infinity	Infinity	Refract	
Stop		Sphere	Infinity	10.0000	Refract	75.0000
2		Conic	-742.8572	-260.0000	Reflect	75.0650
3		Conic	-290.2328	479.9995	Reflect	25.3544
Image		Sphere	Infinity	0.0000	Refract	16.9106

As described in Chapter 5, and in the introductory comments to this Chapter, two categories of FFDs are to be generated and used for comparison. The first category of plots is that based on real ray tracing OPD data (obtained from CODE V®), a GQ method and the double Zernike expansion of the aberration function through 12th order in field and pupil parameters, Eq. (4.53). The results of applying the GQ method are the $U_{n_H, n_\rho}^{m_H, m_\rho}$ expansion coefficients for the double Zernike expansion of the wavefront aberration function. These coefficients, together with the double Zernike expansion

equation, are used to generate the data to be displayed in the first category of FFDs. The second category of FFDs are generated by use of the H. H. Hopkins' W_{klm} expansion coefficients, the single Zernike expansion of the wavefront aberration function (usually through 6th or 8th order) and, if the optical system is not rotationally symmetric, the sigma offset vectors. The FFDs from these two categories are displayed side by side so that a visual qualitative comparison can be made. The goal is to determine if the single Zernike expansion using the Hopkins' W_{klm} expansion coefficients, commonly used by optical designers, together with the sigma offset vectors for rotationally nonsymmetric optical models, can reproduce the FFDs of the double Zernike expansion. The double Zernike expansion is considered by the author to be a closer representation to the true field dependence of the optical model and so it is used as the basis for the comparison. Quantitatively checking the mathematical development is performed by calculating the difference between these two categories of FFD plot data.

The $W_{\text{klm;j}}$ expansion coefficients were calculated as described in Chapter 4 and will be tabulated below. An estimate of the computational error in the $U^{m}_{n_H,n_\rho}$ coefficient values used to calculate the W_{klm} values was performed. Absent from the literature is a discussion of estimations of the error in the calculation of the W_{klm} expansion coefficients. It was found that the method used in this research allows for the calculation of an upper bound for the estimate of the computational error by the relation $U^{+m_H,+m_\rho}_{n_H,n_\rho} = U^{-m,-m}_{n_H,n_\rho} \equiv U^{m}_{n_H,n_\rho}$. The symbol "$\equiv$" is to be read as "is defined to be." This follows by considering both double Zernike expansions for the wavefront aberration function; for the rotationally symmetric case, Eq. (4.52), and for the rotationally

nonsymmetric case, Eq. (4.53). For example, for the rotationally nonsymmetric case, the aberration function expansion will contain the terms $Z_3^{+1}(\vec{\rho})$ and $Z_3^{-1}(\vec{\rho})$ with field functions shown here as,

$$W = \cdots + \left(\cdots + U_{5,3}^{+1,+1} Z_5^{+1}(\vec{H}) + \cdots \right) Z_3^{+1}(\vec{\rho}) + \left(\cdots + U_{5,3}^{-1,-1} Z_5^{-1}(\vec{H}) + \cdots \right) Z_3^{-1}(\vec{\rho}) + \cdots . \quad (6.5)$$

However, when the optical system is rotationally symmetric, Eq. (6.5) will reduce to the form, according to Eq. (4.52),

$$W = \cdots + \left[\cdots + U_{5,3}^{+1,+1} Z_5^{+1}(\vec{H}) + \cdots , \cdots + U_{5,3}^{-1,-1} Z_5^{-1}(\vec{H}) + \cdots \right] \bullet \left[Z_3^{+1}(\vec{\rho}), Z_3^{-1}(\vec{\rho}) \right] + \cdots , (6.6)$$
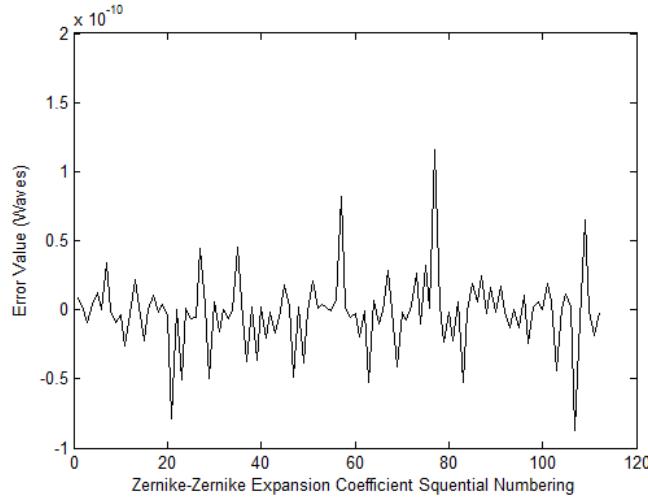
$$W = \cdots + \left(\cdots + U_{5,3}^{1} \left[Z_5^{+1}(\vec{H}), Z_5^{-1}(\vec{H}) \right] + \cdots \right) \bullet \left[Z_3^{+1}(\vec{\rho}), Z_3^{-1}(\vec{\rho}) \right] + \cdots , \quad (6.7)$$

$$W = \cdots + \left(\cdots + U_{5,3}^{1} \vec{Z}_5^{1}(\vec{H}) + \cdots \right) \bullet \vec{Z}_3^{1}(\vec{\rho}) + \cdots . \quad (6.8)$$

Eq. (6.7) is obtained because it is a requirement that $U_{5,3}^{+1,+1} = U_{5,3}^{-1,-1} \equiv U_{5,3}^{1}$ in order to form the Zernike vector $\vec{Z}_5^{1}(\vec{H})$ in Eq. (6.8). So, in general, for a rotationally symmetric optical imaging system, an estimation of the computational error is given by the difference

$$\Delta = U_{n_H,n_\rho}^{+m_H,+m_\rho} - U_{n_H,n_\rho}^{-m,-m} . \quad (6.9)$$

The following chart shows that the worst error Δ in any of the $U_{n_H,n_\rho}^{m_H,m_\rho}$ coefficient values for this optical model is approximately 1.2×10^{-10} waves.

Figure 6.8 Estimated error values for the double Zernike expansion coefficients calculated as the difference $\Delta = U_{n_H,n_\rho}^{+m_H,+m_\rho} - U_{n_H,n_\rho}^{-m,-m}$.

Using this value for the error estimation of the double Zernike expansion coefficients, the equations Eq. (4.66) through Eq. (4.80) for the W_{klm} expansion coefficients, and the equations for error propagation, an estimate for the upper bound of the computational error in the W_{klm} values was calculated and is displayed in Figure 6.9 for the through 8th order coefficients.
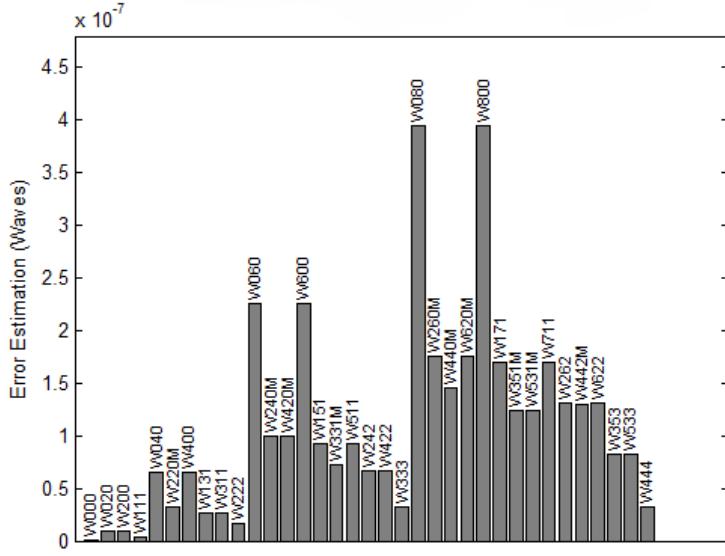
Figure 6.9 Error estimations for the W_{klm} expansion coefficient values.

The bar chart of Figure 6.9 shows that the 4th order W_{klm} values have an upper bound error estimate of approximately $\pm 5 \times 10^{-8}$ waves while most of the 6th order coefficients have an upper bound error estimate of approximately $\pm 10^{-7}$ waves or less. These error estimates provide some confidence that the W_{klm} values used in checking the theoretical development (the resulting values calculated using the W_{klm} values) will not be undermined by their computational error values. From a practical point of view, recalling that values of the W_{klm} coefficients less than about 1/100th of a wave are negligible, it is then confirmed that the values reported have calculation errors well below this criteria.

The through 6th order W_{klm} coefficients for the Zernike astigmatism term $\vec{Z}_2^2(\vec{\rho})$ are presented in Table 6.4 and Table 6.5, rounded to the 4th decimal position. Although not used in this research, the CODE V® fifthdef/FORDER calculated values are provided for comparison.

Table 6.4 Through 6th order intrinsic W_{klm} coefficients for the Zernike astigmatism term.

GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER. (Units = waves.)

Surface Number	W_{klm}	Intrinsic (GQ)			Intrinsic (FF)		
		Sphere	Asphere	Total	Sphere	Asphere	Total
2	W_{222}	1.2772	−0.0003	1.2769	1.2770	−0.0003	1.2767
2	W_{242}	−0.0251	−0.0002	−0.0253	0.0354	−0.0028	0.0326
2	W_{422}	−0.0002	0.0	−0.0002	0.0	0.0	0.0
3	W_{222}	−1.0676	0.6206	−0.4470	−1.0676	0.6206	−0.4470
3	W_{242}	0.0587	−0.0034	0.0553	0.0046	−0.0203	−0.0157
3	W_{422}	−0.0001	0.0	−0.0001	−0.0004	0.0008	0.0004
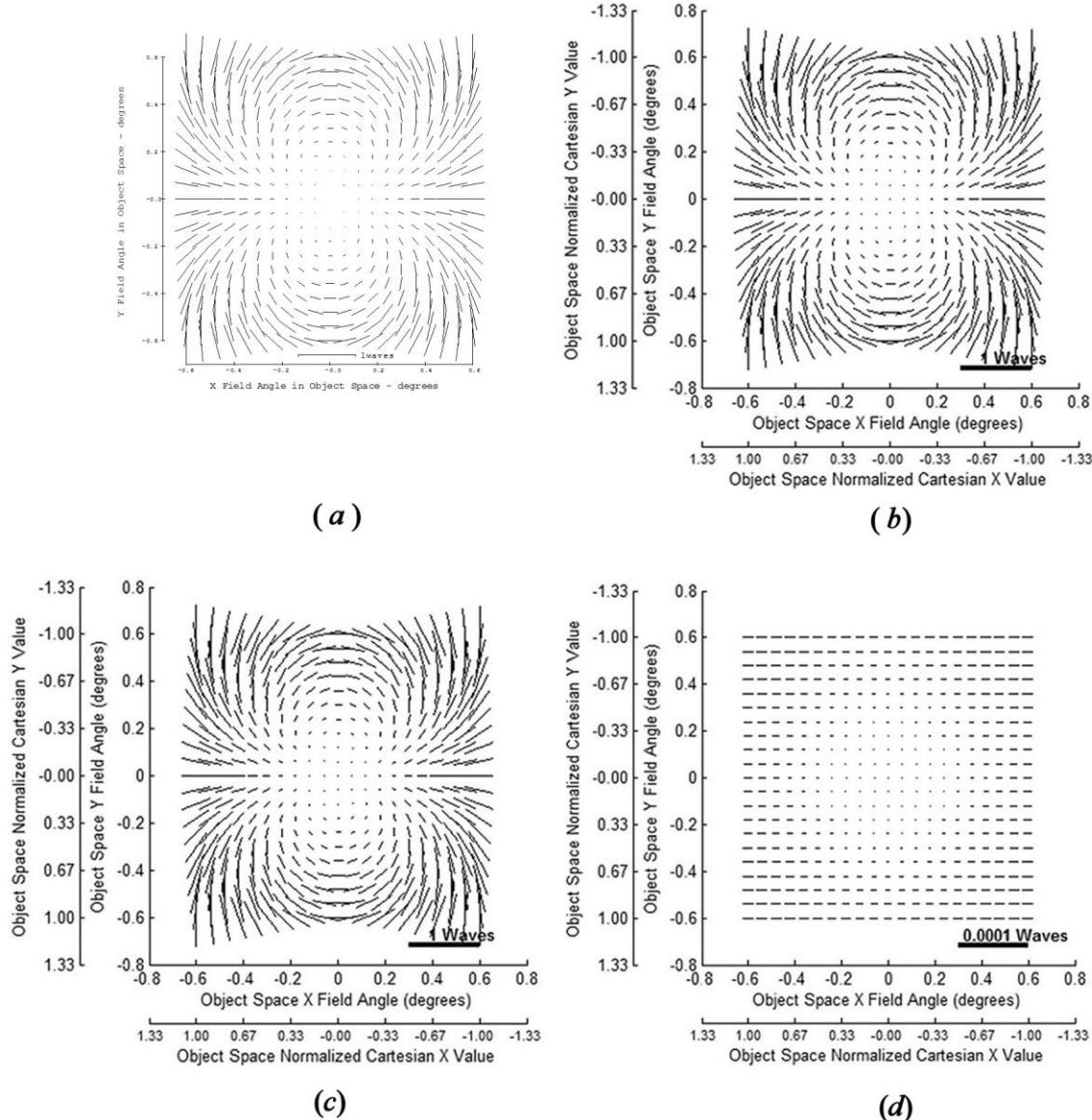
Table 6.5 Through 6th order extrinsic W_{klm} coefficients for the Zernike astigmatism term.

GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER. (Units = waves.)

Surface Number	W_{klm}	Extrinsic(GQ)			Extrinsic(FF)
		Sphere	Asphere	Total	Total
2	W_{222}	0.0	0.0	0.0	0.0
2	W_{242}	0.0	0.0	0.0	0.0
2	W_{422}	0.0	0.0	0.0	0.0
3	W_{222}	0.0	0.0	0.0	0.0
3	W_{242}	−0.0360	0.0019	−0.0341	−0.0184
3	W_{422}	0.0	0.0	0.0	0.0009

As seen in Table 6.4 and Table 6.5, there is a good match between the GQ W_{klm} values and the CODE V® fifthdef/FORDER calculated values for the 4th order coefficients, but there are differences for the 6th order coefficients. Attempts to reconcile these differences were not performed.

Figure 6.10 provides the comparison of the Zernike astigmatism term's FFDs. Figure 6.10(*a*) is a FFD plot for the optical model generated by CODE V® version 10.5 and is included for visual comparison with the FFD plot Figure 6.10(*b*) generated by using the double Zernike expansion and GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.10(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's

data, (b) − (c), is also shown in Figure 6.10(d). The difference data's magnitude values have a maximum difference of 0.00002 waves, a mean of 0.000009 waves and a standard deviation (STD) of 0.000005 waves. The difference data's angle values have a maximum difference of 26.8 degrees, a mean of 0.06 degrees and a STD of 1.28 degrees.

Figure 6.10 Comparison of Zernike astigmatism FFDs. (*a*) **Generated by CODE V®.** (*b*) **Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method.** (*c*) **Generated by using the single Zernike expansion through 6th order and the W_{klm} expansion coefficients.** (*d*) **The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).**

The through 6th order W_{klm} expansion coefficients for the Zernike coma term

$\vec{Z}_3^1(\vec{\rho})$ are presented in Table 6.6 and Table 6.7.

Table 6.6 Through 6th order intrinsic W_{klm} coefficients for the Zernike coma term.
GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER. (Units = waves.)
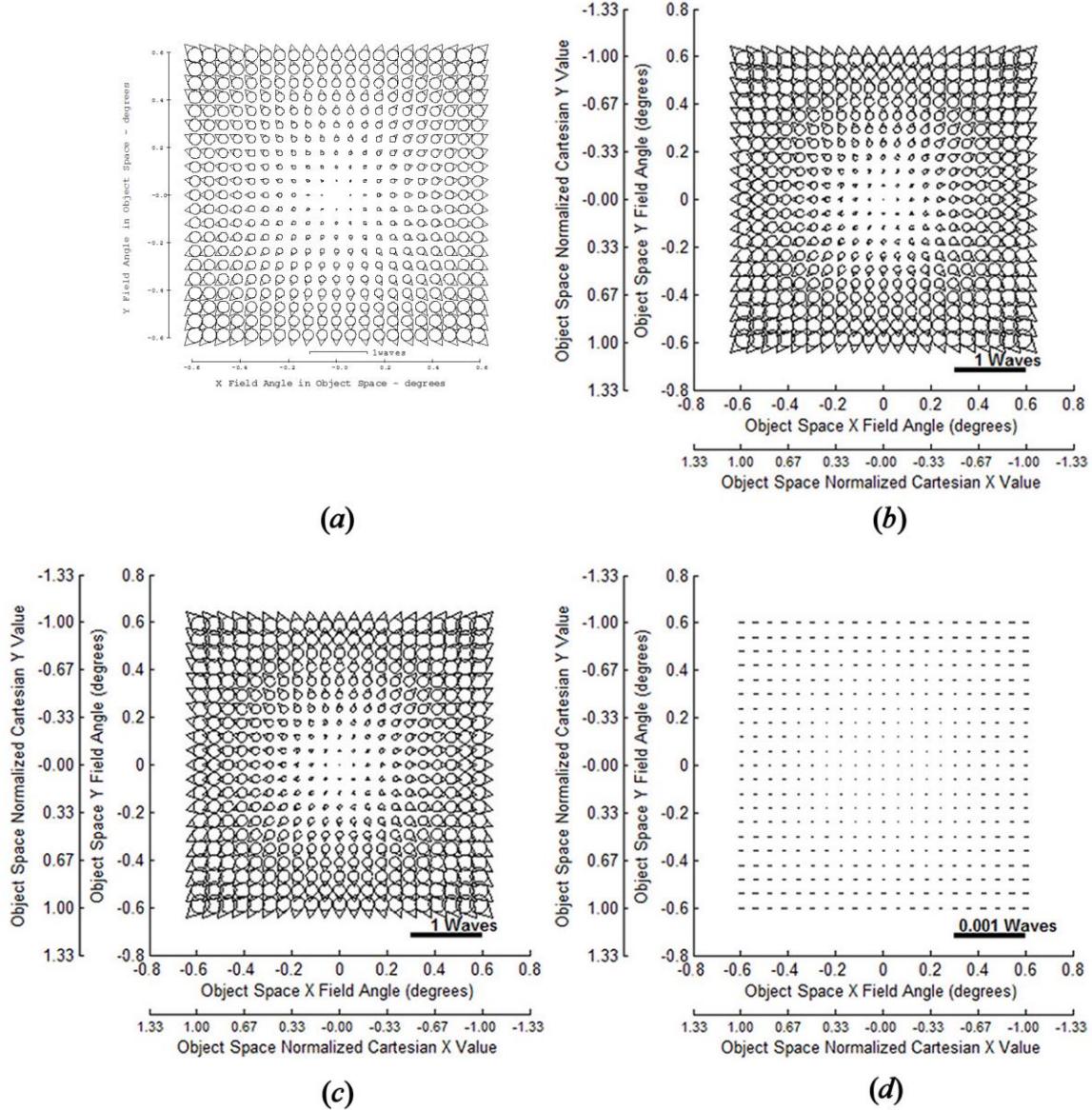
Surface Number	W_{klm}	Intrinsic(GQ)			Intrinsic(FF)		
		Sphere	Asphere	Total	Sphere	Asphere	Total
2	W_{131}	12.4815	0.1873	12.6688	−12.4815	−0.1873	−12.6688
2	W_{151}	−0.0903	−0.0427	−0.1330	−0.2347	0.1448	−0.0899
2	W_{331M}	−0.0026	−0.0	−0.0026	−0.0019	0.0	−0.0019
3	W_{131}	−6.7483	−5.0697	−11.8180	6.7483	5.0697	11.818
3	W_{151}	0.3409	0.1494	0.4903	−0.1045	−0.1523	−0.2568
3	W_{331M}	0.0004	−0.0030	−0.0026	0.0062	0.0008	0.0070

Table 6.7 Through 6th order extrinsic W_{klm} coefficients for the Zernike coma term.
GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER. (Units = waves.)

Surface Number	W_{klm}	Extrinsic(GQ)			Extrinsic(FF)
		Sphere	Asphere	Total	Total
2	W_{131}	0.0	0.0	0.0	0.0
2	W_{151}	0.0	0.0	0.0	0.0
2	W_{331M}	0.0	0.0	0.0	0.0
3	W_{131}	0.0	0.0	0.0	0.0
3	W_{151}	−0.25119	−0.11103	0.36222	0.3702
3	W_{331M}	0.00019	0.00198	−0.00217	−0.0119

Of particular interest, note that the W_{131} value from the GQ method has the opposite sign to the values calculated by fifthdef/FORDER. One possible explanation for this has been offered at the end of Chapter 4. That is, fifthdef/FORDER may be defined to use the normalized object angle field parameters rather than the normalized Cartesian field parameters. There is a sign difference between these two ways of defining the normalized field parameters.

Figure 6.11 Comparison of Zernike coma FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 6th order and the W_{klm} expansion coefficients. (*d*) The difference between the Zernike coma FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

Figure 6.11 provides the comparison of the Zernike coma term's FFDs. Figure

6.11(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.11(*b*) is generated

by using the double Zernike expansion and the GQ method for determining the expansion

coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.11(c) was generated using the single Zernike

expansion and the W_{klm} expansion coefficients. The difference between the two FFD's

data, (b) $-$ (c), is also shown in Figure 6.11(d). The difference data's magnitude values

have a maximum difference of 0.00009 waves, a mean of 0.00005 waves and a STD of

0.00002 waves. The difference data's angle values have a maximum difference of $-$ 158

degrees, a mean of $-$ 0.36 degrees and a STD of 7.5 degrees. The large maximum angle

difference occurs near the node where the data values are very small.
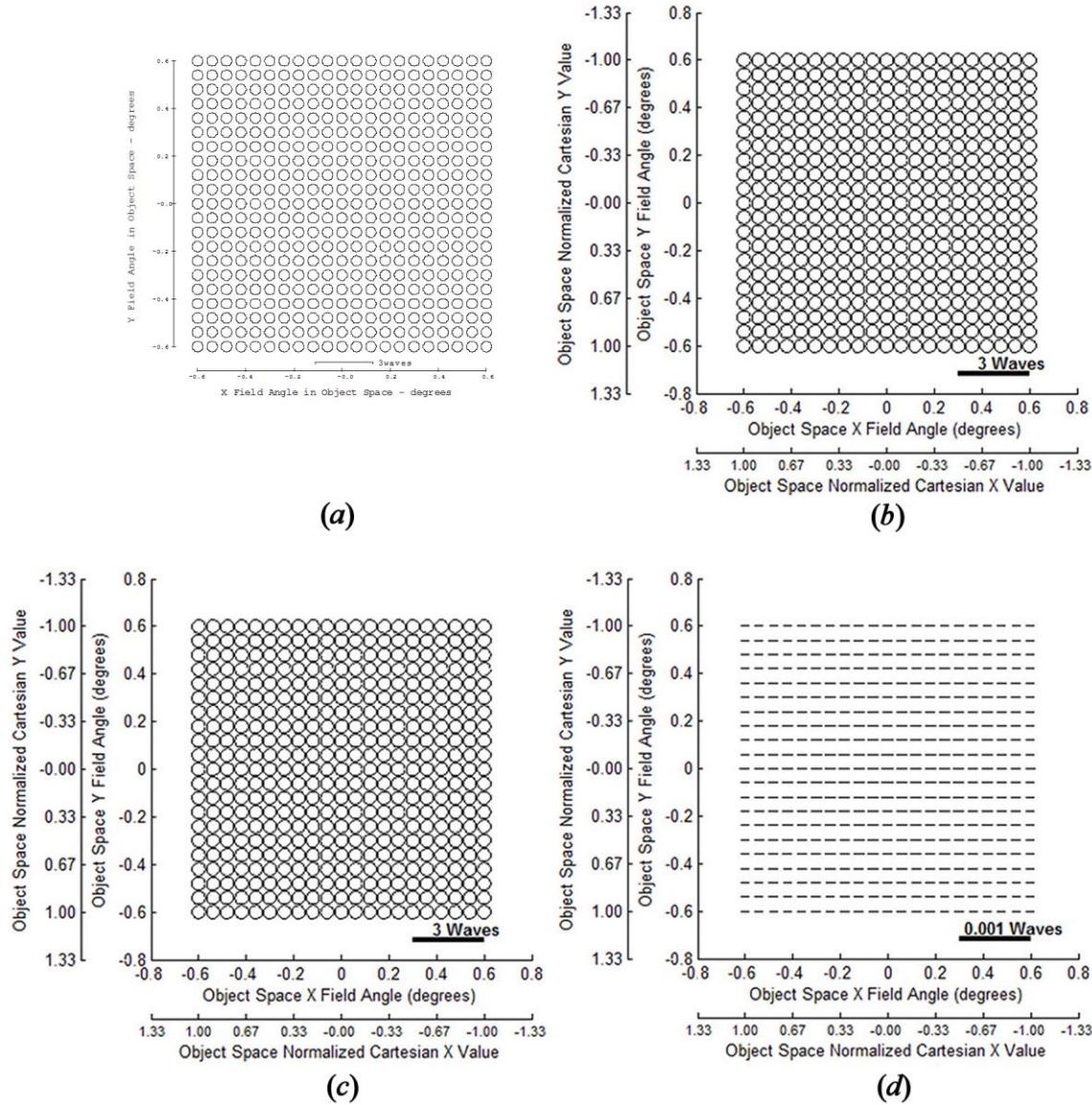
The through 6$^{\text{th}}$ order W_{klm} coefficients for the Zernike spherical term $\vec{Z}_4^0(\vec{\rho})$ are

presented in Table 6.8 and Table 6.9.

Table 6.8 Through 6$^{\text{th}}$ order intrinsic W_{klm} coefficients for the Zernike spherical term.
GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER. (Units = waves.)

Surface Number	W_{klm}	Intrinsic(GQ)			Intrinsic(FF)		
		Sphere	Asphere	Total	Sphere	Asphere	Total
2	W_{040}	30.4932	−33.5425	−3.0493	30.4932	−33.5425	−3.0493
2	W_{060}	−0.1470	0.1798	0.0328	0.4662	−0.4958	−0.0296
2	W_{240M}	−0.0107	0.0018	−0.0089	0.0030	0.0001	0.0031
3	W_{040}	−10.6642	10.3539	−0.3103	−10.6642	10.3539	−0.3103
3	W_{060}	0.5566	−0.5406	0.0160	0.2377	−0.2285	0.0092
3	W_{240M}	0.0053	−0.0020	0.0033	−0.0117	−0.0089	−0.0206

Table 6.9 Through 6$^{\text{th}}$ order extrinsic W_{klm} coefficients for the Zernike spherical term.
GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER. (Units = waves.)

Surface Number	W_{klm}	Extrinsic(GQ)			Extrinsic(FF)
		Sphere	Asphere	Total	Total
2	W_{040}	0.0	0.0	0.0	0.0
2	W_{060}	0.0	0.0	0.0	0.0
2	W_{240M}	0.0	0.0	0.0	0.0
3	W_{040}	0.0	0.0	0.0	0.0
3	W_{060}	−0.43422	0.42159	−0.01263	0.0518
3	W_{240M}	−0.00181	0.00191	0.0001	0.0002

Figure 6.12 Comparison of Zernike spherical FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 6th order and the W_{klm} expansion coefficients. (*d*) The difference between the Zernike spherical FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).

Figure 6.12 provides the comparison of the Zernike spherical term's FFDs.

Figure 6.12(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.12(*b*) is

generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.12(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.12(*d*). The difference data's magnitude values have a maximum difference of 0.0002 waves, a mean of 0.00014 waves and a STD of 0.00001 waves. The angular data is not reported because the *y*-component of the Zernike vector $\vec{Z}_4^0(\vec{\rho})$ is zero by construction.

Figure 6.10 through Figure 6.12 and the quantitative data all show good agreement between the double Zernike generated FFDs and the single Zernike generated displays using the W_{klm} values provided in Table 6.4 through Table 6.9.

The optical model was perturbed to produce a rotationally nonsymmetric case. Specifically, the primary mirror was decentered along the *y*-axis by +0.05 *mm* and the secondary mirror was decentered along the *x*-axis by +2 *mm*. The sigma offset vectors were calculated using Eq. (5.26) and Eq. (5.27) and their values are shown in Table 6.10.

Table 6.10 Normalized sigma offset vector values.
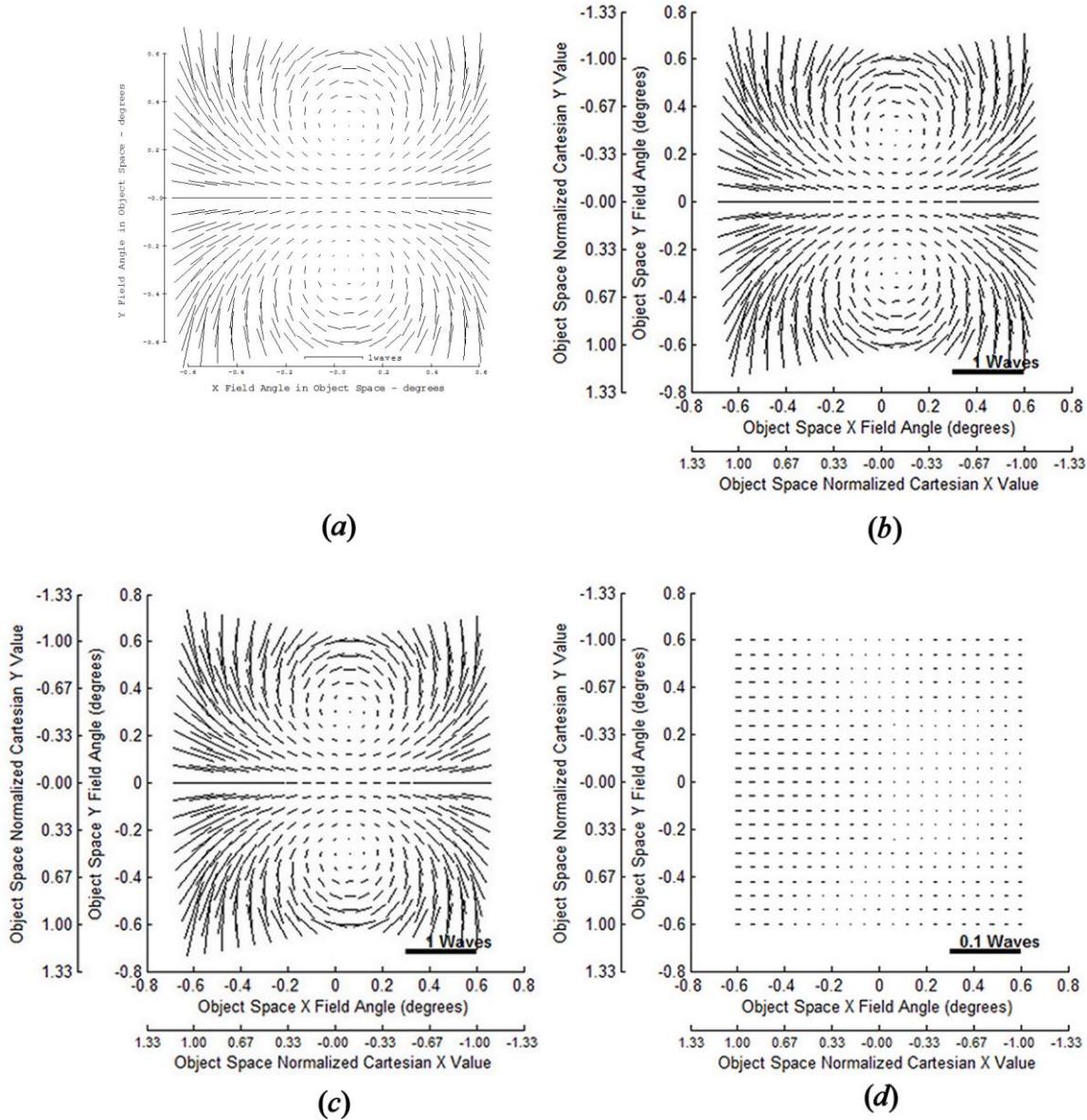
Mirror #	σ_x^S	σ_y^S	σ_x^A	σ_y^A
1	0.0	0.006515	0.0	−0.477447
2	−0.350143	0.012968	−0.726156	0.012708

The FFDs for the Zernike astigmatism term comparison for this rotationally nonsymmetric optical model are presented in Figure 6.13. Figure 6.13(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.13(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients
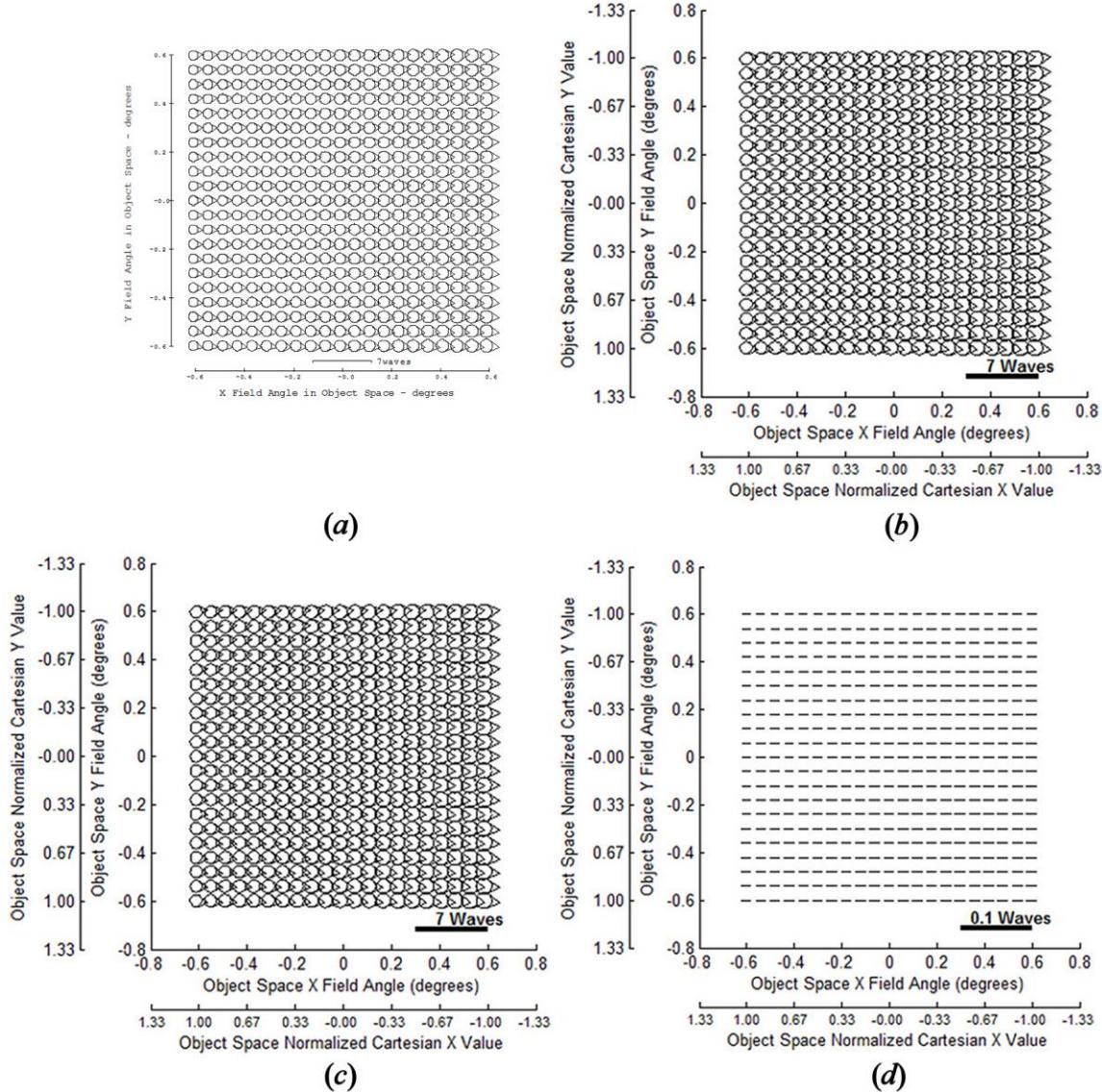
$U^{m_H, m_\rho}_{n_H, n_\rho}$. The FFD plot Figure 6.13(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.13(*d*). The difference data's magnitude values have a maximum difference of − 0.01 waves, a mean of − 0.001 waves and a STD of 0.0046 waves. The difference data's angle values have a maximum difference of − 13.89 degrees, a mean of 0.008 degrees and a STD of 1.38 degrees.

As can be seen in Figure 6.13, the difference data shows a good agreement between the two FFDs. The binodal split of the original single node of the rotationally symmetric case, as discussed in Chapter 5, is clearly observed in Figure 6.13. Due to the decenter of the secondary mirror along the *x*-axis, the nodes are also seen to be shifted slightly to the right in the plots.

The FFDs for the Zernike coma term comparison for this rotationally nonsymmetric optical model are presented in Figure 6.14. Figure 6.14(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.14(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U^{m_H, m_\rho}_{n_H, n_\rho}$. The FFD plot Figure 6.14(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.14(*d*). The difference data's magnitude values have a maximum difference of − 0.016 waves, a mean of 0.014 waves and a STD of 0.001 waves. The difference data's angle values have a maximum difference of − 0.31 degrees, a mean of − 0.22 degrees and a STD of 0.04 degrees.

Figure 6.13 Comparison of Zernike astigmatism FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 6th order the W_{klm} expansion coefficients, and the sigma offset vectors. (*d*) The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).
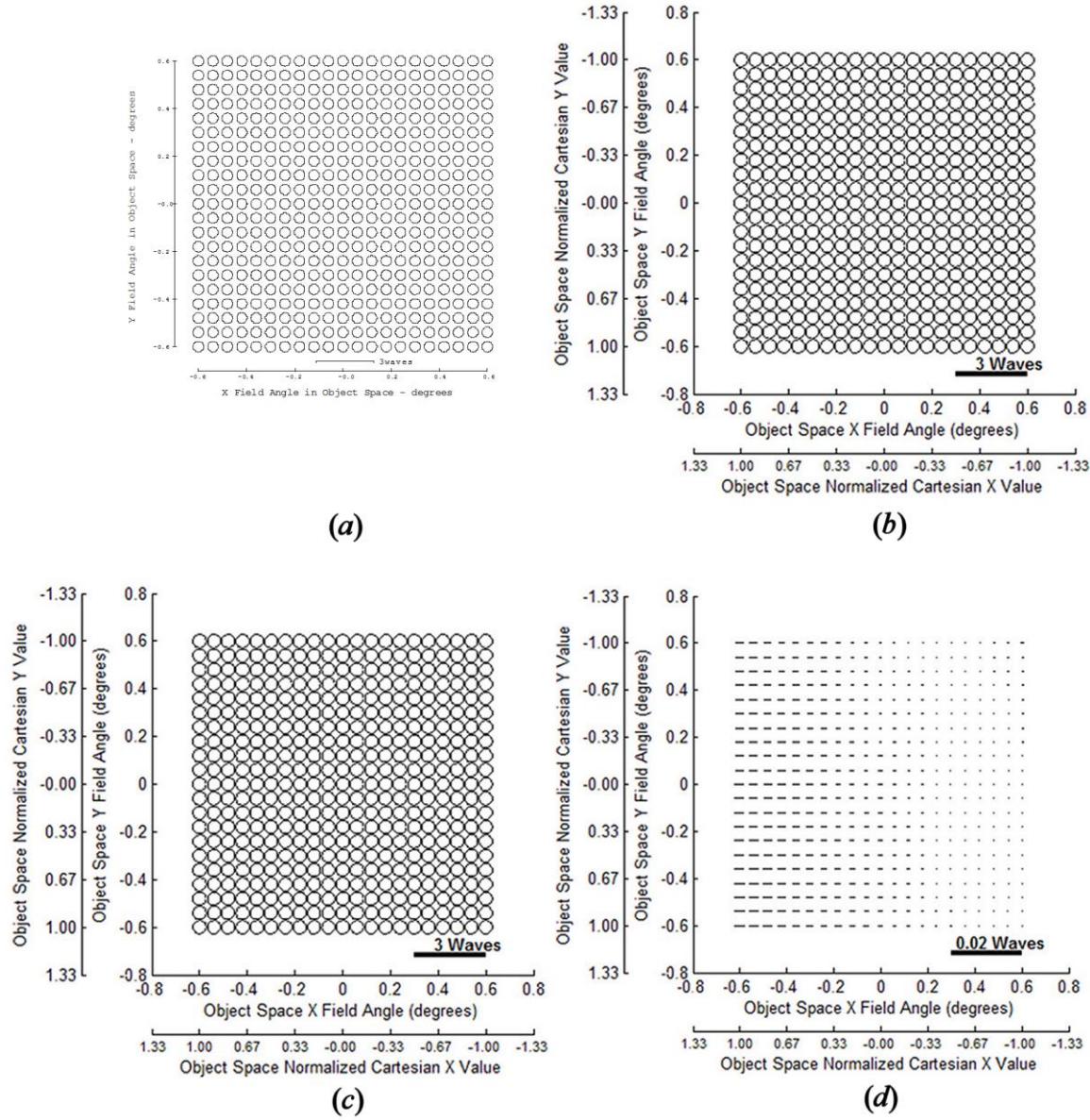
Figure 6.14 Comparison of Zernike coma FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 6th order the W_{klm} expansion coefficients, and the sigma offset vectors.   (*d*) The difference between the Zernike coma FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

It is seen that the perturbations of the mirrors in this model produce a constant coma field dependence.

The FFDs for the Zernike spherical term comparison for this rotationally nonsymmetric optical model are presented in Figure 6.15. Figure 6.15(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.15(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.15(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.15(*d*). The difference data's magnitude values have a maximum difference of 0.003 waves, a mean of 0.001 waves and a STD of 0.001 waves.

From the plots in Figure 6.15 it is seen that the perturbed optical model's spherical field dependence is also dominated by a constant spherical dependence across the field.

Figure 6.15 Comparison of Zernike spherical FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 6th order the W_{klm} expansion coefficients, and the sigma offset vectors. (*d*) The difference between the Zernike spherical FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

Figure 6.13 through Figure 6.15 and the quantitative difference data show that there is a good agreement between the double Zernike expansion of the wavefront aberration function and the single Zernike expansion using the calculated W_{klm} values through 6th order, and the NAT sigma offset vectors of Table 6.10, for this perturbed optical model case.

As another check of the mathematical development, the optical model was again perturbed to include tilts. In this case, the primary mirror was decentered along the x-axis by +0.1 *mm*, and along the y-axis by +0.1 *mm*, and tilted by alpha = +0.25 degrees and beta = − 0.1 degrees. An alpha tilt is a rotation of the mirror about the x-axis. Looking from the negative x-axis toward the positive x-axis, a positive alpha tilt is a counter clockwise rotation. A beta tilt is a rotation about the y-axis. Looking from the negative y-axis toward the positive y-axis, a positive beta tilt is a counter clockwise rotation. The secondary mirror was decentered along the y-axis by − 0.2 *mm*, and tilted by an alpha amount of − 0.15 degrees. The sigma offset vectors were calculated using Eq. (5.26) and Eq. (5.27) and their values are shown in Table 6.11.
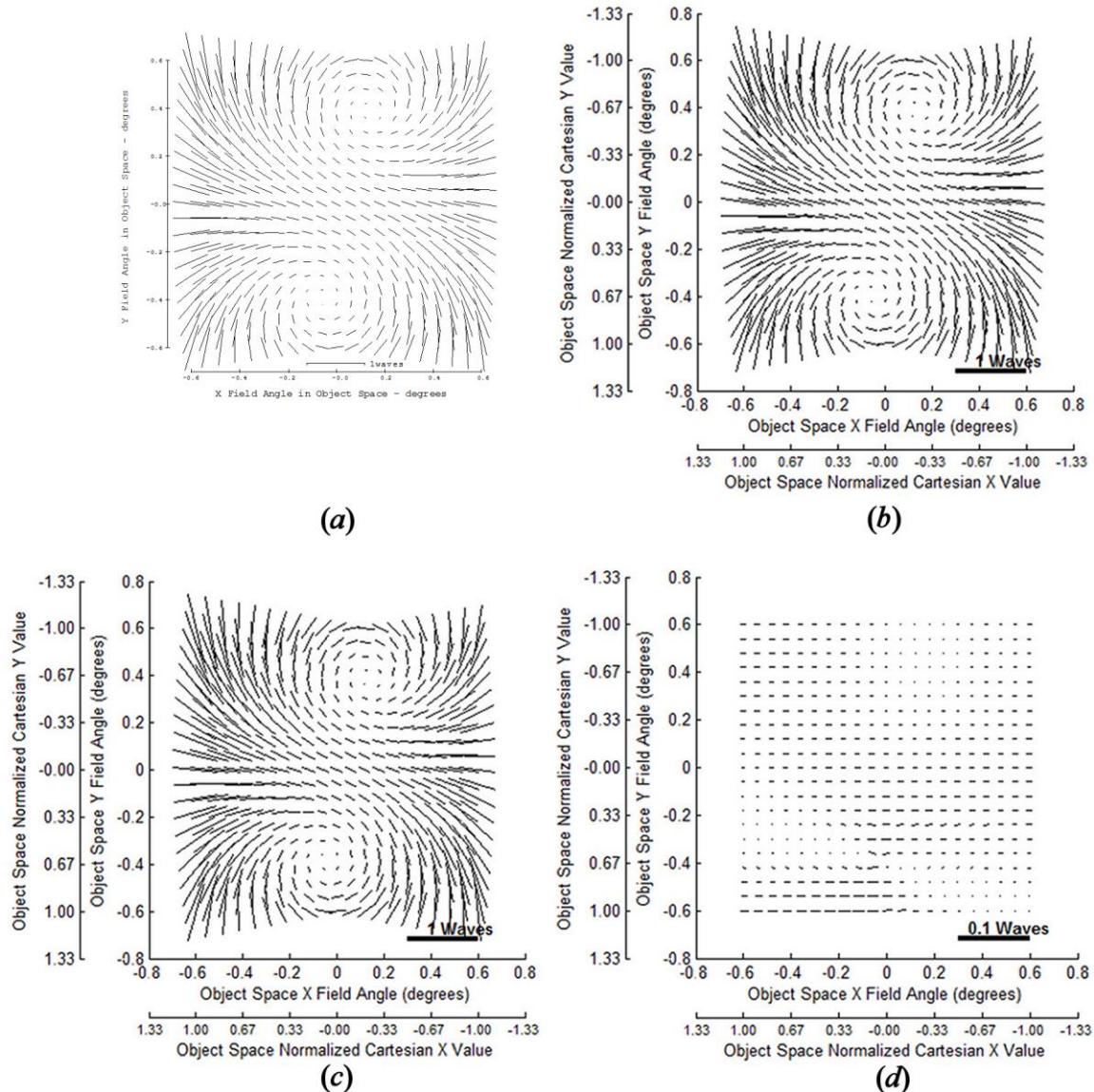
Table 6.11 Normalized sigma offset vector values.

