

FIELD-DEPENDENT ABERRATIONS FOR MISALIGNED  
REFLECTIVE OPTICAL SYSTEMS

by

Anastacia Marie Manuel

---

Copyright © Anastacia Marie Manuel 2009

A Dissertation Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2009

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Anastacia Marie Manuel entitled Field-dependent aberrations for misaligned reflective optical systems and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

\_\_\_\_\_  
James H. Burge

Date: 4 December 2009

\_\_\_\_\_  
Russell Chipman

Date: 4 December 2009

\_\_\_\_\_  
Hubert Martin

Date: 4 December 2009

\_\_\_\_\_  
José Sasián

Date: 4 December 2009

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

\_\_\_\_\_  
Dissertation Director: James H. Burge

Date: 4 December 2009

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Anastacia Marie Manuel

## ACKNOWLEDGEMENTS

I would like to thank everyone who helped me through my graduate career at the College of Optical Sciences at the University of Arizona.

Thanks to all of my fellow students at the UA for sharing your knowledge and offering your friendship. I look forward to being lifetime colleagues in optics.

Thanks to Gail Varin and Cindy Gardner for all the valuable assistance you provided.

Thank you to my most recent officemate Dr. Chunyu Zhao for graciously lending an ear and encouraging me during the last stretch of my dissertation writing.

Thank you to my cyber-colleague Régis Tessieres for answering my questions about your past work and helping me with my dissertation research and papers.

Thank you to Dae Wook Kim, René Zehnder, Tom Zobrist, and Ping Zhou for being great lab mates when we were all together in Lab 125. Sincere thanks are especially due to René, Tom, and Ping for regularly reading my dissertation and offering helpful comments, as well as for your encouragement.

Thanks to my dissertation committee: Dr. James Burge, Dr. Russell Chipman, Dr. José Sasian, and Dr. Buddy Martin for the hours spent reading my dissertation and for offering constructive criticisms and insightful suggestions.

It was a privilege to work with my advisor, Dr. Jim Burge. I am grateful for everything that you taught me about optical engineering. This dissertation would not have been possible without your guidance.

Thanks to my family for your moral support and above all, I wish to express my gratitude to my husband Anthony for his love, support and immense patience while I finished writing.

DEDICATION

To my husband, Anthony

## TABLE OF CONTENTS

LIST OF FIGURES . . . . .	11
LIST OF TABLES . . . . .	18
ABSTRACT . . . . .	22
CHAPTER 1 INTRODUCTION, MOTIVATION AND BACKGROUND . .	24
1.1 INTRODUCTION - ABERRATIONS DEGRADE IMAGE QUALITY	24
1.2 MOTIVATION - MINIMIZING ALIGNMENT ABERRATIONS . . .	25
1.3 BACKGROUND . . . . .	28
1.4 OVERVIEW OF DISSERTATION . . . . .	31
CHAPTER 2 EXAMPLE OF AN EXPECTED OPTICAL PERFORMANCE CALCULATION BASED ON THE FORWARD ANALYSIS OF SYSTEM TOLERANCES . . . . .	33
2.1 INTRODUCTION TO THE HOBBY-EBERLY TELESCOPE . . . . .	35
2.2 INTRODUCTION TO MONTE CARLO SIMULATIONS . . . . .	38
2.3 CONTRIBUTIONS . . . . .	40
2.3.1 MIRROR FABRICATION TOLERANCES . . . . .	42
2.3.2 ALIGNMENT TOLERANCES . . . . .	43
2.3.3 ASSEMBLY AND OPERATION TOLERANCES . . . . .	44
2.3.4 COMPENSATOR TOLERANCES . . . . .	45
2.3.5 POLISHING AND SUPPORT FIGURE ERRORS (INCLUDING TRANSMITTED WAVEFRONT ERROR FOR WINDOWS) . . . . .	46
2.3.6 TEMPERATURE EFFECTS . . . . .	47
2.3.7 CORRELATED MISALIGNMENTS THAT DEPEND ON THE TILT OF THE ENTIRE WFC . . . . .	48
2.3.8 WEDGE AND THICKNESS VARIATION OF THE EN- TRANCE AND EXIT WINDOWS . . . . .	50
2.4 RESULTS . . . . .	50
2.5 OTHER CONFIDENCE LEVEL CURVES . . . . .	52
2.6 IMPROVEMENT WHEN MORE COMPENSATORS ARE USED .	53
2.7 EFFECTIVE FOCAL LENGTH CHANGE DUE TO CORRELATED MIRROR MISALIGNMENTS . . . . .	55
2.8 RESULTING COMPENSATOR RANGES OF MOTION . . . . .	57