Mirror #	σ_x^S	σ_y^S	σ_x^A	σ_y^A
1	−0.155915	−0.409326	−0.954895	−0.954895
2	−0.310435	−0.913054	−0.304113	−0.725789
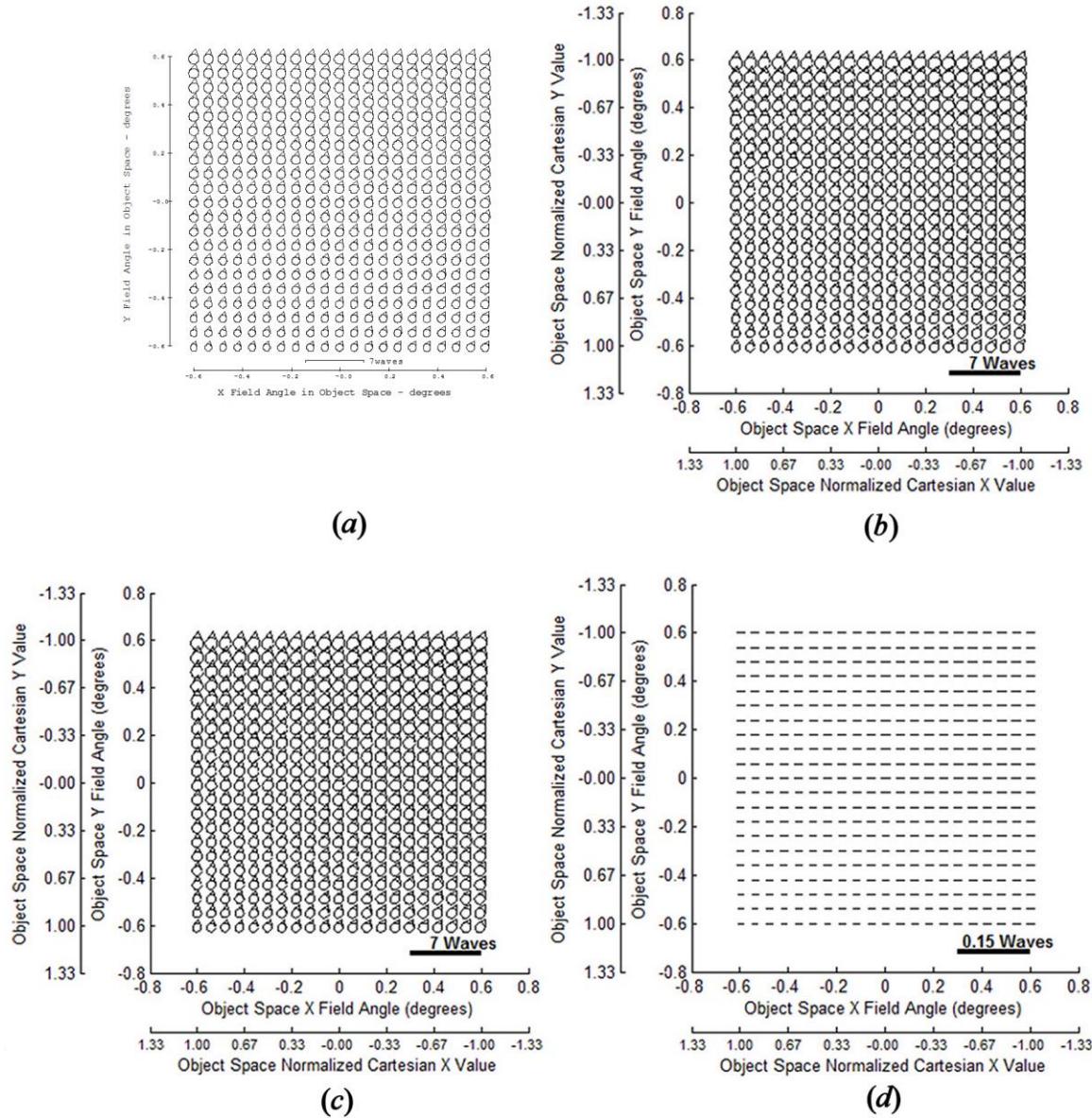
The FFDs for the Zernike astigmatism term comparison for this rotationally nonsymmetric optical model are presented in Figure 6.16. Figure 6.16(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.16(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients

$U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.16(*c*) was generated using the single Zernike expansion

and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) −

(*c*), is also shown in Figure 6.16(*d*). The difference data's magnitude values have a

maximum difference of − 0.017 waves, a mean of 0.004 waves and a STD of 0.006

waves. The difference data's angle values have a maximum difference of − 16.53

degrees, a mean of − 0.17 degrees and a STD of 2.13 degrees. The maximum angle

difference occurs near the lower nodal point where the data values are very small.
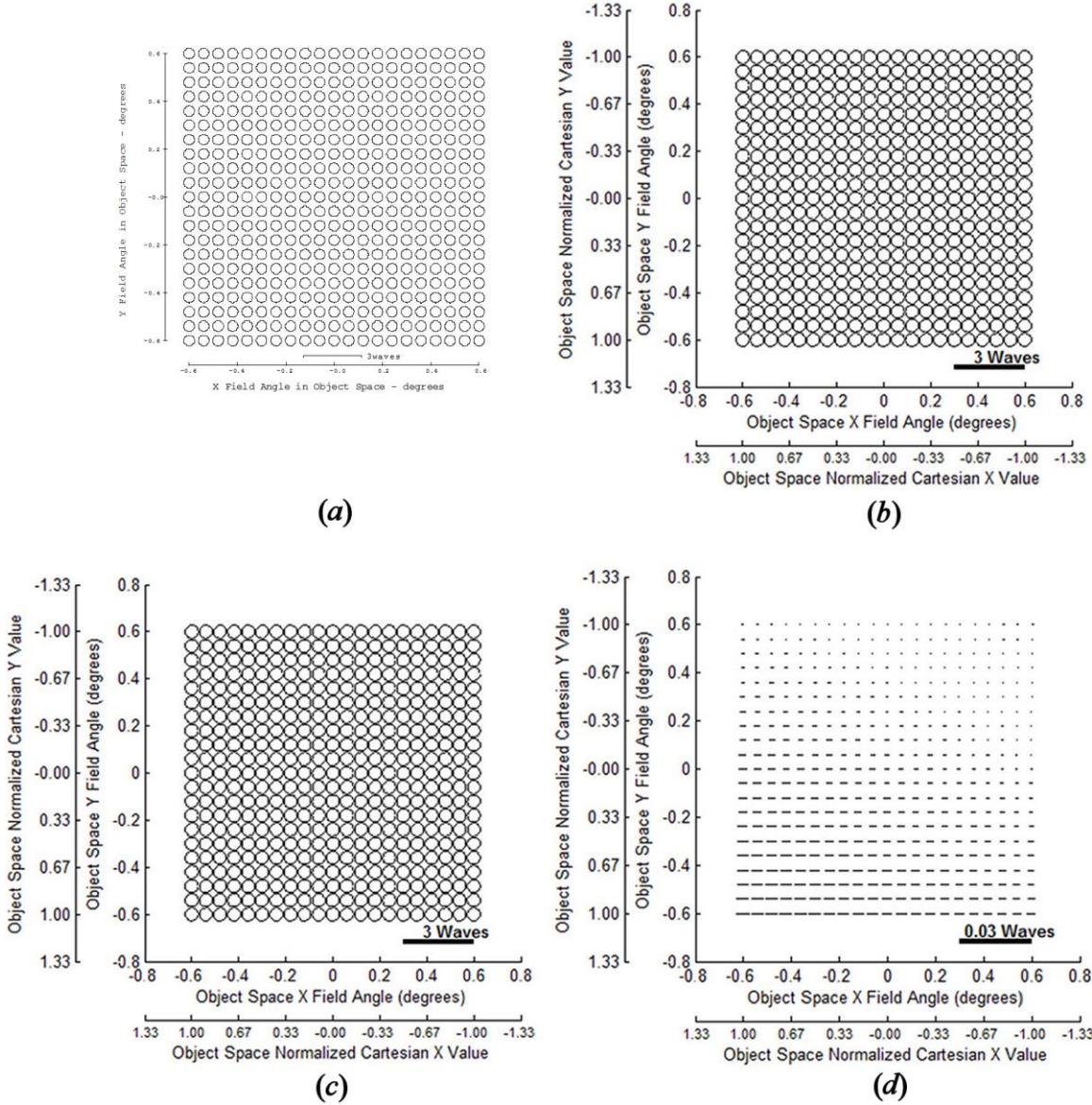
The FFDs for the Zernike coma term comparison for this rotationally

nonsymmetric optical model are presented in Figure 6.17. Figure 6.17(*a*) is a FFD plot

generated by CODE V® version 10.5. Figure 6.17(*b*) is generated by using the double

Zernike expansion and the GQ method for determining the expansion coefficients

$U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.17(*c*) was generated using the single Zernike expansion

and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) −

(*c*), is also shown in Figure 6.17(*d*). The difference data's magnitude values have a

maximum difference of − 0.020 waves, a mean of − 0.018 waves and a STD of 0.001

waves. The difference data's angle values have a maximum difference of 0.37 degrees, a

mean of 0.27 degrees and a STD of 0.06 degrees.

Figure 6.16 Comparison of Zernike astigmatism FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 6th order the W_{klm} expansion coefficients, and the sigma offset vectors. (*d*) The difference between the Zernike spherical FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

Figure 6.17 Comparison of Zernike coma FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 6th order the W_{klm} expansion coefficients, and the sigma offset vectors. (*d*) The difference between the Zernike spherical FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

Figure 6.18   Comparison of Zernike spherical FFDs.   (*a*) Generated by CODE V®.   (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method.   (*c*) Generated by using the single Zernike expansion through 6th order the W_{klm} expansion coefficients, and the sigma offset vectors.   (*d*) The difference between the Zernike spherical FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).

The FFDs for the Zernike spherical term comparison for this rotationally nonsymmetric optical model are presented in Figure 6.18. Figure 6.18(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.18(*b*) is generated by using the double

Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.18(c) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (b) − (c), is also shown in Figure 6.18(d). The difference data's magnitude values have a maximum difference of 0.006 waves, a mean of 0.002 waves and a STD of 0.002 waves.

Figure 6.16 through Figure 6.18 and the quantitative difference data show that there is a good agreement between the double Zernike expansion of the wavefront aberration function and the single Zernike expansion using the calculated W_{klm} values through 6th order, and the NAT sigma offset vectors in this perturbed optical model case. Therefore, this initial example has produced some confidence in the development presented in Chapters 2 through 5.

6.3 A Three Mirror Anastigmat Telescope

A three mirror anastigmat (TMA) telescope model was used to further test the developments detailed in Chapters 2 through 5. The model is defined to have a HFOV of 0.1 degrees, and uses a wavelength of $632.8\,nm$. The entrance pupil is defined to have a diameter of 6603.5 mm. All three mirrors are defined to be conics. The primary, secondary, and tertiary mirrors have conic constants -0.99666, -1.65981, and -0.65954, respectively. Figure 6.19 from CODE V® (modified) provides an illustration of the layout.

1388.89 MM
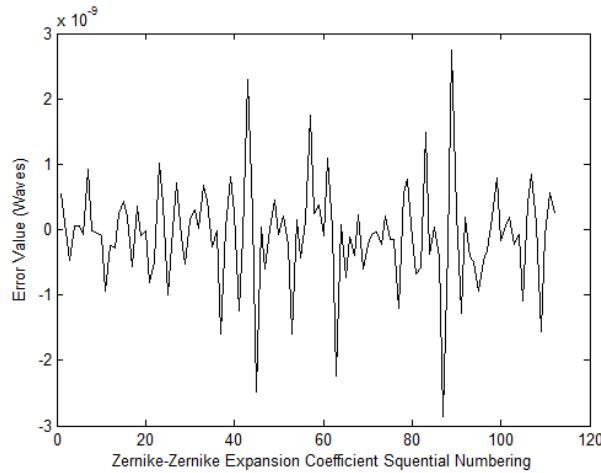
Figure 6.19 TMA model. Surface 1 is the aperture stop and entrance pupil. Surface 2 is the primary mirror. Surface 3 is the secondary mirror. Surface 4 is the tertiary mirror. Surface 5 is the image plane. All mirrors are conics.

Additional model data are provided in Table 6.12 generated by CODE V®.

Table 6.12 Additional surface and model definitions for the TMA.
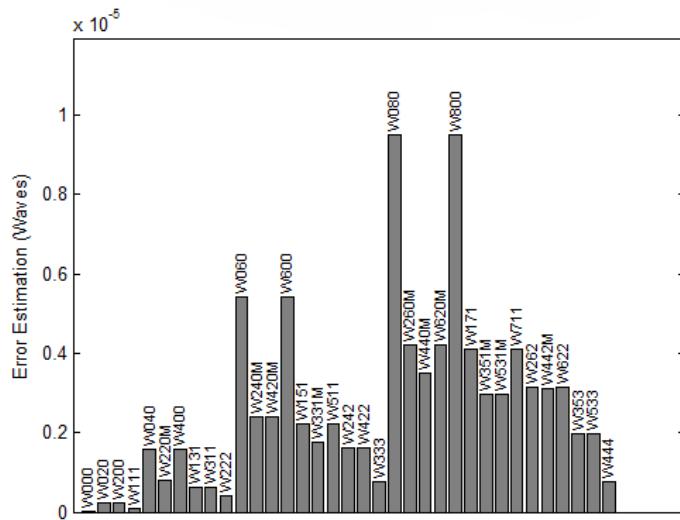
Surface #	Surface Name	Surface Type	Y Radius	Thickness	Refract Mode	Y Semi-Aperture
Object		Sphere	Infinity	Infinity	Refract	
Stop		Sphere	Infinity	2200.0000	Refract	3301.7500
2	PRIMARY	Conic	−1.5880e+004	−7169.0420	Reflect	3304.9894
3	SECONDAR	Conic	−1778.9130	7965.3130	Reflect	335.6752
4	TERTIARY	Conic	−3016.2270	−4861.5912	Reflect	259.7598
Image		Sphere	Infinity	0.0000	Refract	231.4716

The $W_{klm;j}$ expansion coefficients were calculated as described in Chapter 4. An estimate of the computational error in the $U_{n_H,n_\rho}^{m_H,m_\rho}$ coefficient values used to calculate the W_{klm} values was performed. Figure 6.20 shows that the worst error is approximately 2.9×10^{-9} waves.

Figure 6.20 Double Zernike expansion coefficient error estimate.

Using this value for the computational error estimation of the double Zernike expansion coefficients, an estimate for the error in the calculated W_{klm} expansion coefficients was made and these are summarized in the bar chart of Figure 6.21.

Figure 6.21 Computational error estimation for the W_{klm} expansion coefficients.

From this bar chart it can be seen that most of the W_{klm} values are good to about $\pm 3 \times 10^{-6}$ waves. Although not as good as the previous model, it is still well below any threshold for concern.

The W_{klm} expansion coefficients associated with the Zernike astigmatism term $\vec{Z}_2^2(\vec{\rho})$ are provided in Table 6.13 and Table 6.14 through 8th order, rounded to the 4th decimal position.

As mentioned previously, the fifthdef/FORDER CODE V® macros do not provide 8th order W_{klm} expansion coefficients. Therefore, the GQ derived W_{klm} coefficients for 8th order Zernike astigmatism are listed separately in Table 6.14.
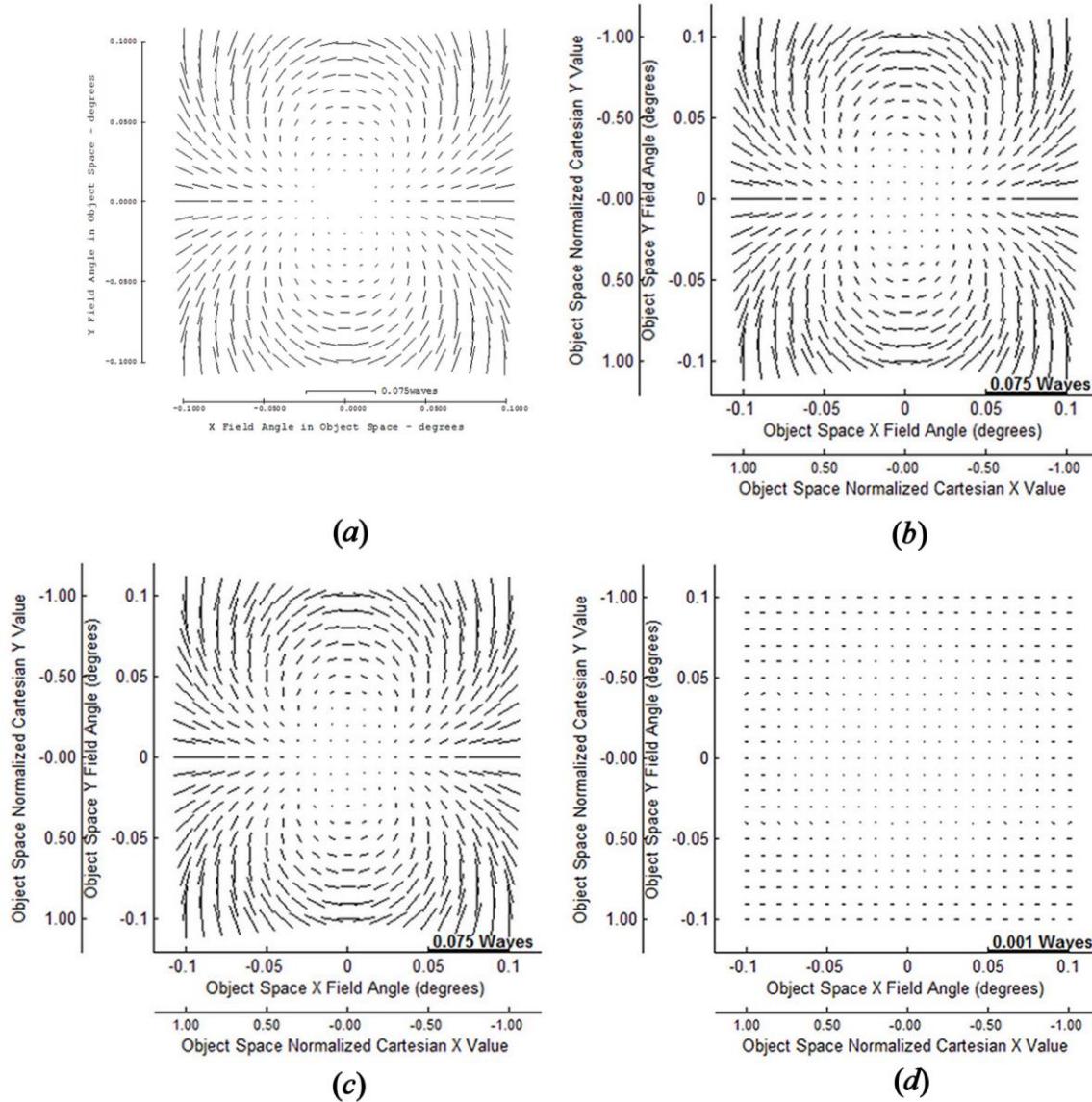
Using these expansion coefficient values the FFD for Zernike astigmatism was generated and is compared to the FFD generated by using the real ray trace OPD data, the double Zernike expansion, and the GQ method. The FFD plots are shown in Figure 6.22. Figure 6.22(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.22(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.22(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.22(*d*).

Table 6.13 W_{klm} expansion coefficients for Zernike astigmatism through 6th order.
GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER.
(Units = waves)

W_{klm}	GQ/FF	Primary	Secondary	Tertiary
W_{222}^{IS}	GQ	2.4525	−6.6022	6.9863
	FF	2.4525	−6.6023	6.9863
W_{222}^{IA}	GQ	−0.0632	7.9483	−10.6876
	FF	−0.0632	7.9482	−10.6876
$W_{222}^{I.Tot}$	GQ	2.3896	1.3461	−3.7013
	FF	2.3893	1.3459	−3.7013
W_{222}^{ES}	GQ	0.0	0.0	0.0
W_{222}^{EA}	GQ	0.0	−0.0001	0.0
$W_{222}^{E.Tot}$	GQ	0.0	−0.0001	0.0
W_{242}^{IS}	GQ	−0.1423	0.9288	−0.0066
	FF	0.2568	−0.2144	−0.0027
W_{242}^{IA}	GQ	0.0149	−0.7030	0.0153
	FF	0.0409	−1.3301	0.0305
$W_{242}^{I.Tot}$	GQ	−0.1274	0.2258	0.0093
	FF	0.2977	−1.5445	0.0278
W_{242}^{ES}	GQ	0.0	−0.7031	0.0216
W_{242}^{EA}	GQ	0.0	0.5447	0.0481
$W_{242}^{E.Tot}$	GQ	0.0	−0.1584	0.0697
	FF	0.0	1.2530	−0.0154
W_{422}^{IS}	GQ	0.0	−0.0004	−0.0133
	FF	0.0	−0.0001	0.0118
W_{422}^{IA}	GQ	0.0	0.0005	0.0115
	FF	0.0	0.0021	−0.0471
$W_{422}^{I.Tot}$	GQ	0.0	0.0001	−0.0018
	FF	0.0	0.0020	−0.0353
W_{422}^{ES}	GQ	0.0	0.0	0.0022
W_{422}^{EA}	GQ	0.0	0.0	−0.0069
$W_{422}^{E.Tot}$	GQ	0.0	0.0	−0.0048
	FF	0.0	0.0003	0.0352

Table 6.14 8th order W_{klm} expansion coefficients for Zernike astigmatism calculated by the GQ method.

| W_{klm}	Primary	Secondary	Tertiary
W_{262}^{IS}	−0.0003	−0.1931	0.0
W_{262}^{IA}	0.0024	−0.1131	0.0003
$W_{262}^{I.Tot}$	0.0021	−0.3062	0.0003
W_{262}^{ES}	0.0	0.2005	−0.0001
W_{262}^{EA}	0.0	0.1038	−0.0005
$W_{262}^{E.Tot}$	0.0	0.3043	−0.0006
W_{442M}^{IS}	0.0	0.0001	0.0
W_{442M}^{IA}	0.0	−0.0006	0.0006
$W_{442M}^{I.Tot}$	0.0	−0.0005	0.0006
W_{442M}^{ES}	0.0	−0.0002	−0.0004
W_{442M}^{EA}	0.0	0.0006	−0.0001
$W_{442M}^{E.Tot}$	0.0	0.0004	−0.0005
W_{622}^{IS}	0.0	0.0	0.0
W_{622}^{IA}	0.0	0.0	0.0
$W_{622}^{I.Tot}$	0.0	0.0	0.0
W_{622}^{ES}	0.0	0.0	−0.0001
W_{622}^{EA}	0.0	0.0	0.0001
$W_{622}^{E.Tot}$	0.0	0.0	0.0

Figure 6.22 Comparison of Zernike astigmatism FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order and the W_{klm} expansion coefficients. (*d*) The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).

The difference data's magnitude values have a maximum difference of − 0.00007 waves, a mean of − 0.00004 waves and a STD of 0.00002 waves. The difference data's

angle values have a maximum difference of 0.93 degrees, a mean of 0.002 degrees and a

STD of 0.04 degrees.

The W_{klm} expansion coefficients associated with the Zernike coma term $\vec{Z}_3^1(\vec{\rho})$

are provided in Table 6.15 and Table 6.16 below, through 8th order, rounded to the 4th

decimal position.

Table 6.15 W_{klm} **expansion coefficients for Zernike coma through 6th order. GQ = value calculated by Gaussian quadrature. FF = fifthdef/FORDER value. (Units = waves)**

W_{klm}	GQ/FF	Primary	Secondary	Tertiary
W_{131}^{IS}	GQ	339.1506	−182.9330	−3.5408
	FF	−339.1506	182.9350	3.5408
W_{131}^{IA}	GQ	54.3607	−197.7262	−9.3698
	FF	−54.3607	197.7269	9.3698
$W_{131}^{I.Tot}$	GQ	393.5113	−380.6592	−12.9106
	FF	−393.5113	380.6592	12.9106
W_{131}^{ES}	GQ	0.0	−0.0020	0.0
W_{131}^{EA}	GQ	0.0	−0.0008	0.0
$W_{131}^{E.Tot}$	GQ	0.0	−0.0028	0.0
W_{151}^{IS}	GQ	−4.9026	35.1951	0.0106
	FF	−26.1816	−9.6005	−0.0045
W_{151}^{IA}	GQ	−12.0678	31.1538	0.0176
	FF	12.5177	−25.8990	−0.0256
$W_{151}^{I.Tot}$	GQ	−16.9704	66.3489	0.0282
	FF	−13.6639	35.4995	−0.0301
W_{151}^{ES}	GQ	0.0	−25.5605	−0.0064
W_{151}^{EA}	GQ	0.0	−23.8655	0.0194
$W_{151}^{E.Tot}$	GQ	0.0	−49.4260	0.0130
	FF	0.0	49.1674	0.0324
W_{331M}^{IS}	GQ	−0.0013	−0.0061	0.0032
	FF	−0.0003	0.0527	0.0253
W_{331M}^{IA}	GQ	−0.0001	−0.0408	−0.0080
	FF	0.0	−0.0224	−0.0058
$W_{331M}^{I.Tot}$	GQ	−0.0014	−0.0469	−0.0048
	FF	−0.0003	0.0303	0.0195
W_{331M}^{ES}	GQ	0.0	0.0037	−0.0123
W_{331M}^{EA}	GQ	0.0	0.0088	0.0844
$W_{331M}^{E.Tot}$	GQ	0.0	0.0125	0.0721
	FF	0.0	−0.0792	−0.0059

Table 6.16 8th order W_{klm} expansion coefficients for Zernike coma. (Units = waves)

W_{klm}	Primary	Secondary	Tertiary
W_{171}^{IS}	−0.0621	−6.5288	0.0
W_{171}^{IA}	0.7929	−4.2792	0.0001
$W_{171}^{I.Tot}$	0.7308	10.8080	0.0001
W_{171}^{ES}	0.0	6.2872	0.0001
W_{171}^{EA}	0.0	3.7902	−0.0002
$W_{171}^{E.Tot}$	0.0	10.0774	−0.0001
W_{351M}^{IS}	0.0	0.0019	0.0
W_{351M}^{IA}	0.0	0.0285	0.0009
$W_{351M}^{I.Tot}$	0.0	0.0304	0.0009
W_{351M}^{ES}	0.0	−0.0029	−0.0001
W_{351M}^{EA}	0.0	−0.0273	−0.0009
$W_{351M}^{E.Tot}$	0.0	−0.0302	−0.0010
W_{531M}^{IS}	0.0	0.0	0.0
W_{531M}^{IA}	0.0	0.0	0.0003
$W_{531M}^{I.Tot}$	0.0	0.0	0.0003
W_{531M}^{ES}	0.0	0.0	−0.0002
W_{531M}^{EA}	0.0	0.0	−0.0001
$W_{531M}^{E.Tot}$	0.0	0.0	−0.0003

It is again seen that the W_{131} values from GQ have an opposite sign from those values calculated by fifthdef/FORDER. (See Chapter 4 for a possible explanation.)

The FFD comparison for Zernike coma terms is provided in Figure 6.23. Figure 6.23(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.23(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion

coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.23(c) was generated using the single Zernike

expansion and the W_{klm} expansion coefficients. The difference between the two FFD's

data, (b) – (c), is also shown in Figure 6.23(d).


(a)

(b)

(c)

(d)
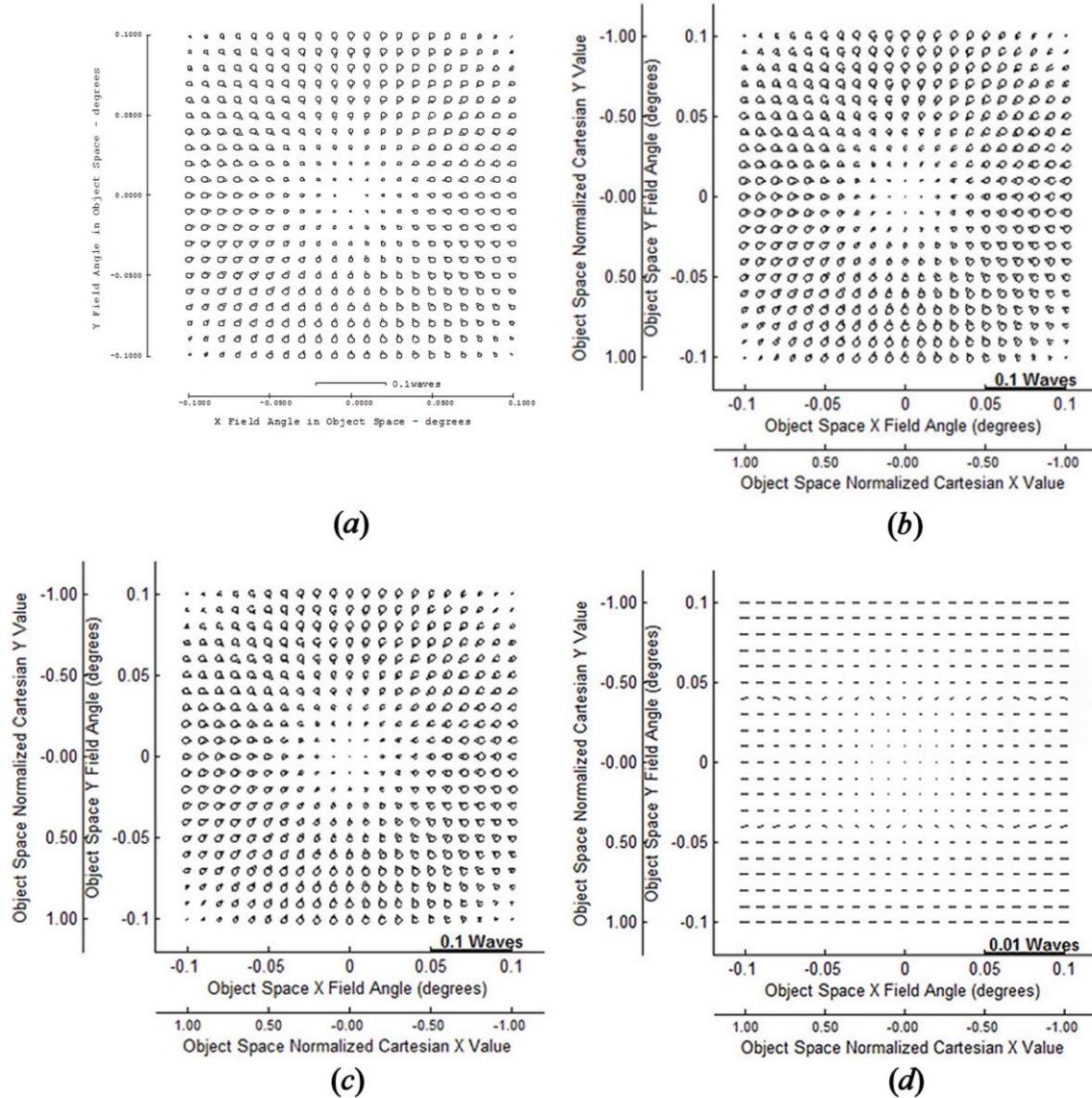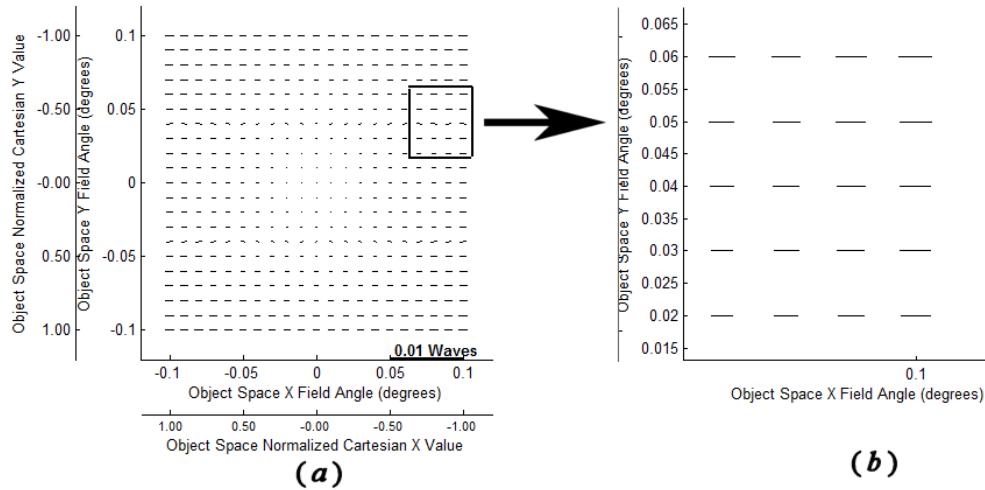
Figure 6.23 Comparison of Zernike coma FFDs. (*a*) **Generated by CODE V.** (*b*) **Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12$^{\text{th}}$ order, and the GQ method.** (*c*) **Generated by using the single Zernike expansion through 8$^{\text{th}}$ order and W_{klm} expansion coefficients.** (d) **The difference between the Zernike coma FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (b) – (c).**

Figure 6.24 Plot line tilts are artifacts of plotting function and not a result of the plot data.
(*a*) **The difference plot appears to show large angular deviations (highlighted box).**   (*b*)
Zooming in on the highlighted boxed area shows these angular deviations are artifacts of the
plot and not an accurate portrayal of the angular difference data.

The difference data's magnitude values have a maximum difference of 0.001 waves, a mean of 0.0007 waves and a STD of 0.0003 waves. The difference data's angle values have a maximum difference of -96.8 degrees, a mean of -0.2 degrees and a STD of 4.6 degrees.

Figure 6.24 shows that the plot of the FFD data difference has lines that appear to have significant amounts of tilt. By zooming in on one area of the plot that has tilted lines (shown as a box in (*a*)), it is seen in (*b*) that the amount of tilt of the lines is much less than suggested by the plot (*a*). These lines are tilted, but not to the extent suggested by the plot (*a*).

As can be seen from Figure 6.23 and the quantitative data, there is good agreement using the W_{klm} expansion coefficients compared with the plot generated by using the double Zernike expansion through 12th order.

Reviewing the W_{klm} values of Table 6.15, it is seen that for the secondary mirror $W_{131}^{ES} = -0.002$ waves. This, being a 4th order expansion coefficient, should have a zero value. Although this and all other 4th order extrinsic W_{klm} coefficients are not included in the equations used to generate the FFDs, it is still concerning. (Recall from Chapter 4, 4th order extrinsic W_{klm} values are calculated in this work only as a debugging aid and they are not used in any of the wavefront aberration expansions programmed to generated the FFD plots. See the discussion immediately after Eq. (4.91).) This may be an indication of the size of the error in the other, higher order, extrinsic coefficient values.

The W_{klm} expansion coefficients for the Zernike spherical term are presented in Table 6.17 and Table 6.18, rounded to the 4th decimal position.
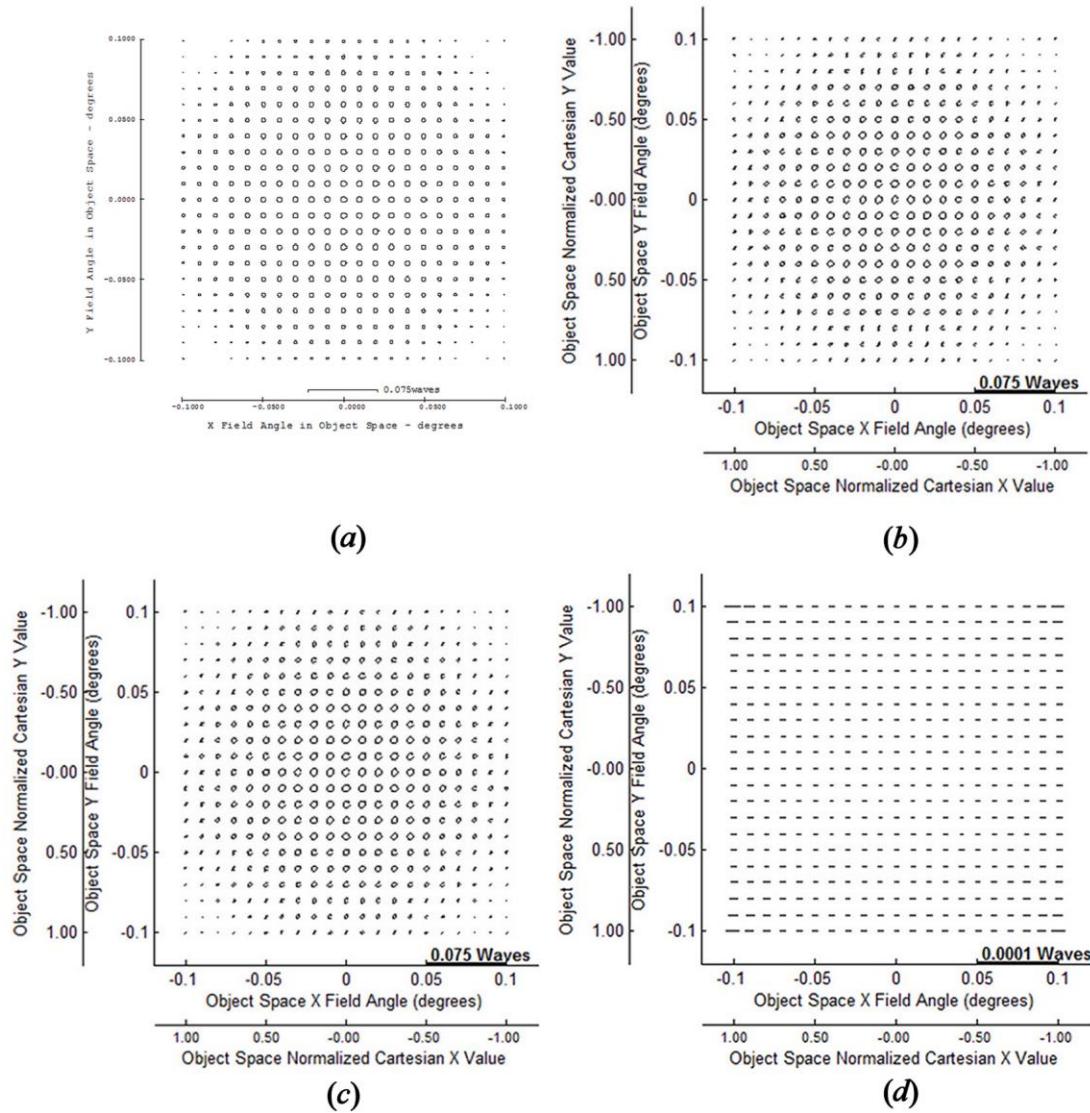
The FFDs for the Zernike spherical term of the wavefront aberration function expansion are presented in Figure 6.25. Figure 6.25(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.25(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.25(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.25(*d*).

Table 6.17 W_{klm} expansion coefficients for Zernike spherical through 6th order. GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER. (Units = waves)

W_{klm}	GQ/FF	Primary	Secondary	Tertiary
W_{040}^{IS}	GQ	11725.2406	−1267.1479	0.4486
	FF	11725.2406	−1267.1890	0.4486
W_{040}^{IA}	GQ	−11686.0841	1229.6641	−2.0536
	FF	−11686.0842	1229.7039	−2.0536
$W_{040}^{I.Tot}$	GQ	39.1565	−37.4838	−1.6050
	FF	39.1564	−37.4851	−1.6050
W_{040}^{ES}	GQ	0.0	−0.0412	0.0
W_{040}^{EA}	GQ	0.0	0.0398	0.0
$W_{040}^{E.Tot}$	GQ	0.0	−0.0014	0.0
W_{060}^{IS}	GQ	−112.9970	296.1752	−0.0019
	FF	760.3538	128.1879	−0.0009
W_{060}^{IA}	GQ	111.3086	−288.0319	0.0065
	FF	−758.6582	−123.3859	0.0050
$W_{060}^{I.Tot}$	GQ	−1.6884	8.1433	0.0046
	FF	1.6956	4.8020	0.0041
W_{060}^{ES}	GQ	0.0	−218.9996	0.0002
W_{060}^{EA}	GQ	0.0	212.5216	−0.0011
$W_{060}^{E.Tot}$	GQ	0.0	−6.4781	0.0008
	FF	0.0	−6.5161	−0.0051
W_{240M}^{IS}	GQ	−0.0373	−0.8171	0.0075
	FF	0.0042	−0.5117	0.0042
W_{240M}^{IA}	GQ	0.0388	−0.4058	0.0048
	FF	0.0042	−0.4873	0.0124
$W_{240M}^{I.Tot}$	GQ	0.0015	−1.2229	0.0123
	FF	0.0084	−0.9990	0.0166
W_{240M}^{ES}	GQ	0.0	0.5726	0.0197
W_{240M}^{EA}	GQ	0.0	0.5445	0.0488
$W_{240M}^{E.Tot}$	GQ	0.0	1.1171	0.0685
	FF	0.0	0.9618	−0.0278

Table 6.18 8th order W_{klm} expansion coefficients for Zernike spherical. (Units = waves)

W_{klm}	Primary	Secondary	Tertiary
W_{080}^{IS}	−1.0735	−57.3475	0.0
W_{080}^{IA}	1.1462	55.8761	0.0
$W_{080}^{I.Tot}$	0.0727	−1.4714	0.0
W_{080}^{ES}	0.0	52.6173	0.0
W_{080}^{EA}	0.0	−51.2179	0.0
$W_{080}^{E.Tot}$	0.0	1.3994	0.0
W_{260M}^{IS}	−0.0006	0.0792	0.0
W_{260M}^{IA}	−0.0023	−0.1157	0.0003
$W_{260M}^{I.Tot}$	−0.0029	−0.0365	0.0003
W_{260M}^{ES}	0.0	−0.0679	−0.0001
W_{260M}^{EA}	0.0	0.1076	−0.0005
$W_{260M}^{E.Tot}$	0.0	0.0397	−0.0006
W_{440M}^{IS}	0.0	0.0	0.0
W_{440M}^{IA}	0.0	−0.0005	0.0005
$W_{440M}^{I.Tot}$	0.0	−0.0005	0.0005
W_{440M}^{ES}	0.0	0.0	−0.0001
W_{440M}^{EA}	0.0	0.0004	−0.0003
$W_{440M}^{E.Tot}$	0.0	0.0004	−0.0004

Figure 6.25 Comparison of Zernike spherical FFDs. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order and W_{klm} expansion coefficients. (*d*) The difference between the Zernike spherical FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

The difference data's magnitude values have a maximum difference of 0.00002 waves, a mean of – 0.000007 waves and a STD of 0.000004 waves.

These FFDs and difference plot of Figure 6.25 and the difference data show a good match between the double Zernike full field display and the FFDs calculated using the W_{klm} expansion coefficients.

As can be seen from Table 6.17, for the secondary mirror, $W_{040}^{ES} = -0.0412$ waves. Being a 4th order expansion coefficient its value should be zero. This is another indication that the calculation of the extrinsic values for this optical model may require further investigation.

It is not clear what the cause of these non-zero extrinsic W_{klm} values for the 4th order coefficients is. Since the estimated errors in the calculated W_{klm} values is at most $\pm 10^{-6}$ waves, and the extrinsic values are calculated (see Chapter 4) by subtraction involving at most 4 W_{klm} values, the error due to error propagation in the calculation is only $\pm 2 \times 10^{-6}$. It is therefore suggested that the large error estimate of the extrinsic values compared to the intrinsic W_{klm} error estimates may be due to small mismatches in setting up the necessary optical subsystems (individual surfaces as independent optical models) as outlined in Chapter 4 for the calculation of the per surface, intrinsic and extrinsic W_{klm} coefficients.
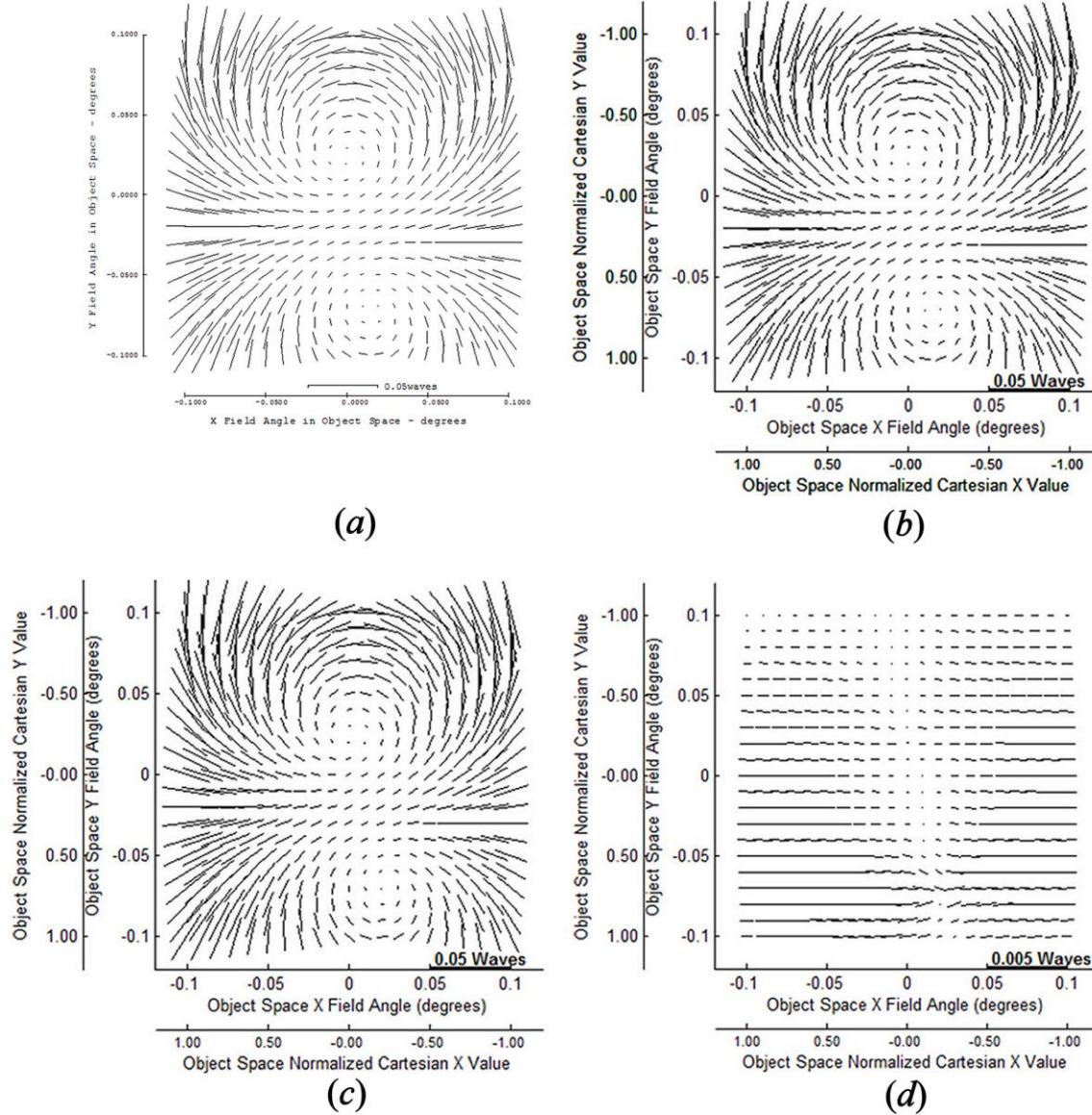
For a first rotationally nonsymmetric test case using this TMA model, the tertiary mirror was decentered along the *x*-axis by +5 *mm*, and along the *y*-axis by the amount +1.0 *mm* and tilted by alpha = +0.0025 degrees, beta = +0.001 degrees. (Only the 3rd mirror was decentered and tilted at this point to provide a less complicated perturbed optical system for debugging purposes. A more complicated set of perturbations will be presented in the next example.) Because the decenters and tilts only concern the last

mirror, only the tertiary mirror will have non-zero sigma vector values. These are shown

in Table 6.19.

Table 6.19 Sigma values for mirror #3.

Mirror #	Sphere Sigma		Asphere Sigma	
	X	Y	X	Y
3	–0.055631	–0.009561	–0.036147	–0.007229

Using these sigma offset vector values, and the W_{klm} values listed in Table 6.13

through Table 6.18 above, the Zernike astigmatism FFD was generated and compared to

the FFD generated from the double Zernike expansion and the GQ method, as shown in

Figure 6.26. Figure 6.26(*a*) is a FFD plot generated by CODE V® version 10.5. Figure

6.26(*b*) is generated by using the double Zernike expansion and the GQ method for

determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.26(*c*) was

generated using the single Zernike expansion and the W_{klm} expansion coefficients. The

difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.26(*d*). The

difference data's magnitude values have a maximum difference of − 0.001 waves, a

mean of − 0.00006 waves and a STD of 0.0009 waves. The difference data's angle

values have a maximum difference of 163.43 degrees, a mean of − 1.25 degrees and a

STD of 9.46 degrees.
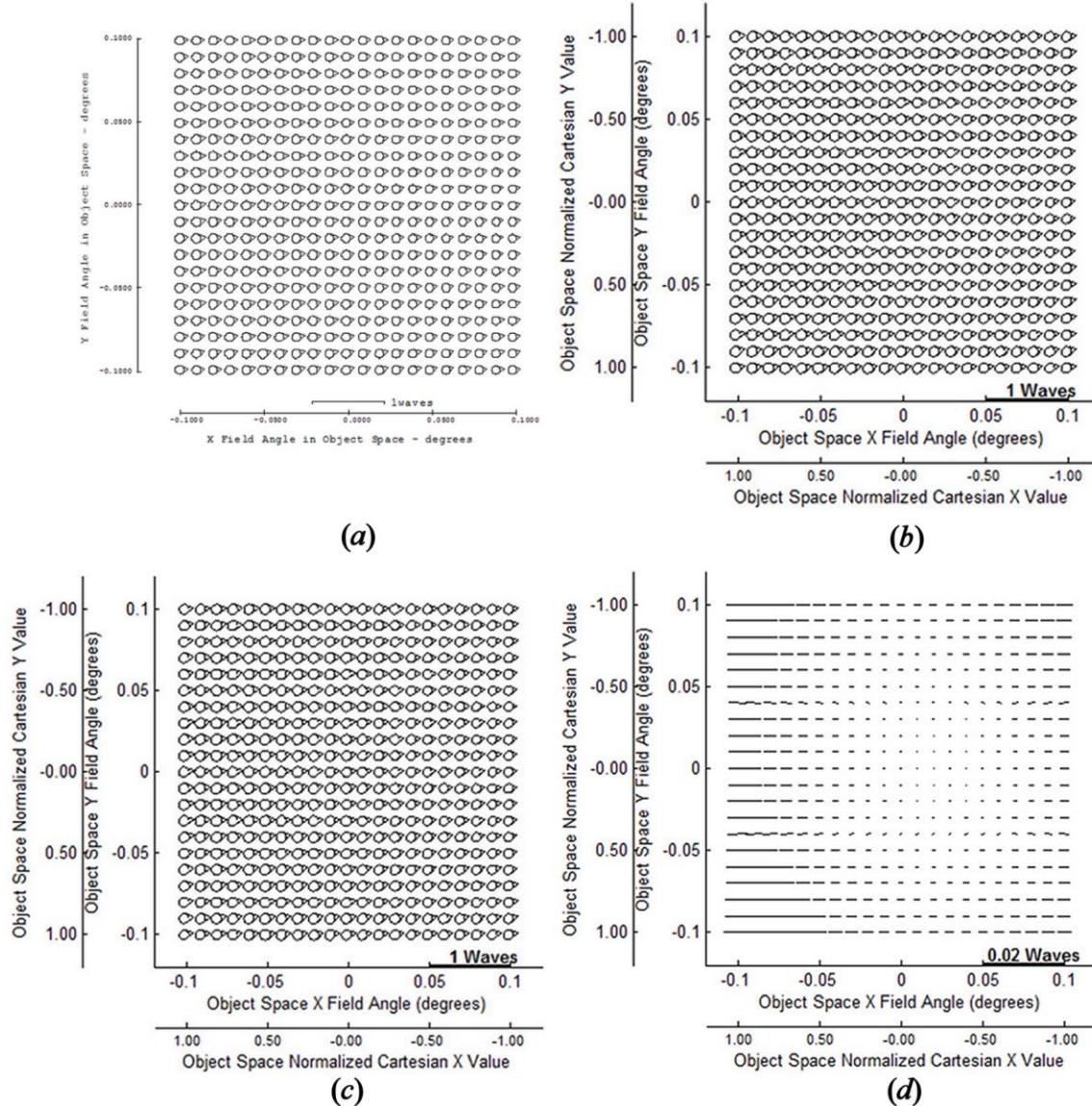
Figure 6.26 Comparison of Zernike astigmatism FFD data. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.19. (*d*) The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

The comparison of the Zernike coma FFD data is shown in Figure 6.27. Figure 6.27(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.27(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion
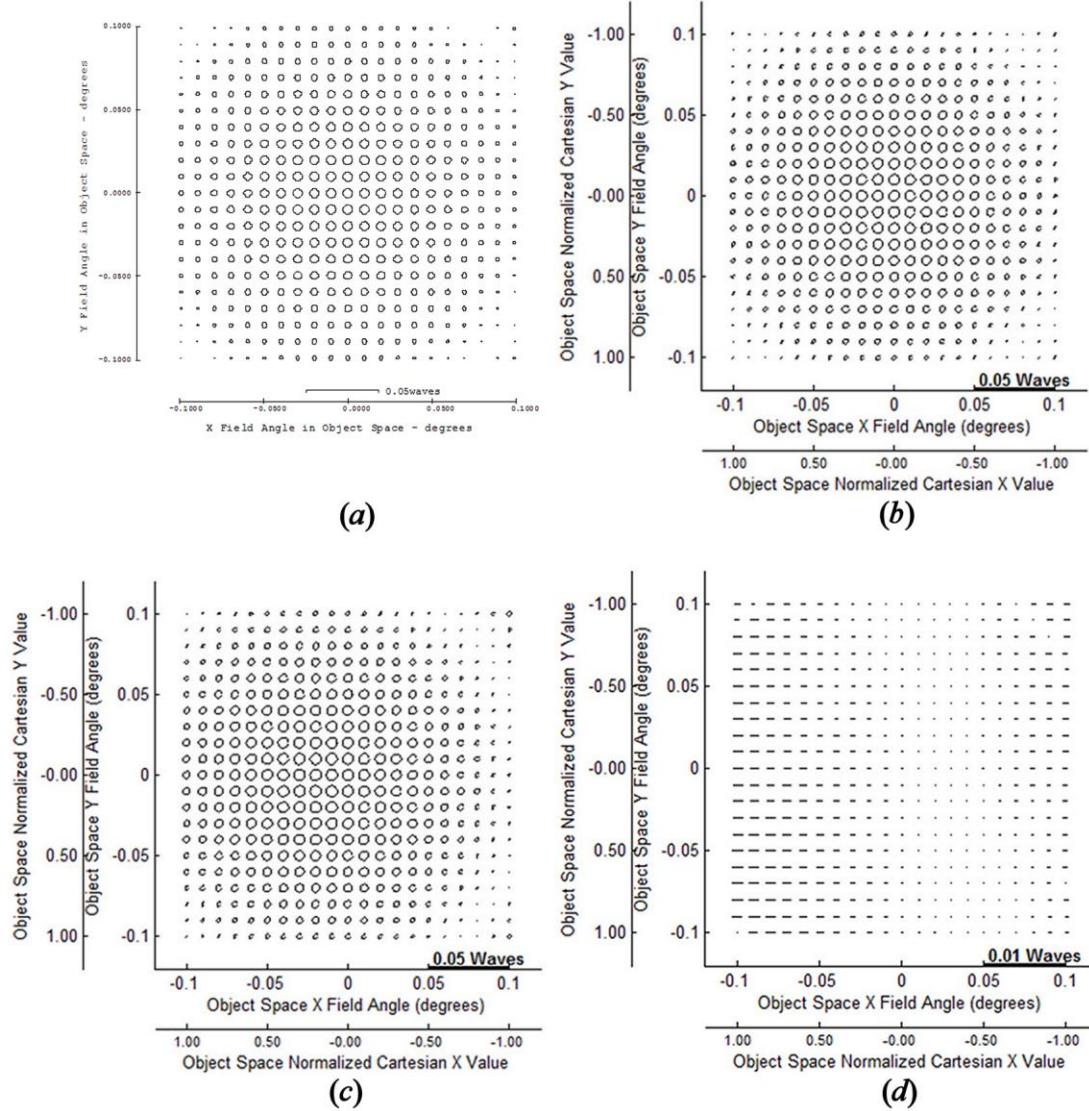
coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.27(c) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (b) − (c), is also shown in Figure 6.27(d). The difference data's magnitude values have a maximum difference of 0.007 waves, a mean of 0.002 waves and a STD of 0.001 waves. The difference data's angle values have a maximum difference of − 0.80 degrees, a mean of − 0.01 degrees and a STD of 0.21 degrees.

The comparison of the Zernike spherical FFD data is shown in Figure 6.28. Figure 6.28(a) is a FFD plot generated by CODE V® version 10.5. Figure 6.28(b) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.28(c) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (b) − (c), is also shown in Figure 6.28(d). The difference data's magnitude values have a maximum difference of 0.001 waves, a mean of − 0.0003 waves and a STD of 0.0006 waves.

Although this may be considered a simple case, it was instrumental in debugging both the mathematical development presented in the previous Chapters and the computer programs used to implement the mathematical development.

Figure 6.27 **Comparison of Zernike coma FFD data.** (*a*) **Generated by CODE V®.** (*b*) **Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method.** (*c*) **Generated by using the single Zernike expansion through 8th order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.19.** (*d*) **The difference between the Zernike coma FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).**

Figure 6.28 Comparison of Zernike spherical FFD data. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.19. (*d*) The difference between the Zernike spherical FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).
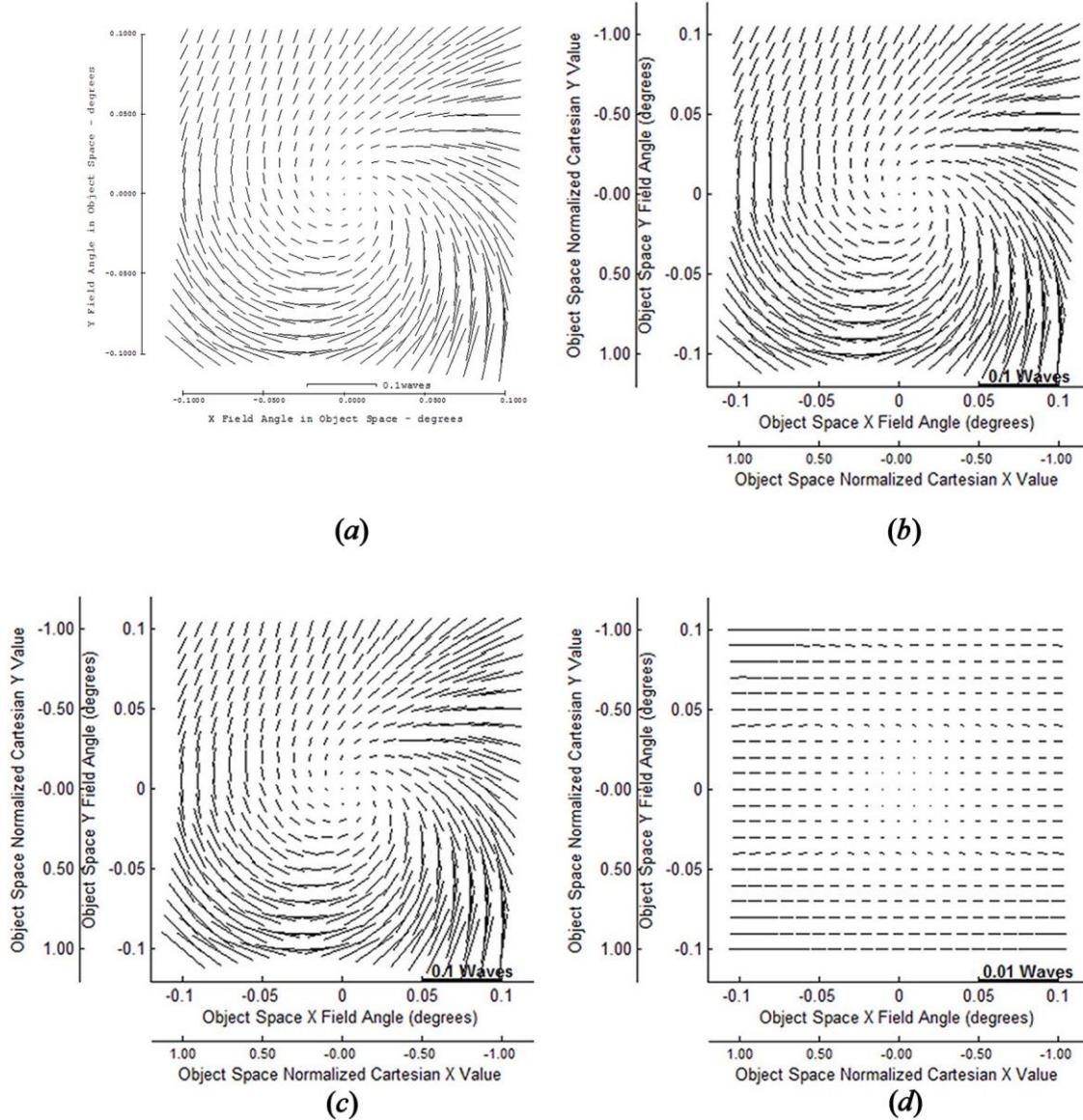
The next case to be used to check the theoretical development decenters and tilts all the surfaces of the optical model. The primary mirror was decentered along the *x*-axis by – 0.5 *mm*, along the *y*-axis by + 0.01 *mm*, and tilted by alpha = – 0.001 degrees. The secondary mirror was decentered along the *x*-axis by + 0.005 *mm*, and tilted by beta = – 0.001 degrees. The tertiary mirror was decentered along the *y*-axis by 1.0 *mm*, and tilted by alpha = – 0.001 degrees. The resulting sigma offset vectors are shown in Table 6.20.

Table 6.20 The sigma offset vectors.
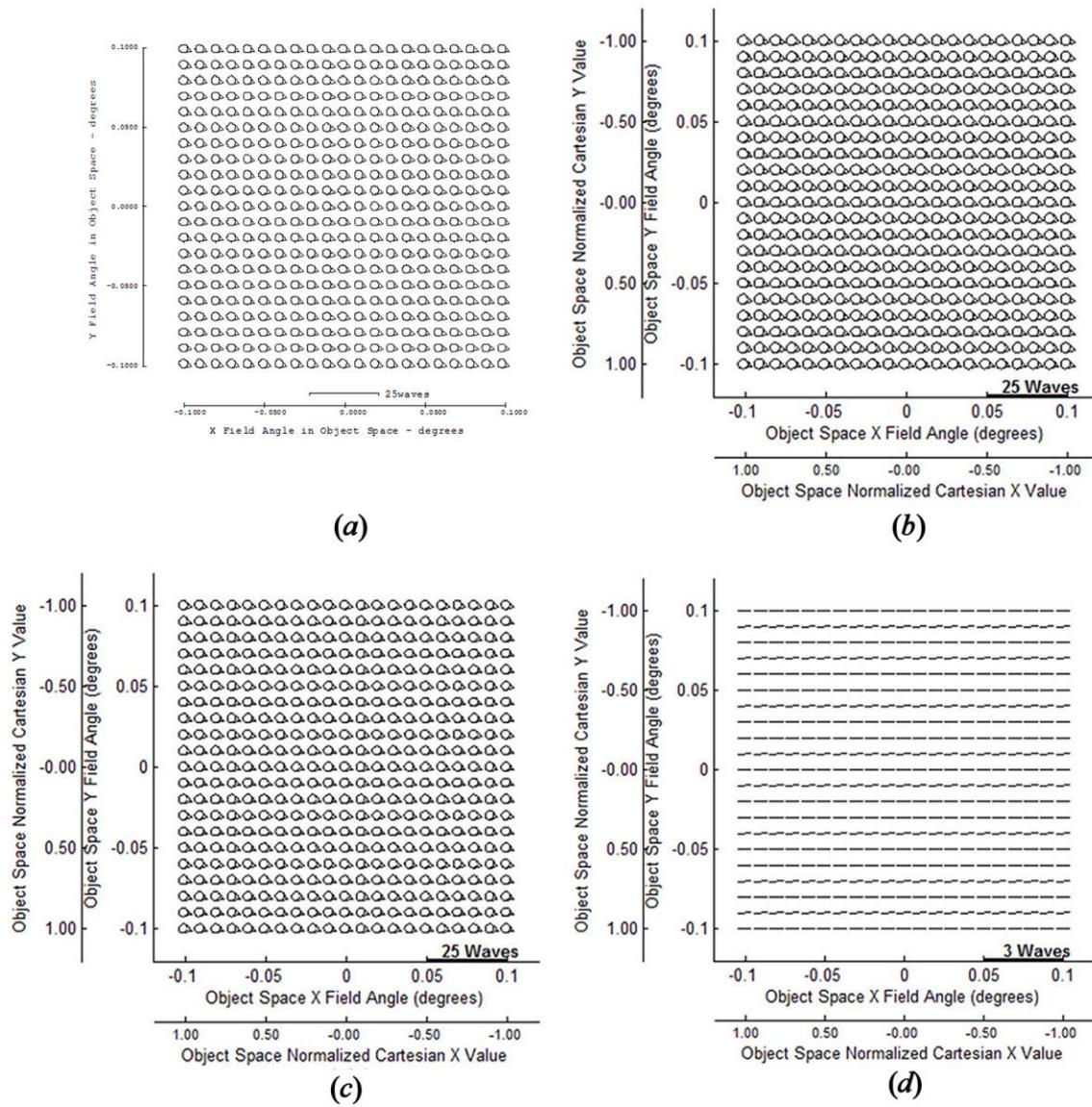
Mirror #	Sphere Sigma		Asphere Sigma	
	X	Y	X	Y
1	–0.020942	0.012027	0.130218	–0.002604
2	–0.035522	0.021389	–0.035425	0.020122
3	–0.034464	0.009120	–0.034431	0.013513

The FFD comparisons for this perturbed model are now presented. Using the sigma offset vector values in Table 6.20, and the W_{klm} values listed in Table 6.13 through Table 6.18, the Zernike astigmatism FFD was generated and compared to the FFD generated from the double Zernike expansion and the GQ method, shown in Figure 6.29. Figure 6.29(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.29(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.29(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) – (*c*), is also shown in Figure 6.29(*d*). The difference data's magnitude values have a maximum difference of 0.003 waves, a mean of 0.001 waves

and a STD of 0.0005 waves. The difference data's angle values have a maximum difference of – 2.6 degrees, a mean of – 0.54 degrees and a STD of 0.83 degrees.

Figure 6.29 Comparison of Zernike astigmatism FFD data. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.20. (*d*) The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).
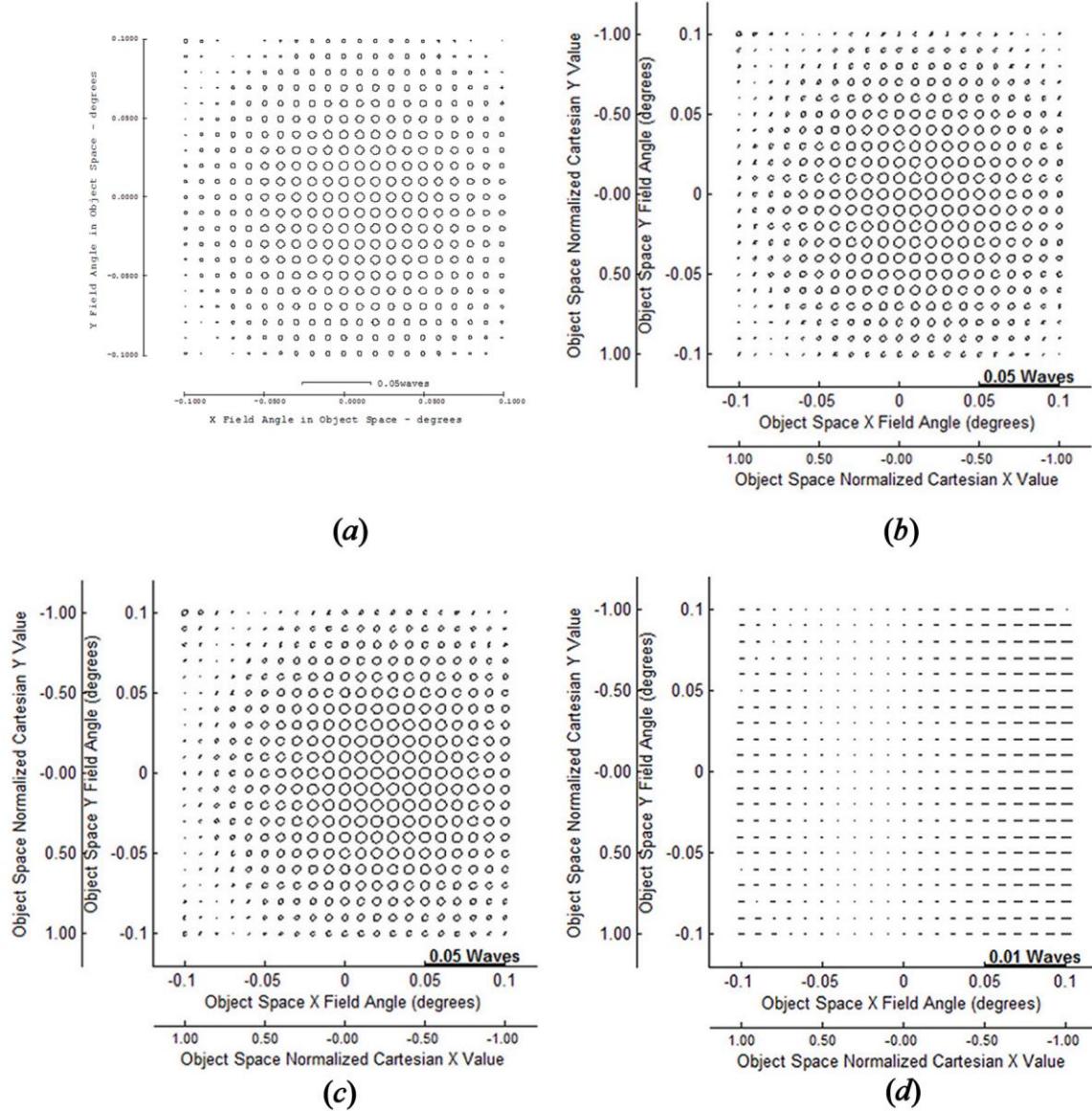
Figure 6.30 Comparison of Zernike coma FFD data. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.20. (*d*) The difference between the Zernike coma FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).

The Zernike coma FFDs are presented in Figure 6.30. Figure 6.30(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.30(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.30(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.30(*d*). The difference data's magnitude values have a maximum difference of 0.5344 waves, a mean of 0.5335 waves and a STD of 0.0005 waves. The difference data's angle values have a maximum difference of 1.61 degrees, a mean of 1.59 degrees and a STD of 0.01 degrees.

The Zernike spherical FFDs are presented in Figure 6.31. Figure 6.31(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.31(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.31(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.31(*d*). The difference data's magnitude values have a maximum difference of 0.001 waves, a mean of − 0.0003 waves and a STD of 0.001 waves.

All these comparisons of the TMA's FFDs and quantitative data show reasonably good agreement in all cases. Further investigation and refinement to the implementation may be needed to better understand the non-zero 4th order extrinsic expansion coefficient values for the secondary mirror of this model.

Figure 6.31 Comparison of Zernike spherical FFD data. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.20. (*d*) The difference between the Zernike spherical FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

6.4 TMA With Coma Free Pivot

This example investigates a TMA optical imaging model having a coma free pivot of the secondary mirror. A coma free pivot is a decenter and tilt that keeps the coma of the total imaging system very small. The TMA of the previous section is used and so its unperturbed data (the W_{klm} expansion coefficients and FFDs for the unperturbed system) will not be repeated here.
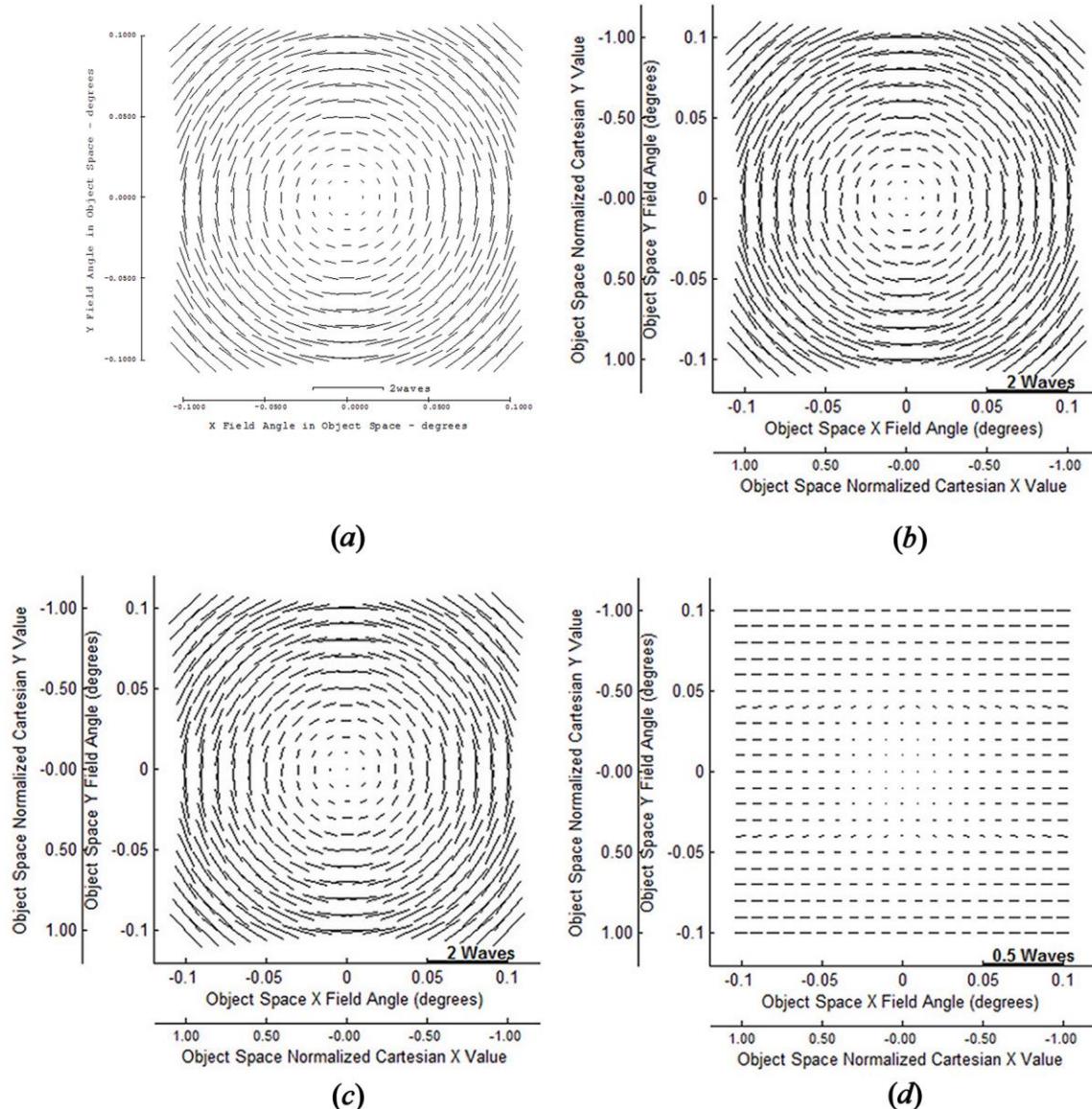
The secondary mirror is decentered along the *y*-axis by +0.59255 *mm* and tilted by alpha = +0.041666 degrees. The sigma offset vectors for this configuration are calculated to be as shown in Table 6.21.

Table 6.21 Sigma offset vectors for coma free pivot model.

Mirror #	Sphere Sigma		Asphere Sigma	
	X	Y	X	Y
1	0.0	0.0	0.0	0.0
2	0.0	0.046339	0.0	−0.045987
3	0.0	0.042951	0.0	0.045389

The FFD comparisons for this model are now presented. Using the sigma offset vector values of Table 6.21, and the W_{klm} values listed in Table 6.13 through Table 6.18 of the previous section, the Zernike astigmatism FFD was generated and compared to the FFD generated from the double Zernike expansion and the GQ method. The FFDs are presented in Figure 6.32. Figure 6.32(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 6.32(*b*) was generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. The FFD plot Figure 6.32(*c*) was generated using the single Zernike expansion and the W_{klm} expansion coefficients. The difference between the two FFD's data, (*b*) − (*c*), is also shown in Figure 6.32(*d*).

The difference data's magnitude values have a maximum difference of 0.08 waves, a mean of 0.04 waves and a STD of 0.02 waves. The difference data's angle values have a maximum difference of – 0.1 degrees, a mean of – 0.000002 degrees and a STD of 0.06 degrees.

Figure 6.32 Comparison of Zernike astigmatism FFD data. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12ᵗʰ order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8ᵗʰ order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.21. (*d*) The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

The Zernike coma and Zernike spherical FFDs were generated and compared to the FFDs generated from the double Zernike expansion and the GQ method. These FFD comparisons are presented in Figure 6.33 and Figure 6.34, respectively.


(*a*)

(*b*)

(*c*)
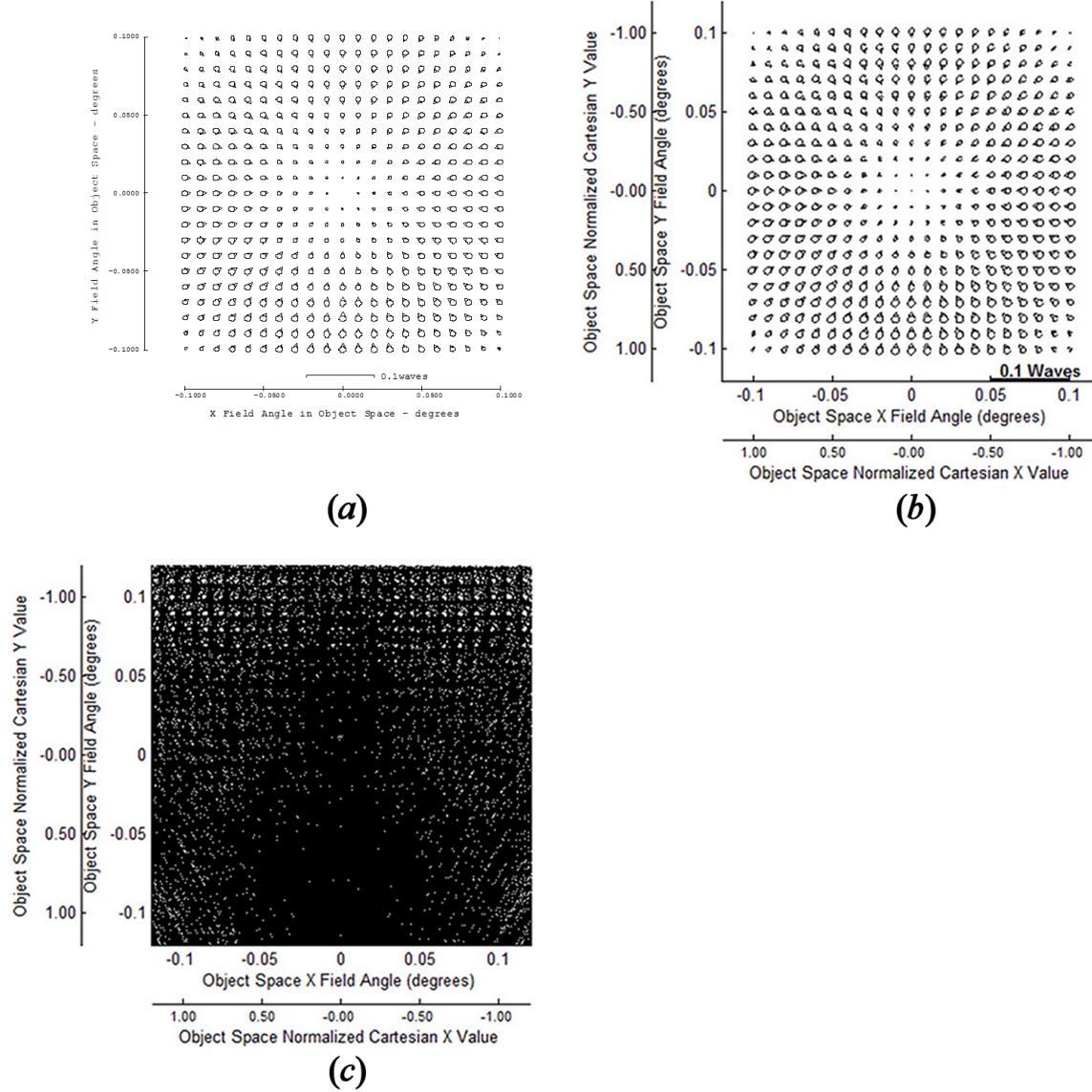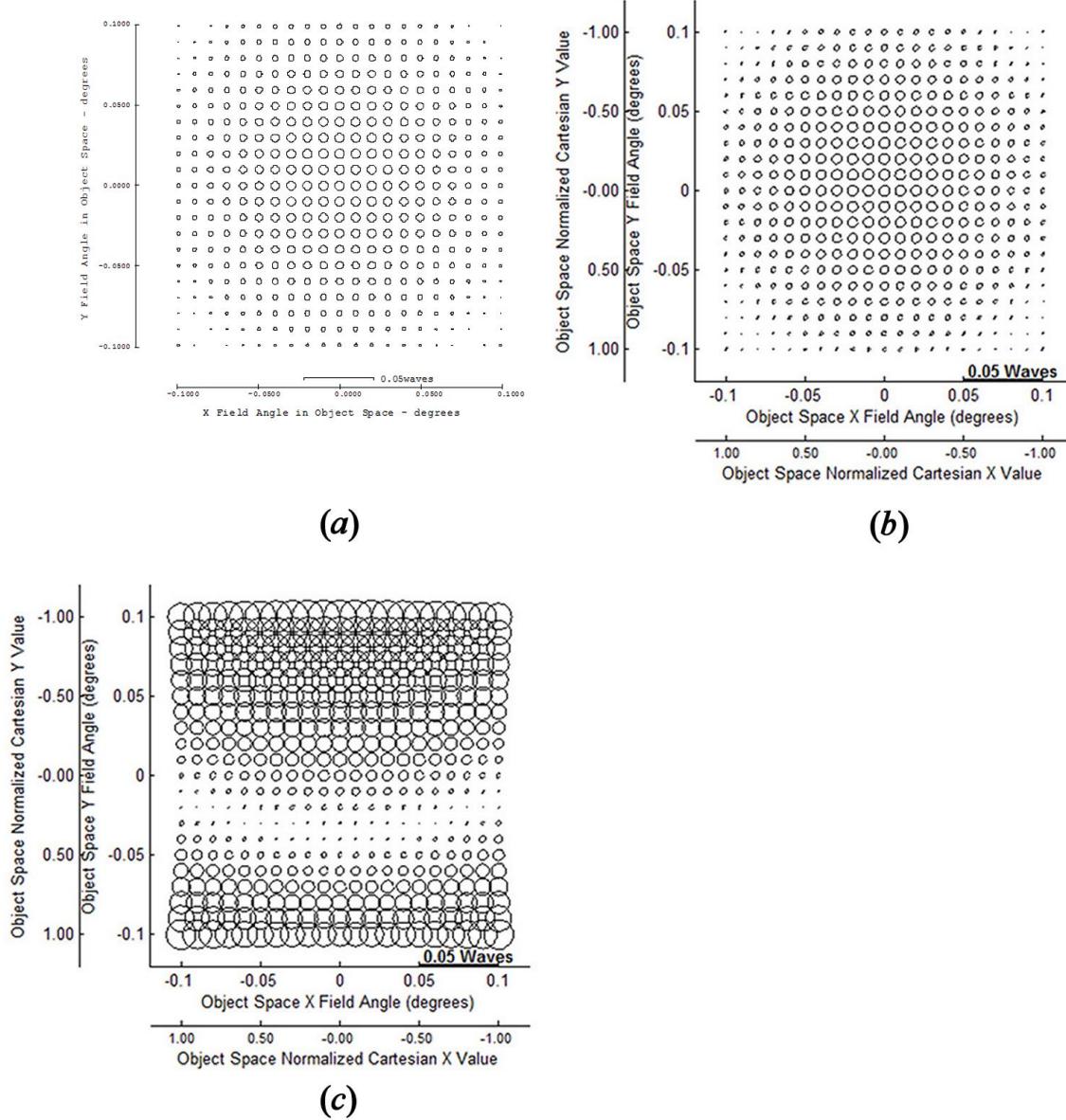
Figure 6.33   Comparison of Zernike coma FFD data. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.21.

Figure 6.34 Comparison of Zernike spherical FFD data. (*a*) Generated by CODE V®. (*b*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*c*) Generated by using the single Zernike expansion through 8th order, the W_{klm} expansion coefficients and the sigma vectors of Table 6.21.

Clearly there is not an agreement between the double Zernike and single Zernike calculated Zernike coma and Zernike spherical FFDs. Since all the preceding examples gave reasonable agreements to the FFDs, including the Zernike astigmatism of this

example shown in Figure 6.32, it is unlikely that these mismatches are due to incorrect sigma offset vector values. As pointed out earlier, it was observed that the 4th order extrinsic W_{klm} values for the secondary mirror were not zero, as they should be. See Table 6.15 and Table 6.17. This may be an indication that some of the greater than 4th order extrinsic W_{klm} values are sufficiently different from their correct values and are the cause of the FFD mismatch issue.

A series of least squares fits were performed using the Zernike coma data of the FFD shown in Figure 6.33(*b*) with the single Zernike expansion's Zernike coma field function through 8th order, Eq. (4.49), the $W_{klm;j}$ values of Table 6.15 and Table 6.16, and the sigma offset values of Table 6.21, in order to determine what changes to the $W_{klm;j}$ values would correct this FFD mismatch issue. Explicitly, the Zernike coma field function programmed for the least squares fit is given by
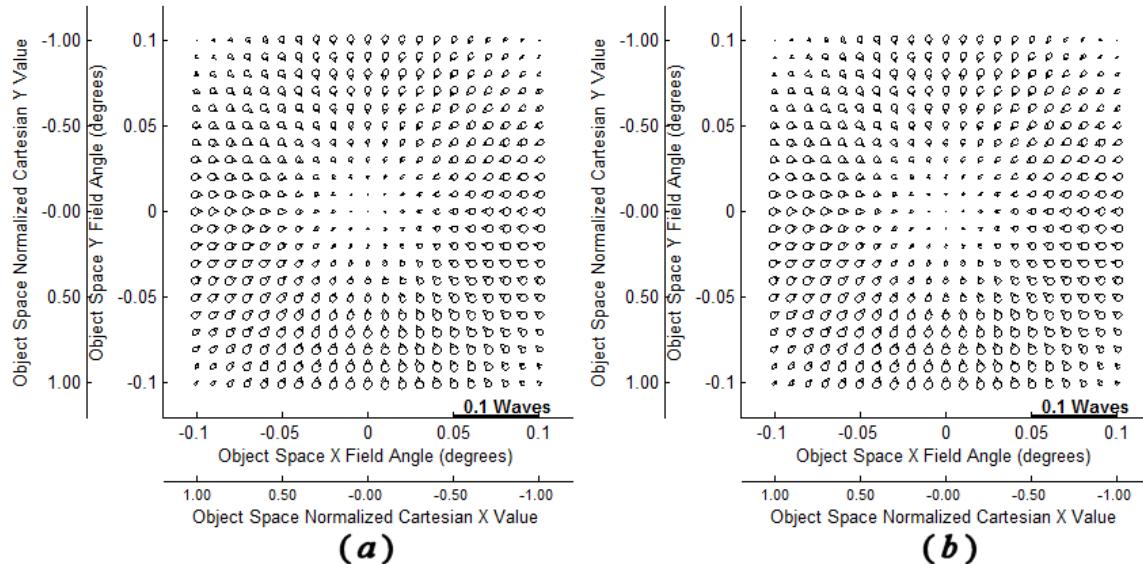
$$
\begin{aligned}
\vec{F}_{coma} = \sum_{j=1}^{3} & \left[\left(\left(\frac{1}{3} W_{131;j}^{IS} + \frac{2}{5} W_{151;j}^{IS} + \frac{2}{5} W_{171;j}^{IS} \right) + \left(\frac{1}{3} W_{331M;j}^{IS} + \frac{2}{5} W_{351M;j}^{IS} \right) \left(\vec{H} - \vec{\sigma}_j^{S} \right) \bullet \left(\vec{H} - \vec{\sigma}_j^{S} \right) + \right. \right. \\
& \quad \frac{1}{3} W_{531M;j}^{IS} \left[\left(\vec{H} - \vec{\sigma}_j^{S} \right) \bullet \left(\vec{H} - \vec{\sigma}_j^{S} \right) \right]^2 \right) \left\{ \vec{H} - \vec{\sigma}_j^{S} \right\}^1 + \\
& \quad \left(\left(\frac{1}{3} W_{131;j}^{IA} + \frac{2}{5} W_{151;j}^{IA} + \frac{2}{5} W_{171;j}^{IA} \right) + \left(\frac{1}{3} W_{331M;j}^{IA} + \frac{2}{5} W_{351M;j}^{IA} \right) \left(\vec{H} - \vec{\sigma}_j^{A} \right) \bullet \left(\vec{H} - \vec{\sigma}_j^{A} \right) + \right. \\
& \quad \frac{1}{3} W_{531M;j}^{IA} \left[\left(\vec{H} - \vec{\sigma}_j^{A} \right) \bullet \left(\vec{H} - \vec{\sigma}_j^{A} \right) \right]^2 \right) \left\{ \vec{H} - \vec{\sigma}_j^{A} \right\}^1 + \\
& \quad \left(\left(\frac{2}{5} W_{151;j}^{ES} + \frac{2}{5} W_{171;j}^{ES} \right) + \left(\frac{1}{3} W_{331M;j}^{ES} + \frac{2}{5} W_{351M;j}^{ES} \right) \left(\vec{H} - \vec{\sigma}_j^{S} \right) \bullet \left(\vec{H} - \vec{\sigma}_j^{S} \right) + \right. \\
& \quad \frac{1}{3} W_{531M;j}^{ES} \left[\left(\vec{H} - \vec{\sigma}_j^{S} \right) \bullet \left(\vec{H} - \vec{\sigma}_j^{S} \right) \right]^2 \right) \left\{ \vec{H} - \vec{\sigma}_j^{S} \right\}^1 + \\
& \quad \left(\left(\frac{2}{5} W_{151;j}^{EA} + \frac{2}{5} W_{171;j}^{EA} \right) + \left(\frac{1}{3} W_{331M;j}^{EA} + \frac{2}{5} W_{351M;j}^{EA} \right) \left(\vec{H} - \vec{\sigma}_j^{S} \right) \bullet \left(\vec{H} - \vec{\sigma}_j^{S} \right) + \right. \\
& \quad \left. \left. \frac{1}{3} W_{531M;j}^{EA} \left[\left(\vec{H} - \vec{\sigma}_j^{S} \right) \bullet \left(\vec{H} - \vec{\sigma}_j^{S} \right) \right]^2 \right) \left\{ \vec{H} - \vec{\sigma}_j^{S} \right\}^1 \right],
\end{aligned}
\tag{6.10}
$$

where the sum over j is a sum over each of the 3 mirrors of the optical system. Different combinations of the $W_{klm;j}$ expansion coefficients in Eq. (6.10) were selected to be determined by the least squares fit, keeping the original values for all other $W_{klm;j}$ coefficients as listed in Table 6.15 and Table 6.16. In this way, it was discovered that only two $W_{klm;j}$ coefficient values, W_{151}^{IA} and W_{151}^{EA}, needed to be changed to produce a FFD plot similar to Figure 6.33(b). The two new values, together with their original values, are presented in Table 6.22.