TABLE OF CONTENTS – *Continued*

2.9	CONCLUSION . . . . .	58
CHAPTER 3 BASIC PRINCIPLES OF ABERRATIONS . . . . .		59
3.1	ABERRATIONS OF ROTATIONALLY SYMMETRIC OPTICAL SYSTEMS . . . . .	61
3.1.1	ABERRATION DEFINITIONS . . . . .	65
3.1.2	DESCRIPTION OF THIRD-ORDER ABERRATIONS USING ZERNIKE POLYNOMIALS . . . . .	73
3.2	ABERRATIONS OF NON-ROTATIONALLY SYMMETRIC OPTICAL SYSTEMS . . . . .	76
3.2.1	NODAL ABERRATION THEORY . . . . .	77
3.2.2	DESCRIPTION OF THIRD-ORDER ABERRATIONS USING ZERNIKE POLYNOMIALS . . . . .	83
3.3	CONCLUSION . . . . .	89
CHAPTER 4 DEMONSTRATION OF COUPLING OF LOW ORDER FIELD-DEPENDENT ABERRATIONS TO ALIGNMENT OF AN OFF- AXIS TWO MIRROR TELESCOPE . . . . .		91
4.1	OPTICAL DESIGN OF THE NEW SOLAR TELESCOPE . . . . .	92
4.2	ABERRATIONS OF THE PARENT ON-AXIS GREGORIAN TELESCOPE . . . . .	94
4.2.1	DEFINITION OF THE DEGREES OF FREEDOM STUDIED FOR NST . . . . .	94
4.2.2	USING ABERRATIONS FOR ALIGNING THE SYSTEM . . . . .	100
4.2.3	FIELD-DEPENDENT ZERNIKE ABERRATIONS IN A MISALIGNED SYSTEM . . . . .	103
4.2.4	SUMMARY OF LEAST-SQUARES FIT COEFFICIENTS . . . . .	109
4.2.5	ALTERNATE DEGREES OF FREEDOM . . . . .	110
4.3	ABERRATIONS OF NST (THE ACTUAL OFF-AXIS TELESCOPE) . . . . .	112
4.3.1	ABERRATION SCALE FACTORS FOR THE OFF-AXIS NST PUPIL . . . . .	112
4.3.2	COMBINING THE RESULTS . . . . .	115
4.4	SPOT DIAGRAMS FOR THE MISALIGNED OFF-AXIS NST . . . . .	118
4.5	EFFECT OF MISALIGNMENTS ON POINTING . . . . .	122
4.6	CONCLUSION . . . . .	124
CHAPTER 5 COMPARISON OF NODAL ABERRATION THEORY AND THE FUNCTIONAL DESCRIPTION OF ABERRATIONS IN A		