Table 6.22 New asphere W_{klm} values for secondary mirror obtained by least squares fit (LSF) of FFD data of Figure 6.33(a) and the field function through 8th order. (Units = waves.)

Mirror #	W_{151}^{IA} (Original)	W_{151}^{IA} (LSF)	W_{151}^{EA} (Original)	W_{151}^{EA} (LSF)
2	31.1538	12.77773	-23.8655	-5.49191

The FFD comparison using the two new W_{klm} values listed in Table 6.22, and all of the other values listed in Table 6.15 and Table 6.16, is shown in Figure 6.35.

Figure 6.35 Zernike coma FFD comparison using new W_{klm} values. (a) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (b) Generated by using the single Zernike expansion through 8th order, the new W_{klm} expansion coefficients and the sigma vectors of Table 6.21.

For the Zernike spherical term, a similar least squares fit process was used to search for the minimum number of $W_{klm;j}$ values that needed to be changed, and their values, that would result in a FFD similar to Figure 6.34(*b*). It was found that changing the following $W_{klm;j}$ values in Table 6.17 to the values shown in Table 6.23 produced a similar FFD to that of Figure 6.34(*b*).

Table 6.23 New asphere W_{klm} values for secondary mirror obtained by least squares fit (LSF) of FFD data of Figure 6.34(*a*) and the field function through 8th order. (Units = waves.)

Mirror #	W_{060}^{IA}	W_{060}^{IA}	W_{060}^{EA}	W_{060}^{EA}
	(Original)	(LSF)	(Original)	(LSF)
2	− 288.0319	− 37.7553	212.5216	− 37.7553
	W_{240M}^{IA}	W_{240M}^{IA}	W_{240M}^{EA}	W_{240M}^{EA}
	(Original)	(LSF)	(Original)	(LSF)
2	− 0.4058	0.1440	0.5445	− 0.0053

The FFD comparison using these W_{klm} values, and the other values in Table 6.17 and Table 6.18, is shown in Figure 6.36.
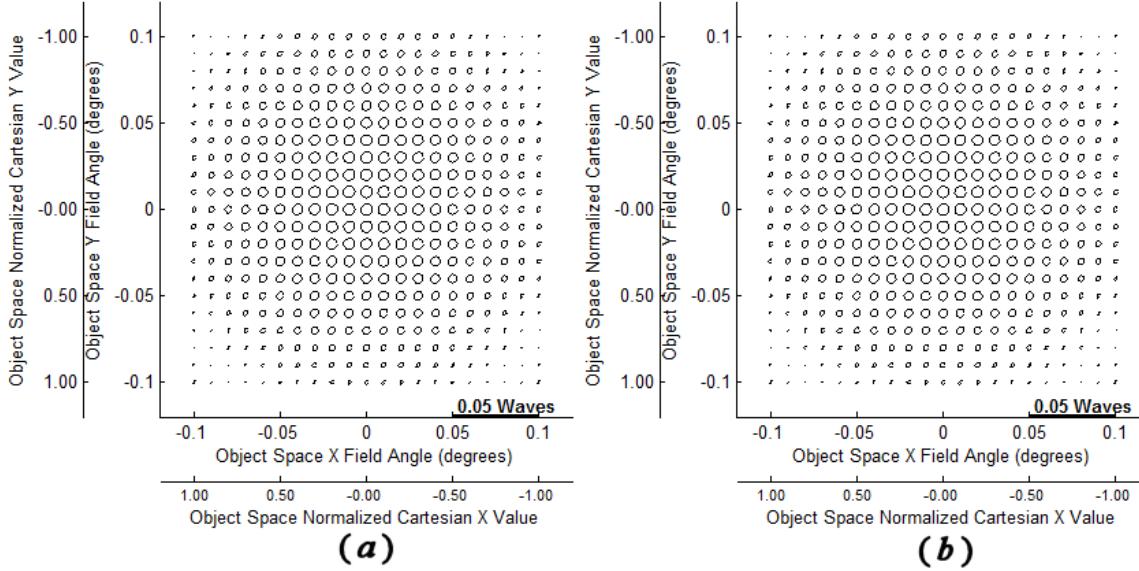
Figure 6.36 Zernike spherical FFD comparison. (*a*) Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method. (*b*) Generated by using the single Zernike expansion through 8th order, the new W_{klm} expansion coefficients and the sigma vectors of Table 6.21.

It should be pointed out that the $W_{klm;j}$ values reported in Table 6.22 and Table 6.23 should not be taken literally. Consider one part of the wavefront expansion's field function used in the least squares fit of Zernike coma (the sum over j and sigma vectors are omitted to keep the equation simple),

$$\vec{F}_{coma} = \left(\left(\frac{1}{3} W_{131}^{IA} + \frac{2}{5} W_{151}^{IA} + \frac{2}{5} W_{171}^{IA} \right) + \left(\frac{1}{3} W_{331M}^{IA} + \frac{2}{5} W_{351M}^{IA} \right) H^2 + \frac{1}{3} W_{531M}^{IA} H^4 \right) \{H\}^1 +$$

$$\left(\left(\frac{2}{5} W_{151}^{EA} + \frac{2}{5} W_{171}^{EA} \right) + \left(\frac{1}{3} W_{331M}^{EA} + \frac{2}{5} W_{351M}^{EA} \right) H^2 + \frac{1}{3} W_{531M}^{EA} H^4 \right) \{H\}^1 + \cdots \quad . \quad (6.11)$$

By replacing W_{151}^{IA} and W_{151}^{EA} by variables C_1 and C_2, to be determined by the least squares fit, the equation becomes

$$\vec{F}_{coma} = \left(\left(\frac{1}{3} W_{131}^{IA} + C_1 + \frac{2}{5} W_{171}^{IA} \right) + \left(\frac{1}{3} W_{331M}^{IA} + \frac{2}{5} W_{351M}^{IA} \right) H^2 + \frac{1}{3} W_{531M}^{IA} H^4 \right) \{H\}^1 +$$

$$\left(\left(C_2 + \frac{2}{5} W_{171}^{EA} \right) + \left(\frac{1}{3} W_{331M}^{EA} + \frac{2}{5} W_{351M}^{EA} \right) H^2 + \frac{1}{3} W_{531M}^{EA} H^4 \right) \{H\}^1 + \cdots \qquad , \quad (6.12)$$

effectively, as far as the fit is concerned, resulting in the equation

$$\vec{F}_{coma} = \left(C_1' + \left(\frac{1}{3} W_{331M}^{IA} + \frac{2}{5} W_{351M}^{IA} \right) H^2 + \frac{1}{3} W_{531M}^{IA} H^4 \right) \{H\}^1 +$$

$$\left(C_2' + \left(\frac{1}{3} W_{331M}^{EA} + \frac{2}{5} W_{351M}^{EA} \right) H^2 + \frac{1}{3} W_{531M}^{EA} H^4 \right) \{H\}^1 + \cdots \; . \qquad (6.13)$$

Therefore, from performing this least squares fit, it cannot be determined if W_{131}^{IA}, W_{171}^{IA}, W_{171}^{EA} are also incorrect or if some combination of these are incorrect. Nor can it be determined by how much these coefficients are incorrect. Similarly for the Zernike spherical W_{klm} values listed in Table 6.23. Therefore, the values listed in Table 6.22 and Table 6.23 cannot be said to be the true values of the W_{klm} coefficients listed.

One question to be asked is: Why didn't this issue manifest with the other TMA models shown previously? The answer may be that the other models were dominated by 4th order coma and spherical aberrations while this model, due to the coma free pivot perturbation, reduces (or eliminates) the 4th order coma and spherical aberration contributions to very small values, resulting in 6th order coma and spherical values being on a par with or greater than 4th order. Thus the issue is seen in this coma free pivot model but not in the preceding models.

The other question to be asked is: What is the cause of the incorrect W_{klm} values for the secondary mirror? It is hypothesized that this is most likely due to a mismatch in
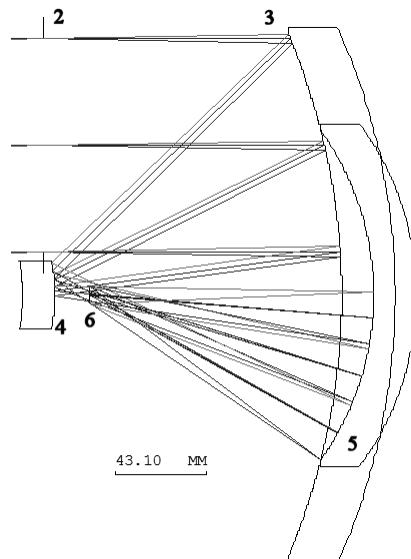
the subsystem and per surface optical models created out of the original optical model that are necessary for the calculation of the per surface, per sphere/asphere, intrinsic/extrinsic W_{klm} values as described in Chapter 4. This mismatch may be due to the way that the OPD values used in the GQ method are calculated for the subsystems and per surface models needed for the calculation of the $W_{klm;j}$ expansion coefficient values. Recall that the OPD values are calculated based on the placement of the image space reference sphere. CODE V® uses the real ray trace of the chief ray intersection with the real ray traced OAR for the location of the exit pupil rather than using the idealized, static, Gaussian exit pupil location. Then, when an aspheric surface is changed from an aspheric shape to a sphere shaped surface, the real chief ray will intersect the OAR ray at a different position. Thus the reference sphere is different for the two surface shapes (asphere and sphere). Additionally, the intersection of the real ray traced chief ray with the image plane will be different. Thus the center of the reference sphere will also be different for the asphere and sphere cases. Because of these differences in the reference sphere for the asphere and sphere cases, the resulting W_{klm} values *will not be referenced to the same reference sphere* and it may therefore not be appropriate to subtract the asphere and sphere W_{klm} values. It is therefore suggested that the W_{klm} issue may be resolved by defining the OPD calculations (and thus defining the wavefront aberration function) to be based on reference spheres defined by the Gaussian exit pupil location and the Gaussian location of the chief rays' intersection points in the Gaussian image plane. In this way, the reference spheres are not altered when going from an

aspheric shape to a sphere shape. The W_{klm} values will then be well defined and will have a common basis for their subtraction.

6.5 A Three Mirror Anastigmat Telescope with Decentered Aperture Stop

This example investigates an optical imaging model having a decentered aperture stop as the entrance pupil. The model is based on that presented by Forbes and Menke [69]. The purpose of this example is to determine whether the mathematical formulation presented in Chapter 5 is suitable for application to decentered aperture optical imaging models, or if some other component or parameterization needs to be added.

Figure 6.37 from CODE V® (modified) presents the layout of the optical model used.

Figure 6.37 Layout of optical imaging model having a decentered aperture stop as entrance pupil.

The model is defined to have a HFOV of 1 degree. The wavelength used with the model is $500\ nm$. The entrance pupil has a diameter of $100\ mm$ and is decentered from the MCA along the positive y-axis by $+70\ mm$. All three mirrors are conics. The primary, secondary and tertiary mirrors have conic constants -0.792, -6.1, and -0.135, respectively. The other parameters for the optical model are presented in Table 6.24 generated by CODE V®. Surface #1 is a dummy surface (having no optical power) used to define the global coordinate system of the model.

Table 6.24 Layout parameters for the optical imaging model.

Surface #	Surface Name	Surface Type	Y Radius	Thickness	Refract Mode	Y Semi-Aperture	Non-Centered Data
Object		Sphere	Infinity	Infinity	Refract		
1		Sphere	Infinity	0.0000	Refract	120.0000	
Stop		Sphere	Infinity	140.0000	Refract	50.0000	Decenter & Return
3		Conic	-300.0000	-135.0000	Reflect	122.0069	
4		Conic	-82.5000	150.0000	Reflect	14.7131	
5		Conic	-130.0000	-133.7791	Reflect	77.4170	
Image		Sphere	Infinity	0.0000	Refract	4.3745	

For the rotationally symmetric case (the aperture stop is not decentered) the W_{klm} expansion coefficients were calculated. The resulting values are presented in Table 6.25 for the Zernike astigmatism term of the wavefront aberration function expansion through 8th order, rounded to the 5th decimal position.
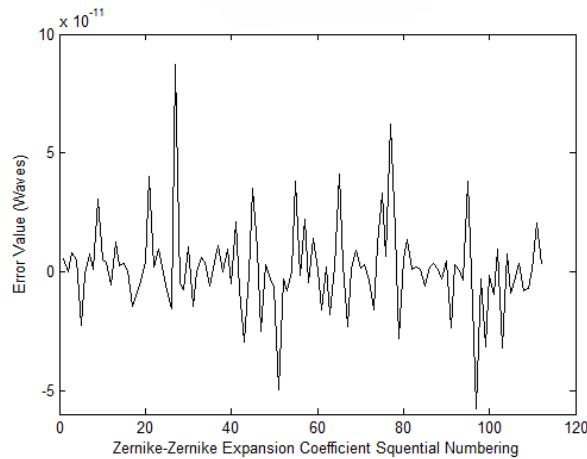
Table 6.25 Per surface expansion coefficients. IS = intrinsic sphere, IA = intrinsic asphere, ES = extrinsic sphere, EA = extrinsic asphere. GQ = values from Gaussian quadrature. FF = values from fifthdef/FORDER. (Units = waves.)

W_{klm}	GQ/FF	Primary	Secondary	Tertiary
W_{222}^{IS}	GQ	1.44441	−0.64760	9.88228
	FF	1.4444	−0.6476	9.8823
W_{222}^{IA}	GQ	−0.87585	3.67412	−13.22655
	FF	−0.8759	3.6741	−13.2265
$W_{222}^{I.Tot}$	GQ	0.56825	3.02652	−3.34427
	FF	0.5686	3.0265	−3.3443
W_{222}^{ES}	GQ	0.0	0.0	0.0
W_{222}^{EA}	GQ	0.0	0.0	0.00002
$W_{222}^{E.Tot}$	GQ	0.0	0.0	0.00002
W_{242}^{IS}	GQ	−0.01837	0.02611	−0.33636
	FF	0.0324	−0.1058	−0.1010
W_{242}^{IA}	GQ	0.18643	−0.39656	1.24415
	FF	0.0896	−0.2750	0.4604
$W_{242}^{I.Tot}$	GQ	0.16806	−0.37045	−0.90779
	FF	0.1220	−0.3807	0.3594
W_{242}^{ES}	GQ	0.0	−0.05981	0.04781
W_{242}^{EA}	GQ	0.0	0.18864	−1.01611
$W_{242}^{E.Tot}$	GQ	0.0	0.12883	−0.96830
	FF	0.0	0.7861	−1.0185
W_{422}^{IS}	GQ	−0.00044	−0.00344	−0.20483
	FF	−0.0004	0.0160	0.0025
W_{422}^{IA}	GQ	0.00064	0.00520	0.26627
	FF	0.0004	−0.0103	−0.0415
$W_{422}^{I.Tot}$	GQ	0.00019	0.00175	0.06144
	FF	0.0	0.0057	−0.0390
W_{422}^{ES}	GQ	0.0	−0.00178	0.10669
W_{422}^{EA}	GQ	0.0	−0.00242	−0.17188
$W_{422}^{E.Tot}$	GQ	0.0	−0.00420	−0.06520
	FF	0.0	−0.0027	0.0356

Table 6.26 Per surface expansion coefficients. IS = intrinsic sphere, IA = intrinsic asphere, ES = extrinsic sphere, EA = extrinsic asphere. (Units = waves.)
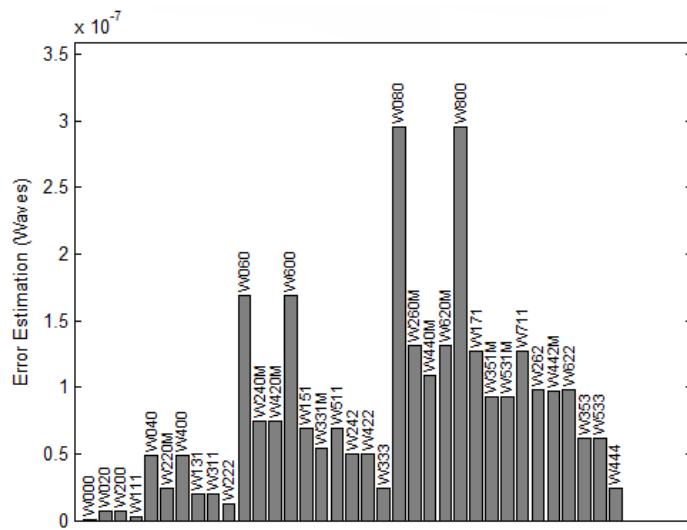
W_{klm}	Primary	Secondary	Tertiary
W_{262}^{IS}	0.00059	−0.00064	0.01086
W_{262}^{IA}	−0.00752	0.04553	0.00589
$W_{262}^{I.Tot}$	−0.00693	0.04489	0.01675
W_{262}^{ES}	0.0	0.02038	−0.03829
W_{262}^{EA}	0.0	−0.04128	−0.00699
$W_{262}^{E.Tot}$	0.0	−0.02090	−0.04528
W_{422}^{IS}	−0.00044	−0.00344	−0.20483
W_{422}^{IA}	0.00064	0.00520	0.26627
$W_{422}^{I.Tot}$	0.00020	0.00176	0.06144
W_{422}^{ES}	0.0	−0.00178	0.10669
W_{422}^{EA}	0.0	−0.00242	−0.17188
$W_{422}^{E.Tot}$	0.0	−0.00420	−0.06519
W_{442M}^{IS}	−0.00001	0.00005	0.02884
W_{442M}^{IA}	−0.00022	−0.00442	−0.03278
$W_{442M}^{I.Tot}$	−0.00023	−0.00437	−0.00394
W_{442M}^{ES}	0.0	0.00082	−0.03858
W_{442M}^{EA}	0.0	0.01016	0.03321
$W_{442M}^{E.Tot}$	0.0	0.01098	−0.00537
W_{622}^{IS}	0.0	0.00001	0.00291
W_{622}^{IA}	0.0	−0.00020	−0.00328
$W_{622}^{I.Tot}$	0.0	−0.00019	−0.00037
W_{622}^{ES}	0.0	0.0	−0.00244
W_{622}^{EA}	0.0	0.00005	0.00336
$W_{622}^{E.Tot}$	0.0	0.00005	0.00092

The following graph shows the estimate of computational error for the $U_{n_H,n_\rho}^{m_H,m_\rho}$ expansion coefficients is at most $\pm 9 \times 10^{-11}$ waves.

Figure 6.38 Upper bound error estimate of the double Zernike expansion coefficients.

Using this value, the estimate of error for the W_{klm} expansion coefficients was calculated and is presented in the bar chart of Figure 6.39.

Figure 6.39 Estimate of errors in the W_{klm} expansion coefficient values.

With some exceptions the through 8th order W_{klm} coefficients have an estimated computational error of approximately $\pm 10^{-7}$ waves or less.

Note that the value of $W_{222;3}^{EA}$, a 4th order extrinsic expansion coefficient, is not exactly zero, as it theoretically should be. (As mentioned in Chapter 4, all 4th order extrinsic W_{klm} values should be identically zero.) This suggests that the accuracy of the calculated extrinsic W_{klm} values may not be better than 10^{-5} waves.

With the aperture decentered, the sigma offset vectors were calculated. Since the decentering is along the positive y-axis, all the x components of the sigma vectors will be zero and are therefore not listed.

Table 6.27 Normalized sigma offset vectors' y components.

Mirror #	$\sigma_y^{\ S}$	$\sigma_y^{\ A}$
1	−26.8322	28.6450
2	−11.8987	2.6975
3	0.2823	−3.2057

It is concerning that most of the normalized sigma offset vectors' y-components are greater than one in absolute value. It may be that the convergence of the aberration function expansion significantly changes when the sigma offset vectors have magnitudes greater than one. Or, due to the limited accuracy of the W_{klm} expansion coefficients, the resulting numeric value of the field terms may be significantly different from their true values. For example, for the Zernike astigmatism term, $\vec{Z}_2^2(\vec{\rho})$, the field dependent factors are all of the form

$$\vec{F} = W_{kl2} H^{2s} \left\{ \vec{H} \right\}_\eta^2 , \tag{6.14}$$

where s is a positive integer or zero. The magnitude of the term goes as $2s+2$. For the through 8th order expansion being used, s is at most 2 and so the maximum power of H is 6. When following the NAT prescription for introducing the sigma offset vectors,

$$\vec{H} \rightarrow \vec{H} - \vec{\sigma}_j ,$$

$$\vec{F} = W_{kl2} \left(\left(H - \vec{\sigma}_j \right) \bullet \left(H - \vec{\sigma}_j \right) \right)^s \left\{ \vec{H} - \vec{\sigma}_j \right\}_\eta^2 . \tag{6.15}$$

Expanding the resulting equation, a term of the form $W_{kl2} \sigma_j^{2s} \left\{ \vec{\sigma}_j \right\}_\eta^2$ will appear. For $s = 2$, the sigma vector magnitude will then also have a power of 6. Using $\sigma_y = 28$, from Table 6.27 of sigma values, the sigma factor will have an order of magnitude of
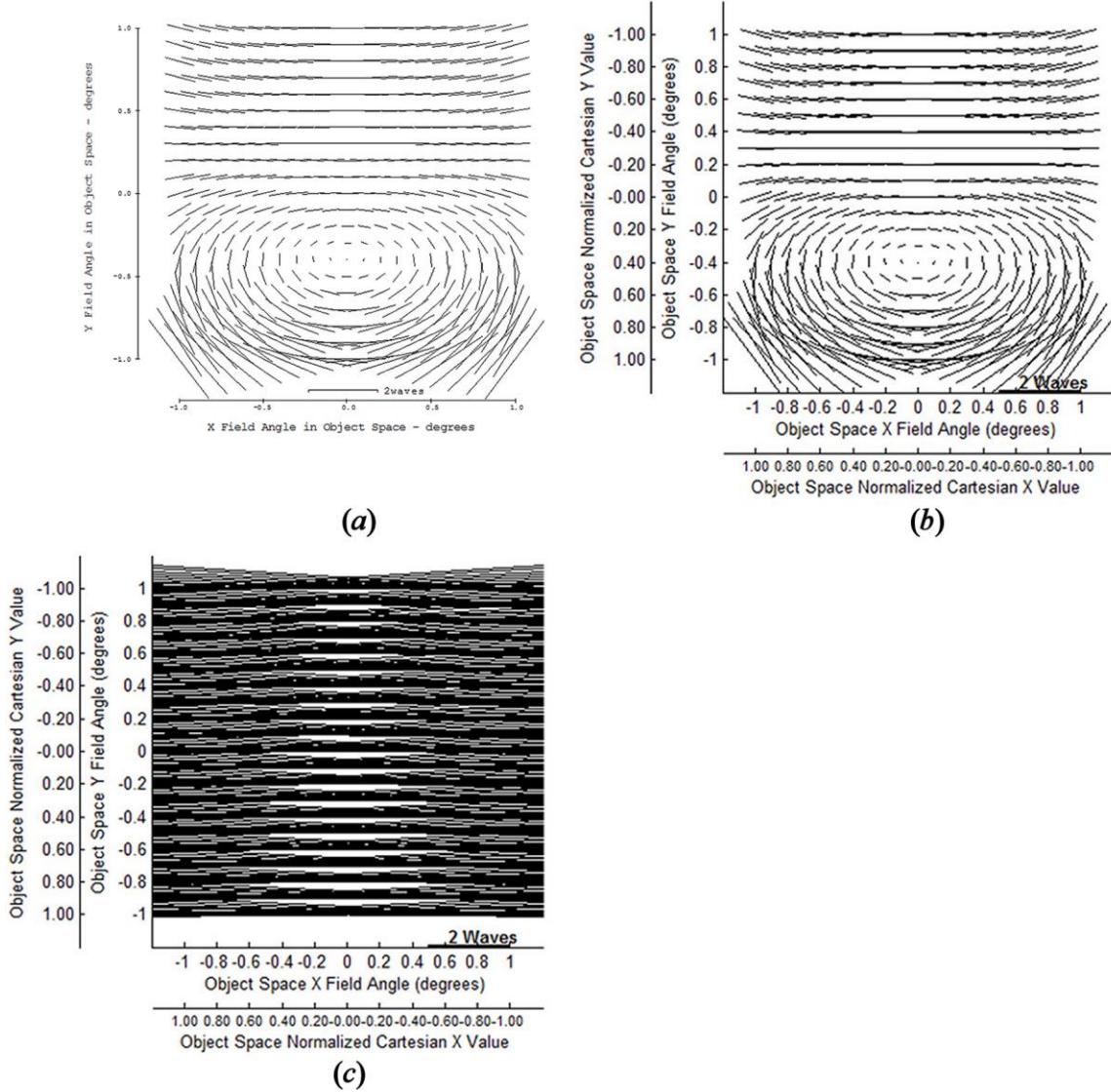
$$O = (28)^6 \approx 10^8 . \tag{6.16}$$

If the W_{klm} values are accurate out to the 7th decimal position, then the resulting contribution to the overall field dependence may be off by tens of waves. A similar result is obtained when using the estimate of error for the extrinsic W_{klm} values ($\pm 10^{-5}$ waves) and mirror #2's sigma y-component value (-12). For such cases, when the sigma offset vector components are large, the practical rule of thumb that $1/100^{th}$ waves is sufficient for W_{klm} values is not valid. Even if the rounded to 5th order W_{klm} values in Table 6.25 are used with the $\sigma_y = 28$ value, the resulting error is of the order of 10^3 waves. That is, thousands of waves in error.

A comparison of the Zernike astigmatism $\vec{Z}_2^2 (\vec{\rho})$ FFDs shows a significant mismatch.

Figure 6.40 **Comparison of Zernike astigmatism** $\vec{Z}_2^2(\vec{\rho})$ **FFDs.** (*a*) **Generated by CODE V®.** (*b*) **Generated from real ray tracing OPD data from CODE V®, double Zernike expansion through 12th order, and the GQ method.** (*c*) **Generated by using the single Zernike expansion through 8th order, W_{klm} expansion coefficients, and sigma offset vectors.**

Whether the mismatch is due to a convergence issue and/or an accuracy issue, or some other factor, is not clear. Note that the W_{klm} values actually used in creating the FFDs were calculated to the 8th decimal position.

A modification to the mathematical framework of Chapter 5 such that two additional sigma offset vectors are defined was investigated. There would then be a total of four sigma offset vectors, one for each type of W_{klm} expansion coefficient (intrinsic sphere, intrinsic asphere, extrinsic sphere, and extrinsic asphere). A Matlab® function was written, utilizing the built-in nonlinear least squares fit function provided with Matlab®, to obtain values for the new per surface extrinsic sphere and extrinsic asphere sigma offset vectors. The input to the fit was the FFD data that produced the FFD shown in Figure 6.40(*b*), the original W_{klm} values (not the rounded values of Table 6.25), the two known sigma offset values of Table 6.27 used as the intrinsic sphere and intrinsic asphere sigma vectors, and the single Zernike wavefront aberration function expansion through 8th order for the Zernike astigmatism term, $\vec{Z}_2^2(\vec{\rho})$. A parameter to the fit program specified the interval within which to search for the 6 new *y*-component extrinsic sigma values (3 surfaces × 2 new sigma vectors = 6 component values to find). The fit was run many times, varying the search interval from ±1 to ±300, in various increments ranging from 0.1 to 10 depending on the interval specified. No values were found that significantly improved the FFD comparison shown in Figure 6.40.

Another attempt was made using the nonlinear least squares fit program, but this time without using any of the known sigma values. That is, the program was modified to find all 12 *y*-components of the sigma vectors (3 surfaces × 4 different sigma types = 12). In this case, values for the sigma vectors that produced better matches to the FFD shown in Figure 6.40(*b*), were obtained. However, the resulting sigma values were not unique. That is, varying the fit interval parameters resulted in greatly differing sigma values while

still producing better FFD matches than that shown in Figure 6.40. Because unique (stable) values could not be determined the resulting sigma values where considered to be unreliable, and so they could not be used to relate the new values obtained to the optical model in any meaningful way. Further investigation concerning the introduction of new sigma vectors was then abandoned.

An alternative approach that has had some success in addressing the issue of optical systems with a decentered pupil has been published in the literature [70, 71]. In their approach, a new decentered pupil vector is introduced into the pupil parameters of the wavefront aberration function expansion rather than attempting to adjust the field parameters. Further details can be found in the papers cited.

Chapter 7 Extending NAT For Freeform Optical Elements

An approach for the development of mathematical expressions for the field dependence associated with freeform optical elements, a result of this research, using the techniques of NAT is proposed in this Chapter. Examples will be shown that utilize the derived expressions and qualitative and quantitative comparisons of the FFDs are made. Early results sufficient to suggest that this approach may be useful are provided leaving a more detailed and exhaustive analysis for future research.

In this work, a freeform optical element will mean a circular optical element that has a rotationally nonsymmetric smooth surface shape beyond the anamorphic shape. Since the Zernike polynomials are complete over a unit radius disk, the mathematical description of such a surface can be represented by a sum of Zernike polynomials.

7.1 NAT and Freeform Optical Surfaces

An asphere is defined by the sag equation

$$z_{Asphere} = \frac{c\rho^2}{1+\sqrt{1-(1+\kappa)c^2\rho^2}} + A_4\rho^4 + A_6\rho^6 + \cdots \,, \tag{7.1}$$

where $c = \dfrac{1}{r}$ is the curvature of the surface, r being the radius of curvature for the surface at the surface's vertex point, κ is the conic constant, $\rho^2 = x^2 + y^2$, and A_4, A_6, etc. are constant coefficients.

In NAT, the contribution to the field dependence of a term in the expansion of the wavefront aberration function due to an aspheric surface is separated into the contribution due to a base sphere shape of the optical element plus the field dependence due to the remaining aspheric cap terms. The sag of a spherical surface $(\kappa = 0)$ is given by

$$z_{sphere} = \frac{c\rho^2}{1 + \sqrt{1 - c^2\rho^2}} \quad . \tag{7.2}$$

Using Eq. (7.2) in Eq. (7.1), the base sphere shape of the asphere can be separated from the aspheric terms as follows,

$$z_{Asphere} = \frac{c\rho^2}{1 + \sqrt{1 - c^2\rho^2}} + \frac{c\rho^2}{1 + \sqrt{1 - (1+\kappa)c^2\rho^2}} - \frac{c\rho^2}{1 + \sqrt{1 - c^2\rho^2}} + A_4\rho^4 + A_6\rho^6 + \cdots . \tag{7.3}$$

Using the expansion

$$\frac{c\rho^2}{1 + \sqrt{1 - (1+\kappa)c^2\rho^2}} = \sum_{n=1}^{\infty} \frac{(2n)!}{(2n-1)(n!)^2(4^n)}(\kappa+1)^{n-1}c^{2n-1}\rho^{2n} , \tag{7.4}$$

the second and third terms of Eq. (7.3) can be written as

$$\frac{c\rho^2}{1 + \sqrt{1 - (1+\kappa)c^2\rho^2}} - \frac{c\rho^2}{1 + \sqrt{1 - c^2\rho^2}} = \sum_{n=1}^{\infty} \frac{(2n)!}{(2n-1)(n!)^2(4^n)}(\kappa+1)^{n-1}c^{2n-1}\rho^{2n}$$
$$- \sum_{n=1}^{\infty} \frac{(2n)!}{(2n-1)(n!)^2(4^n)}c^{2n-1}\rho^{2n} . \tag{7.5}$$

Matching like terms on the right hand side gives

$$\frac{c\rho^2}{1 + \sqrt{1 - (1+\kappa)c^2\rho^2}} - \frac{c\rho^2}{1 + \sqrt{1 - c^2\rho^2}} = \sum_{n=2}^{\infty} \frac{(2n)!}{(2n-1)(n!)^2(4^n)}\left((\kappa+1)^{n-1} - 1\right)c^{2n-1}\rho^{2n} . \tag{7.6}$$

The asphere sag equation can then be written as

$$z_{Asphere} = \frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}} + \sum_{n=2}^{\infty} \frac{(2n)!}{(2n-1)(n!)^2(4^n)}\left((\kappa+1)^{n-1}-1\right)c^{2n-1}\rho^{2n} + A_4\rho^4 + A_6\rho^6 + \cdots . \quad (7.7)$$

This can be written as

$$z_{Asphere} = \frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}} + \sum_{n=2}^{\infty}\left(\frac{(2n)!}{(2n-1)(n!)^2(4^n)}\left((\kappa+1)^{n-1}-1\right)c^{2n-1} + A_{2n}\right)\rho^{2n} , \quad (7.8)$$

and as

$$z_{Asphere} = \frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}} + \sum_{n=2}^{\infty} A'_{2n}\rho^{2n} , \quad (7.9)$$

where

$$A'_{2n} = \frac{(2n)!}{(2n-1)(n!)^2(4^n)}\left((\kappa+1)^{n-1}-1\right)c^{2n-1} + A_{2n} . \quad (7.10)$$

Since Eq. (7.9) has no azimuthal angle dependence it can also be written as

$$z_{Asphere} = \frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}} + \sum_{n=0}^{\infty} C^0_{2n} Z^0_{2n}\left(\rho/\rho_{max}\right) , \quad (7.11)$$

where C^0_{2n} are the Zernike polynomial expansion coefficients, and ρ_{max} is the maximum radial extent of the surface. Note that because only the $m = 0$ Zernike polynomials are used in Eq. (7.11), there is no azimuthal angle dependence.

For a freeform optical element, the remaining Zernike polynomials, those having an azimuthal angle dependence, are added to either Eq. (7.9) or to Eq. (7.11). That is

$$z_{freeform} = \frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}} + \sum_{n=2}^{\infty} A'_{2n}\rho^{2n} + \sum_{n=1}^{\infty} \sum_{\substack{m=-n \\ m=m+2 \\ m\neq 0}}^{n} C^m_n Z^m_n\left(\vec{\rho}/\rho_{max}\right) , \quad (7.12)$$

or

$$z_{freeform} = \frac{c\rho^2}{1+\sqrt{1-c^2\rho^2}} + \sum_{n=0}^{\infty} C_{2n}^{0} Z_{2n}^{0}\left(\rho/\rho_{\max}\right) + \sum_{n=1}^{\infty} \sum_{\substack{m=-n \\ m=m+2 \\ m\neq 0}}^{n} C_n^m Z_n^m\left(\vec{\rho}/\rho_{\max}\right) , \quad (7.13)$$

where the summation over m is such that $-n \leq m \leq n$, $m \neq 0$, and m changes in increments of $+2$.

To aid in the description of the method to be presented in the remainder of this section, and for explicit illustration purposes, the field function for the $\vec{Z}_2^2(\vec{\rho})$ term in the wavefront aberration function expansion Eq. (4.49) will be used, through 6th order. (The "through 6th order" restriction is too restrictive for some calculations, as will be pointed out below.) The field function is then written, for some fixed but arbitrary optical element in the optical imaging system (suppressing the surface number "j" subscript), as

$$\vec{F}_2^2 = \left(\left(\frac{1}{2}W_{222} + \frac{3}{8}W_{242}\right) + \frac{1}{2}W_{422}H^2\right)\{\vec{H}\}_\eta^2 . \quad (7.14)$$
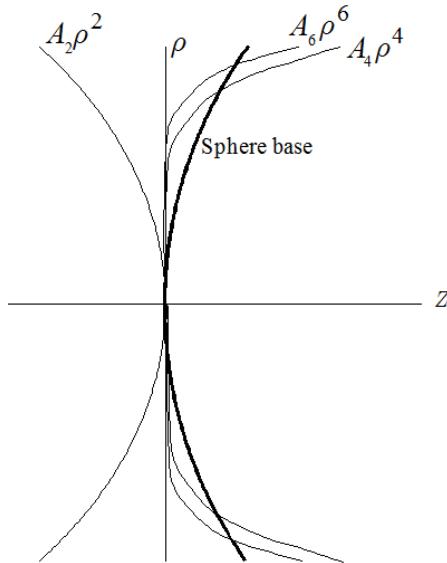
After relabeling the expansion coefficients and distributing the Shack vector product, this can be written as

$$\vec{F}_2^2 = V_1\{\vec{H}\}_\eta^2 + V_2 H^2 \{\vec{H}\}_\eta^2 . \quad (7.15)$$

If the surface is an asphere, then, according to NAT, the field function is given in two parts, one for the base sphere shape contribution and one for the aspheric cap shape contribution,

$$\vec{F}_2^2 = \left[V_1^S\{\vec{H}\}_\eta^2 + V_2^S H^2\{\vec{H}\}_\eta^2\right] + \left[V_1^A\{\vec{H}\}_\eta^2 + V_2^A H^2\{\vec{H}\}_\eta^2\right] . \quad (7.16)$$

Continuing in a similar way to separate out different surface shape contributions to the field function, an assumption is made that the field contribution due to any number of rotationally symmetric terms added to the base sphere shape can be separated into their individual term's field contributions. See Figure 7.1 where terms of the form $A_2\rho^2$, $A_4\rho^4$, and $A_6\rho^6$ are added to the base sphere. Any rotationally symmetric term of the form ρ^n where n is a real number, $n \geq 2$ could be added to the base sphere.

Figure 7.1 Separation of the surface shape into a sphere base shape plus individual rotationally symmetric terms. The additional terms will contribute different amounts to the wavefront aberration function for the surface considered.   The A_2, A_4, A_6 are the shape coefficients used to defined the additional term shapes. $\left(\rho = \sqrt{x^2 + y^2}\right)$
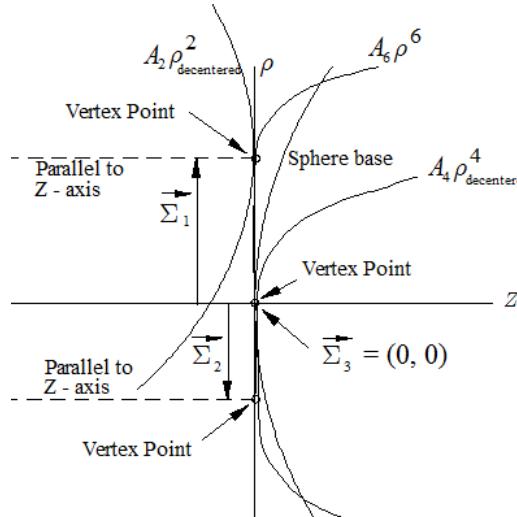
Each separate additional term will contribute a different amount to the overall surface's field function, and each term having its own expansion coefficient. Assuming that the surface has been defined with n additional terms, there are then n rotationally symmetric term contributions to the field function. The field function is then written as

$$\vec{F}_2^2 = \left[V_1^S \left\{ \vec{H} \right\}_\eta^2 + V_2^S H^2 \left\{ \vec{H} \right\}_\eta^2 \right] +$$

$$\left[V_{222(1)} \left\{ \vec{H} \right\}_\eta^2 + V_{422(1)} H^2 \left\{ \vec{H} \right\}_\eta^2 \right] + \cdots + \left[V_{222(n)} \left\{ \vec{H} \right\}_\eta^2 + V_{422(n)} H^2 \left\{ \vec{H} \right\}_\eta^2 \right] , \quad (7.17)$$

where the n $V_{222(i)}$'s and n $V_{422(i)}$'s are new expansion coefficients. The subscript notation for the expansion coefficients $V_{klm(i)}$ is such that k is the total power of the field parameter H, l is the total power of the pupil parameter ρ for which, in this example, $\vec{F}_2^2$ is its field function, m is the power of SVP in field and pupil parameter and i is a sequential numbering of the rotationally symmetric additional terms being added to the sphere base shape. Note that there may be multiple copies of terms having the same ρ dependence. For example, the additional terms may include two separate surface shape contributions, $2\rho^4 - \rho^4$, or even $2\rho^4 - 2\rho^4$. In the latter case, there is no net surface shape change, but each term is still to be considered as a separate surface added to the base sphere and contributing equal but opposite amounts to the field function.

It is known from NAT that if an aspheric cap is decentered, the resulting field function is obtained by introducing a $\vec{\sigma}$ offset vector, defined in the total system's object plane, into the field dependence of the vector field function. The decenter vector of the aspheric cap at surface j is denoted by the $\vec{\Sigma}$ symbol. The $\vec{\sigma}$ offset vector is obtained by normalizing and projecting the $\vec{\Sigma}$ decenter vector to the local image plane of surface j and then potentially adjusting the sign to account for the possible orientation difference between the system's object plane and the local image plane. This potential sign adjustment moves the offset vector to the total system's object plane. See Chapter 5 for details.

It is now assumed that the individual additional surface shape terms being added to the sphere base shape may be separately and independently decentered radially away from the MCA (the z-axis) in a similar way that an aspheric cap is decentered in NAT. That is, the vertex of each of the additional terms may be decentered away from the MCA, remaining in the xy-plane, by different amounts and in different directions given by $\vec{\Sigma}_i$ decenter vectors in the surface's vertex plane. Note that the case that one or more of the additional terms are not decentered is allowed by setting the corresponding $\vec{\Sigma}_i$ decenter vectors to be zero vectors. This is illustrated in Figure 7.2.

Figure 7.2 **Some, not all, of the additional surface shape terms, considered as individual surface contributions, are decentered in the xy-plane. The amount of and direction of the decentering is specified by the $\vec{\Sigma}_i$ offset vectors, which is the zero vector for the additional surface shape terms that are not decentered.**

Letting $\vec{\sigma}_{Ai}$ represent the normalized offset vector in the total system's object plane for additional surface shape term i, the field function can then be written as

$$\vec{F}_2^2 = \left[V_1^S \left\{ \vec{H} \right\}_\eta^2 + V_2^S H^2 \left\{ \vec{H} \right\}_\eta^2 \right] +$$

$$\left[V_{222(1)} \left\{ \vec{H} - \vec{\sigma}_{A1} \right\}_\eta^2 + V_{422(1)} \left(\vec{H} - \vec{\sigma}_{A1} \right) \bullet \left(\vec{H} - \vec{\sigma}_{A1} \right) \left\{ \vec{H} - \vec{\sigma}_{A1} \right\}_\eta^2 \right] + \cdots +$$

$$\left[V_{222(n)} \left\{ \vec{H} - \vec{\sigma}_{An} \right\}_\eta^2 + V_{422(n)} \left(\vec{H} - \vec{\sigma}_{An} \right) \bullet \left(\vec{H} - \vec{\sigma}_{An} \right) \left\{ \vec{H} - \vec{\sigma}_{An} \right\}_\eta^2 \right] \qquad . \qquad (7.18)$$

Eq. (7.18) describes the field dependent contribution of a surface having multiple additional surface shape terms, some of which may be individually decentered from the MCA, some of which may not be decentered, but for which as a whole the optical element is not decentered from the MCA.

As a specific example, consider a sphere shaped surface defined with an additional $Z_3^{+3}(\vec{\rho})$ "cap", where $\vec{\rho} = (x, y)$ is the normalized radial vector from the z-axis to the surface. The $Z_3^{+3}(\vec{\rho})$ cap may represent the effect of a 3 point mechanical mounting system's deformation of an optical element's surface shape as described in [72]. Using the x-axis as the axis of reference, and from Table IV.5 in Appendix IV, this Zernike polynomial can be written as a sum of decentered rotationally symmetric surface shape terms as follows.
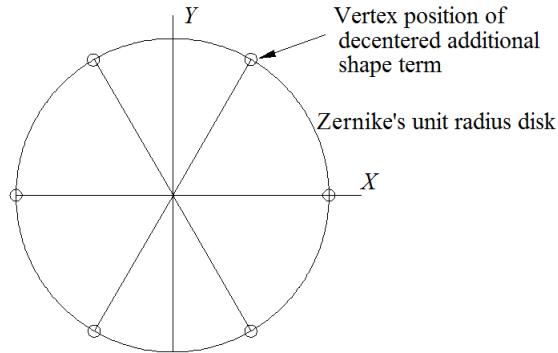
$$Z_3^{+3}(\vec{r}) = \frac{1}{12} \left[P^3(-1,0) - P^3(1,0) + P^3\left(\frac{1}{2}, \frac{\sqrt{3}}{2} \right) - P^3\left(\frac{-1}{2}, \frac{-\sqrt{3}}{2} \right) + P^3\left(\frac{1}{2}, \frac{-\sqrt{3}}{2} \right) - P^3\left(\frac{-1}{2}, \frac{\sqrt{3}}{2} \right) \right], \quad (7.19)$$

where, as discussed in Appendix IV,

$$P^n(\delta x, \delta y) \equiv \left[(x - \delta x)^2 + (y - \delta y)^2 \right]^n, \qquad (7.20)$$

is the equation for the decentered shape terms. There are then 6 decentered terms to define the Zernike polynomial $Z_3^{+3}(\vec{\rho})$. Note that because this decomposition of the

Zernike polynomial decenters each term radially out to the perimeter of the Zernike polynomial's unit radius disk, that is, because $\delta x^2 + \delta y^2 = 1$, then $\vec{\Sigma}_{norm} = (\delta x, \delta y)$ are normalized decenter vectors at the optical surface. The position of each of the 6 decentered terms that defines this Zernike polynomial is illustrated in Figure 7.3.

Figure 7.3 Position of 6 decentered additional shape terms around the unit radius Zernike disk to define $Z_3^{+3}(\vec{\rho})$.

To calculate the decentered additional shape terms' normalized sigma vectors to be used in the wavefront aberration function's expansion field functions, the sign of the chief ray height at the surface with respect to the sign of the chief ray height at the system's object plane needs to be taken into consideration. Then, as with NAT's sigma offset vectors described in Chapter 5,

$$\vec{\sigma}_{Ai;j} = \text{sgn}\left(HCY_{Tot}^{Obj}\right)\text{sgn}\left(HCY_j^{Surf}\right)\vec{\Sigma}_{norm}$$
$$= \text{sgn}\left(HCY_{Tot}^{Obj}\right)\text{sgn}\left(HCY_j^{Surf}\right)(\delta x, \delta y) , \qquad (7.21)$$

where subscript "i" identifies which decentered additional term, subscript "j" identifies the surface, HCY_{Tot}^{Obj} is the chief ray height at the total system's object plane (which may be infinite) and HCY_j^{Surf} is the chief ray height at surface j.

Further noticing that the decentered additional surface shape terms occur in diametrically opposite positions around the perimeter of the unit radius disk, there are then only 3 $\vec{\sigma}_{Ai}$ offset vectors up to sign. All the $\vec{\sigma}_{Ai}$ offset vectors are known from Eq. (7.19) and Eq. (7.21). Not only are the positions of the decentered additional term pairs positioned diametrically opposite from one another, they have opposing sag shapes due to the alternating "+" and "−" signs in front of the terms in Eq. (7.19). Since the shapes are otherwise the same, for the specific example of Eq. (7.19), their contributions to the field dependence is assumed to be the same up to sign. The field function then becomes

$$
\begin{aligned}
\vec{F}_2^2 = & \left[V_1^S \left\{ \vec{H} \right\}_\eta^2 + V_2^S H^2 \left\{ \vec{H} \right\}_\eta^2 \right] + \\
& \left[V_{222(1)} \left\{ \vec{H} - \vec{\sigma}_{A1} \right\}_\eta^2 + V_{422(1)} \left(\vec{H} - \vec{\sigma}_{A1} \right) \cdot \left(\vec{H} - \vec{\sigma}_{A1} \right) \left\{ \vec{H} - \vec{\sigma}_{A1} \right\}_\eta^2 \right] + \\
& \left[-V_{222(1)} \left\{ \vec{H} + \vec{\sigma}_{A1} \right\}_\eta^2 - V_{422(1)} \left(\vec{H} + \vec{\sigma}_{A1} \right) \cdot \left(\vec{H} + \vec{\sigma}_{A1} \right) \left\{ \vec{H} + \vec{\sigma}_{A1} \right\}_\eta^2 \right] + \\
& \left[V_{222(2)} \left\{ \vec{H} - \vec{\sigma}_{A2} \right\}_\eta^2 + V_{422(2)} \left(\vec{H} - \vec{\sigma}_{A2} \right) \cdot \left(\vec{H} - \vec{\sigma}_{A2} \right) \left\{ \vec{H} - \vec{\sigma}_{A2} \right\}_\eta^2 \right] + \\
& \left[-V_{222(2)} \left\{ \vec{H} + \vec{\sigma}_{A2} \right\}_\eta^2 - V_{422(2)} \left(\vec{H} + \vec{\sigma}_{A2} \right) \cdot \left(\vec{H} + \vec{\sigma}_{A2} \right) \left\{ \vec{H} + \vec{\sigma}_{A2} \right\}_\eta^2 \right] + \\
& \left[V_{222(3)} \left\{ \vec{H} - \vec{\sigma}_{A3} \right\}_\eta^2 + V_{422(3)} \left(\vec{H} - \vec{\sigma}_{A3} \right) \cdot \left(\vec{H} - \vec{\sigma}_{A3} \right) \left\{ \vec{H} - \vec{\sigma}_{A3} \right\}_\eta^2 \right] + \\
& \left[-V_{222(3)} \left\{ \vec{H} + \vec{\sigma}_{A3} \right\}_\eta^2 - V_{422(3)} \left(\vec{H} + \vec{\sigma}_{A3} \right) \cdot \left(\vec{H} + \vec{\sigma}_{A3} \right) \left\{ \vec{H} + \vec{\sigma}_{A3} \right\}_\eta^2 \right]
\end{aligned}
\quad , \quad (7.22)
$$

where, for this specific example, from Eq. (7.19),

$$
\vec{\sigma}_{A1} = \text{sgn}\left(HCY_{Tot}^{Obj} \right) \text{sgn}\left(HCY_j^{Surf} \right)(-1, 0) \ , \tag{7.23}
$$

$$
\vec{\sigma}_{A2} = \text{sgn}\left(HCY_{Tot}^{Obj} \right) \text{sgn}\left(HCY_j^{Surf} \right)\left(\frac{1}{2}, \frac{\sqrt{3}}{2} \right) , \tag{7.24}
$$

$$\vec{\sigma}_{A3} = \text{sgn}\left(HCY_{Tot}^{Obj}\right)\text{sgn}\left(HCY_j^{Surf}\right)\left(\frac{1}{2}, \frac{-\sqrt{3}}{2}\right) . \tag{7.25}$$

In this, Eq. (7.22), form, there are 6 unknown expansion coefficients, $V_{222(1)}$ through $V_{222(3)}$ and $V_{422(1)}$ through $V_{422(3)}$. A least squares fit to FFD data obtained by real ray tracing has been performed to obtain these 6 coefficients for the example being described. Prior to performing the fit, and for this specific case of Eq. (7.19) defining the 6 additional decentered surface shape terms, it was hypothesized that all the $V_{222(i)}$ coefficients should be the same and all the $V_{422(i)}$ coefficients should be the same. This again is based on the idea that each of the 6 decentered surface shape terms are the same shape and should therefore contribute the same amount to the field dependence, up to sign. The resulting 6 values obtained from the least squares fit for the coefficients confirmed this hypothesis. (Explicit results will be shown in the next section.) Therefore, the field dependent vector function for the $\vec{Z}_2^2\left(\vec{\rho}\right)$ pupil term of the wavefront aberration function expansion for a sphere shaped optical element having an additional $Z_3^{+3}\left(\vec{\rho}\right)$ cap shape is given by,

$$\vec{F}_2^2 = \left[V_1^S \left\{ \vec{H} \right\}_\eta^2 + V_2^S H^2 \left\{ \vec{H} \right\}_\eta^2 \right] +$$

$$\left[V_{222} \left\{ \vec{H} - \vec{\sigma}_{A1} \right\}_\eta^2 + V_{422} \left(\vec{H} - \vec{\sigma}_{A1} \right) \bullet \left(\vec{H} - \vec{\sigma}_{A1} \right) \left\{ \vec{H} - \vec{\sigma}_{A1} \right\}_\eta^2 \right] +$$

$$\left[-V_{222} \left\{ \vec{H} + \vec{\sigma}_{A1} \right\}_\eta^2 - V_{422} \left(\vec{H} + \vec{\sigma}_{A1} \right) \bullet \left(\vec{H} + \vec{\sigma}_{A1} \right) \left\{ \vec{H} + \vec{\sigma}_{A1} \right\}_\eta^2 \right] +$$

$$\left[V_{222} \left\{ \vec{H} - \vec{\sigma}_{A2} \right\}_\eta^2 + V_{422} \left(\vec{H} - \vec{\sigma}_{A2} \right) \bullet \left(\vec{H} - \vec{\sigma}_{A2} \right) \left\{ \vec{H} - \vec{\sigma}_{A2} \right\}_\eta^2 \right] +$$

$$\left[-V_{222} \left\{ \vec{H} + \vec{\sigma}_{A2} \right\}_\eta^2 - V_{422} \left(\vec{H} + \vec{\sigma}_{A2} \right) \bullet \left(\vec{H} + \vec{\sigma}_{A2} \right) \left\{ \vec{H} + \vec{\sigma}_{A2} \right\}_\eta^2 \right] +$$

$$\left[V_{222} \left\{ \vec{H} - \vec{\sigma}_{A3} \right\}_\eta^2 + V_{422} \left(\vec{H} - \vec{\sigma}_{A3} \right) \bullet \left(\vec{H} - \vec{\sigma}_{A3} \right) \left\{ \vec{H} - \vec{\sigma}_{A3} \right\}_\eta^2 \right] +$$

$$\left[-V_{222} \left\{ \vec{H} + \vec{\sigma}_{A3} \right\}_\eta^2 - V_{422} \left(\vec{H} + \vec{\sigma}_{A3} \right) \bullet \left(\vec{H} + \vec{\sigma}_{A3} \right) \left\{ \vec{H} + \vec{\sigma}_{A3} \right\}_\eta^2 \right] \quad . \tag{7.26}$$

This equation has only 2 unknown expansion coefficients V_{222} and V_{422}. These 2 coefficients are related to the strength (a.k.a. amplitude or amount) of the Zernike polynomial that is added to the base sphere surface. An explicit equation for coefficients V_{222} and V_{422} in terms of the Zernike polynomial strength (and other system and surface specific parameter values) would be very useful, allowing for the determination of the coefficients without the need for additional ray tracing and fitting. The determination of coefficients V_{222} and V_{422} may be incorporated into the determination of the W_{klm} expansion coefficients since both, in this work, require real ray tracing data and some form of data fitting. Alternatively, the paper by Fuerschbach, et al., [72] provides an alternative way to calculate the coefficients. This then allows for the quantitative prediction of the change in the aberration function's field dependence when one or more Zernike polynomials are added to the shape of an optical element. Even without knowing the values of the V_{222} and V_{422} coefficients, the form of the contribution to the field dependence due to an additional Zernike cap has here been determined.

When the surface to which the Zernike polynomial is to be added is not a spherical shape, i.e. aspheric shape, then the field dependence for the $\vec{Z}_n^m(\vec{\rho})$ aberration expansion term would be written as

$$\vec{F}_n^m = \left(\left[Sphere\ Base\right] + \left[Aspheric\ Cap\right]\right)_{original\ surface} + \left[Zernike\right], \qquad (7.27)$$

where the [*Sphere Base*] + [*Aspheric Cap*] components are treated as in previous Chapters and the [*Zernike*] component is treated as described in this section, adapted for the specific field function of the $\vec{Z}_n^m(\vec{\rho})$ aberration expansion term of interest. It is to be understood that the [*Zernike*] component does not include the rotationally symmetric Zernike polynomials. The rotationally symmetric Zernike polynomials are included in the [*Aspheric Cap*] as discussed during the development of Eq. (7.12) and Eq. (7.13).

It may be the case that the change in the field dependence due to the added Zernike polynomial(s) is not well reproduced by the through 6th order of the total wavefront aberration limitation imposed in the above example (the field function of the $\vec{Z}_2^2(\vec{\rho})$ term). In that case, the next higher order field dependent term can be included into the field function just for the Zernike surface shape contributions. This would, for example, introduce a third expansion coefficient, "V_{622}", into the above example.

When the freeform optical element as a whole is decentered (meaning the base sphere, aspheric terms, and the additional non-rotationally symmetric Zernike polynomial terms are decentered as a single element) the field dependence of the field function is shifted, as prescribed by NAT, by the sigma offset vectors. The sphere sigma offset vector, $\vec{\sigma}^S$, is used for the base sphere's field terms and the aspheric cap sigma offset

vector, $\vec{\sigma}^A$, is used for the aspheric caps and for the decentered term's field contributions associated with the Zernike caps written as decentered surface shape terms. Continuing with the example above, $Z_3^{+3}(\vec{\rho})$ added to a base sphere shaped element, this gives

$$
\begin{aligned}
\vec{F}_2^2 =& \left[V_1^S \left\{ \vec{H} - \vec{\sigma}^S \right\}_\eta^2 + V_2^S H^2 \left\{ \vec{H} - \vec{\sigma}^S \right\}_\eta^2 \right] + \\
& \left[V_{222} \left\{ \vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A1} \right\}_\eta^2 + V_{422} \left(\vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A1} \right) \bullet \left(\vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A1} \right) \left\{ \vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A1} \right\}_\eta^2 \right] + \\
& \left[-V_{222} \left\{ \vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A1} \right\}_\eta^2 - V_{422} \left(\vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A1} \right) \bullet \left(\vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A1} \right) \left\{ \vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A1} \right\}_\eta^2 \right] + \\
& \left[V_{222} \left\{ \vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A2} \right\}_\eta^2 + V_{422} \left(\vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A2} \right) \bullet \left(\vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A2} \right) \left\{ \vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A2} \right\}_\eta^2 \right] + \\
& \left[-V_{222} \left\{ \vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A2} \right\}_\eta^2 - V_{422} \left(\vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A2} \right) \bullet \left(\vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A2} \right) \left\{ \vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A2} \right\}_\eta^2 \right] + \\
& \left[V_{222} \left\{ \vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A3} \right\}_\eta^2 + V_{422} \left(\vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A3} \right) \bullet \left(\vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A3} \right) \left\{ \vec{H} - \vec{\sigma}^A - \vec{\sigma}_{A3} \right\}_\eta^2 \right] + \\
& \left[-V_{222} \left\{ \vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A3} \right\}_\eta^2 - V_{422} \left(\vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A3} \right) \bullet \left(\vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A3} \right) \left\{ \vec{H} - \vec{\sigma}^A + \vec{\sigma}_{A3} \right\}_\eta^2 \right] \quad .(7.28)
\end{aligned}
$$

Note that in writing Eq. (7.28) a restriction was implicitly made that the surface to which the Zernike cap is being added to is the first optical surface in the optical system. This was done to keep the example simple at this point. That is, all the extrinsic contributions are zero. Including the extrinsic contributions is mentioned later in this section.

These developments should, at present, be considered speculative rather than rigorously derived. However, as the examples in the next section of this Chapter will show, the resulting FFDs using the above development match the FFDs obtained by real ray tracing to a high degree.

This example (adding a $Z_3^{+3}(\vec{\rho})$ surface shape to a base sphere shaped optical element) also illustrates a potential practical limitation to this method. Each additional

non-rotationally symmetric Zernike polynomial used to define the optical surface will

add additional decentered terms and new expansion coefficients. The number of terms

could become overwhelming quickly for surfaces requiring several Zernike polynomials

for their shape description. However, by using GA of Chapter 2, and for this specific

case of adding a $Z_3^{+3}(\vec{\rho})$ cap shape to an optical surface, it can be shown (see Appendix

VI) that Eq. (7.26) greatly simplifies because

$$
\begin{aligned}
\Big[V_{222}\left\{\vec{H}-\vec{\sigma}_{A1}\right\}_{\eta}^{2} - V_{222}\left\{\vec{H}+\vec{\sigma}_{A1}\right\}_{\eta}^{2} + V_{222}\left\{\vec{H}-\vec{\sigma}_{A2}\right\}_{\eta}^{2} - V_{222}\left\{\vec{H}+\vec{\sigma}_{A2}\right\}_{\eta}^{2} + \\
V_{222}\left\{\vec{H}-\vec{\sigma}_{A3}\right\}_{\eta}^{2} - V_{222}\left\{\vec{H}+\vec{\sigma}_{A3}\right\}_{\eta}^{2} \Big] = \vec{0},
\end{aligned}
\tag{7.29}
$$

and

$$
\begin{aligned}
\Big[V_{422}\left(\vec{H}-\vec{\sigma}_{A1}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A1}\right)\left\{\vec{H}-\vec{\sigma}_{A1}\right\}_{\eta}^{2} - V_{422}\left(\vec{H}+\vec{\sigma}_{A1}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A1}\right)\left\{\vec{H}+\vec{\sigma}_{A1}\right\}_{\eta}^{2} + \\
V_{422}\left(\vec{H}-\vec{\sigma}_{A2}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A2}\right)\left\{\vec{H}-\vec{\sigma}_{A2}\right\}_{\eta}^{2} - V_{422}\left(\vec{H}+\vec{\sigma}_{A2}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A2}\right)\left\{\vec{H}+\vec{\sigma}_{A2}\right\}_{\eta}^{2} + \\
V_{422}\left(\vec{H}-\vec{\sigma}_{A3}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A3}\right)\left\{\vec{H}-\vec{\sigma}_{A3}\right\}_{\eta}^{2} - V_{422}\left(\vec{H}+\vec{\sigma}_{A3}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A3}\right)\left\{\vec{H}+\vec{\sigma}_{A3}\right\}_{\eta}^{2} \Big] = 6V_{422}\vec{H}^{*},
\end{aligned}
\tag{7.30}
$$

where the following notation is used

$$
\vec{H}^{*} \equiv \left(H_{x}, -H_{y}\right) .
\tag{7.31}
$$

That is, Eq. (7.26) reduces to

$$
\vec{F}_{2}^{2} = \left[V_{1}^{S}\left\{\vec{H}\right\}_{\eta}^{2} + V_{2}^{S} H^{2}\left\{\vec{H}\right\}_{\eta}^{2} \right] + 6V_{422}\vec{H}^{*}
\tag{7.32}
$$

This field dependence due to the addition of a $Z_3^{+3}(\vec{\rho})$ cap shape onto an optical element

has been named "field linear, field conjugate astigmatism" in the literature [72]. In the

2D vector plane, a vector of the form as given in Eq. (7.31) is simply a reflection of the

vector $\vec{H} = (H_x, H_y)$ about the *x*-axis, accomplished using GA, as shown in Chapter 2, by

$$\vec{H}^* = \hat{e}_1 \vec{H} \hat{e}_1 \ . \tag{7.33}$$

In a similar way, the field contribution to the Zernike astigmatism term's field dependence for other Zernike cap shapes can be calculated. A small selection is presented in Table 7.1.

Table 7.1 Field contributions to Zernike astigmatism term due to Zernike cap.

Zernike Cap $Z_n^{\pm m}(\vec{\rho})$	Vector Field Contribution To Zernike Astigmatism Term	Shorthand Notation
$Z_2^{+2}(\vec{\rho})$	$4(V_{222}+V_{422},0)+12V_{422}(H_x^2+H_y^2,0)$	$\vec{V}_{const}+12V_{422}\vec{H}\star_x\vec{H}^*$
$Z_2^{-2}(\vec{\rho})$	$4(0,V_{222}+V_{422})+12V_{422}(0,H_x^2+H_y^2)$	$\vec{V}_{const}+12V_{422}\ \vec{H}\star_x\vec{H}^*$
$Z_3^{+3}(\vec{\rho})$	$6V_{422}(H_x,-H_y)$	$6V_{422}\vec{H}^* = 6V_{422}\left(\left\{\vec{H}\right\}_x^1\right)^*$
$Z_3^{-3}(\vec{\rho})$	$6V_{422}(H_y,H_x)$	$6V_{422}\vec{H} = 6V_{422}\left\{\vec{H}\right\}_x^1$
$Z_4^{+4}(\vec{\rho})$	$8V_{622}(H_x^2-H_y^2,-2H_xH_y)$	$8V_{622}\left(\left\{\vec{H}\right\}_x^2\right)^*$
$Z_4^{-4}(\vec{\rho})$	$8V_{622}(2H_xH_y,H_x^2-H_y^2)$	$8V_{622}\left\{\vec{H}\right\}_x^2$

In Table 7.1, the notation $\vec{H} \equiv (H_y, H_x)$ has been used. Note that for $Z_4^{+4}(\vec{\rho})$ and $Z_4^{-4}(\vec{\rho})$, the restriction of using through 6$^{\text{th}}$ order expansion needed to be removed to obtain a non-zero field contribution. The through 8$^{\text{th}}$ order Zernike astigmatism

expansion term has been used for $Z_4^{+4}(\vec{\rho})$ and $Z_4^{-4}(\vec{\rho})$, hence the appearance of the expansion coefficient "V_{622}" mentioned previously.

Note that the coefficients in each row of Table 7.1 are independent of any coefficients in any of the other rows in the Table. That is "V_{422}" in the $Z_3^{+3}(\vec{\rho})$ row is not the same as the "V_{422}" coefficient in the $Z_3^{-3}(\vec{\rho})$ row. Therefore, if both $Z_3^{+3}(\vec{\rho})$ and $Z_3^{-3}(\vec{\rho})$ Zernike caps are to be added to the shape of an optical element, then the resulting additional field contribution should be written using an apostrophe on one of the "V_{442}" coefficients. For example, this could be written as

$$\vec{F}_{Z.Cap} = 6V_{422}\vec{H}^* + 6V'_{422}\vec{H} \; , \tag{7.34}$$

where V_{422} would not necessarily have the same value as V'_{422}. Of course, any other notational device could be employed to make it clear that the coefficients in one row of Table 7.1 are not the same as the coefficients in any other row.

By defining the vector

$$\vec{C} \equiv \left(6V_{422}, \; 6V'_{422} \right) \; , \tag{7.35}$$

it can be shown that (see Appendix VI)

$$\vec{F}_{Z.Cap} = 6V_{422}\vec{H}^* + 6V'_{422}\vec{H} = \vec{C} \star_x \vec{H}^* \; . \tag{7.36}$$

This is the form of the field linear, field conjugate astigmatism field contribution presented in [72]. However, in [72] the y-axis is used as the reference axis for SVP whereas Eq. (7.36) is written using the x-axis as the reference axis.

The approach being described in this Chapter may not be limited to using Zernike polynomials for the description of the freeform surface shape. It may be that any rotationally symmetric shape could be utilized as the surface shape terms added to the sphere base shape of the optical element. However, at this point, this is only speculation and left for future research to explore.

It may be the case that there is an extrinsic aberration contribution to be considered due to the Zernike cap added to an otherwise rotationally symmetric optical element. This may occur when the optical element to which the Zernike cap is to be added is not the first optical element in the imaging system. Following the NAT treatment of an aspheric cap leads to the possibility that the Zernike cap aberration contribution may split into an intrinsic and an extrinsic contribution. Then the application of NAT's sphere and aspheric sigma offset vectors may be grouped as

$$\vec{F}_n^m = \left([Sphere\ Base]^{Intrinsic} + [Sphere\ Base]^{Extrinsic} + [Aspheric\ Cap]_{Original}^{Extrinsic} + [Zernike]^{Extrinsic}\right)^{Sphere\ Sigma} +$$
$$\left([Aspheric\ Cap]_{Original}^{Intrinsic} + [Zernike]^{Intrinsic}\right)^{Asphere\ Sigma}. \tag{7.37}$$
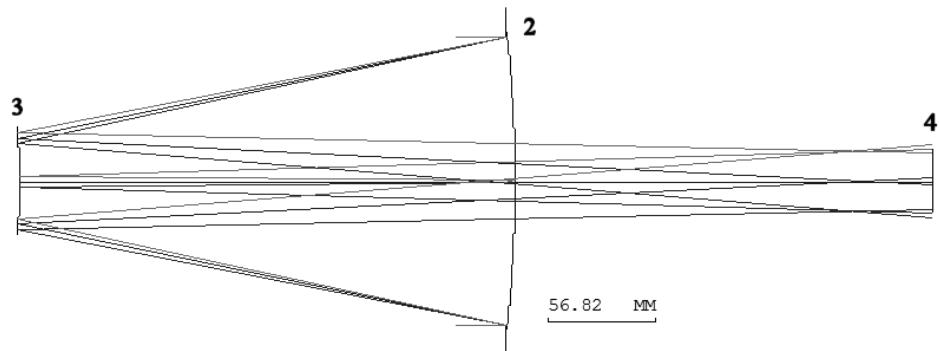
It is to be understood that the "$[Zernike]$" mentioned in Eq. (7.37) excludes any of the rotationally symmetric Zernike polynomials because the rotationally symmetric Zernike polynomials are to be considered part of the surface's aspheric cap. See the discussion associated with the development of Eq. (7.12) and Eq. (7.13) for a freeform optical surface. The implementation of Eq. (7.37) would require repeating the terms associated with the Zernike cap in the field function and additional coefficients for the $[Zernike]^{Extrinsic}$ contributions that are obtained during the fit for the other Zernike-as-decentered-terms expansion coefficients.

7.2 Examples of NAT and Zernike Defined Freeform Optical Surfaces

In this section, several examples introducing a new way of applying NAT to optical imaging systems having an optical element with a Zernike freeform surface definition are provided. It is suggested that this may be a fruitful way to determine the form of the field functions for freeform optical elements.

Two Mirror Optical System with a Zernike Primary

Figure 7.4 from CODE V® (modified) shows the configuration of the two mirror telescope model used. The model is initially rotationally symmetric. A Fringe Zernike polynomial $Z_{10}(\vec{\rho}) = Z_3^{+3}(\vec{\rho})$ is added to the spherical primary mirror, surface #2. The Zernike coefficient is 0.075 *mm* and the Zernike normalization radius is 75 *mm*.
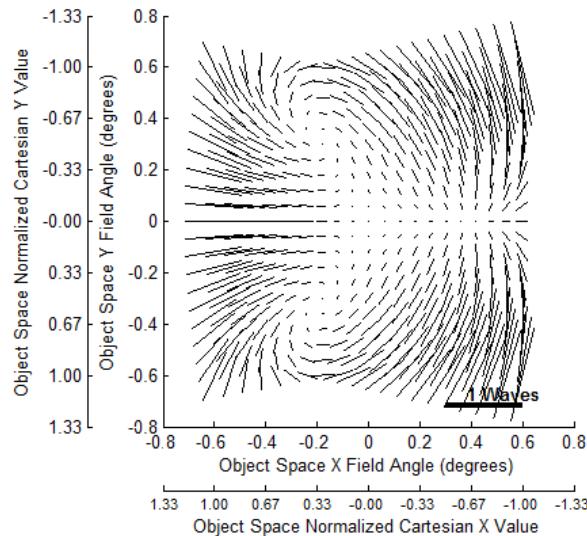
Figure 7.4 Two mirror system with Fringe Zernike surface on primary mirror. Surface 2 is the primary mirror having a base sphere shape plus a Zernike polynomial cap, surface 3 is the secondary mirror, and surface 4 is the image plane.

The entrance pupil has a diameter of 150 *mm*. The HFOV is 0.6 degrees. The secondary

mirror is a conic with a conic constant -2.915. The wavelength is 632.8 *nm*. Table 7.2

produced by CODE V® presents the other system parameters.

Table 7.2 Two mirror optical system's parameters.

Surface #	Surface Name	Surface Type	Y Radius	Thickness	Refract Mode	Y Semi-Aperture
Object		Sphere	Infinity	Infinity	Refract	
Stop		Sphere	Infinity	5.0000	Refract	75.0000
2		Fringe Zer	-742.8572	-260.0000	Reflect	75.0134
3		Conic	-290.2328	479.9995	Reflect	26.5895
Image		Sphere	Infinity	0.0000	Refract	28.0919

The resulting Zernike astigmatism, $\left(Z_2^{+2}(\vec{\rho}), Z_2^{-2}(\vec{\rho}) \right)$, FFD is presented in Figure

7.5.

Figure 7.5 FFD of Zernike astigmatism term created by using real ray tracing, GQ, and double Zernike equation Eq. (4.53).

Removing the Zernike cap from the primary mirror, the W_{klm} coefficients were

calculated for the rotationally symmetric optical model. Those associated with the

Zernike astigmatism term through 6th order are given in Table 7.3 showing that only the

4th order coefficient values are significant for this model.

Table 7.3 Zernike astigmatism related W_{klm} coefficients through 6th order. (Units = waves.)

Surface Number	W_{klm}	Intrinsic		Extrinsic	
		Sphere	Asphere	Sphere	Asphere
2	W_{222}	1.2947	0.0	0.0	0.0
2	W_{242}	−0.0260	0.0	0.0	0.0
2	W_{422}	−0.0002	0.0	0.0	0.0
3	W_{222}	−1.0770	0.7154	0.0	0.0
3	W_{242}	0.0595	−0.0034	−0.0364	0.0023
3	W_{422}	−0.0001	0.0	0.0	0.0

The normalized $\vec{\sigma}_{Ai}$ decenter vectors, introduced in the preceding section, for

$Z_3^{+3}(\vec{\rho})$ using the *x*-axis as the reference axis, are obtained from Table IV.5 in Appendix

IV and are

$$\vec{\sigma}_{A1} = (-1,0) \ , \tag{7.38}$$

$$\vec{\sigma}_{A2} = \left(\frac{1}{2}, \frac{\sqrt{3}}{2} \right) , \tag{7.39}$$

$$\vec{\sigma}_{A3} = \left(\frac{1}{2}, \frac{-\sqrt{3}}{2} \right) . \tag{7.40}$$

Using the plot data that produced the FFD of Figure 7.5 and a least squares fit

with Eq. (7.26) for the field dependent vector function for a sphere surface having an

additional $Z_3^{+3}(\vec{\rho})$ cap shape provided, as explained in previous section, the six total

$V_{222(i)}$, $V_{422(i)}$ expansion coefficients. As can be seen in Table 7.4, there are actually only
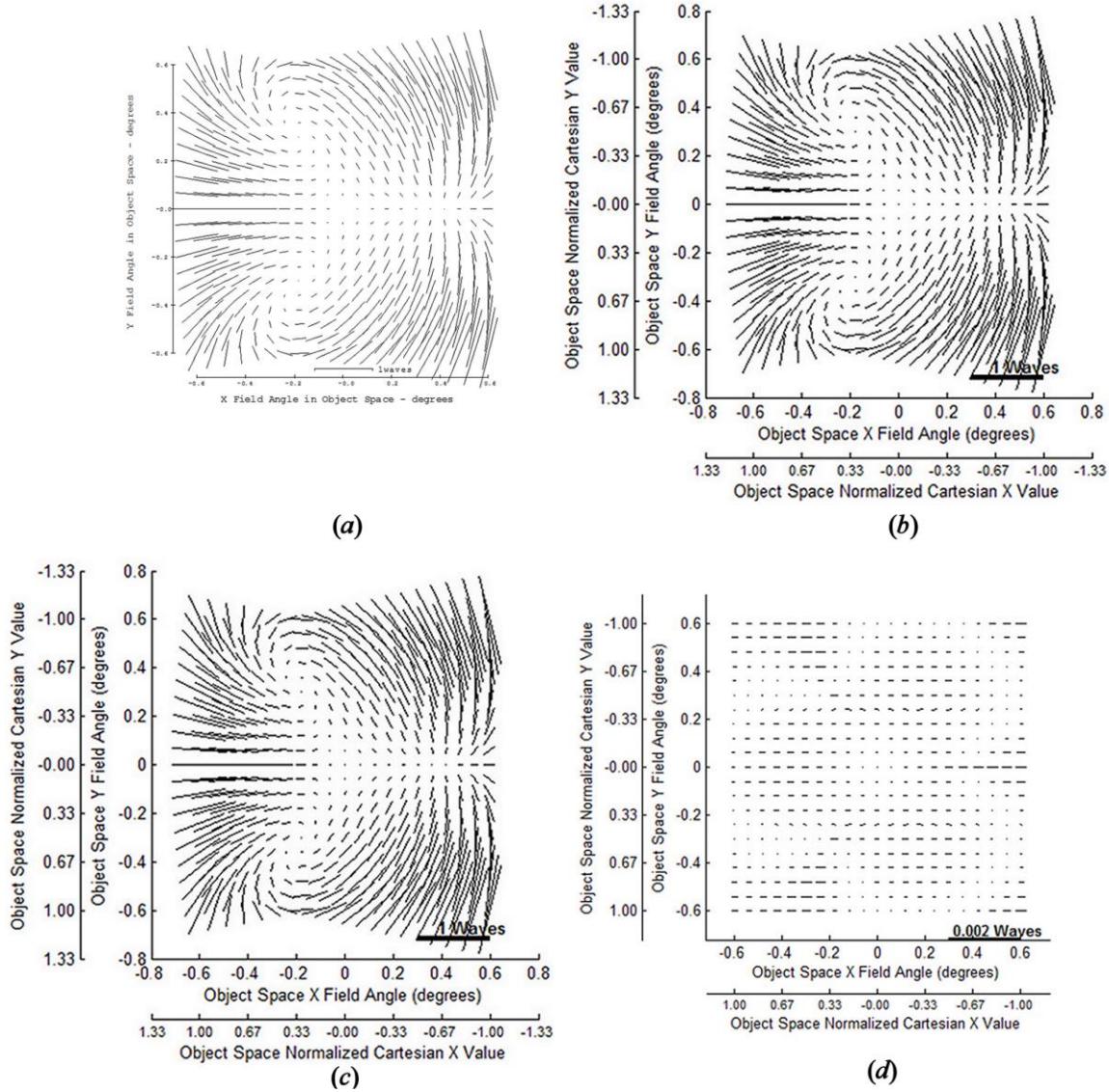
2 different coefficient values, providing some confirmation of the development presented

in the previous section.

Table 7.4 Least squares fit determined expansion coefficients. (Units = waves)

	$i = 1$	$i = 2$	$i = 3$
V_{222}	0.0197	0.0197	0.0197
V_{422}	0.0493	0.0493	0.0493

Using the field vector function Eq. (7.26) developed in the previous section of this

Chapter for the application of NAT to optical systems with Zernike defined freeform

optical surfaces, the FFD shown in Figure 7.6(c) was produced. Figure 7.6(a) is a FFD

plot generated by CODE V® version 10.5. Figure 7.6(b) is generated by using the

double Zernike expansion and the GQ method for determining the expansion coefficients

$U_{n_H,n_\rho}^{m_H,m_\rho}$.

Figure 7.6 Comparison of Zernike astigmatism FFDs for two mirror optical system having a Zernike freeform surface defined primary mirror. (*a*) Generated by CODE V®. (*b*) Repeat of Figure 7.5, (*c*) Plot calculated by Eq. (7.26) using W_{klm} and $\vec{\sigma}_{Ai}$ decenter vectors. (*d*) The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) – (*c*).

This plot, Figure 7.6(*c*), visually matches the FFD of Figure 7.5 (repeated as plot (*b*) in Figure 7.6). For a quantitative comparison, the difference between the two FFD's

data is also shown in Figure 7.6(*d*). The difference data's magnitude values have a maximum difference value of -0.003 waves, a mean of 0.00003 waves, and a STD of 0.0001 waves. The difference data's angle values have a maximum difference of -5.5 degrees, a mean of 0.01 degrees, and a STD of 0.27 degrees.
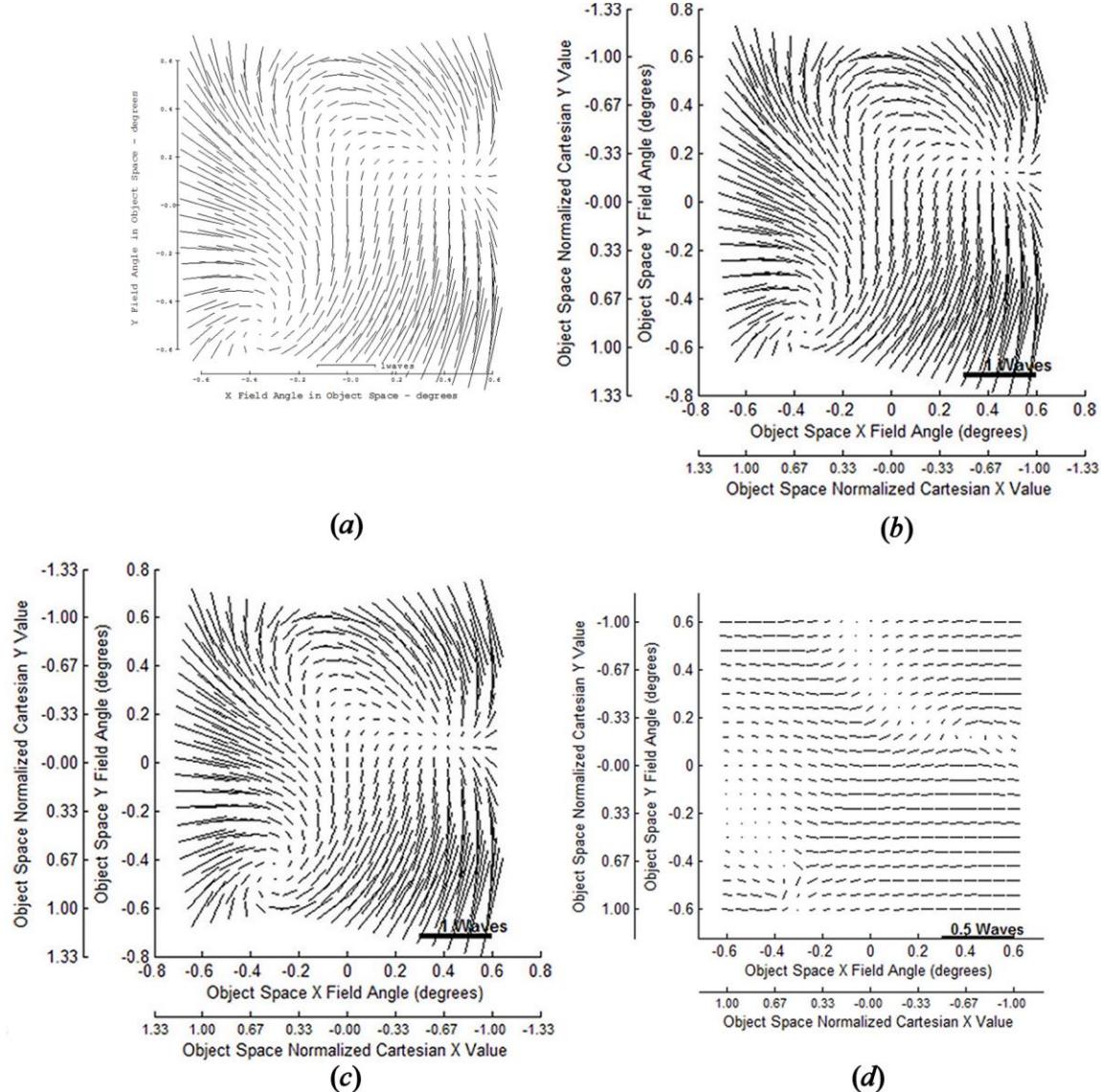
Two Mirror Optical System with a Zernike Primary and Decentered

The primary mirror of the previous optical model, having a $Z_3^{+3}(\vec{\rho})$ defined shape, was decentered as a whole optical element along the *y*-axis by an amount $+0.025$ *mm*. The NAT sigma offset vectors were calculated by Eq. (5.28) and Eq. (5.29) presented in Chapter 5 and their values are given in Table 7.5.

Table 7.5 Sigma offset vector values.

Surf.#	$\vec{\sigma}^{Sphere}$		$\vec{\sigma}^{Asphere}$	
	X	Y	X	Y
2	0.0	0.003235	0.0	−0.477447
3	0.0	0.006455	0.0	0.00639

With these sigma offset values, and Eq. (7.28), the FFD for the Zernike astigmatism term of the wavefront aberration function expansion was recreated and presented in Figure 7.7. Figure 7.7(*a*) is a FFD plot generated by CODE V® version 10.5. Figure 7.7(*b*) is generated by using the double Zernike expansion and the GQ method for determining the expansion coefficients $U_{n_H,n_\rho}^{m_H,m_\rho}$. Figure 7.7(c) is generated by using the W_{klm} expansion coefficients, the $\vec{\sigma}_{Ai}$ vectors of Eq. (7.38) through Eq. (7.40), the sigma offset vectors of Table 7.5, and the single Zernike expansion.

Figure 7.7   FFD comparison for decentered primary mirror case. (*a*) **Generated by CODE V®.** (*b*) **By using real ray tracing, GQ, and double Zernike equation Eq. (4.53).** (*c*) **By using Eq. (5.43), W_{klm} values, $\vec{\sigma}_{Ai}$ decenter and sigma offset vectors.** (*d*) **The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).**

The difference between the two FFD's data is also shown in Figure 7.7(*d*). The

difference data's magnitude values have a maximum difference value of 0.09 waves, a

mean of 0.04 waves, and a STD of 0.06 waves. The difference data's angle values have a
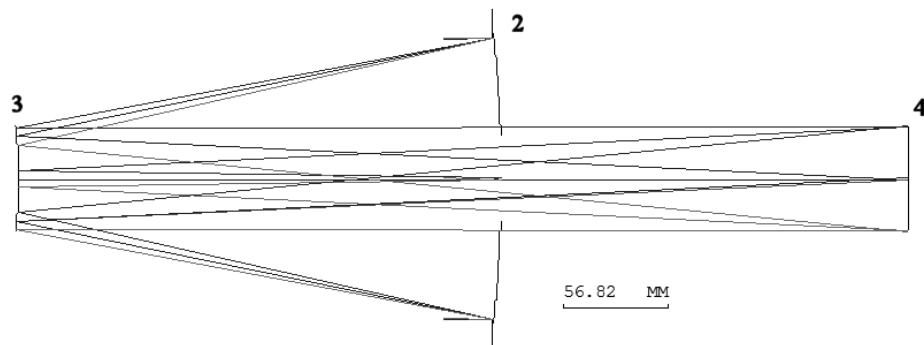
maximum difference of − 175.9 degrees, a mean of 1.3 degrees, and a STD of 18.9 degrees. The large maximum angle difference occurs near the node locations where the data values are very small.

Further research into the field contribution due to decentering of freeform surfaces in an optical imaging system is left for future research to address.

Two Mirror Optical System with a Zernike Defined Secondary

In this next example, a two mirror telescope model is used onto which an additional $Z_3^{+3}(\vec{\rho})$ cap is added to the secondary mirror. This example was investigated to determine if the success of the previous examples in which the Zernike cap was added to a spherical primary mirror could be repeated for the original aspheric secondary mirror.

For this model, the entrance pupil has a diameter of 150 mm. The HFOV is 1.0 degree. The primary mirror is a conic with a conic constant of -1.046192. The secondary mirror is a conic with a conic constant -2.915.

Figure 7.8 Two mirror telescope model with Zernike cap added to the secondary mirror. Surface 2 is the primary mirror having a conic shape, surface 3 is the secondary mirror having a conic shape plus a Zernike polynomial cap, and surface 4 is the image plane.

The Zernike strength coefficient is 0.00075 mm and the Zernike normalization radius is 27.1573 mm. The wavelength for the model is 632.8 nm. Table 7.6 produced by CODE V® presents the other system parameters.

Table 7.6 The two mirror optical system's parameters.

Surface #	Surface Name	Surface Type	Y Radius	Thickness	Refract Mode	Y Semi-Aperture
Object		Sphere	Infinity	Infinity	Refract	
Stop		Sphere	Infinity	5.0000	Refract	75.0000
2		Conic	-742.8572	-260.0000	Reflect	75.0000
3		Fringe Zer	-290.2328	479.9995	Reflect	27.1575
Image		Sphere	Infinity	0.0000	Refract	28.1628

The rotationally symmetric version of this optical imaging system (removal of the Zernike cap) was used to obtain the W_{klm} expansion coefficients. Table 7.7 provides the W_{klm} values though 6th order associated with the Zernike astigmatism $\vec{Z}_2^2(\vec{\rho})$ term.

Table 7.7 Zernike astigmatism related W_{klm} coefficients through 6th order. (Units = waves)

Surface Number	W_{klm}	Intrinsic		Extrinsic	
		Sphere	Asphere	Sphere	Asphere
2	W_{222}	3.5969	−0.0002	0.0	0.0
2	W_{242}	−0.0721	−0.0002	0.0	0.0
2	W_{422}	−0.0016	0.0	0.0	0.0
3	W_{222}	−2.9922	1.9874	0.0	0.0
3	W_{242}	0.1652	−0.0095	−0.1007	0.0060
3	W_{422}	−0.0008	−0.0001	0.0	0.0

Since the Zernike cap is added to the secondary mirror, which is a conic, it is reasoned, as detailed in the previous section (see Eq. (7.12) and Eq. (7.13)), that the field dependent vector functions of the terms in the Zernike (in pupil) expansion of the wavefront aberration function will have the following form

$$\vec{F}_n^m = \left[Sphere\ Base\right]_{j=2} + \left[Asphere\ Cap\right]_{j=2}$$
$$\left[Sphere\ Base\right]_{j=3} + \left[Asphere\ Cap\right]_{j=3} + \left[Zernike\right]_{j=3} , \qquad (7.41)$$

where $j = 2$ signifies the primary mirror contribution and $j = 3$ signifies the secondary mirror contribution. The "[$Zernike$]" refers to only non-rotationally symmetric Zernike polynomials as described previously with the development of Eq. (7.12) and Eq. (7.13) for the freeform optical surface shape. The $\left[Zernike\right]_{j=3}$ contribution to the field dependence for the Zernike astigmatism $\vec{Z}_2^2(\vec{\rho})$ term, through 6$^{\text{th}}$ order, is given by (see Eq. (7.22) in the previous section of this Chapter)

$$\begin{aligned}
\left[Zernike\right]_{j=3} = &\left[V_{222(1)}\left\{\vec{H}-\vec{\sigma}_{A1}\right\}_\eta^2 + V_{422(1)}\left(\vec{H}-\vec{\sigma}_{A1}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A1}\right)\left\{\vec{H}-\vec{\sigma}_{A1}\right\}_\eta^2 \right. \\
&- V_{222(1)}\left\{\vec{H}+\vec{\sigma}_{A1}\right\}_\eta^2 - V_{422(1)}\left(\vec{H}+\vec{\sigma}_{A1}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A1}\right)\left\{\vec{H}+\vec{\sigma}_{A1}\right\}_\eta^2 + \\
&+ V_{222(2)}\left\{\vec{H}-\vec{\sigma}_{A2}\right\}_\eta^2 + V_{422(2)}\left(\vec{H}-\vec{\sigma}_{A2}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A2}\right)\left\{\vec{H}-\vec{\sigma}_{A2}\right\}_\eta^2 \\
&- V_{222(2)}\left\{\vec{H}+\vec{\sigma}_{A2}\right\}_\eta^2 - V_{422(2)}\left(\vec{H}+\vec{\sigma}_{A2}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A2}\right)\left\{\vec{H}+\vec{\sigma}_{A2}\right\}_\eta^2 \\
&+ V_{222(3)}\left\{\vec{H}-\vec{\sigma}_{A3}\right\}_\eta^2 + V_{422(3)}\left(\vec{H}-\vec{\sigma}_{A3}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A3}\right)\left\{\vec{H}-\vec{\sigma}_{A3}\right\}_\eta^2 \\
&\left. - V_{222(3)}\left\{\vec{H}+\vec{\sigma}_{A3}\right\}_\eta^2 - V_{422(3)}\left(\vec{H}+\vec{\sigma}_{A3}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A3}\right)\left\{\vec{H}+\vec{\sigma}_{A3}\right\}_\eta^2 \right] . \qquad (7.42)
\end{aligned}$$

The FFD data for the Zernike astigmatism term was calculated by using the double Zernike aberration function expansion and the $U_{n_H,n_\rho}^{m_H,m_\rho}$ expansion coefficients. Using this FFD data, and the proposed field function based on Eq. (7.41) and Eq. (7.42), as well as the W_{klm} values of Table 7.7 and the $\vec{\sigma}_{Ai}$ decenter vector values of Eq. (7.38) through Eq. (7.40), a least squares fit, in 6 variables, was performed to obtain the V_{222} and V_{422} expansion coefficients shown in Table 7.8.

Table 7.8 Least squares fit determined expansion coefficients. (Units = waves)

	1	2	3
V_{222}	−0.0055	−0.0055	−0.0055
V_{422}	−0.1373	−0.1373	−0.1373

The obtained V_{222} and V_{422} expansion coefficient values again confirmed the hypothesis that there are actually only 2 coefficients for the additional field dependence due to the addition of the Zernike cap.

Figure 7.9 provides a comparison of the Zernike astigmatism term of the wavefront aberration function expansions. In Figure 7.9, plot (*a*) is a FFD plot generated by CODE V® version 10.5, plot (*b*) was generated using the double Zernike expansion of the wavefront aberration function and the associated $U_{n_H,n_\rho}^{m_H,m_\rho}$ expansion coefficients, while plot (*c*) was generated by using the single Zernike (in pupil) expansion of the wavefront aberration function, the W_{klm} expansion coefficients, the V_{222} and V_{422} expansion coefficients, and the $\vec{\sigma}_{Ai}$ decenter vectors. The difference between the two FFD's data is shown in Figure 7.9(*d*). The difference data's magnitude values have a maximum difference value of 0.03 waves, a mean of − 0.0003 waves, and a STD of 0.016 waves. The difference data's angle values have a maximum difference of − 29.2 degrees, a mean of − 0.07 degrees, and a STD of 2.4 degrees.

(*a*)

(*b*)

(*c*)

(*d*)

Figure 7.9 Comparison of Zernike astigmatism FFDs for two mirror optical system having a Zernike freeform surface defined secondary mirror. (*a*) Generated by CODE V®. (*b*) By using real ray tracing, GQ, and double Zernike equation Eq. (4.53). (*c*) Plot calculated using W_{klm} , V_{222} & V_{422} coefficients, and $\vec{\sigma}_{Ai}$ decenter vectors. (*d*) The difference between the Zernike astigmatism FFD data generated by the double Zernike expansion and the single Zernike expansion using the calculated W_{klm} values, (*b*) − (*c*).

This lends support to the approach outlined in the previous section for using NAT to develop the field dependent contributions due to the addition of Zernike caps to the optical elements' shapes, defining freeform optical elements, has merit. However, as mentioned before, these are only the initial, tentative steps to a complete theory of the aberration function's field dependence for optical imaging systems defined with freeform optical elements.

An alternative approach, based on utilizing stop shift equations, has been presented in the literature [72]. Future research may combine the two approaches and may provide a theory having analytic capabilities to provide the additional field contributions to the wavefront aberration function expansion terms due to the additional of Zernike caps to the base surface shape to define freeform shaped optical elements.

Chapter 8   Conclusion and Future Research

This work has demonstrated how to form a connection between three forms of the wavefront aberration function expansion for rotationally symmetric optical imaging systems: 1) the extended H. H. Hopkins' form, 2) the single Zernike in pupil form, and 3) the double Zernike, in field and pupil, form. Optical designers typically use some form the Hopkins' expansion involving the W_{klm} expansion coefficients, while for testing and measurement the Zernike in pupil form is typically used. A common feature of these expansions, as developed in this research, for rotationally symmetric optical imaging systems, is the use of SVP.

In this work, SVP has been developed for the first time in terms of GA. As shown, SVP fits naturally into GA and avoids the use of the exponential complex number form for the representation of 2D vectors. An important result of using the GA approach is that it explicitly shows how SVP involves three 2D vectors, not just two vectors, thus clarifying the nature of SVP. New notation for SVP has been developed that explicitly captures this important fact.

The various definitions of the real number Zernike polynomials that occur in the literature was briefly reviewed. An Appendix provided the derivation of the orthogonality property for the discrete sampling of the real Zernike polynomials over the unit radius disk. This property has been utilized in this work, together with the double Zernike expansion of the wavefront aberration function, to obtain the double Zernike's $U_{n_H,n_\rho}^{m_H,m_\rho}$ expansion coefficients. New equations have been developed for calculating the

W_{klm} expansion coefficients from the $U_{n_H,n_\rho}^{m_H,m_\rho}$ coefficients for rotationally symmetric optical imaging systems. It has further been demonstrated how the per surface, sphere/asphere, and intrinsic/extrinsic W_{klm} coefficients can be obtained.

A new equation for the expansion of the wavefront aberration function using the medial focal surface based expansion coefficients has been developed. This equation explicitly provides the separation of the field and pupil parameters, to any order, into a simple form similar to that of the traditional vector form developed by Shack for which the field and pupil parameters are not separated (are not factored by a vector dot product). Once again, it is seen that SVP plays a crucial role in accomplishing this result.

A brief introduction to NAT was presented. The meaning of the full field displays (FFDs), the basic tools for the qualitative and quantitative comparisons performed in this work, has been discussed. The sigma offset vectors were introduced. Their meaning, equations for the calculation of their values, and method for introducing them into the field functions has been clarified. It was shown how the equations traditionally presented in the literature for the sigma offset vectors can be modified to clarify the issue involving their definition in the surface's local image space and their use in the wavefront aberration function expansion defined using either the system's object space field parameters or the total system's image space field parameters.

Validation of the mathematical development has been provided by using two and three mirror telescope optical models. The models were perturbed and the resulting FFDs were generated using the W_{klm} expansion coefficients and sigma offset vectors of nodal aberration theory. Comparison (both qualitative and quantitative) to the FFDs based on

real ray trace data and a double Zernike expansion of the wavefront aberration function through 12th order has been presented. In most cases, the two different methods for generating the FFDs produced very good matches. However, for the case of the coma free pivot model and for the case of the aperture offset model, the developed equations did not produce good matches to the double Zernike generated FFDs. This indicates that there is something new and interesting still to learn in this field.

A proposed method for using NAT to obtain the contribution to the field functions due to freeform optical elements has been presented. A property of Zernike polynomials, discovered while performing this research and presented in Appendix IV, that they can be mathematically represented as a sum of decentered rotationally symmetric terms, was used to obtain the contribution to the field functions due to the addition of a Zernike cap to an otherwise rotationally symmetric optical surface. Explicit examples have been provided.

There are many avenues that could be followed in future work. Of particular necessity for any future research is the need for a more practical way to obtain the W_{klm} expansion coefficients, including their per surface, sphere/asphere, and intrinsic/extrinsic values. The current method developed in this research works well but is far too slow and involves too many steps for any interactive optical design work.

Additionally, different definitions for the placement of the reference sphere should be implemented and results compared. For example, using the real exit pupil location versus the Gaussian exit pupil location, and using the real chief ray intersection point in the image plane as the center of ther reference sphere versus the Gaussian image

point. These should be investigated to determine if different definitions of the reference sphere make a difference to the higher then 4th order W_{klm} coefficients, particularly for the extrinsic coefficient values.

This research has proposed one way to apply NAT to freeform optical elements. The initial success of this approach for both non-decentered freeform surfaces and decentered freeform surfaces, as well as when the freeform surface is not the first surface in the system, is very encouraging. The development also suggests that the approach may also be suitable for use with other than Zernike definitions for freeform optical surfaces. Specifically, the sum of the decentered rotationally symmetric terms need not define a single Zernike polynomial. It may be possible to use the method with other rotationally symmetric shaped "caps" for the definition of the freeform surface shapes and for determination of their field contributions. Further testing and evaluation is needed to develop these ideas into a reliable approach for the determination of the field contribution of freeform optical elements to the wavefront aberration function.

Bibliography

1. Gray, R. W., Dunn, C., Thompson, K. P., and Rolland, J. R. *Opt. Express* **20**, 16436–16449 (2012).
2. Gray, R. W. and Rolland, J. P. *J. Opt. Soc. Am. A* **32**(10), 1836–1847 (2015).
3. Hopkins, H. H. *Wave Theory of Aberrations*. Clarendon Press, (1950).
4. Shack, R. V. and Thompson, K. P. *Proc. SPIE* **251**, 146–153 (1980).
5. Thompson, K. P. *J. Opt. Soc. Am. A* **22**(7), 1389–1401 (2005).
6. Needham, T. *Visual Complex Analysis*. Oxford University Press, Inc., (2002).
7. Hestenes, D. *New Foundations for Classical Mechanics*. Kluwer Academic Publishers, 2nd edition, (1999).
8. Macdonald, A. *Linear and Geometric Algebra*. CreateSpace Independent Publishing Platform, (2011).
9. Dorst, L., Fontijne, D., and Mann, S. *Geometric Algebra for Computer Science*. Elsvier, (2007).
10. Doran, C. and Lasenby, A. *Geometric Algebra for Physicists*. Cambridge University Press, (2007).
11. Born, M. and Wolf, E. *Principles of Optics*. Cambridge University Press, 7th edition, (2003).
12. Mahajan, V. N. *Optical Shop Testing*, chapter "Zernike Polynomials and Wavefront Fitting", 498–546. John Wiley & Sons (2007).
13. Wyant, J. C. and Creath, K. *Applied Optics and Optical Engineering*, volume 11, chapter Basic Wavefront Aberration Theory for Optical Metrology, 1–39. Academic Press, Inc. (1992).
14. McLeod, B. A. *PASP* **108**, 217–219 (1996).
15. Rakich, A. *SPIE-OSA* **7652** (2010).
16. Burch, C. R. *Mon. Not. R. Astron. Soc.* **103**, 159–165 (1942).
17. Noethe, L. and Guisard, S. In *Proc. SPIE*, volume 4003, (2000).
18. Matsuyama, T. and Ujike, T. *Opt. Rev.* **11**, 199–207 (2004).
19. Kim, S., Yang, H.-S., Lee, Y.-W., and Kim, S.-W. *Opt. Express* **15**(8), 5059–5068 (2007).
20. Lee, H., Dalton, G. B., Tosh, I. A. J., and Kim, S.-W. *Opt. Express* **15**(6), 3127–3138 (2007).
21. Schechter, P. L. and Levinson, R. S. *PASP* **123**(905), 812–832 (2011).
22. Hanlon, J. and Ziok, H. Technical Report LA-UR-97-721, Los Alamos National Lab, (1997).
23. Crowe, M. J. *A History of Vector Analysis: The Evolution of the Idea of a Vectorial System*. Dover Publications, (1994).

24. http://en.wikipedia.org/wiki/Gerolamo_Cardano, Accessed September (2014).
25. http://en.wikipedia.org/wiki/Cauchy, Accessed September (2014).
26. Hamilton, S. W. R. *Proceedings of the Royal Irish Academy* **2**, 424–434 (1844).
27. Kuipers, J. B. *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*. Princeton University Press, (2002).
28. Grassmann, H. *Extension Theory*. American Mathematical Society, London Mathematical Society, (2000).
29. Hitzer, E. M. S. *International Symposium 2003 of Advanced Mechanical Engineering* , 22–25 (2003). Pukyong National University, Busan, Korea.
30. Hestenes, D. *Space-Time Algebra*. Gordon and Breach, (1966).
31. Irving, R. *Beyond the Quadratic Formula*. The Mathematical Association of America, (2013).
32. Zernike, F. *Physica* **1**, 689–704 (1934).
33. Dai. *J. Opt. Soc. Am. A* **23**(7), 1657–1668 (2006).
34. Noll, R. J. *J. Opt. Soc. Am.* **66**(3), 207–211 (1976).
35. Unknown. Technical Report ISO 10110-5:1996(E), (1996).
36. Chapra, S. C. *Applied Numerical Methods with Matlab for Engineers and Scientists*. The McGraw-Hill Companies, Inc., 3rd edition, (2012).
37. Kaya, I. *Mathematical and Computational Methods for Freeform Optical Shape Description*. PhD thesis, University of Central Florida, (2013).
38. Forbes, G. W. *Optics Express* **18**(13), 13851–13862 (2010).
39. Ruoff, J. and Totzeck, M. *J. Micro/Nanolith MEMS MOEMS* **8**(3) (2009).
40. Gavrielides, A. *Opt. Lett.* **7**(11), 526–528 (1982).
41. Zhao, C. and Burge, J. H. *Opt. Express* **15**(26), 18014–18024 (2007).
42. Zhao, C. and Burge, J. H. *Opt. Express* **16**(9), 6586–6591 (2008).
43. Pap, M. and Schipp, F. *Mathematica Pannonica* **16**(1), 137–144 (2005).
44. Fazekas, Z., Soumelidis, A., and Schipp, F. *Proc. World Cong. Eng.* **1** (2009).
45. Soumelidis, A., Fazekas, Z., Pap, M., and Schipp, F. *Electronic Engineering and Computing Technology*, chapter Discrete Orthogonality of Zernike Functions and Its Application to Corneal Measurements, 455–469. Springer (2010).
46. Mahajan, V. N. and Dai, G.-m. *J. Opt. Soc. Am. A* **24**(9), 2994–3016 (2007).
47. Mahajan, V. N. *Supplement to Applied Optics* , 8125–8127 (1994).
48. Gray, http://www.mathworks.com/matlabcentral/fileexchange/33330-zernikecalc, Accessed (2011).
49. Mahajan, V. N. *Optical Imaging and Aberrations: Part I. Ray Geometrical Optics*. SPIE Publications, (1998).
50. Hoffman, J. *Induced aberrations in optical systems*. PhD thesis, The University of Arizona, (1993).
51. Hopkins, H. H. *Japanese Journal of Applied Physics* **4**, 31–35 (1965).
52. Sasian, J. *Introduction to Aberrations in Optical Imaging Systems*. Cambridge University Press, (2013).
53. Cox, A. *A System Of Optical Design*. The Focal Press, (1964).
54. Thompson, K. P. *Aberration Fields in Tilted and Decentered Optical Systems*. PhD thesis, The University of Arizona, (1980).

55. Kwee, I. W. and Braat, J. J. M. *Pure Appl. Opt.* **2**, 21–32 (1993).
56. Agurok, I. *SPIE* **3430**, 80–87 (1998).
57. Manuel, A. M., Burge, J. H., and Tessieres, R. *OSA FiO* (2009).
58. Matsuzawa, T. *J. Opt. Soc. Am. A* **28**(2), 96–110 (2011).
59. Joseph J. M. Braat, A. J. E. M. J. *J. Opt. Soc. Am. A* **30**(6), 1213–1222 (2013).
60. Ming-Sen Tsao, C.-W. L. *Proc. of SPIE* **8844**(88440B-1) (2013).
61. Thompson, K. P. *J. Opt. Soc. Am. A* **26**(3), 699 (2009).
62. Thompson, K. P. *J. Opt. Soc. Am. A* **26**(5), 1090–1100 (2009).
63. Thompson, K. P. *J. Opt. Soc. Am. A* **27**(6), 1490–1504 (2010).
64. Thompson, K. P. *J. Opt. Soc. Am. A* **28**(5), 821–836 (2011).
65. Buchroeder, R. A. *Tilted component optical systems*. PhD thesis, The University of Arizona, (1976).
66. Thompson, K. P., Schmid, T., Cakmakci, O., and Rolland, J. P. *J. Opt. Soc. Am. A* **26**(6), 1503–1517 (2009).
67. Buchdahl, H. A. *Optical Aberration Coefficients*. Dover Publications Inc., (1969).
68. Rimmer, M. P. Master's thesis, University of Rochester, (1963).
69. Forbes, G. F. and Menke, C. *Proc. of SPIE* **88841C** (2013).
70. Wang, J., Guo, B., Sun, Q., and Lu, Z. *Opt. Express* **20**(11), 11652–11658 (2012).
71. Hu, H., Liu, J., and Fan, Z. *Opt. Express* **21**(15), 17986–17998 (2013).
72. Kyle Fuerschbach, Jannick P. Rolland, K. P. T. *Opt. Express* **20**(8), 20139–20155 (2012).
73. Shi, Z., Sui, Y., Liu, Z., Peng, J., and Yang, H. *Applied Optics* **51**(18), 4210–4214 (2012).
74. http://en.wikipedia.org/wiki/Jacobi_polynomials, Accessed (2013).
75. Szego, G. *Orthogonal Polynomials*. American Mathematical Society, 4th edition, (1978).
76. http://en.wikipedia.org/wiki/Legendre_polynomials, (2013).
77. Cruickshank, F. D. and Hills, G. A. *Journal Of The Optical Society Of America* **50**(4), 379–387 (1960).

Appendix I. Geometric Algebra Forms of Shack's Vector Product

During the initial development of SVP in terms of Geometric Algebra (GA), several different procedural interpretations of this vector product were considered and equations for each interpretation developed. This led to many different ways of expressing SVP in GA. Eventually, the simplest form was discovered. This Appendix provides several of the other forms developed during this investigation.

In the following equations the y-axis is used as the reference axis. The vector $\hat{n}$ is defined to be a unit vector halfway between $\hat{e}_2$ (the unit vector along the positive y-axis) and an arbitrary 2D vector $\vec{B}$. The vector $\hat{a}$ is a unit vector along the arbitrary 2D vector $\vec{A}$. The angles α and β are the angles of the vectors $\vec{A}$ and $\vec{B}$ with respect to the y-axis, respectively. Assuming all angles are positive, then all rotations indicated in the equations below are clockwise (from the y-axis toward x-axis). R_j is used to indicate different rotors for different values of j. The vectors in component form are given by

$$\vec{A} = a_x \hat{e}_1 + a_y \hat{e}_2, \quad \vec{B} = b_x \hat{e}_1 + b_y \hat{e}_2. \tag{I.1}$$

In 2D GA, the pseudoscalar I is given by

$$I = \hat{e}_1 \hat{e}_2 . \tag{1.2}$$

The pseudoscalar I is used in the equations below. Shack's vector product, $\vec{A} \star \vec{B}$, can then be interpreted and written as shown below.

As vector components:

$$\vec{A} \star \vec{B} = \left(a_y b_x + a_x b_y \right) \hat{e}_1 + \left(a_y b_y - a_x b_x \right) \hat{e}_2. \tag{I.3}$$

All vector elements explicitly listed:

$$\vec{A} \star \vec{B} = \left\| \vec{A} \right\| \left\| \vec{B} \right\| \hat{n} \hat{e}_2 \, \hat{a} \hat{e}_2 \, \hat{n}. \tag{I.4}$$

As two reflections of $\vec{A}$:

$$\vec{A} \star \vec{B} = \left\| \vec{B} \right\| \left(\hat{n} \left(\hat{e}_2 \vec{A} \hat{e}_2 \right) \hat{n} \right). \tag{I.5}$$

As a double sided rotor rotating $\vec{A}$:

$$\vec{A} \star \vec{B} = \left\| \vec{B} \right\| \left(\hat{n} \hat{e}_2 \right) \vec{A} \left(\hat{e}_2 \hat{n} \right) = \left\| \vec{B} \right\| R_1 \vec{A} R_1^{-1}. \tag{I.6}$$

As double sided GA exponentials rotating $\vec{A}$:

$$\vec{A} \star \vec{B} = \left\| \vec{B} \right\| e^{I\beta/2} \vec{A} e^{-I\beta/2}. \tag{I.7}$$

As a single, right sided rotor rotating $\vec{A}$:

$$\vec{A} \star \vec{B} = \vec{A} \left(\hat{e}_2 \vec{B} \right) = \vec{A} \left(\hat{e}_2 \hat{B} \right) \left\| \vec{B} \right\| = \vec{A} R_2 \left\| \vec{B} \right\|. \tag{I.8}$$

As a single, left sided rotor rotating $\vec{B}$:

$$\vec{A} \star \vec{B} = \vec{A} \hat{e}_2 \vec{B} = \left(\vec{A} \hat{e}_2 \right) \vec{B} = \left\| \vec{A} \right\| R_3 \vec{B} \tag{I.9}$$

As a single, right sided GA exponential rotating $\vec{A}$:

$$\vec{A} \star \vec{B} = \left\| \vec{B} \right\| \vec{A} e^{-I\beta} \tag{I.10}$$

As a single, left sided GA exponential rotating $\vec{A}$:

$$\vec{A} \star \vec{B} = \left\| \vec{B} \right\| e^{I\beta} \vec{A} \tag{I.11}$$

As a single left sided GA rotor rotating $\vec{B}$:

$$\vec{A} \star \vec{B} = \left(\vec{A} \hat{e}_2 \right) \vec{B} = \left\| \vec{A} \right\| \left(\hat{A} \hat{e}_2 \right) \vec{B} = \left\| \vec{A} \right\| R_3 \vec{B} \tag{I.12}$$

As a single left sided GA exponential rotating $\vec{B}$:

$$\vec{A} \star \vec{B} = \left\| \vec{A} \right\| e^{I\alpha} \vec{B} \tag{I.13}$$

As a single right sided GA exponential rotating $\vec{B}$:

$$\vec{A} \star \vec{B} = \left\| \vec{A} \right\| \vec{B} e^{-I\alpha} \tag{I.14}$$

As a single right sided GA exponential rotating $\hat{e}_2$:

$$\vec{A} \star \vec{B} = \left\| \vec{A} \right\| \left\| \vec{B} \right\| \hat{e}_2 e^{-I\alpha} e^{-I\beta} = \left\| \vec{A} \right\| \left\| \vec{B} \right\| \hat{e}_2 e^{-I(\alpha+\beta)} \tag{I.15}$$

As a single left sided GA exponential rotating $\hat{e}_2$:

$$\vec{A} \star \vec{B} = \left\| \vec{A} \right\| \left\| \vec{B} \right\| e^{I\alpha} e^{I\beta} \hat{e}_2 = \left\| \vec{A} \right\| \left\| \vec{B} \right\| e^{I(\alpha+\beta)} \hat{e}_2 \tag{I.16}$$

Simplest GA expressions:

$$\vec{A} \star \vec{B} = \vec{A} \hat{e}_2 \vec{B} = \vec{B} \hat{e}_2 \vec{A} = \vec{B} \star \vec{A}. \tag{I.17}$$

Appendix II.   Implementation of Recurrence Relations For Zernike Polynomials

The following is an implementation of recurrence relations [36, 38] for the computation of the radial factor of the Zernike polynomials.   The function returns the Zernike polynomial values at the specified points.

```
function Zv=calculateZernikeForbes(n,m,rho,phi,NormType)
  %  Implementation of recursion relations for calculating the
  %  Zernike Polynomial values.
  %
  %  Based on G.W. Forbes, "Robust and fast computation for the polynomials of optics"
  %  Optics Express, 18(13), 13851-13862, 2010.  See in particular Forbes'
  %  Eq. (4.1a,b,c) in this paper.
  %
  %  Calculate the Zernike polynomial for the rho, phi values specified,
  %    n = positive integer radial order,
  %    m = positive or negative or zero azimuthal order
  %    rho = vector of normalized rho values
  %    phi = vector of angles (radians) to go with the rho values.
  %    NormType = '0P' = Zero-to-Peak normalization,
  %               'RMS' = root-mean-square normalized (optional).
  %
  % Updates:
  %   2014-08-12:  Initially written by Robert W. Gray
  %

  if nargin < 5
    NormType = '0P';  % default is zero-to-peak normalization
  end

  % 0P normalized (default)
  N = 1;

  if strcmpi(NormType,'RMS')
    % RMS normalized
    N = sqrt(2*(n+1)/(1 + (m==0)));
  end % if statement

  R = calcR(n,abs(m),rho);

  % Calculate the azimuthal factor (cs = cosine or sine factor)
  cs = ones(size(phi));
  if m > 0
    cs = cos(m.*phi);
  end
  if m < 0
    cs = sin(abs(m).*phi);
```

```
  end

  Zv =  N * (R .* cs);

end % function calculateZernike2

function R=calcR(n,m,rho)
  % Use recurrence relations to calculate Zernike radial part

  nf = (n-m)/2;
  R = rho.^m .* Znf(nf,m,rho.^2) ;

end % calcR

function result = Znf(k, mf, u2)
      z(1,:) = ones(size(u2));
      z(2,:) = (mf + 2) * u2 - (mf + 1);

      if (k >= 2)
            for nf=1:(k-1)
        s = mf + 2*nf;
        a =  -(s+1)*((s-nf)^2 + nf*nf + s)/(s*(nf + 1)*(s - nf + 1));
        b =   (s + 1)*(s + 2)/((nf+1)*(s-nf+1));
        c =   (s+2)*(s-nf)*nf/(s*(nf+1)*(s-nf+1));
         z(nf+2,:) = (a+b*u2).*z(nf+1,:) - c*z(nf,:);
       end % for statement
    end % if statement

  result = z(k+1,:);

end  % function Znf
```

Appendix III.   Discrete Orthogonality of Zernike Polynomials

In a 2005 paper [43], Pap and Schipp published a result showing that a finite set of complex number Zernike polynomials are orthogonal over a finite set of discrete points across a unit radius disk. Prior to this work, it was known that the Zernike polynomials were orthogonal to one another over the continuous unit radius disk. This orthogonality property is very desirable because it means that the coefficients of the expansion of any function defined over the continuous unit disk (wavefront aberration function or an optic element's surface shape, etc.) in terms of the Zernike polynomials would be independent of one another, and therefore, would not change if additional terms of the expansion were latter included. It was thought that when only a discrete set of data points over the unit radius disk were used (discrete sampling of the function) the Zernike polynomials were no longer orthogonal, and the derived coefficients would then not be independent to one another. To overcome this issue, many data points over the disk were typically used in an attempt to obtain a good approximation to the exact coefficient values. With the Pap and Schipp result [43, 44, 45], it is possible to select a finite number of data points over the unit radius disk such that all the Zernike functions of order n_{max} or less remain orthogonal over these discrete data points, provided that the function's data values can be exactly represented by a sum of Zernike polynomials of order less than or equal to n_{max}. The number of data points needed is dependent on the maximum radial order n_{max}

of the Zernike polynomials needed to exactly define the function over the unit radius disk.

One drawback to this result is that the value of n_{max} that exactly defines the function of interest over the disk is not in general known. With the discrete sampling and finite subset of the Zernike polynomials, including the next higher order Zernike polynomial will change all the lower order coefficients. However, the change is of the order of the coefficient of the next highest order Zernike polynomial included. Then, assuming the function converges for low values of n_{max}, only a small number of Zernike polynomials need to be considered for an acceptable approximation of the function expressed as an expansion in low order Zernike polynomials. Another potential drawback is that the highest order Zernike needed to accurately represent the function (to expand a given function) may be so large that the number of data points across the unit disk is too large to be practicable. Additionally, the higher the value of n_{max} the more concern there is for the numeric accuracy of the calculated Zernike polynomial values.

On the other hand, a significant advantage to the method of Pap and Schipp is that there is no data fitting operation involved. The coefficients are calculated directly from the equations to be developed below by use of the Gaussian Quadrature (GQ) technique.

In this appendix, the results of Pap and Schipp [43], incorporating an improvement pointed out by Shi, *et al*. [73], are derived for the real number Zernike polynomials. Pap and Schipp used a complex exponential form for the Zernike polynomials while the research presented in this dissertation uses the real number form of the Zernike polynomials. Therefore, this derivation, in terms of real number Zernike

polynomials, was necessary in order to obtain the equations needed for this research project.

The notation employed in this Appendix is not necessarily the same as in the main text of this dissertation.

A Zernike polynomial, using the 0-P normalization, can be defined as two factors: 1) The radial dependence $R_n^{|m|}(\rho)$, and 2) the azimuthal dependence $\Phi_m(\varphi)$. Then a Zernike polynomial $Z_n^m(\rho,\varphi)$ may be written as

$$Z_n^m(\rho,\varphi) = R_n^{|m|}(\rho)\Phi_m(\varphi), \qquad (\text{III.1})$$

where n is a positive integer, m is a positive or negative integer, $|m| \leq n$ and $n - |m|$ is even. The radial dependence is given by

$$R_n^{|m|}(\rho) = \sum_{k=0}^{\frac{n-|m|}{2}} \frac{(-1)^k (n-k)!}{k!\left(\frac{n+|m|}{2}-k\right)!\left(\frac{n-|m|}{2}-k\right)!}\rho^{n-2k} \; . \qquad (\text{III.2})$$

This can also be written as

$$R_n^{|m|}(\rho) = \rho^{|m|} P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)}\left(2\rho^2-1\right), \qquad (\text{III.3})$$

where $P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)}\left(2\rho^2-1\right)$ are Jacobi polynomials [74]. This is an important relation to be used later. Explicit forms of the Jacobi polynomials will not be needed here. The azimuthal dependence is defined as

$$\Phi_m(\varphi) = \begin{cases} \cos(m\varphi) & m \geq 0 \\ \sin(|m|\varphi) & m < 0 \end{cases}. \qquad (\text{III.4})$$

The orthogonality relation for these Zernike polynomials is written as

$$\frac{2n+2}{\pi\left(1+\delta_{m,0}\right)}\int_0^{2\pi}\int_0^1 Z_n^m\left(\rho,\varphi\right)Z_{n'}^{m'}\left(\rho,\varphi\right)\rho\,d\rho\,d\varphi=\delta_{n,n'}\,\delta_{m,m'}.\tag{III.5}$$

This double integral separates into

$$\frac{2n+2}{\pi\left(1+\delta_{m,0}\right)}\int_0^1 R_n^{|m|}\left(\rho\right)R_{n'}^{|m'|}\left(\rho\right)\rho\,d\rho\int_0^{2\pi}\Phi_m\left(\varphi\right)\Phi_{m'}\left(\varphi\right)d\varphi=\delta_{n,n'}\,\delta_{m,m'}.\tag{III.6}$$

Moving the constant factor to the right hand side of the equation,

$$\int_0^1 R_n^{|m|}\left(\rho\right)R_{n'}^{|m'|}\left(\rho\right)\rho\,d\rho\int_0^{2\pi}\Phi_m\left(\varphi\right)\Phi_{m'}\left(\varphi\right)d\varphi=\frac{\pi\delta_{n,n'}\,\delta_{m,m'}\left(1+\delta_{m,0}\right)}{2n+2}.\tag{III.7}$$

Consider the integral

$$\int_0^{2\pi}\Phi_m\left(\varphi\right)\Phi_{m'}\left(\varphi\right)d\varphi.\tag{III.8}$$

There are four cases to consider.

CASE I: $m\geq 0$, and $m'\geq 0$

$$\int_0^{2\pi}\Phi_m\left(\varphi\right)\Phi_{m'}\left(\varphi\right)d\varphi=\int_0^{2\pi}\cos\left(m\varphi\right)\cos\left(m'\varphi\right)d\varphi=\pi\left(\delta_{m,m'}\delta_{m,0}+\delta_{m,m'}\right).\tag{III.9}$$

CASE II: $m\geq 0$, and $m'<0$

$$\int_0^{2\pi}\Phi_m\left(\varphi\right)\Phi_{m'}\left(\varphi\right)d\varphi=\int_0^{2\pi}\cos\left(m\varphi\right)\sin\left(|m'|\varphi\right)d\varphi=0.\tag{III.10}$$

CASE III: $m<0$, and $m'\geq 0$

$$\int_0^{2\pi}\Phi_m\left(\varphi\right)\Phi_{m'}\left(\varphi\right)d\varphi=\int_0^{2\pi}\sin\left(|m|\varphi\right)\cos\left(m'\varphi\right)d\varphi=0.\tag{III.11}$$

CASE IV: $m<0$, and $m'<0$

$$\int_0^{2\pi} \Phi_m(\varphi)\Phi_{m'}(\varphi)d\varphi = \int_0^{2\pi} \sin(|m|\varphi)\sin(|m'|\varphi)d\varphi = \pi\delta_{m,m'}. \tag{III.12}$$

These four cases can be combined into the single equation

$$\int_0^{2\pi} \Phi_m(\varphi)\Phi_{m'}(\varphi)d\varphi = \pi\left(\delta_{m,m'}\delta_{m,0} + \delta_{m,m'}\right) \tag{III.13}$$

The strategy at this point is to develop alternative equations for the Kronecker deltas that occur in the above equations in terms of sums. This is accomplished as follows.

CASE I: $m \geq 0$, and $m' \geq 0$

For this case write

$$\frac{2}{N_\varphi}\sum_{j=0}^{N_\varphi-1} \Phi_m(\varphi_j)\Phi_{m'}(\varphi_j) = \frac{2}{N_\varphi}\sum_{j=0}^{N_\varphi-1} \cos(m\varphi_j)\cos(m'\varphi_j), \tag{III.14}$$

where N_φ is an integer to be determined below. Express the cosines in terms of Euler's equations gives

$$\frac{2}{N_\varphi}\sum_{j=0}^{N_\varphi-1} \cos(m\varphi_j)\cos(m'\varphi_j) = \frac{1}{2N_\varphi}\sum_{j=0}^{N_\varphi-1}\left(e^{i(m+m')\varphi_j} + e^{i(m-m')\varphi_j} + e^{i(-m+m')\varphi_j} + e^{i(-m-m')\varphi_j}\right). \tag{III.15}$$

A relation from discrete Fourier analysis can now be used, namely

$$\frac{1}{N_\varphi}\sum_{j=0}^{N_\varphi-1} e^{\frac{\pm i(m-m')2\pi j}{N_\varphi}} = \begin{cases} 1 & \text{for } m-m' = pN_\varphi, \text{ where } p = \text{integer} \\ 0 & \text{otherwise} \end{cases}. \tag{III.16}$$

Comparing the exponentials in Eq. (III.15) to Eq. (III.16) the following association can be made

$$\varphi_j = \frac{2\pi j}{N_\varphi}. \tag{III.17}$$

With m and m' maximal when $m = n$ and $m' = n'$, set

$$N_\varphi = 2n_{\max} + 1. \tag{III.18}$$

where $n_{\max}$ is the maximum of $\{n, n'\}$. In this way, the angle φ_j cannot exceed 2π.

$$\varphi_j = \frac{2\pi j}{N_\varphi} \leq \frac{2\pi(N_\varphi - 1)}{2n_{\max} + 1} \leq \frac{2\pi(2n_{\max})}{2n_{\max} + 1} < 2\pi. \tag{III.19}$$

Then, with (III.18) it follows that the condition for the first relation of (III.16) to hold, that is, for $m - m' = pN_\varphi$ to hold, requires that p be zero and therefore that $m - m' = 0$.

With these definitions and restrictions, the following relation is obtained.

$$\frac{1}{N_\varphi} \sum_{j=0}^{N_\varphi - 1} e^{\frac{\pm i(m-m')2\pi j}{N_\varphi}} = \delta_{m,m'}. \tag{III.20}$$

This can be used to write Eq. (III.15) as

$$\frac{2}{N_\varphi} \sum_{j=0}^{N_\varphi - 1} \cos(m\varphi_j)\cos(m'\varphi_j) = \frac{1}{2}\left(\delta_{m,0}\delta_{m,m'} + \delta_{m,m'} + \delta_{m,m'} + \delta_{m,0}\delta_{m,m'}\right). \tag{III.21}$$

Then, for this case

$$\frac{2}{N_\varphi} \sum_{j=0}^{N_\varphi - 1} \Phi_m(\varphi_j)\Phi_{m'}(\varphi_j) = \left(\delta_{m,0}\delta_{m,m'} + \delta_{m,m'}\right). \tag{III.22}$$

For cases II and III, it can be shown that

$$\frac{2}{N_\varphi} \sum_{j=0}^{N_\varphi - 1} \Phi_m(\varphi_j)\Phi_{m'}(\varphi_j) = 0. \tag{III.23}$$

And for case IV,

$$\frac{2}{N_\varphi} \sum_{j=0}^{N_\varphi - 1} \Phi_m(\varphi_j)\Phi_{m'}(\varphi_j) = \delta_{m,m'}. \tag{III.24}$$

These four cases can be combined to give the relation

$$\frac{2}{N_\varphi}\sum_{j=0}^{N_\varphi-1}\Phi_m\left(\varphi_j\right)\Phi_{m'}\left(\varphi_j\right)=\delta_{m,m'}\delta_{m,0}+\delta_{m,m'}. \tag{III.25}$$

Comparing this result with Eq. (III.13) gives the equation

$$\int_0^{2\pi}\Phi_m\left(\varphi\right)\Phi_{m'}\left(\varphi\right)d\varphi=\pi\left(\delta_{m,m'}\delta_{m,0}+\delta_{m,m'}\right)=\frac{2\pi}{N_\varphi}\sum_{j=0}^{N_\varphi-1}\Phi_m\left(\varphi_j\right)\Phi_{m'}\left(\varphi_j\right). \tag{III.26}$$

To convert the radial integral into a discrete, finite sum, the Gaussian Quadrature theorem can be used. One form of the theorem is presented in [75] as Theorem 3.4.1 (p. 47). The following form of the theorem is from [75] with a few changes of the text, as well as notational changes, for clarity. The changes are not explicitly indicated. Since the space over which the theorem is to be applied is a flat Euclidean space such that the measure of the space is $\mu(x)=x$, the integral is changed from the original Stieltjes-Lebesque integral to a Riemann integral. From [75] (edited):

<u>Theorem 3.4.1</u> If $x_1<x_2<...<x_N$ denote the zeros (roots) of a polynomial function

$f_N(x)$ of order N, then there exits real numbers $A_1,A_2,...,A_N$ (called Christoffel

numbers) such that

$$\int_a^b g(x)dx=A_1 g(x_1)+A_2 g(x_2)+...+A_N g(x_N), \tag{III.27}$$

whenever $g(x)$ is an arbitrary polynomial function of order ***less than*** $2N$.

In the Pap and Schipp paper [43], the theorem is stated as follows. The notation and text have again been changed from their paper for clarity.

<u>Theorem A</u> Let λ_1, λ_2, ..., λ_N with $\lambda_i \in (-1, 1)$, be the N roots of the Legendre

polynomial $P_N(x)$ of order N. Define

$$\Lambda_i(x) \equiv \frac{(x-\lambda_1)\cdots(x-\lambda_{i-1})(x-\lambda_{i+1})\cdots(x-\lambda_N)}{(\lambda_i-\lambda_1)\cdots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\cdots(\lambda_i-\lambda_N)} \tag{III.28}$$

to be the fundamental polynomials of the Lagrange interpolation of the Legendre

polynomial. We define the corresponding Christoffel numbers by

$$A_k \equiv \int_{-1}^{1} \Lambda_k(x)\,dx \tag{III.29}$$

with $1 \leq k \leq N$ an integer. Then for every polynomial $g(x)$ of order ***less than*** $2N$

we can write

$$\int_{-1}^{1} g(x)\,dx = \sum_{k=1}^{N} g(\lambda_k) A_k \tag{III.30}$$

Although explicit equations for the Legendre polynomials will not be needed, several

forms from [76] are provided here as explicit examples.

$$P_n(x) = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k}^2 (x-1)^{n-k}(x+1)^k \tag{III.31}$$

$$P_n(x) = \sum_{k=0}^{n} \binom{n}{k}\binom{-n-1}{k}\left(\frac{1-x}{2}\right)^k \tag{III.32}$$

$$P_n(x) = 2^n \sum_{k=0}^{n} \binom{n}{k}\binom{\frac{n+k-1}{2}}{n} x^k \tag{III.33}$$

The next step is to convert the radial integral [11]

$$\int_0^1 R_n^{|m|}(\rho) R_{n'}^{|m|}(\rho) \rho \, d\rho = \frac{1}{2(n+1)} \delta_{n,n'} \tag{III.34}$$

into a form suitable for application of the above Theorem. Inserting Eq. (III.3), the alternative form for expressing $R_n^{|m|}(\rho)$ in terms of the Jacobi polynomials, we have

$$\int_0^1 \left(\rho^{|m|} P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)} \left(2\rho^2 - 1\right) \right) \left(\rho^{|m|} P_{\left(\frac{n'-|m|}{2}\right)}^{(0,|m|)} \left(2\rho^2 - 1\right) \right) \rho \, d\rho = \frac{\delta_{n,n'}}{2n+2}, \tag{III.35}$$

which can be written as

$$\int_{\rho=0}^{\rho=1} \rho^{2|m|} P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)} \left(2\rho^2 - 1\right) P_{\left(\frac{n'-|m|}{2}\right)}^{(0,|m|)} \left(2\rho^2 - 1\right) \rho \, d\rho = \frac{\delta_{n,n'}}{2n+2}. \tag{III.36}$$

Performing a change of variables, let

$$u = 2\rho^2 - 1, \tag{III.37}$$

then

$$du = 4\rho \, d\rho. \tag{III.38}$$

For $\rho = 0$, $u = -1$ while for $\rho = 1$, $u = 1$. Solving for ρ^2 gives

$$\rho^2 = \frac{u+1}{2}. \tag{III.39}$$

Using these results in equation (III.36) results in the equation

$$\frac{1}{4} \int_{u=-1}^{u=1} \left(\frac{u+1}{2}\right)^{|m|} P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)}(u) P_{\left(\frac{n'-|m|}{2}\right)}^{(0,|m|)}(u) \, du = \frac{\delta_{n,n'}}{2n+2}. \tag{III.40}$$

Let

$$g(u) = \left(\frac{u+1}{2}\right)^{|m|} P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)}(u) P_{\left(\frac{n'-|m|}{2}\right)}^{(0,|m|)}(u) \tag{III.41}$$

then Eq. (III.40) takes the form

$$\frac{1}{4}\int\limits_{u=-1}^{u=1} g\left(u\right)du = \frac{\delta_{n,n'}}{2n+2}.$$ (III.42)

This is now in the correct form for applying the Gaussian Quadrature Theorem. This results in the equation

$$\frac{1}{4}\int\limits_{u=-1}^{u=1} g\left(u\right)du = \frac{1}{4}\sum_{k=1}^{N_\rho} g\left(\lambda_k\right)A_k = \frac{\delta_{n,n'}}{2n+2},$$ (III.43)

where λ_k are the roots of the Legendre polynomial of order N_ρ. Note that the function

$g\left(u\right)$ is of order $\frac{1}{2}\left(n+n'\right)$ in u. This is from considering the powers of u, which are

$$|m|+\left(\frac{n-|m|}{2}\right)+\left(\frac{n'-|m|}{2}\right) = \frac{1}{2}\left(n+n'\right).$$ (III.44)

Since there is a Kronecker delta on the right hand side of Eq. (III.43), the maximum order

of $g\left(u\right)$ is given by $\frac{1}{2}\left(n_{max}+n_{max}\right)=n_{max}$. The Theorem states that the equality of Eq.

(III.43) holds when $g\left(u\right)$ is a polynomial of order less then $2N_\rho$. Then there is the restriction that

$$n_{max} < 2N_\rho.$$ (III.45)

This means that

$$N_\rho = \frac{n_{max}}{2}+1.$$ (III.46)

This is important because it specifies the number of radial parameter values, ρ, needed

for the fixed but arbitrary maximum order n_{max} of the subset of Zernike polynomials

involved. Since n_{max} may be an odd integer, the following relation is used, where $\lfloor \bullet \rfloor$ denotes the integer part of the real number

$$N_\rho = \left\lfloor \frac{n_{max}}{2} \right\rfloor + 1 . \tag{III.47}$$

To summarize,

$$\int_0^1 R_n^m(\rho) R_{n'}^m(\rho) \rho \, d\rho = \frac{1}{4} \int_{u=-1}^{u=1} g(u) \, du = \frac{1}{4} \sum_{k=1}^{N_\rho} g(\lambda_k) A_k = \frac{\delta_{n,n'}}{2n+2} \tag{III.48}$$

where λ_k are the roots of the Legendre polynomial of order N_ρ and A_k are the associated Christoffel numbers calculated by Eqs. (III.28) and (III.29).

Since u is replaced with λ_k in the above equation, and using Eq. (III.39) the relation

$$u \to \lambda_k = 2\rho_k^2 - 1 \tag{III.49}$$

is obtained. This relation defines the discrete radial values to be used to be

$$\rho_k = \sqrt{\frac{\lambda_k + 1}{2}} . \tag{III.50}$$

Eq. (III.49) is used to rewrite Eq. (III.34) into a summation explicitly involving the radial function. Using Eq. (III.41) in Eq. (III.48) gives

$$\int_0^1 R_n^m(\rho) R_{n'}^m(\rho) \rho \, d\rho = \frac{1}{4} \sum_{k=1}^{N_\rho} \left(\frac{\lambda_k + 1}{2} \right)^{|m|} P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)}(\lambda_k) P_{\left(\frac{n'-|m|}{2}\right)}^{(0,|m|)}(\lambda_k) A_k = \frac{\delta_{n,n'}}{2n+2} \tag{III.51}$$

It then follows that

$$\int_0^1 R_n^m(\rho) R_{n'}^m(\rho) \rho \, d\rho = \frac{1}{4} \sum_{k=1}^{N_\rho} \left(\frac{2\rho_k^2 - 1 + 1}{2} \right)^{|m|} P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)}(2\rho_k^2 - 1) P_{\left(\frac{n'-|m|}{2}\right)}^{(0,|m|)}(2\rho_k^2 - 1) A_k , \tag{III.52}$$

and

$$\int_0^1 R_n^m(\rho) R_{n'}^m(\rho)\rho\, d\rho = \frac{1}{4}\sum_{k=1}^{N_\rho}\left(\rho_k^{|m|}P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)}\left(2\rho_k^2-1\right)\right)\left(\rho_k^{|m|}P_{\left(\frac{n'-|m|}{2}\right)}^{(0,|m|)}\left(2\rho_k^2-1\right)\right)A_k \ . \quad \text{(III.53)}$$

Therefore,

$$\frac{1}{4}\sum_{k=1}^{N_\rho}\left(\rho_k^{|m|}P_{\left(\frac{n-|m|}{2}\right)}^{(0,|m|)}\left(2\rho_k^2-1\right)\right)\left(\rho_k^{|m|}P_{\left(\frac{n'-|m|}{2}\right)}^{(0,|m|)}\left(2\rho_k^2-1\right)\right)A_k = \frac{\delta_{n,n'}}{2n+2} \ . \quad \text{(III.54)}$$

And so the integration over the radial part of the Zernike polynomials is replaced by the discrete sum

$$\frac{1}{4}\sum_{k=1}^{N_\rho}R_n^m(\rho_k) R_{n'}^{m'}(\rho_k) A_k = \frac{\delta_{n,n'}}{2n+2} \ . \quad \text{(III.55)}$$

Combining this result with the sum over the azimuthal factors gives

$$\left(\frac{1}{4}\sum_{k=1}^{N_\rho}R_n^m(\rho_k) R_{n'}^{m'}(\rho_k) A_k\right)\left(\frac{2}{N_\varphi}\sum_{j=0}^{N_\varphi-1}\Phi_m(\varphi_j)\Phi_{m'}(\varphi_j)\right) = \left(\frac{\delta_{n,n'}}{2n+2}\right)\left(\delta_{m,m'}\delta_{m,0}+\delta_{m,m'}\right). \quad \text{(III.56)}$$

Collecting common factors and distributing the summation over the azimuthal parameters gives

$$\frac{2n+2}{2}\left(\frac{1}{\left(1+\delta_{m,0}\right)N_\varphi}\right)\left(\sum_{k=1}^{N_\rho}\sum_{j=0}^{N_\varphi-1}R_n^m(\rho_k)\Phi_m(\varphi_j)R_{n'}^{m'}(\rho_k)\Phi_{m'}(\varphi_j)A_k\right) = \delta_{n,n'}\delta_{m,m'} \ . \quad \text{(III.57)}$$

The final expression for the orthogonality of the Zernike polynomials over a discrete sampling of data points can now be written as

$$\left(\frac{n+1}{\left(1+\delta_{m,0}\right)N_\varphi}\right)\left(\sum_{k=1}^{N_\rho}\sum_{j=0}^{N_\varphi-1}Z_n^m(\rho_k,\varphi_j)Z_{n'}^{m'}(\rho_k,\varphi_j)A_k\right) = \delta_{n,n'}\delta_{m,m'} \ . \quad \text{(III.58)}$$

To illustrate how these equations are used in the research, and to provide a short summary of the important equations, assume that a function $f(\rho,\varphi)$ over the unit radius disk can be written as

$$f(\rho,\varphi) = \sum_{n=0}^{n_{max}} \sum_{\substack{m=-n \\ m=m+2}}^{n} C_n^m Z_n^m(\rho,\varphi) + O_Z(n_{max}+1), \qquad \text{(III.59)}$$

where C_n^m are the expansion coefficients to be determined and $O_Z(n_{max}+1)$ signifies the order of magnitude of the real Zernike polynomial terms of Zernike order higher than n_{max}. The expansion coefficients can be determined approximately by

$$C_n^m \approx \frac{(n+1)}{(1+\delta_{m,0})N_\varphi} \sum_{k=1}^{N_\rho} \sum_{j=0}^{N_\varphi-1} f(\rho_k,\varphi_j) Z_n^m(\rho_k,\varphi_j) A_k, \qquad \text{(III.60)}$$

where the approximation is good to the order of magnitude of the real Zernike polynomial terms of Zernike order higher than n_{max}. Note that the expansion coefficients are given exactly by Eq. (III.60) when $O_Z(n_{max}+1)=0$. Continuing with the notation used,

$$N_\rho = \left\lfloor \frac{n_{max}}{2} \right\rfloor + 1, \qquad \text{(III.61)}$$

$$\rho_k = \sqrt{\frac{1+\lambda_k}{2}}. \qquad \text{(III.62)}$$

λ_k are the roots of the Legendre polynomial of order N_ρ, $1 \le k \le N_\rho$, and the Christoffel numbers A_k are given by

$$A_k \equiv \int_{-1}^{1} \Lambda_k(x)\,dx\ , \tag{III.63}$$

where

$$\Lambda_k(x) \equiv \frac{\left(x - \lambda_1\right)\cdots\left(x - \lambda_{k-1}\right)\left(x - \lambda_{k+1}\right)\cdots\left(x - \lambda_{N_\rho}\right)}{\left(\lambda_k - \lambda_1\right)\cdots\left(\lambda_k - \lambda_{k-1}\right)\left(\lambda_k - \lambda_{k+1}\right)\cdots\left(\lambda_k - \lambda_{N_\rho}\right)}\ , \tag{III.64}$$

are the fundamental polynomials of the Lagrange interpolation of the Legendre polynomial. Further,

$$N_\varphi = 2n_{max} + 1\ , \tag{III.65}$$

$$\varphi_j = \frac{2\pi j}{N_\varphi} = \frac{2\pi j}{2n_{max} + 1}\ , \tag{III.66}$$

with $0 \le j \le N_\varphi - 1$.

The research conducted in this dissertation used Zernike polynomials through $n = 12$. Using the above equations, there are then $N_\rho = 7$ normalized radial values and $N_\varphi = 25$ azimuthal values need for the application of the GQ method. The actual values used in this research, rounded to 20th decimal position, are listed in Table III.1 and Table III.2.

Table III.1 Radial values used with GQ method.

#	Normalized Radial Values
1	0.15951816143819091089
2	0.35949187362206503717
3	0.54504809357643057125
4	0.70710678118654752440
5	0.83840478033507095662
6	0.93314821587982325196
7	0.98719499399631239331

Table III.2 Azimuthal values used with GQ method.

#	Azimuthal Values (Radians)
1	0.0
2	0.25132741228718345908
3	0.50265482457436691815
4	0.75398223686155037723
5	1.00530964914873383631
6	1.25663706143591729539
7	1.50796447372310075446
8	1.75929188601028421354
9	2.01061929829746767262
10	2.26194671058465113169
11	2.51327412287183459077
12	2.76460153515901804985
13	3.01592894744620150892
14	3.26725635973338496800
15	3.51858377202056842708
16	3.76991118430775188616
17	4.02123859659493534523
18	4.27256600888211880431
19	4.52389342116930226339
20	4.77522083345648572246
21	5.02654824574366918154
22	5.27787565803085264062
23	5.52920307031803609969
24	5.78053048260521955877
25	6.03185789489240301785

The validation of the above GQ method for obtaining the coefficients of an expression comprising a sum of Zernike polynomials through $n = 12$ was conducted by implementing the GQ method as Matlab® functions, presented below. Random numbers were used for the initial Zernike coefficient values. The GQ method was then used in an

attempt to determine the initial random number coefficient values. The difference between the initial coefficient values and the GQ calculated coefficients were calculated and was seen to be zero. That is, the GQ method reproduced the initial random coefficient values.

The following Matlab® functions are the functions used for the validation testing.

Function: verifyRealZernike

This function is the main GQ validation Matlab® function. It generates random coefficient values and calculates an expression of the sum of the Zernike polynomials through $n = 12$. It then uses the GQ method described above to obtain the coefficients from the calculated expression's values at the GQ prescribed Zernike parameter values. The difference between the original random coefficient values and the GQ obtained coefficient values is calculated and displayed.

```
function verifyRealZernike
% Verifies the real Zernike orthogonality equations by obtaining the
% random coefficients of a polynomial composed of real Zernike polynomials
% through 12th order.
%
% Written by: R. W. Gray 2014
%
  syms R aP W f G coeff real;
  sym('pi');

  maxn = 12; % MUST BE EVEN POSITIVE INTEGER.

  NR = maxn/2+1;
  NA = 2*maxn+1;

  % build the W function with random coefficients.
  % on interval [a,b]
  a = -10;
  b = +10;
  p = 0;
  W = sym(0);

  for n=0:maxn
    for m=n:-2:-n
      p = p + 1;
      Wc(p) = a + (b-a).*rand(1,1);
```

```
        Mn(p) = n;
        Mm(p) = m;
        fprintf(1,'P%2d: %2d,%3d: %16.12f \n',p,n,m,Wc(p));

        Z{p} = makeZernike(n,m,'R','aP');

        W = W + Wc(p)*Z{p};    % real W

    end % for m
  end % for n
  maxp = p;

  % At this point, we have an equation W of thru 12th order Zernikes and
  % random coefficient values.

  % Now do Gaussian Quadrature (GQ) to retreive the coefficient values.

  % Obtain the Christoffel coefficients and the Legendre roots.
  [Ak lambdak] = calcAks(NR);

  % Now obtain the coefficients from W.
  for p=1:maxp
    coeff = sym('0');

    for k=1:NR
      for j=0:(NA-1)
        % evaluate the function at the appropriate GQ prescribed position.
        f = subs(W,{R,aP},{sqrt((1+lambdak(k))/2),sym(2*pi*j/NA)});

        % get the Zernike GQ value
        G = f*subs(Z{p},{R,aP},{sqrt((1+lambdak(k))/2),sym(2*pi*j/NA)})*Ak{k};

        coeff = coeff + G;
      end % for j
    end % for k

    % Adjust value by the normalization value.  For real zernikes.
    coeff2 = double(vpa(coeff * (Mn(p) + 1)/((1+(Mm(p)==0))*NA)));

    % Calc and print the difference between original coefficient value and
    % GQ calculated coefficient value.
    fprintf(1,'P%2d: %2d,%3d: %30.27f %30.27f %30.27f \n', ...
              p,Mn(p),Mm(p),Wc(p),coeff2,Wc(p)-coeff2);

  end % for p statement

end % function verifyRealZernike
```

<u>Function: calcAks</u>

The following is a Matlab® function to calculate the Nth order roots of the Legendre polynomial and the associated Christoffel numbers. This is called from the main validation function presented above.

```matlab
function [Ak theRoots]=calcAks(N)
%CALCAKS Roots of the Nth order Legendre polynomial and Christoffel numbers.
%  To obtain accurate results out to the DIGITS decimal position, Matlab's
%  symbolic processing is used in the calculations.
%
%  INPUT:
%     N = The order of the Legendre polynomial.
%
%  OUTPUT:
%     Ak       = a Matlab cell containing the symbolic Christoffel numbers.
%     theRoots = a matrix containing the N roots of the Legendra polynomial.
%
%  Written by Robert W. Gray, 08-22-2013
%

  syms x lk real;

  % Calculate the N roots of Legendre polynomial.
  % From: http://www.mathworks.com/help/symbolic/mupad_ref/orthpoly-legendre.html

  strEqu = sprintf('QRoots := numeric::gldata(%d, DIGITS)[2]',N);
  evalin(symengine, strEqu);
  theRoots = evalin(symengine,'PRoots := map(QRoots, y -> 2*y - 1)');

  % Preallocate the cell
  Ak = cell(N,1);

  % Form the Legendre interpolation fundamental polynomials
  % and then calculate the Christoffel numbers by integration.
  for k=1:N
    lk = sym(1);

    for j=1:N
      if j == k
        % skip the j=k term in the formation of the polynomial.
        continue;
      end % if statement
      lk = lk * (x - theRoots(j)) / (theRoots(k)-theRoots(j));
    end % for j

    % Do the integeration of the Lagrange polynomial to obtain Ak.
    Ak{k} = int(lk,x,-1,1);
    %fprintf(1,'%d: %s \n',k,char(Ak{k}));   % Optional printing of values
  end % for k

end % function calcAks
```

<div style="text-align: center;">

Function: makeZernike

</div>

      This function returns a Matlab® symbolic expression for the specified Zernike polynomial. Using symbolic expressions for the Zernike polynomials avoids computational inaccuracies in the numerical representation of the Zernike polynomials. This is called from the main validation function presented above.

```
function Znm = makeZernike(n,m,sR,sA)
% Returns a symbolic express for the specified real Zernike polynomial
% INPUT:
%  n  = positive integer or zero, order of the Zernike polynomial.
%  m  = integer (positive or negative or zero), azimuthal parameter.
%      m >= 0 will ALWAYS mean the cosine factor.
%      m  < 0 will ALWAYS mean the sine factor.
%  sR = string containning the radial variable name.
%      For example 'H' or 'R', etc.
%  sA = string containning the angle variable name.
%      For example 'aPx' or 'aPy' etc.
%      aPx = angle, Phi, x-axis reference
%      aPy = angle, Phi, y-axis reference
%      aTx = angle, Theta, x-axis reference
%      aTy = angle, Theta, y-axis reference
%
% OUTPUT:
%  Znm  = symbolic expression of the specified Zernike polynomial.
%

  syms Znm R aP numerator denominator real;

  if n == 0
    Znm = sym(1);
    return
  end % if statement

  R = sym(sR);
  aP = sym(sA);

  Rnm = sym(0);
  positivem = abs(m);
  for k=0:((n-positivem)/2)
    numerator = sym((-1)^k*factorial(n-k));
    denominator = sym(factorial(k)*factorial((n+positivem)/2-k)*factorial((n-
positivem)/2-k));
    Rnm = Rnm + numerator/denominator * R^(n-2*k);
  end % for k statement

  % Put in the real azimuthal dependence.

  if m >= 0
    Znm = Rnm * cos(positivem*aP);
  else
    Znm = Rnm * sin(positivem*aP);
  end % if statement

end % function makeZernike
```
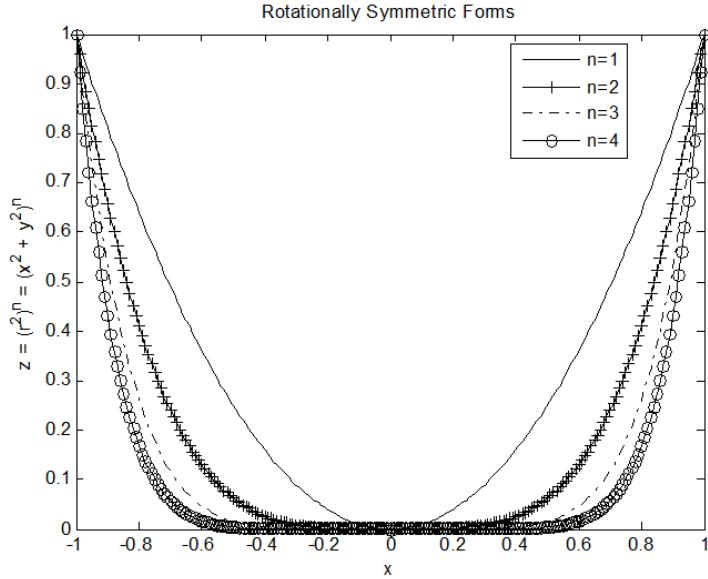
Appendix IV.   Zernike Polynomials as Decentered Rotationally Symmetric Terms

While working with the real number Zernike polynomials and aspheric optical element definitions, it was discovered that the low order Zernike polynomials could be expressed as a sum of decentered rotationally symmetric terms. A literature search did not reveal any reference to this property of Zernike polynomials. Since one way to define freeform surface shaped optical elements is by the Zernike polynomials, and since NAT can be used to analyze the field dependence of decentered optical elements (including aspheres), one wonders if there might be a way to use this sum of decentered rotationally symmetric term composition property of the Zernike polynomials to include surfaces defined by Zernike polynomials (or by sums of monomials) into NAT.

Consider the following sag equation for a component of a surface

$$z = c\left[\left(x - \delta x\right)^2 + \left(y - \delta y\right)^2\right]^n,$$ (IV.1)

where δx and δy are fixed but arbitrary real constants and where n is a positive integer and c is a real constant. This describes a rotationally symmetric shape (called a decentered rotationally symmetric shape term) that is decentered to the position $(\delta x, \delta y)$ in the $z = 0$ plane. The case for (δx=0, δy=0), $c = 1$, and $n = 1,2,3,4$ are shown in Figure IV.1.

Figure IV.1 Cross-section of several rotationally symmetric forms.

Hypothesis: Any *xy*-polynomial surface having a component sag equation of the form

$$z = x^m y^n, \tag{IV.2}$$

where *m* and *n* are positive (or zero) integers, can be written as

$$z = \sum_{i=0}^{i_{max}} c_i P^{n_i}\left(\delta x_i, \delta y_i\right), \tag{IV.3}$$

where i_{max} is a finite integer, n_i are positive integers, c_i are real constants and

$$P^{n_i}\left(\delta x_i, \delta y_i\right) = \left[\left(x - \delta x_i\right)^2 + \left(y - \delta y_i\right)^2\right]^{n_i}. \tag{IV.4}$$

A general proof of this hypothesis has not been found. However, several explicit examples have been worked out and are tabulated below.

If the hypothesis is true, then since any Zernike polynomial is a finite sum of monomials of the form $x^m y^n$, the following Corollary is implied.

Corollary I: The Zernike polynomials are composed of finite sums of decentered rotationally symmetric shape terms.

Note that a decenter amount of zero is allowed. This is needed for the Zernike polynomials that have no azimuthal angle dependence.

A few monomials and their decentered shape term equivalence are tabulated in Table IV.1 and Table IV.2. The only reason that the table stops at the point shown is because of the time involved in working out the correct decenter values. Aside from time, there appears to be nothing limiting the table from being expanded to any monomial required. Table IV.1 and Table IV.2 are suggestive that all monomials can be represented as sums of decentered shape terms.

A potential down side to expressing the Zernike polynomials as sums of decentered shape terms is that the higher the order of the Zernike polynomial, the more decentered shape terms required for its equivalent expression.

Table IV.1 Monomials and their decentered shape term equivalent.

Monom.	Sum of Decentered Shape Terms
1	$P^0(0,0)$
x	$\dfrac{1}{4}\big(P(-1,0)-P(1,0)\big)$
y	$\dfrac{1}{4}\big(P(0,-1)-P(0,1)\big)$
x^2	$\dfrac{1}{2}\big(P(0,0)\big)+\dfrac{1}{16}\big(P^2(-1,0)+P^2(1,0)-P^2(0,-1)-P^2(0,1)\big)$
xy	$\dfrac{1}{16}\left(P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$
y^2	$\dfrac{1}{2}\big(P(0,0)\big)+\dfrac{1}{16}\big(P^2(0,-1)+P^2(0,1)-P^2(-1,0)-P^2(1,0)\big)$
x^3	$\dfrac{3}{16}\big(P(1,0)-P(-1,0)\big)+\dfrac{3}{32}\big(P^2(-1,0)-P^2(1,0)\big)$ $+\dfrac{1}{48}\left(P^3(-1,0)-P^3(1,0)+P^3\left(\dfrac{1}{2},\dfrac{\sqrt{3}}{2}\right)-P^3\left(\dfrac{-1}{2},\dfrac{-\sqrt{3}}{2}\right)+P^3\left(\dfrac{1}{2},\dfrac{-\sqrt{3}}{2}\right)-P^3\left(\dfrac{-1}{2},\dfrac{\sqrt{3}}{2}\right)\right)$
x^2y	$\dfrac{1}{16}\big(P(0,1)-P(0,-1)\big)+\dfrac{1}{32}\big(P^2(0,-1)-P^2(0,1)\big)$ $+\dfrac{1}{48}\left(P^3(0,1)-P^3(0,-1)+P^3\left(\dfrac{-\sqrt{3}}{2},\dfrac{-1}{2}\right)-P^3\left(\dfrac{\sqrt{3}}{2},\dfrac{1}{2}\right)+P^3\left(\dfrac{\sqrt{3}}{2},\dfrac{-1}{2}\right)-P^3\left(\dfrac{-\sqrt{3}}{2},\dfrac{1}{2}\right)\right)$
xy^2	$\dfrac{1}{16}\big(P(1,0)-P(-1,0)\big)+\dfrac{1}{32}\big(P^2(-1,0)-P^2(1,0)\big)$ $+\dfrac{1}{48}\left(P^3(1,0)-P^3(-1,0)+P^3\left(\dfrac{-1}{2},\dfrac{\sqrt{3}}{2}\right)-P^3\left(\dfrac{1}{2},\dfrac{-\sqrt{3}}{2}\right)+P^3\left(\dfrac{-1}{2},\dfrac{-\sqrt{3}}{2}\right)-P^3\left(\dfrac{1}{2},\dfrac{\sqrt{3}}{2}\right)\right)$
y^3	$\dfrac{3}{16}\big(P(0,1)-P(0,-1)\big)+\dfrac{3}{32}\big(P^2(0,-1)-P^2(0,1)\big)$ $+\dfrac{1}{48}\left(P^3(0,-1)-P^3(0,1)+P^3\left(\dfrac{\sqrt{3}}{2},\dfrac{1}{2}\right)-P^3\left(\dfrac{-\sqrt{3}}{2},\dfrac{-1}{2}\right)+P^3\left(\dfrac{-\sqrt{3}}{2},\dfrac{1}{2}\right)-P^3\left(\dfrac{\sqrt{3}}{2},\dfrac{-1}{2}\right)\right)$

Table IV.2 Monomials and their decentered shape term equivalent, cont.

Monom.	Sum of Decentered Shape Terms
x^4	$\dfrac{3}{8}P^2(0,0)+\dfrac{1}{16}\left(P^2(0,-1)+P^2(0,1)-P^2(-1,0)-P^2(1,0)\right)$ $+\dfrac{1}{48}\left(P^3(-1,0)+P^3(1,0)-P^3(0,-1)-P^3(0,1)\right)$ $+\dfrac{1}{128}\left(P^4(0,-1)+P^4(0,1)+P^4(-1,0)+P^4(1,0)\right.$ $\left.-P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$
x^3y	$\dfrac{1}{32}\left(-P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)+P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)\right)$ $+\dfrac{1}{96}\left(P^3\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^3\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^3\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^3\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)\right)$ $+\dfrac{1}{128}\left(P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{2},\dfrac{-\sqrt{2-\sqrt{2}}}{2}\right)-P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{2},\dfrac{-\sqrt{2+\sqrt{2}}}{2}\right)+P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{2},\dfrac{-\sqrt{2+\sqrt{2}}}{2}\right)-P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{-\sqrt{2-\sqrt{2}}}{2}\right)\right.$ $\left.+P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)-P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)+P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)-P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)\right)$
x^2y^2	$\dfrac{1}{8}P^2(0,0)+\dfrac{1}{128}\left(-P^4(-1,0)-P^4(1,0)-P^4(0,-1)-P^4(0,1)\right.$ $\left.+P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)+P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)\right)$
xy^3	$\dfrac{1}{32}\left(-P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)+P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)+P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$ $+\dfrac{1}{96}\left(P^3\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^3\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^3\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^3\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$ $+\dfrac{1}{128}\left(P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{2},\dfrac{-\sqrt{2+\sqrt{2}}}{2}\right)-P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{2},\dfrac{-\sqrt{2-\sqrt{2}}}{2}\right)+P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)-P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)\right.$ $\left.+P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)-P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)+P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{-\sqrt{2-\sqrt{2}}}{2}\right)-P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{2},\dfrac{-\sqrt{2+\sqrt{2}}}{2}\right)\right)$
y^4	$\dfrac{3}{8}P^2(0,0)+\dfrac{1}{16}\left(P^2(-1,0)+P^2(1,0)-P^2(0,-1)-P^2(0,1)\right)$ $+\dfrac{1}{48}\left(P^3(0,-1)+P^3(0,1)-P^3(-1,0)-P^3(1,0)\right)$ $+\dfrac{1}{128}\left(P^4(-1,0)+P^4(1,0)+P^4(0,-1)+P^4(0,1)\right.$ $\left.-P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)\right)$

Table IV.3 lists the Zernike polynomials that have been expressed using the monomials listed in Table IV.1 and Table IV.2.

Table IV.3 0-P normalized Zernike polynomials and their binomial representation. The y-axis is being used as the reference axis for the Zernike azimuthal parameter.

Stnd.	Polar Form	Sum of Monomials
Z_0^0	1	1
Z_1^{-1}	$\rho \sin(\theta)$	x
Z_1^1	$\rho \cos(\theta)$	y
Z_2^{-2}	$\rho^2 \sin(2\theta)$	$2xy$
Z_2^0	$2\rho^2 - 1$	$2x^2 + 2y^2 - 1$
Z_2^2	$\rho^2 \cos(2\theta)$	$-x^2 + y^2$
Z_3^{-3}	$\rho^3 \sin(3\theta)$	$-x^3 + 3xy^2$
Z_3^{-1}	$(3\rho^3 - 2\rho)\sin(\theta)$	$3x^3 - 2x + 3xy^2$
Z_3^1	$(3\rho^3 - 2\rho)\cos(\theta)$	$3y^3 - 2y + 3x^2 y$
Z_3^3	$\rho^3 \cos(3\theta)$	$-3x^2 y + y^3$
Z_4^{-4}	$\rho^4 \sin(4\theta)$	$-4x^3 y + 4xy^3$
Z_4^{-2}	$(4\rho^4 - 3\rho^2)\sin(2\theta)$	$8x^3 y - 6xy + 8xy^3$
Z_4^0	$6\rho^4 - 6\rho^2 + 1$	$6x^4 - 6x^2 - 6y^2 + 12x^2 y^2 + 6y^4 + 1$
Z_4^2	$(4\rho^4 - 3\rho^2)\cos(2\theta)$	$-4x^4 + 3x^2 - 3y^2 + 4y^4$
Z_4^4	$\rho^4 \cos(4\theta)$	$x^4 - 6x^2 y^2 + y^4$

Table IV.4 gives explicitly the Zernike polynomials in terms of the decentered shape terms. Table IV.4 uses the y-axis as the Zernike reference axis.

Table IV.4 0-P normalized Zernike polynomials composed of decentered shape terms. The _y_-axis is the reference axis.
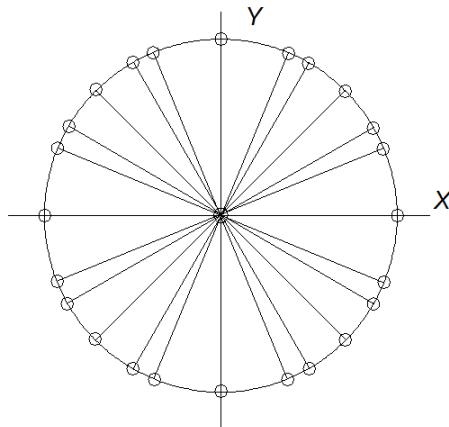
Stnd.	Sum of Decentered Shape Terms
Z_0^0	$P^0(0,0)$
Z_1^{-1}	$\dfrac{1}{4}\big(P(-1,0)-P(1,0)\big)$
Z_1^1	$\dfrac{1}{4}\big(P(0,-1)-P(0,1)\big)$
Z_2^{-2}	$\dfrac{1}{8}\left(P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$
Z_2^0	$-P^0(0,0)+2P(0,0)$
Z_2^2	$\dfrac{1}{8}\big(-P^2(-1,0)-P^2(1,0)+P^2(0,1)+P^2(0,-1)\big)$
Z_3^{-3}	$\dfrac{1}{12}\left(P^3(1,0)-P^3(-1,0)+P^3\left(\dfrac{-1}{2},\dfrac{-\sqrt{3}}{2}\right)-P^3\left(\dfrac{1}{2},\dfrac{\sqrt{3}}{2}\right)+P^3\left(\dfrac{-1}{2},\dfrac{\sqrt{3}}{2}\right)-P^3\left(\dfrac{1}{2},\dfrac{-\sqrt{3}}{2}\right)\right)$
Z_3^{-1}	$\dfrac{5}{4}\big(P(1,0)-P(-1,0)\big)+\dfrac{3}{8}\big(P^2(-1,0)-P^2(1,0)\big)$
Z_3^1	$\dfrac{5}{4}\big(P(0,1)-P(0,-1)\big)+\dfrac{3}{8}\big(P^2(0,-1)-P^2(0,1)\big)$
Z_3^3	$\dfrac{1}{12}\left(P^3(0,-1)-P^3(0,1)+P^3\left(\dfrac{\sqrt{3}}{2},\dfrac{1}{2}\right)-P^3\left(\dfrac{-\sqrt{3}}{2},\dfrac{-1}{2}\right)+P^3\left(\dfrac{-\sqrt{3}}{2},\dfrac{1}{2}\right)-P^3\left(\dfrac{\sqrt{3}}{2},\dfrac{-1}{2}\right)\right)$
Z_4^{-4}	$\dfrac{1}{16}\left(P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{2},\dfrac{-\sqrt{2+\sqrt{2}}}{2}\right)-P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{2},\dfrac{-\sqrt{2-\sqrt{2}}}{2}\right)+P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)-P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)\right.$ $\left.+P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)-P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)+P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{-\sqrt{2-\sqrt{2}}}{2}\right)-P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{2},\dfrac{-\sqrt{2+\sqrt{2}}}{2}\right)\right)$
Z_4^{-2}	$\dfrac{1}{6}\left(P^3\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^3\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^3\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^3\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$ $+\dfrac{7}{8}\left(-P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)+P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)+P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$
Z_4^0	$6P^2(0,0)-6P(0,0)+P^0(0,0)$
Z_4^2	$\dfrac{1}{6}\big(-P^3(-1,0)-P^3(1,0)+P^3(0,-1)+P^3(0,1)\big)+\dfrac{7}{8}\big(P^2(1,0)+P^2(-1,0)-P^2(0,1)-P^2(0,-1)\big)$
Z_4^4	$\dfrac{1}{16}\left(P^4(0,1)+P^4(0,-1)+P^4(1,0)+P^4(-1,0)-P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)\right)$

Table IV.5 provides the same decentered shape term tabulation of the Zernike polynomials now using the _x_-axis as the reference axis.

Table IV.5 0-P normalized Zernike polynomials composed of decentered shape term using the *x*-axis as reference axis.

Stnd.	Decentered Shape Term Equivalent
Z_0^0	$P^0(0,0)$
Z_1^{-1}	$\dfrac{1}{4}\left(P(0,-1)-P(0,1)\right)$
Z_1^1	$\dfrac{1}{4}\left(P(-1,0)-P(1,0)\right)$
Z_2^{-2}	$\dfrac{1}{8}\left(P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$
Z_2^0	$-P^0(0,0)+2P(0,0)$
Z_2^2	$\dfrac{1}{8}\left(P^2(-1,0)+P^2(1,0)-P^2(0,1)-P^2(0,-1)\right)$
Z_3^{-3}	$\dfrac{1}{12}\left(P^3(0,1)-P^3(0,-1)+P^3\left(\dfrac{-\sqrt{3}}{2},\dfrac{-1}{2}\right)-P^3\left(\dfrac{\sqrt{3}}{2},\dfrac{1}{2}\right)+P^3\left(\dfrac{\sqrt{3}}{2},\dfrac{-1}{2}\right)-P^3\left(\dfrac{-\sqrt{3}}{2},\dfrac{1}{2}\right)\right)$
Z_3^{-1}	$\dfrac{5}{4}\left(P(0,1)-P(0,-1)\right)+\dfrac{3}{8}\left(P^2(0,-1)-P^2(0,1)\right)$
Z_3^1	$\dfrac{5}{4}\left(P(1,0)-P(-1,0)\right)+\dfrac{3}{8}\left(P^2(-1,0)-P^2(1,0)\right)$
Z_3^3	$\dfrac{1}{12}\left(P^3(-1,0)-P^3(1,0)+P^3\left(\dfrac{1}{2},\dfrac{\sqrt{3}}{2}\right)-P^3\left(\dfrac{-1}{2},\dfrac{-\sqrt{3}}{2}\right)+P^3\left(\dfrac{1}{2},\dfrac{-\sqrt{3}}{2}\right)-P^3\left(\dfrac{-1}{2},\dfrac{\sqrt{3}}{2}\right)\right)$
Z_4^{-4}	$\dfrac{1}{16}\left(P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)-P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)+P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{2},\dfrac{-\sqrt{2+\sqrt{2}}}{2}\right)-P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{2},\dfrac{-\sqrt{2-\sqrt{2}}}{2}\right)\right.$ $\left.+P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{2},\dfrac{-\sqrt{2-\sqrt{2}}}{2}\right)-P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{2},\dfrac{-\sqrt{2+\sqrt{2}}}{2}\right)+P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{2},\dfrac{\sqrt{2+\sqrt{2}}}{2}\right)-P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{2},\dfrac{\sqrt{2-\sqrt{2}}}{2}\right)\right)$
Z_4^{-2}	$\dfrac{1}{6}\left(P^3\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)+P^3\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^3\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^3\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$ $+\dfrac{7}{8}\left(-P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)+P^2\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)+P^2\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$
Z_4^0	$6P^2(0,0)-6P(0,0)+P^0(0,0)$
Z_4^2	$\dfrac{1}{6}\left(-P^3(0,1)-P^3(0,-1)+P^3(1,0)+P^3(-1,0)\right)+\dfrac{7}{8}\left(P^2(0,-1)+P^2(0,1)-P^2(-1,0)-P^2(1,0)\right)$
Z_4^4	$\dfrac{1}{16}\left(P^4(-1,0)+P^4(1,0)+P^4(0,-1)+P^4(0,1)-P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)-P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^4\left(\dfrac{-1}{\sqrt{2}},\dfrac{1}{\sqrt{2}}\right)-P^4\left(\dfrac{1}{\sqrt{2}},\dfrac{-1}{\sqrt{2}}\right)\right)$

Note that for each case shown in Table IV.4 and Table IV.5 either $\delta x^2 + \delta y^2 = 1$

or $\delta x^2 + \delta y^2 = 0$. That is, the decentered shape terms are all decentered to the perimeter

of a unit radius disk, or are not decentered at all. This is shown in Figure IV.2.

Figure IV.2 The small circles indicate the positions of the decentered shape terms for the Zernike Polynomials listed in Table IV.4.

Upon further investigation, it has been found that the position of the decentered

shape terms need not be restricted to the perimeter of a unit radius disk. It appears that

any convenient radial position can be selected. For example, selecting a radial

decentering of $1/10^{th}$ of the unit radius results in the following decentered shape term

sums for the Zernike polynomials of Table IV.6.

Table IV.6 0-P normalized Zernike polynomials composed of decentered shape terms using the *x*-axis as reference axis, and a decenter amount of 1/10.

Stnd.	Decentered Shape Terms Equivalent
Z_0^0	$P^0(0,0)$
Z_1^1	$\dfrac{5}{2}\left(P\left(\dfrac{-1}{10},0\right)-P\left(\dfrac{1}{10},0\right)\right)$
Z_1^{-1}	$\dfrac{5}{2}\left(P\left(0,\dfrac{-1}{10}\right)-P\left(0,\dfrac{1}{10}\right)\right)$
Z_2^0	$-P^0(0,0)+2P(0,0)$
Z_2^2	$\dfrac{25}{2}\left(P^2\left(\dfrac{-1}{10},0\right)+P^2\left(\dfrac{1}{10},0\right)-P^2\left(0,\dfrac{1}{10}\right)-P^2\left(0,\dfrac{-1}{10}\right)\right)$
Z_2^{-2}	$\dfrac{25}{2}\left(P^2\left(\dfrac{-1}{10\sqrt{2}},\dfrac{-1}{10\sqrt{2}}\right)+P^2\left(\dfrac{1}{10\sqrt{2}},\dfrac{1}{10\sqrt{2}}\right)-P^2\left(\dfrac{-1}{10\sqrt{2}},\dfrac{1}{10\sqrt{2}}\right)-P^2\left(\dfrac{1}{10\sqrt{2}},\dfrac{-1}{10\sqrt{2}}\right)\right)$
Z_3^1	$\dfrac{203}{40}\left(P\left(\dfrac{1}{10},0\right)-P\left(\dfrac{-1}{10},0\right)\right)+\dfrac{15}{4}\left(P^2\left(\dfrac{-1}{10},0\right)-P^2\left(\dfrac{1}{10},0\right)\right)$
Z_3^{-1}	$\dfrac{203}{40}\left(P\left(0,\dfrac{1}{10}\right)-P\left(0,\dfrac{-1}{10}\right)\right)+\dfrac{15}{4}\left(P^2\left(0,\dfrac{-1}{10}\right)-P^2\left(0,\dfrac{1}{10}\right)\right)$
Z_3^3	$\dfrac{250}{3}\left(P^3\left(\dfrac{-1}{10},0\right)-P^3\left(\dfrac{1}{10},0\right)+P^3\left(\dfrac{1}{20},\dfrac{\sqrt{3}}{20}\right)-P^3\left(\dfrac{-1}{20},\dfrac{-\sqrt{3}}{20}\right)+P^3\left(\dfrac{1}{20},\dfrac{-\sqrt{3}}{20}\right)-P^3\left(\dfrac{-1}{20},\dfrac{\sqrt{3}}{20}\right)\right)$
Z_3^{-3}	$\dfrac{250}{3}\left(P^3\left(0,\dfrac{1}{10}\right)-P^3\left(0,\dfrac{-1}{10}\right)+P^3\left(\dfrac{-\sqrt{3}}{20},\dfrac{-1}{20}\right)-P^3\left(\dfrac{\sqrt{3}}{20},\dfrac{1}{20}\right)+P^3\left(\dfrac{\sqrt{3}}{20},\dfrac{-1}{20}\right)-P^3\left(\dfrac{-\sqrt{3}}{20},\dfrac{1}{20}\right)\right)$
Z_4^0	$6P^2(0,0)-6P(0,0)+P^0(0,0)$
Z_4^2	$\dfrac{50}{3}\left(-P^3\left(0,\dfrac{1}{10}\right)-P^3\left(0,\dfrac{-1}{10}\right)+P^3\left(\dfrac{1}{10},0\right)+P^3\left(\dfrac{-1}{10},0\right)\right)+38\left(P^2\left(0,\dfrac{-1}{10}\right)+P^2\left(0,\dfrac{1}{10}\right)-P^2\left(\dfrac{-1}{10},0\right)-P^2\left(\dfrac{1}{10},0\right)\right)$
Z_4^{-2}	$\dfrac{50}{3}\left(P^3\left(\dfrac{-1}{10\sqrt{2}},\dfrac{-1}{10\sqrt{2}}\right)+P^3\left(\dfrac{1}{10\sqrt{2}},\dfrac{1}{10\sqrt{2}}\right)-P^3\left(\dfrac{-1}{10\sqrt{2}},\dfrac{1}{10\sqrt{2}}\right)-P^3\left(\dfrac{1}{10\sqrt{2}},\dfrac{-1}{10\sqrt{2}}\right)\right)$ $+38\left(-P^2\left(\dfrac{-1}{10\sqrt{2}},\dfrac{-1}{10\sqrt{2}}\right)-P^2\left(\dfrac{1}{10\sqrt{2}},\dfrac{1}{10\sqrt{2}}\right)+P^2\left(\dfrac{-1}{10\sqrt{2}},\dfrac{1}{10\sqrt{2}}\right)+P^2\left(\dfrac{1}{10\sqrt{2}},\dfrac{-1}{10\sqrt{2}}\right)\right)$
Z_4^4	$625\left(P^4\left(\dfrac{-1}{10},0\right)+P^4\left(\dfrac{1}{10},0\right)+P^4\left(0,\dfrac{-1}{10}\right)+P^4\left(0,\dfrac{1}{10}\right)-P^4\left(\dfrac{-1}{10\sqrt{2}},\dfrac{-1}{10\sqrt{2}}\right)-P^4\left(\dfrac{1}{10\sqrt{2}},\dfrac{1}{10\sqrt{2}}\right)-P^4\left(\dfrac{-1}{10\sqrt{2}},\dfrac{1}{10\sqrt{2}}\right)-P^4\left(\dfrac{1}{10\sqrt{2}},\dfrac{-1}{10\sqrt{2}}\right)\right)$
Z_4^{-4}	$625\left(P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{20},\dfrac{\sqrt{2-\sqrt{2}}}{20}\right)-P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{20},\dfrac{\sqrt{2+\sqrt{2}}}{20}\right)+P^4\left(\dfrac{\sqrt{2-\sqrt{2}}}{20},\dfrac{-\sqrt{2+\sqrt{2}}}{20}\right)-P^4\left(\dfrac{\sqrt{2+\sqrt{2}}}{20},\dfrac{-\sqrt{2-\sqrt{2}}}{20}\right)\right.$ $\left.+P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{20},\dfrac{-\sqrt{2-\sqrt{2}}}{20}\right)-P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{20},\dfrac{-\sqrt{2+\sqrt{2}}}{20}\right)+P^4\left(\dfrac{-\sqrt{2-\sqrt{2}}}{20},\dfrac{\sqrt{2+\sqrt{2}}}{20}\right)-P^4\left(\dfrac{-\sqrt{2+\sqrt{2}}}{20},\dfrac{\sqrt{2-\sqrt{2}}}{20}\right)\right)$

An example of utilizing this property of the Zernike polynomials together with NAT to describe the field dependence of the wavefront aberration function for optical elements having a freeform surface shape defined by the Zernike polynomials is presented in Chapter 7.

Appendix V. Wavefront and Transverse Ray Coefficients

In his book [53] Cox provides a derivation for the expansion of the transverse ray aberration function, $\vec{\epsilon}$, through 5th order that is independent of the wavefront aberration function. Buchdahl [67] provides a method for calculating the resulting transverse ray aberration function's expansion coefficients. In an attempt to obtain equations for the 6th order wavefront aberration function's expansion coefficients, W_{klm}, it might be thought that the equation

$$\vec{\epsilon}\left(H_x, H_y, \rho_x, \rho_y\right) = \frac{1}{n'u'} \nabla_{\rho_x, \rho_y} W\left(H_x, H_y, \rho_x, \rho_y\right) , \tag{V.1}$$

may be used to obtain expressions for the transverse ray aberration function expansion in terms of the wavefront aberration function expansion. One would then only have to associate Cox's results with the gradient equation results to obtain equations for the W_{klm} expansion coefficients in terms of the transverse ray expansion coefficients. However, it is shown here that this approach leads to inconsistent equations for some of the 6th order W_{klm} coefficients.

Cox's equations can be stated as follows

$$\begin{aligned}
\epsilon_x = {} & A_1 \rho^3 \sin(\varphi) + A_2 H \rho^2 \sin(2\varphi) + A_3 H^2 \rho \sin(\varphi) \\
& + B_1 \rho^5 \sin(\varphi) + B_2 H \rho^4 \sin(2\varphi) + B_3 H^2 \rho^3 \sin(\varphi) \\
& + B_4 H^2 \rho^3 \cos^2(\varphi) \sin(\varphi) + B_5 H^3 \rho^2 \sin(2\varphi) + B_6 H^4 \rho \sin(\varphi) ,
\end{aligned} \tag{V.2}$$

$$\epsilon_y = A_1\rho^3\cos(\varphi) + A_2\left(2+\cos(2\varphi)\right)H\rho^2 + A_4 H^2\rho\cos(\varphi) + A_5 H^3$$
$$+ B_1\rho^5\cos(\varphi) + B_2\left(\frac{3}{2}+\cos(2\varphi)\right)H\rho^4 - \frac{1}{2d'^2}(A_1+A_2)H\rho^4$$
$$B_4 H^2\rho^3\left(\cos^2(\varphi)+1\right)\cos(\varphi)$$
$$+ B_3 H^2\rho^3\cos(\varphi) + \frac{1}{d'^2}\left(A_1 - \frac{1}{2}A_4 + \frac{1}{2}A_3\right)H^2\rho^3\cos(\varphi)$$
$$+ \left[(2B_5+B_7)+(B_5+B_7)\cos(2\varphi)\right]H^3\rho^2 + \frac{1}{d'^2}\left(A_2+\frac{1}{2}A_4-\frac{1}{2}A_3\right)H^3\rho^2$$
$$+ B_8 H^4\rho\cos(\varphi) + B_9 H^5 , \tag{V.3}$$

where A_i are the 3rd order transverse ray expansion coefficients, B_i are the 5th order transverse ray expansion coefficients, $\vec{\rho}=(\rho,\varphi)=(\rho_x,\rho_y)$ are the pupil parameters with the angle φ measured from the y-axis, H is the field parameter, and d' is the distance from the exit pupil to the Gaussian image plane. The notation used for Eq. (V.2) and Eq. (V.3) is substantially different from that employed by Cox in order to match the notation used throughout this dissertation.

Buchdahl [67, 68] employs a different set of expansion coefficients. He uses σ_i for the 3rd order coefficients and μ_i for the 5th order coefficients. The conversions are as follows:

$$\sigma_1 = A_1 , \tag{V.4}$$

$$\sigma_2 = A_2 , \tag{V.5}$$

$$\sigma_3 = \frac{1}{2}(A_4 - A_3), \tag{V.6}$$

$$\sigma_4 = \frac{1}{2}(3A_3 - A_4), \tag{V.7}$$

$$\sigma_5 = A_5 \ , \tag{V.8}$$

$$\mu_1 = B_1 \ , \tag{V.9}$$

$$\mu_2 = \frac{3}{2}B_2 - \frac{1}{2d'^2}\left(A_1 + A_2\right) \ , \tag{V.10}$$

$$\mu_3 = B_2 \ , \tag{V.11}$$

$$\mu_4 = B_4 + B_3 + \frac{1}{d'^2}\left(A_1 - \frac{1}{2}A_4 + \frac{1}{2}A_3\right) \ , \tag{V.12}$$

$$\mu_5 = B_3 \ , \tag{V.13}$$

$$\mu_6 = B_4 \ , \tag{V.14}$$

$$\mu_7 = 2B_5 + B_7 + \frac{1}{d'^2}\left(A_2 + \frac{1}{2}A_4 - \frac{1}{2}A_3\right) \ , \tag{V.15}$$

$$\mu_8 = \left(B_5 + B_7\right) \ , \tag{V.16}$$

$$\mu_9 = B_5 \ , \tag{V.17}$$

$$\mu_{10} = B_8 \ , \tag{V.18}$$

$$\mu_{11} = B_6 \ , \tag{V.19}$$

$$\mu_{12} = B_9 \ . \tag{V.20}$$

It is important to note that some of the 5th order expansion coefficients of Buchdahl depend on 3rd order expansion coefficients.

Using Buchdahl's expansion coefficients, the transverse ray aberration expansion functions can be written as [67, 68, 77]

$$\epsilon_x = \sigma_1 \sin(\varphi)\rho^3 + \sigma_2 \sin(2\varphi)\rho^2 H + (\sigma_3 + \sigma_4)\sin(\varphi)\rho H^2$$
$$+ \mu_1 \sin(\varphi)\rho^5 + \mu_3 \sin(2\varphi)\rho^4 H + (\mu_5 + \mu_6 \cos^2(\varphi))\sin(\varphi)\rho^3 H^2$$
$$+ \mu_9 \sin(2\varphi)\rho^2 H^3 + \mu_{11}\sin(\varphi)\rho H^4 , \tag{V.21}$$

$$\epsilon_y = \sigma_1 \cos(\varphi)\rho^3 + \sigma_2 \left(2 + \cos(2\varphi)\right)\rho^2 H + (3\sigma_3 + \sigma_4)\cos(\varphi)\rho H^2 + \sigma_5 H^3$$
$$+ \mu_1 \cos(\varphi)\rho^5 + (\mu_2 + \mu_3 \cos(2\varphi))\rho^4 H + (\mu_4 + \mu_6 \cos^2(\varphi))\cos(\varphi)\rho^3 H^2$$
$$+ (\mu_7 + \mu_8 \cos(2\varphi))\rho^2 H^3 + \mu_{10}\cos(\varphi)\rho H^4 + \mu_{12}H^5 . \tag{V.22}$$

For a rotationally symmetric optical imaging system, restricting the field parameter to be along the y-axis, the wavefront aberration function expansion through 6th order can be written as

$$W = W_{040}\rho^4 + W_{131}H\rho^3\cos(\varphi_y) + W_{222}H^2\rho^2\cos^2(\varphi_y) + W_{220S}H^2\rho^2 + W_{311}H^3\rho\cos(\varphi_y)$$
$$+ W_{060}\rho^6 + W_{151}H\rho^5\cos(\varphi_y) + W_{242}H^2\rho^4\cos^2(\varphi_y) + W_{240S}H^2\rho^4 + W_{333}H^3\rho^3\cos^3(\varphi_y)$$
$$+ W_{331}H^3\rho^3\cos(\varphi_y) + W_{420S}H^4\rho^2 + W_{422}H^4\rho^2\cos^2(\varphi_y) + W_{511}H^5\rho\cos(\varphi_y) , \tag{V.23}$$

where

$$\rho^2 = \rho_x^2 + \rho_y^2 . \tag{V.24}$$

The derivatives of W with respect to ρ_x and ρ_y can be calculated to be

$$\frac{\partial W}{\partial \rho_x} = 4W_{040}\sin(\varphi_y)\rho^3 + W_{131}\sin(2\varphi_y)\rho^2 H + 2W_{220S}\sin(\varphi_y)\rho H^2$$
$$+ 6W_{060}\sin(\varphi_y)\rho^5 + 2W_{151}\sin(2\varphi_y)\rho^4 H + (4W_{240S} + 2W_{242}\cos^2(\varphi_y))\sin(\varphi_y)\rho^3 H^2$$
$$+ W_{331}\sin(2\varphi_y)\rho^2 H^3 + 2W_{420S}\sin(\varphi_y)\rho H^4 , \tag{V.25}$$

and

$$\frac{\partial W}{\partial \rho_y} = 4W_{040}\cos\left(\varphi_y\right)\rho^3 + W_{131}\left(2+\cos\left(2\varphi_y\right)\right)\rho^2 H + 2\left(W_{222}+W_{220S}\right)\cos\left(\varphi_y\right)\rho H^2$$

$$+ W_{311}H^3 + 6W_{060}\cos\left(\varphi_y\right)\rho^5 + W_{151}\left(3+2\cos\left(2\varphi_y\right)\right)\rho^4 H$$

$$+ \left(2W_{242}+4W_{240S}+2W_{242}\cos^2\left(\varphi_y\right)\right)\cos\left(\varphi_y\right)\rho^3 H^2$$

$$+ \left(\frac{3}{2}W_{333}+2W_{331}+\left(\frac{3}{2}W_{333}+W_{331}\right)\cos\left(2\varphi_y\right)\right)\rho^2 H^3$$

$$+ 2\left(W_{420S}+W_{422}\right)\cos\left(\varphi_y\right)\rho H^4 + W_{511}H^5 . \tag{V.26}$$

Using these results in Eq. (V.1) gives

$$\epsilon_x = \frac{1}{n_k' u_k'}\left[4W_{040}\sin\left(\varphi_y\right)\rho^3 + W_{131}\sin\left(2\varphi_y\right)\rho^2 H + 2W_{220S}\sin\left(\varphi_y\right)\rho H^2\right.$$

$$+ 6W_{060}\sin\left(\varphi_y\right)\rho^5 + 2W_{151}\sin\left(2\varphi_y\right)\rho^4 H$$

$$+ \left(4W_{240S}+2W_{242}\cos^2\left(\varphi_y\right)\right)\sin\left(\varphi_y\right)\rho^3 H^2$$

$$\left. + W_{331}\sin\left(2\varphi_y\right)\rho^2 H^3 + 2W_{420S}\sin\left(\varphi_y\right)\rho H^4\right], \tag{V.27}$$

and

$$\epsilon_y = \frac{1}{n_k' u_k'}\left[4W_{040}\cos\left(\varphi_y\right)\rho^3 + W_{131}\left(2+\cos\left(2\varphi_y\right)\right)\rho^2 H + 2\left(W_{222}+W_{220S}\right)\cos\left(\varphi_y\right)\rho H^2\right.$$

$$+ W_{311}H^3 + 6W_{060}\cos\left(\varphi_y\right)\rho^5 + W_{151}\left(3+2\cos\left(2\varphi_y\right)\right)\rho^4 H$$

$$+ \left(2W_{242}+4W_{240S}+2W_{242}\cos^2\left(\varphi_y\right)\right)\cos\left(\varphi_y\right)\rho^3 H^2$$

$$+ \left(\frac{3}{2}W_{333}+2W_{331}+\left(\frac{3}{2}W_{333}+W_{331}\right)\cos\left(2\varphi_y\right)\right)\rho^2 H^3$$

$$\left. + 2\left(W_{420S}+W_{422}\right)\cos\left(\varphi_y\right)\rho H^4 + W_{511}H^5\right]. \tag{V.28}$$

Comparing like terms in Eq. (V.21) and Eq. (V.27) as well as in Eq. (V.22) and Eq. (V.28), for the case of W_{151} the following associations are made

$$\mu_3\sin\left(2\varphi\right)\rho^4 H = \frac{2}{n_k' u_k'}W_{151}\sin\left(2\varphi_y\right)\rho^4 H , \tag{V.29}$$

$$\left(\mu_2 + \mu_3 \cos\left(2\varphi\right)\right)\rho^4 H = \frac{1}{n_k' u_k'} W_{151}\left(3 + 2\cos\left(2\varphi_y\right)\right)\rho^4 H \quad . \tag{V.30}$$

From the first equation, the following relation is obtained,

$$W_{151} = \frac{n_k' u_k'}{2}\mu_3 \quad . \tag{V.31}$$

The second equation can be written as

$$\frac{3}{n_k' u_k'}W_{151} + \frac{2}{n_k' u_k'}W_{151}\cos\left(2\varphi_y\right) = \mu_2 + \mu_3\cos\left(2\varphi\right) \quad . \tag{V.32}$$

It then follows that

$$W_{151} = \frac{n_k' u_k'}{3}\mu_2 \quad . \tag{V.33}$$

Combining Eq. (V.31) and Eq. (V.33) leads to

$$\mu_2 = \frac{3}{2}\mu_3 \quad . \tag{V.34}$$

This relation is only true for special cases and is not true in general. Therefore, the equations developed for W_{151} above only hold for special cases and are not consistent generally. Buchdahl, p. 47 [67] writes:

"If the primary coefficients are negligible, we have

$$\left.\begin{array}{l} \mu_2 = \dfrac{3}{2}\mu_3 \\[2mm] \mu_4 = \mu_5 + \mu_6 \\[2mm] \mu_7 = \mu_8 + \mu_9 \end{array}\right\} \tag{31.81}$$

and these identities will be approximately valid for well-corrected systems."

By "primary coefficients" Buchdahl is referring to the 3rd order transverse ray aberration coefficients. From Eq. (V.10) and Eq. (V.11) it can be seen that when the 3rd order coefficients A_1 and A_2 are zero then Eq. (V.34) will hold. However, this is a special case and is not true for all optical imaging systems.

In a similar way, equations for other 6th order W_{klm} coefficients can be derived that lead to inconsistent equations. Therefore, this approach for obtaining equations for the 6th order W_{klm} expansion coefficients cannot be used.

Appendix VI. Calculating Field Linear, Field Conjugate Astigmatism

In this Appendix, the derivation of field linear, field conjugate astigmatism is presented. This field dependence for the Zernike astigmatism term of the wavefront aberration function's expansion occurs when a $Z_3^{+3}(\vec{\rho})$ cap shape is added to a spherical shaped optical element, as discussed in Chapter 7. Additionally, it is shown that

$$6V_{422}\vec{H}^* + 6V'_{422}\vec{H} = \vec{C} \star_x \vec{H}^* \,, \tag{VI.1}$$

where $6V_{422}\vec{H}^*$ is the field contribution due to adding a $Z_3^{+3}(\vec{\rho})$ shape to an optical element's surface, $6V'_{422}\vec{H}$ is the field contribution due to adding a $Z_3^{-3}(\vec{\rho})$ shape to an optical element's surface, and where $\vec{H}^* \equiv (H_x, -H_y)$, $\vec{H} \equiv (H_y, H_x)$, and $\vec{C} \equiv (6V_{422}, 6V'_{422})$. This provides a connection to the form for field linear, field conjugate astigmatism reported in [72].

It is first shown that

$$\left[V_{222}\left\{\vec{H} - \vec{\sigma}_{A1}\right\}_\eta^2 - V_{222}\left\{\vec{H} + \vec{\sigma}_{A1}\right\}_\eta^2 + V_{222}\left\{\vec{H} - \vec{\sigma}_{A2}\right\}_\eta^2 - V_{222}\left\{\vec{H} + \vec{\sigma}_{A2}\right\}_\eta^2 + \right.$$
$$\left. V_{222}\left\{\vec{H} - \vec{\sigma}_{A3}\right\}_\eta^2 - V_{222}\left\{\vec{H} + \vec{\sigma}_{A3}\right\}_\eta^2 \right] = \vec{0}, \tag{VI.2}$$

and

$$\left[V_{422}\left(\vec{H}-\vec{\sigma}_{A1}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A1}\right)\left\{\vec{H}-\vec{\sigma}_{A1}\right\}_{\eta}^{2}-V_{422}\left(\vec{H}+\vec{\sigma}_{A1}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A1}\right)\left\{\vec{H}+\vec{\sigma}_{A1}\right\}_{\eta}^{2}+ \right.$$

$$V_{422}\left(\vec{H}-\vec{\sigma}_{A2}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A2}\right)\left\{\vec{H}-\vec{\sigma}_{A2}\right\}_{\eta}^{2}-V_{422}\left(\vec{H}+\vec{\sigma}_{A2}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A2}\right)\left\{\vec{H}+\vec{\sigma}_{A2}\right\}_{\eta}^{2}+$$

$$\left. V_{422}\left(\vec{H}-\vec{\sigma}_{A3}\right)\bullet\left(\vec{H}-\vec{\sigma}_{A3}\right)\left\{\vec{H}-\vec{\sigma}_{A3}\right\}_{\eta}^{2}-V_{422}\left(\vec{H}+\vec{\sigma}_{A3}\right)\bullet\left(\vec{H}+\vec{\sigma}_{A3}\right)\left\{\vec{H}+\vec{\sigma}_{A3}\right\}_{\eta}^{2}\right]=6V_{422}\vec{H}^{*}, \quad \text{(VI.3)}$$

where

$$\vec{H}^{*}\equiv\left(H_{x},-H_{y}\right)=\hat{e}_{1}\vec{H}\hat{e}_{1} , \quad \text{(VI.4)}$$

(juxtaposition of vectors indicates the geometric product of GA) and where

$$\vec{\sigma}_{A1}=\left(-1,0\right) , \quad \text{(VI.5)}$$

$$\vec{\sigma}_{A2}=\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right) , \quad \text{(VI.6)}$$

$$\vec{\sigma}_{A3}=\left(\frac{1}{2},\frac{-\sqrt{3}}{2}\right) . \quad \text{(VI.7)}$$

In the derivations below, the following 2D GA relations are often used without explicitly indicating when they are used:

$$\hat{e}_{1}\hat{e}_{1}=\hat{e}_{2}\hat{e}_{2}=1 , \quad \text{(VI.8)}$$

$$\hat{e}_{1}\hat{e}_{2}=-\hat{e}_{2}\hat{e}_{1} . \quad \text{(VI.9)}$$

By Eq. (2.64) and Eq. (2.65)

$$\left\{\vec{H}\pm\vec{\sigma}_{Ai}\right\}_{\eta}^{2}=\left(\vec{H}\pm\vec{\sigma}_{Ai}\right)\hat{e}_{\eta}\left(\vec{H}\pm\vec{\sigma}_{Ai}\right) , \quad \text{(VI.10)}$$

where $\eta=1$ or 2 specifying which axis is used as the reference axis. This can be written as

$$\left\{\vec{H} \pm \vec{\sigma}_{Ai}\right\}^2_\eta = \vec{H}\hat{e}_\eta\left(\vec{H} \pm \vec{\sigma}_{Ai}\right) \pm \vec{\sigma}_{Ai}\hat{e}_\eta\left(\vec{H} \pm \vec{\sigma}_{Ai}\right)$$

$$= \vec{H}\hat{e}_\eta\vec{H} \pm \vec{H}\hat{e}_\eta\vec{\sigma}_{Ai} \pm \vec{\sigma}_{Ai}\hat{e}_\eta\vec{H} + \vec{\sigma}_{Ai}\hat{e}_\eta\vec{\sigma}_{Ai} \quad . \qquad \text{(VI.11)}$$

Since SVP is commutative,

$$\left\{\vec{H} \pm \vec{\sigma}_{Ai}\right\}^2_\eta = \vec{H}\hat{e}_\eta\vec{H} \pm 2\vec{\sigma}_{Ai}\hat{e}_\eta\vec{H} + \vec{\sigma}_{Ai}\hat{e}_\eta\vec{\sigma}_{Ai} \quad . \qquad \text{(VI.12)}$$

Letting $\vec{A}$ stand for the left hand side of Eq. (VI.2), and using Eq. (VI.12), Eq. (VI.2) can now be written as

$$\begin{aligned}
\vec{A} &= V_{222}\left\{\vec{H} - \vec{\sigma}_{A1}\right\}^2_\eta - V_{222}\left\{\vec{H} + \vec{\sigma}_{A1}\right\}^2_\eta \\
&\quad + V_{222}\left\{\vec{H} - \vec{\sigma}_{A2}\right\}^2_\eta - V_{222}\left\{\vec{H} + \vec{\sigma}_{A2}\right\}^2_\eta \\
&\quad + V_{222}\left\{\vec{H} - \vec{\sigma}_{A3}\right\}^2_\eta - V_{222}\left\{\vec{H} + \vec{\sigma}_{A3}\right\}^2_\eta \\
&= V_{222}\left(\vec{H}\hat{e}_\eta\vec{H} - 2\vec{\sigma}_{A1}\hat{e}_\eta\vec{H} + \vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1}\right) - V_{222}\left(\vec{H}\hat{e}_\eta\vec{H} + 2\vec{\sigma}_{A1}\hat{e}_\eta\vec{H} + \vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1}\right) \\
&\quad + V_{222}\left(\vec{H}\hat{e}_\eta\vec{H} - 2\vec{\sigma}_{A2}\hat{e}_\eta\vec{H} + \vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2}\right) - V_{222}\left(\vec{H}\hat{e}_\eta\vec{H} + 2\vec{\sigma}_{A2}\hat{e}_\eta\vec{H} + \vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2}\right) \\
&\quad + V_{222}\left(\vec{H}\hat{e}_\eta\vec{H} - 2\vec{\sigma}_{A3}\hat{e}_\eta\vec{H} + \vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\right) - V_{222}\left(\vec{H}\hat{e}_\eta\vec{H} + 2\vec{\sigma}_{A3}\hat{e}_\eta\vec{H} + \vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\right) \quad . \quad \text{(VI.13)}
\end{aligned}$$

Eq. (VI.13) reduces to

$$\begin{aligned}
\vec{A} &= V_{222}\left(-2\vec{\sigma}_{A1}\hat{e}_\eta\vec{H}\right) - V_{222}\left(2\vec{\sigma}_{A1}\hat{e}_\eta\vec{H}\right) \\
&\quad + V_{222}\left(-2\vec{\sigma}_{A2}\hat{e}_\eta\vec{H}\right) - V_{222}\left(2\vec{\sigma}_{A2}\hat{e}_\eta\vec{H}\right) \\
&\quad + V_{222}\left(-2\vec{\sigma}_{A3}\hat{e}_\eta\vec{H}\right) - V_{222}\left(2\vec{\sigma}_{A3}\hat{e}_\eta\vec{H}\right) \quad . \qquad \text{(VI.14)}
\end{aligned}$$

Regrouping the terms, Eq. (VI.14) can be written as

$$\vec{A} = -4V_{222}\left(\vec{\sigma}_{A1} + \vec{\sigma}_{A2} + \vec{\sigma}_{A3}\right)\hat{e}_\eta\vec{H} \quad . \qquad \text{(VI.15)}$$

And by Eq. (VI.5) through Eq. (VI.7)

$$\vec{\sigma}_{A1} + \vec{\sigma}_{A2} + \vec{\sigma}_{A3} = \vec{0} \quad . \qquad \text{(VI.16)}$$

Therefore, by using Eq. (VI.16) in Eq. (VI.15),

$$\vec{A} = \left[V_{222} \left\{ \vec{H} - \vec{\sigma}_{A1} \right\}_\eta^2 - V_{222} \left\{ \vec{H} + \vec{\sigma}_{A1} \right\}_\eta^2 + V_{222} \left\{ \vec{H} - \vec{\sigma}_{A2} \right\}_\eta^2 - V_{222} \left\{ \vec{H} + \vec{\sigma}_{A2} \right\}_\eta^2 + \right.$$

$$\left. V_{222} \left\{ \vec{H} - \vec{\sigma}_{A3} \right\}_\eta^2 - V_{222} \left\{ \vec{H} + \vec{\sigma}_{A3} \right\}_\eta^2 \right] = \vec{0}. \tag{VI.17}$$

To derive Eq. (VI.3), let $\vec{B}$ stand for the left hand side of the equation and write out the dot products as

$$\left(\vec{H} \pm \vec{\sigma}_{Ai} \right) \bullet \left(\vec{H} \pm \vec{\sigma}_{Ai} \right) = D_i \pm 2 \vec{\sigma}_{Ai} \bullet \vec{H} , \tag{VI.18}$$

where

$$D_i = \vec{H} \bullet \vec{H} + \vec{\sigma}_{Ai} \bullet \vec{\sigma}_{Ai} . \tag{VI.19}$$

To simplify a little, Eq. (VI.12) is written as

$$\left\{ \vec{H} \pm \vec{\sigma}_{Ai} \right\}_\eta^2 = \vec{E}_i \pm 2 \vec{\sigma}_{Ai} \hat{e}_\eta \vec{H} , \tag{VI.20}$$

where

$$\vec{E}_i = \vec{H} \hat{e}_\eta \vec{H} + \vec{\sigma}_{Ai} \hat{e}_\eta \vec{\sigma}_{Ai} . \tag{VI.21}$$

Then

$$\vec{B} = \left[V_{422} \left(\vec{H} - \vec{\sigma}_{A1} \right) \bullet \left(\vec{H} - \vec{\sigma}_{A1} \right) \left\{ \vec{H} - \vec{\sigma}_{A1} \right\}_\eta^2 - V_{422} \left(\vec{H} + \vec{\sigma}_{A1} \right) \bullet \left(\vec{H} + \vec{\sigma}_{A1} \right) \left\{ \vec{H} + \vec{\sigma}_{A1} \right\}_\eta^2 + \right.$$

$$V_{422} \left(\vec{H} - \vec{\sigma}_{A2} \right) \bullet \left(\vec{H} - \vec{\sigma}_{A2} \right) \left\{ \vec{H} - \vec{\sigma}_{A2} \right\}_\eta^2 - V_{422} \left(\vec{H} + \vec{\sigma}_{A2} \right) \bullet \left(\vec{H} + \vec{\sigma}_{A2} \right) \left\{ \vec{H} + \vec{\sigma}_{A2} \right\}_\eta^2 +$$

$$\left. V_{422} \left(\vec{H} - \vec{\sigma}_{A3} \right) \bullet \left(\vec{H} - \vec{\sigma}_{A3} \right) \left\{ \vec{H} - \vec{\sigma}_{A3} \right\}_\eta^2 - V_{422} \left(\vec{H} + \vec{\sigma}_{A3} \right) \bullet \left(\vec{H} + \vec{\sigma}_{A3} \right) \left\{ \vec{H} + \vec{\sigma}_{A3} \right\}_\eta^2 \right]$$

$$= V_{422} \left(D_1 - 2 \vec{\sigma}_{A1} \bullet \vec{H} \right) \left(\vec{E}_1 - 2 \vec{\sigma}_{A1} \hat{e}_\eta \vec{H} \right) - V_{422} \left(D_1 + 2 \vec{\sigma}_{A1} \bullet \vec{H} \right) \left(\vec{E}_1 + 2 \vec{\sigma}_{A1} \hat{e}_\eta \vec{H} \right) +$$

$$V_{422} \left(D_2 - 2 \vec{\sigma}_{A2} \bullet \vec{H} \right) \left(\vec{E}_2 - 2 \vec{\sigma}_{A2} \hat{e}_\eta \vec{H} \right) - V_{422} \left(D_2 + 2 \vec{\sigma}_{A2} \bullet \vec{H} \right) \left(\vec{E}_2 + 2 \vec{\sigma}_{A2} \hat{e}_\eta \vec{H} \right) +$$

$$V_{422} \left(D_3 - 2 \vec{\sigma}_{A3} \bullet \vec{H} \right) \left(\vec{E}_3 - 2 \vec{\sigma}_{A3} \hat{e}_\eta \vec{H} \right) - V_{422} \left(D_3 + 2 \vec{\sigma}_{A3} \bullet \vec{H} \right) \left(\vec{E}_3 + 2 \vec{\sigma}_{A3} \hat{e}_\eta \vec{H} \right) \tag{VI.22}$$

Expanding gives

$$\vec{B} = V_{422}\left(D_1\vec{E}_1 - 2D_1\vec{\sigma}_{A1}\hat{e}_\eta\vec{H} - 2\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_1 + 4\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{\sigma}_{A1}\hat{e}_\eta\vec{H}\right)$$

$$+ V_{422}\left(-D_1\vec{E}_1 - 2D_1\vec{\sigma}_{A1}\hat{e}_\eta\vec{H} - 2\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_1 - 4\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{\sigma}_{A1}\hat{e}_\eta\vec{H}\right)$$

$$+ V_{422}\left(D_2\vec{E}_2 - 2D_2\vec{\sigma}_{A2}\hat{e}_\eta\vec{H} - 2\left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_2 + 4\left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{\sigma}_{A2}\hat{e}_\eta\vec{H}\right)$$

$$+ V_{422}\left(-D_2\vec{E}_2 - 2D_2\vec{\sigma}_{A2}\hat{e}_\eta\vec{H} - 2\left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_2 - 4\left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{\sigma}_{A2}\hat{e}_\eta\vec{H}\right)$$

$$+ V_{422}\left(D_3\vec{E}_3 - 2D_3\vec{\sigma}_{A3}\hat{e}_\eta\vec{H} - 2\left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_3 + 4\left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{\sigma}_{A3}\hat{e}_\eta\vec{H}\right)$$

$$+ V_{422}\left(-D_3\vec{E}_3 - 2D_3\vec{\sigma}_{A3}\hat{e}_\eta\vec{H} - 2\left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_3 - 4\left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{\sigma}_{A3}\hat{e}_\eta\vec{H}\right) \quad . \quad \text{(VI.23)}$$

And simplifying gives

$$\vec{B} = V_{422}\left(-4D_1\vec{\sigma}_{A1}\hat{e}_\eta\vec{H} - 4\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_1\right) + V_{422}\left(-4D_2\vec{\sigma}_{A2}\hat{e}_\eta\vec{H} - 4\left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_2\right)$$

$$+ V_{422}\left(-4D_3\vec{\sigma}_{A3}\hat{e}_\eta\vec{H} - 4\left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_3\right). \quad \text{(VI.24)}$$

Rearranging gives

$$\vec{B} = -4V_{422}D_1\vec{\sigma}_{A1}\hat{e}_\eta\vec{H} - 4V_{422}D_2\vec{\sigma}_{A2}\hat{e}_\eta\vec{H} - 4V_{422}D_3\vec{\sigma}_{A3}\hat{e}_\eta\vec{H}$$

$$- 4V_{422}\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_1 - 4V_{422}\left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_2 - 4V_{422}\left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_3 \ , \quad \text{(VI.25)}$$

and

$$\vec{B} = -4V_{422}\left(D_1\vec{\sigma}_{A1} + D_2\vec{\sigma}_{A2} + D_3\vec{\sigma}_{A3}\right)\hat{e}_\eta\vec{H}$$

$$- 4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_1 + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_2 + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_3\right] . \quad \text{(VI.26)}$$

Because the $\vec{\sigma}_{Ai}$ are unit vectors, (see Eq. (VI.5) through Eq. (VI.7))

$$D_i = \vec{H}\bullet\vec{H} + \vec{\sigma}_{Ai}\bullet\vec{\sigma}_{Ai} = \vec{H}\bullet\vec{H} + 1 \ . \quad \text{(VI.27)}$$

Then

$$\vec{B} = -4V_{422}\left(\left(\vec{H}\bullet\vec{H} + 1\right)\vec{\sigma}_{A1} + \left(\vec{H}\bullet\vec{H} + 1\right)\vec{\sigma}_{A2} + \left(\vec{H}\bullet\vec{H} + 1\right)\vec{\sigma}_{A3}\right)\hat{e}_\eta\vec{H}$$

$$- 4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_1 + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_2 + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_3\right] \quad , \quad \text{(VI.28)}$$

and

$$\vec{B} = -4V_{422}\left(\left(\vec{H}\bullet\vec{H}\right)\vec{\sigma}_{A1} + \vec{\sigma}_{A1} + \left(\vec{H}\bullet\vec{H}\right)\vec{\sigma}_{A2} + \vec{\sigma}_{A2} + \left(\vec{H}\bullet\vec{H}\right)\vec{\sigma}_{A3} + \vec{\sigma}_{A3}\right)\hat{e}_{\eta}\vec{H}$$
$$- 4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_{1} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_{2} + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_{3}\right], \tag{VI.29}$$

$$\vec{B} = -4V_{422}\left(\left(\vec{H}\bullet\vec{H}\right)\left(\vec{\sigma}_{A1} + \vec{\sigma}_{A2} + \vec{\sigma}_{A3}\right) + \left(\vec{\sigma}_{A1} + \vec{\sigma}_{A2} + \vec{\sigma}_{A3}\right)\right)\hat{e}_{\eta}\vec{H}$$
$$- 4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_{1} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_{2} + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_{3}\right]. \tag{VI.30}$$

By Eq. (VI.16) this is

$$\vec{B} = -4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{E}_{1} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{E}_{2} + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{E}_{3}\right]. \tag{VI.31}$$

Inserting the $\vec{E}_i$ gives

$$\vec{B} = -4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\left(\vec{H}\hat{e}_{\eta}\vec{H} + \vec{\sigma}_{A1}\hat{e}_{\eta}\vec{\sigma}_{A1}\right) + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\left(\vec{H}\hat{e}_{\eta}\vec{H} + \vec{\sigma}_{A2}\hat{e}_{\eta}\vec{\sigma}_{A2}\right)\right.$$
$$\left. + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\left(\vec{H}\hat{e}_{\eta}\vec{H} + \vec{\sigma}_{A3}\hat{e}_{\eta}\vec{\sigma}_{A3}\right)\right], \tag{VI.32}$$

$$\vec{B} = -4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{H}\hat{e}_{\eta}\vec{H} + \left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{\sigma}_{A1}\hat{e}_{\eta}\vec{\sigma}_{A1} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{H}\hat{e}_{\eta}\vec{H} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{\sigma}_{A2}\hat{e}_{\eta}\vec{\sigma}_{A2}\right.$$
$$\left. + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{H}\hat{e}_{\eta}\vec{H} + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{\sigma}_{A3}\hat{e}_{\eta}\vec{\sigma}_{A3}\right]. \tag{VI.33}$$

Rearranging gives

$$\vec{B} = -4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{H}\hat{e}_{\eta}\vec{H} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{H}\hat{e}_{\eta}\vec{H} + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{H}\hat{e}_{\eta}\vec{H}\right.$$
$$\left. + \left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{\sigma}_{A1}\hat{e}_{\eta}\vec{\sigma}_{A1} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{\sigma}_{A2}\hat{e}_{\eta}\vec{\sigma}_{A2} + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{\sigma}_{A3}\hat{e}_{\eta}\vec{\sigma}_{A3}\right], \tag{VI.34}$$

$$\vec{B} = -4V_{422}\left[\left(\vec{\sigma}_{A1} + \vec{\sigma}_{A2} + \vec{\sigma}_{A3}\right)\bullet\vec{H}\left(\vec{H}\hat{e}_{\eta}\vec{H}\right)\right.$$
$$\left. + \left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{\sigma}_{A1}\hat{e}_{\eta}\vec{\sigma}_{A1} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{\sigma}_{A2}\hat{e}_{\eta}\vec{\sigma}_{A2} + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{\sigma}_{A3}\hat{e}_{\eta}\vec{\sigma}_{A3}\right], \tag{VI.35}$$

$$\vec{B} = -4V_{422}\left[\left(\vec{\sigma}_{A1}\bullet\vec{H}\right)\vec{\sigma}_{A1}\hat{e}_{\eta}\vec{\sigma}_{A1} + \left(\vec{\sigma}_{A2}\bullet\vec{H}\right)\vec{\sigma}_{A2}\hat{e}_{\eta}\vec{\sigma}_{A2} + \left(\vec{\sigma}_{A3}\bullet\vec{H}\right)\vec{\sigma}_{A3}\hat{e}_{\eta}\vec{\sigma}_{A3}\right]. \tag{VI.36}$$

Using the GA relation Eq. (2.7) to write the dot products as

$$\vec{\sigma}_{Ai} \bullet \vec{H} = \frac{1}{2}\left(\vec{\sigma}_{Ai}\vec{H} + \vec{H}\vec{\sigma}_{Ai}\right) , \tag{VI.37}$$

Eq. (VI.36) can then be written as

$$\vec{B} = -4V_{422}\left[\frac{1}{2}\left(\vec{\sigma}_{A1}\vec{H} + \vec{H}\vec{\sigma}_{A1}\right)\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1} + \frac{1}{2}\left(\vec{\sigma}_{A2}\vec{H} + \vec{H}\vec{\sigma}_{A2}\right)\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2}\right.$$
$$\left. + \frac{1}{2}\left(\vec{\sigma}_{A3}\vec{H} + \vec{H}\vec{\sigma}_{A3}\right)\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\right] , \tag{VI.38}$$

$$\vec{B} = -2V_{422}\left[\left(\vec{\sigma}_{A1}\vec{H}\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1} + \vec{H}\vec{\sigma}_{A1}\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1}\right) + \left(\vec{\sigma}_{A2}\vec{H}\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2} + \vec{H}\vec{\sigma}_{A2}\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2}\right)\right.$$
$$\left. + \left(\vec{\sigma}_{A3}\vec{H}\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3} + \vec{H}\vec{\sigma}_{A3}\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\right)\right] , \tag{VI.39}$$

Since

$$\vec{\sigma}_{Ai}\vec{\sigma}_{Ai} = \vec{\sigma}_{Ai}\bullet\vec{\sigma}_{Ai} + \vec{\sigma}_{Ai} \wedge \vec{\sigma}_{Ai} = 1 + 0 = 1 , \tag{VI.40}$$

Eq. (VI.39) can be written as

$$\vec{B} = -2V_{422}\left[\vec{\sigma}_{A1}\vec{H}\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1} + \vec{H}\hat{e}_\eta\vec{\sigma}_{A1} + \vec{\sigma}_{A2}\vec{H}\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2} + \vec{H}\hat{e}_\eta\vec{\sigma}_{A2}\right.$$
$$\left. + \vec{\sigma}_{A3}\vec{H}\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3} + \vec{H}\hat{e}_\eta\vec{\sigma}_{A3}\right] . \tag{VI.41}$$

Regrouping the terms gives

$$\vec{B} = -2V_{422}\left[\vec{\sigma}_{A1}\vec{H}\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1} + \vec{\sigma}_{A2}\vec{H}\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2} + \vec{\sigma}_{A3}\vec{H}\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\right.$$
$$\left. + \vec{H}\hat{e}_\eta\vec{\sigma}_{A1} + \vec{H}\hat{e}_\eta\vec{\sigma}_{A2} + \vec{H}\hat{e}_\eta\vec{\sigma}_{A3}\right] , \tag{VI.42}$$

$$\vec{B} = -2V_{422}\left[\vec{\sigma}_{A1}\vec{H}\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1} + \vec{\sigma}_{A2}\vec{H}\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2} + \vec{\sigma}_{A3}\vec{H}\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3} + \vec{H}\hat{e}_\eta\left(\vec{\sigma}_{A1} + \vec{\sigma}_{A2} + \vec{\sigma}_{A3}\right)\right] , \tag{VI.43}$$

$$\vec{B} = -2V_{422}\left(\vec{\sigma}_{A1}\vec{H}\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1} + \vec{\sigma}_{A2}\vec{H}\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2} + \vec{\sigma}_{A3}\vec{H}\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\right) , \tag{VI.44}$$

$$\vec{B} = -2V_{422}\left(\hat{e}_\eta\hat{e}_\eta\vec{\sigma}_{A1}\vec{H}\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1} + \hat{e}_\eta\hat{e}_\eta\vec{\sigma}_{A2}\vec{H}\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2} + \hat{e}_\eta\hat{e}_\eta\vec{\sigma}_{A3}\vec{H}\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\right) , \tag{VI.45}$$

$$\vec{B} = -2V_{422}\left(\hat{e}_\eta\vec{H}\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1}\hat{e}_\eta\vec{\sigma}_{A1} + \hat{e}_\eta\vec{H}\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2}\hat{e}_\eta\vec{\sigma}_{A2} + \hat{e}_\eta\vec{H}\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\hat{e}_\eta\vec{\sigma}_{A3}\right) , \tag{VI.46}$$

$$\vec{B} = -2V_{422}\left(\hat{e}_\eta \vec{H} \hat{e}_\eta \hat{e}_\eta \vec{\sigma}_{A1} \hat{e}_\eta \vec{\sigma}_{A1} \hat{e}_\eta \vec{\sigma}_{A1} + \hat{e}_\eta \vec{H} \hat{e}_\eta \hat{e}_\eta \vec{\sigma}_{A2} \hat{e}_\eta \vec{\sigma}_{A2} \hat{e}_\eta \vec{\sigma}_{A2} + \hat{e}_\eta \vec{H} \hat{e}_\eta \hat{e}_\eta \vec{\sigma}_{A3} \hat{e}_\eta \vec{\sigma}_{A3} \hat{e}_\eta \vec{\sigma}_{A3}\right). \quad \text{(VI.47)}$$

The reference axis selected in this dissertation is the x-axis, so $\hat{e}_\eta$ is set to $\hat{e}_1$. Then

$$\vec{B} = -2V_{422}\left(\vec{H}^* \hat{e}_1 \vec{\sigma}_{A1} \hat{e}_1 \vec{\sigma}_{A1} \hat{e}_1 \vec{\sigma}_{A1} + \vec{H}^* \hat{e}_1 \vec{\sigma}_{A2} \hat{e}_1 \vec{\sigma}_{A2} \hat{e}_1 \vec{\sigma}_{A2} + \vec{H}^* \hat{e}_1 \vec{\sigma}_{A3} \hat{e}_1 \vec{\sigma}_{A3} \hat{e}_1 \vec{\sigma}_{A3}\right). \quad \text{(VI.48)}$$

Since the $\vec{\sigma}_{Ai}$ are known explicitly, the following calculations can be performed to show

that

$$\vec{\sigma}_{A1} \hat{e}_1 \vec{\sigma}_{A1} \hat{e}_1 \vec{\sigma}_{A1} = -\hat{e}_1 \ , \quad \text{(VI.49)}$$

$$\vec{\sigma}_{A2} \hat{e}_1 \vec{\sigma}_{A2} \hat{e}_1 \vec{\sigma}_{A2} = -\hat{e}_1 \ , \quad \text{(VI.50)}$$

and

$$\vec{\sigma}_{A3} \hat{e}_1 \vec{\sigma}_{A3} \hat{e}_1 \vec{\sigma}_{A3} = -\hat{e}_1 \ . \quad \text{(VI.51)}$$

Explicitly, the calculations are

$$\vec{\sigma}_{A1} = \left(-1,0\right) = -\hat{e}_1 \ , \quad \text{(VI.52)}$$

$$\vec{\sigma}_{A1} \hat{e}_1 \vec{\sigma}_{A1} \hat{e}_1 \vec{\sigma}_{A1} = \left(-\hat{e}_1\right)\hat{e}_1\left(-\hat{e}_1\right)\hat{e}_1\left(-\hat{e}_1\right) = -\hat{e}_1\hat{e}_1\hat{e}_1\hat{e}_1\hat{e}_1 = -\hat{e}_1 \ , \quad \text{(VI.53)}$$

Next,

$$\vec{\sigma}_{A2} = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) = \frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2 \ , \quad \text{(VI.54)}$$

$$\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2} = \left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right)\hat{e}_1\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right)\hat{e}_1\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right)$$

$$= \left(\frac{1}{2}\hat{e}_1\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\hat{e}_1\right)\left(\frac{1}{2}\hat{e}_1\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\hat{e}_1\right)\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right)$$

$$= \left(\frac{1}{2}\left(\frac{1}{2} + \frac{\sqrt{3}}{2}\hat{e}_2\hat{e}_1\right) + \frac{\sqrt{3}}{2}\hat{e}_2\hat{e}_1\left(\frac{1}{2} + \frac{\sqrt{3}}{2}\hat{e}_2\hat{e}_1\right)\right)\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right)$$

$$= \left(\frac{1}{4} + \frac{\sqrt{3}}{4}\hat{e}_2\hat{e}_1 + \frac{\sqrt{3}}{4}\hat{e}_2\hat{e}_1 + \frac{3}{4}\hat{e}_2\hat{e}_1\hat{e}_2\hat{e}_1\right)\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right) \qquad , \text{(VI.55)}$$

$$\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2} = \frac{1}{4}\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right) + \frac{\sqrt{3}}{4}\hat{e}_2\hat{e}_1\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right) + \frac{\sqrt{3}}{4}\hat{e}_2\hat{e}_1\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right)$$

$$+ \frac{3}{4}\hat{e}_2\hat{e}_1\hat{e}_2\hat{e}_1\left(\frac{1}{2}\hat{e}_1 + \frac{\sqrt{3}}{2}\hat{e}_2\right)$$

$$= \frac{1}{8}\hat{e}_1 + \frac{\sqrt{3}}{8}\hat{e}_2 + \frac{\sqrt{3}}{8}\hat{e}_2\hat{e}_1\hat{e}_1 + \frac{3}{8}\hat{e}_2\hat{e}_1\hat{e}_2 + \frac{\sqrt{3}}{8}\hat{e}_2\hat{e}_1\hat{e}_1 + \frac{3}{8}\hat{e}_2\hat{e}_1\hat{e}_2$$

$$+ \left(\frac{3}{8}\hat{e}_2\hat{e}_1\hat{e}_2\hat{e}_1\hat{e}_1 + \frac{3\sqrt{3}}{8}\hat{e}_2\hat{e}_1\hat{e}_2\hat{e}_1\hat{e}_2\right) \qquad , \text{ (VI.56)}$$

$$\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2} = \frac{1}{8}\hat{e}_1 + \frac{\sqrt{3}}{8}\hat{e}_2 + \frac{\sqrt{3}}{8}\hat{e}_2 - \frac{3}{8}\hat{e}_1 + \frac{\sqrt{3}}{8}\hat{e}_2 - \frac{3}{8}\hat{e}_1 - \frac{3}{8}\hat{e}_1 - \frac{3\sqrt{3}}{8}\hat{e}_2 \ , \text{ (VI.57)}$$

$$\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2} = \frac{1}{8}\hat{e}_1 - \frac{3}{8}\hat{e}_1 - \frac{3}{8}\hat{e}_1 - \frac{3}{8}\hat{e}_1 \ , \qquad \text{(VI.58)}$$

$$\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2}\hat{e}_1\vec{\sigma}_{A2} = -\hat{e}_1 \ . \qquad \text{(VI.59)}$$

And for Eq. (VI.51)

$$\vec{\sigma}_{A3} = \left(\frac{1}{2}, \frac{-\sqrt{3}}{2}\right) = \frac{1}{2}\hat{e}_1 + \frac{-\sqrt{3}}{2}\hat{e}_2 \qquad \text{(VI.60)}$$

So,

$$\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3} = \left(\frac{1}{2}\hat{e}_1 + \frac{-\sqrt{3}}{2}\hat{e}_2\right)\hat{e}_1\left(\frac{1}{2}\hat{e}_1 + \frac{-\sqrt{3}}{2}\hat{e}_2\right)\hat{e}_1\left(\frac{1}{2}\hat{e}_1 + \frac{-\sqrt{3}}{2}\hat{e}_2\right) \ , \qquad \text{(VI.61)}$$

$$\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3} = \left(\frac{1}{2}\left(\frac{1}{2}+\frac{-\sqrt{3}}{2}\hat{e}_2\hat{e}_1\right)+\frac{-\sqrt{3}}{2}\hat{e}_2\hat{e}_1\left(\frac{1}{2}+\frac{-\sqrt{3}}{2}\hat{e}_2\hat{e}_1\right)\right)\left(\frac{1}{2}\hat{e}_1+\frac{-\sqrt{3}}{2}\hat{e}_2\right), \quad \text{(VI.62)}$$

$$\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3} = \left(\frac{1}{4}+\frac{-\sqrt{3}}{4}\hat{e}_2\hat{e}_1+\frac{-\sqrt{3}}{4}\hat{e}_2\hat{e}_1+\frac{3}{4}\hat{e}_2\hat{e}_1\hat{e}_2\hat{e}_1\right)\left(\frac{1}{2}\hat{e}_1+\frac{-\sqrt{3}}{2}\hat{e}_2\right), \quad \text{(VI.63)}$$

$$\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3} = \left(\frac{1}{4}\left(\frac{1}{2}\hat{e}_1+\frac{-\sqrt{3}}{2}\hat{e}_2\right)+\frac{-\sqrt{3}}{2}\hat{e}_2\hat{e}_1\left(\frac{1}{2}\hat{e}_1+\frac{-\sqrt{3}}{2}\hat{e}_2\right)+\frac{3}{4}\hat{e}_2\hat{e}_1\hat{e}_2\hat{e}_1\left(\frac{1}{2}\hat{e}_1+\frac{-\sqrt{3}}{2}\hat{e}_2\right)\right), \text{(VI.64)}$$

$$\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3} = \left(\frac{1}{8}\hat{e}_1+\frac{-\sqrt{3}}{8}\hat{e}_2+\frac{-2\sqrt{3}}{8}\hat{e}_2\hat{e}_1\hat{e}_1+\frac{6}{8}\hat{e}_2\hat{e}_1\hat{e}_2+\frac{3}{8}\hat{e}_2\hat{e}_1\hat{e}_2\hat{e}_1\hat{e}_1+\frac{-3\sqrt{3}}{8}\hat{e}_2\hat{e}_1\hat{e}_2\hat{e}_1\hat{e}_2\right), \text{(VI.65)}$$

$$\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3} = \left(\frac{1}{8}\hat{e}_1+\frac{-\sqrt{3}}{8}\hat{e}_2+\frac{-2\sqrt{3}}{8}\hat{e}_2-\frac{6}{8}\hat{e}_1-\frac{3}{8}\hat{e}_1+\frac{3\sqrt{3}}{8}\hat{e}_2\right), \quad \text{(VI.66)}$$

$$\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3} = \frac{1}{8}\hat{e}_1-\frac{6}{8}\hat{e}_1-\frac{3}{8}\hat{e}_1, \quad \text{(VI.67)}$$

$$\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3}\hat{e}_1\vec{\sigma}_{A3} = -\hat{e}_1. \quad \text{(VI.68)}$$

Then Eq. (VI.48) can be written as

$$\vec{B} = -2V_{422}\left(\vec{H}^*\hat{e}_1\left(-\hat{e}_1\right)+\vec{H}^*\hat{e}_1\left(-\hat{e}_1\right)+\vec{H}^*\hat{e}_1\left(-\hat{e}_1\right)\right), \quad \text{(VI.69)}$$

$$\vec{B} = 2V_{422}\left(\vec{H}^*\hat{e}_1\hat{e}_1+\vec{H}^*\hat{e}_1\hat{e}_1+\vec{H}^*\hat{e}_1\hat{e}_1\right). \quad \text{(VI.70)}$$

And finally,

$$\vec{B} = 6V_{422}\vec{H}^*. \quad \text{(VI.71)}$$

This is the field linear, field conjugate field contribution described in Chapter 7 for the case that a $Z_3^{+3}\left(\vec{\rho}\right)$ Zernike cap shape is added to an optical element.

When the Zernike cap shape added to an optical element's shape is of the form $aZ_3^{+3}(\vec{\rho})+bZ_3^{-3}(\vec{\rho})$ where a and b are real constants, then the field contribution to the Zernike astigmatism term of the wavefront aberration function expansion takes the form (see Table 7.1)

$$\vec{F}_{Z.Cap} = 6V_{422}\vec{H}^* + 6V'_{422}\vec{H} \ , \tag{VI.72}$$

where

$$\vec{H} \equiv \left(H_y, H_x\right) . \tag{VI.73}$$

It is here shown that (VI.72) can be written as

$$\vec{F}_{Z.Cap} = 6V_{422}\vec{H}^* + 6V'_{422}\vec{H} = \vec{C} \star_x \vec{H}^* \ , \tag{VI.74}$$

where $\vec{C} \equiv \left(6V_{422}, 6V'_{422}\right)$.

From Table 7.1

$$6V_{422}\vec{H}^* = 6V_{422}\left(H_x, -H_y\right) , \tag{VI.75}$$

and

$$6V'_{422}\vec{H} = 6V'_{422}\left(H_y, H_x\right) . \tag{VI.76}$$

Then

$$\vec{F}_{Z.Cap} = 6V_{422}\vec{H}^* + 6V'_{422}\vec{H} = 6V_{422}\left(H_x, -H_y\right) + 6V'_{422}\left(H_y, H_x\right) . \tag{VI.77}$$

Combining into a single vector, this becomes

$$\vec{F}_{Z.Cap} = \left(6V_{422}H_x + 6V'_{422}H_y, 6V'_{422}H_x - 6V_{422}H_y\right) \tag{VI.78}$$

Define

$$\vec{C} \equiv \left(6V_{422}, 6V'_{422}\right) . \tag{VI.79}$$

Then

$$\vec{C} \star_x \vec{H}^* = \left(6V_{422}, 6V'_{422}\right) \star_x \left(H_x, -H_y\right) = \left(6V_{422}\hat{e}_1 + 6V'_{422}\hat{e}_2\right)\hat{e}_1\left(H_x\hat{e}_1 - H_y\hat{e}_2\right) . \tag{VI.80}$$

Expanding gives

$$\vec{C} \star_x \vec{H}^* = 6V_{422}\left(H_x\hat{e}_1 - H_y\hat{e}_2\right) + 6V'_{422}\hat{e}_2\hat{e}_1\left(H_x\hat{e}_1 - H_y\hat{e}_2\right) , \tag{VI.81}$$

$$\vec{C} \star_x \vec{H}^* = 6V_{422}H_x\hat{e}_1 - 6V_{422}H_y\hat{e}_2 + 6V'_{422}H_x\hat{e}_2\hat{e}_1\hat{e}_1 - 6V'_{422}H_y\hat{e}_2\hat{e}_1\hat{e}_2 , \tag{VI.82}$$

$$\vec{C} \star_x \vec{H}^* = 6V_{422}H_x\hat{e}_1 - 6V_{422}H_y\hat{e}_2 + 6V'_{422}H_x\hat{e}_2 + 6V'_{422}H_y\hat{e}_1 , \tag{VI.83}$$

$$\vec{C} \star_x \vec{H}^* = \left(6V_{422}H_x + 6V'_{422}H_y\right)\hat{e}_1 + \left(6V'_{422}H_x - 6V_{422}H_y\right)\hat{e}_2 , \tag{VI.84}$$

$$\vec{C} \star_x \vec{H}^* = \left(6V_{422}H_x + 6V'_{422}H_y, \ 6V'_{422}H_x - 6V_{422}H_y\right) . \tag{VI.85}$$

The right hand side of Eq. (VI.85) is the same as the right hand side of Eq. (VI.78). Therefore,

$$\vec{F}_{Z.Cap} = 6V_{422}\vec{H}^* + 6V'_{422}\vec{H} = \vec{C} \star_x \vec{H}^* . \tag{VI.86}$$

This $\vec{C} \star_x \vec{H}^*$ form is the field linear, field conjugate astigmatism field contribution reported in [72]. However, in [72] the *y*-axis is used as the reference axis for SVP whereas Eq. (VI.86) is written using the *x*-axis as the reference axis.