TABLE OF CONTENTS – *Continued*

MISALIGNED RITCHEY-CHRETIEN TELESCOPE . . . . .	125
5.1 INTRODUCTION TO THE RITCHEY-CHRETIEN TELESCOPE	
ASTIGMATISM EXAMPLES . . . . .	126
5.1.1 QUADRATIC ASTIGMATISM IN THE NOMINAL	
OPTICAL DESIGN . . . . .	129
5.1.2 CONSTANT ASTIGMATISM FROM PRIMARY MIRROR	
BENDING . . . . .	134
5.1.3 LINEAR ASTIGMATISM FROM SECONDARY MIRROR	
MISALIGNMENT . . . . .	140
5.2 DISCUSSION OF NODAL ABERRATION THEORY . . . . .	149
CHAPTER 6 AN ORTHOGONAL SET OF FIELD-DEPENDENT	
ABERRATIONS . . . . .	153
6.1 INTRODUCTION TO ORTHOGONAL POLYNOMIALS . . . . .	154
6.2 ORTHOGONALITY OF ZERNIKE POLYNOMIALS . . . . .	157
6.3 ABERRATIONS FOR MISALIGNED SYSTEMS USING ZERNIKE	
POLYNOMIALS . . . . .	158
6.4 EXAMPLE DERIVATIONS OF ORTHOGONAL FIELD-	
DEPENDENT ABERRATIONS . . . . .	161
6.5 SUMMARY OF ORTHOGONAL FIELD-DEPENDENT	
ABERRATIONS . . . . .	163
6.6 COMPLETING THE BASIS . . . . .	168
6.7 ORTHOGONAL ABERRATIONS INDUCED BY SHAPE ERRORS	170
6.7.1 COMPENSATION OF SHAPE ERRORS ON ONE MIRROR	
BY SHAPE ERRORS ON OTHER MIRRORS . . . . .	177
6.8 SUMMARY OF FIELD DEPENDENT ABERRATIONS THAT	
DEPEND LINEARLY ON MISALIGNMENT . . . . .	178
6.9 NUMERICAL SIMULATIONS OF MISALIGNED SYSTEMS . . . .	180
6.10 FINAL COMMENTS ON CHOICE OF BASIS FUNCTIONS . . . .	185
CHAPTER 7 LINEAR ANALYSIS OF SYSTEMS WITH MULTIPLE	
DEGREES OF FREEDOM . . . . .	187
7.1 INTRODUCTION TO USING SINGULAR VALUE	
DECOMPOSITIONS IN LINEAR SYSTEMS ANALYSIS . . . . .	188
7.2 PHYSICAL INSIGHT INTO THE SINGULAR VALUE	
DECOMPOSITION . . . . .	192
7.2.1 A MECHANICAL SYSTEM SVD EXAMPLE . . . . .	194
7.2.2 AN SVD EXAMPLE USING A GREGORIAN TELESCOPE	205



TABLE OF CONTENTS – *Continued*

CHAPTER 8	SVD OF THE INFLUENCE MATRIX FOR THE HOBBY- EBERLY TELESCOPE WIDE FIELD CORRECTOR . . . . .	214
8.1	OVERVIEW OF METHOD . . . . .	214
8.2	INFLUENCE MATRIX . . . . .	218
8.2.1	PERTURBATIONS . . . . .	218
8.2.2	ABERRATIONS IN TERMS OF ZERNIKE COEFFICIENTS	221
8.3	SINGULAR VALUE DECOMPOSITION OF THE ALIGNMENT SENSITIVITY MATRIX . . . . .	222
8.3.1	U MATRIX . . . . .	222
8.3.2	S MATRIX . . . . .	229
8.3.3	V MATRIX . . . . .	230
8.4	SUMMARY OF RESULTING MODES FROM SVD . . . . .	235
8.5	VARIATIONS ON THE SENSITIVITY MATRIX . . . . .	235
8.6	WHAT SHOULD GO IN THE SENSITIVITY MATRIX? . . . . .	241
8.7	CONCLUSIONS . . . . .	244
CHAPTER 9	CONCLUSION . . . . .	245
APPENDIX A	HET WFC MONTE CARLO PROCEDURE . . . . .	248
APPENDIX B	ZEMAX MACRO FOR HET WFC . . . . .	250
APPENDIX C	MATLAB CODE FOR THE HET WFC FORWARD PROBLEM . . . . .	252
C.1	GETHET_EE.M MATLAB CODE . . . . .	252
C.2	READTEXT2.M MATLAB CODE . . . . .	256
C.3	HET_MC_ANALYSIS_OCT09.M MATLAB CODE . . . . .	257
APPENDIX D	OPTICAL SPECIFICATIONS FOR HET . . . . .	263
APPENDIX E	HET WFC FINITE ELEMENT ANALYSIS RESULTS . . . . .	274
APPENDIX F	ZERNIKE POLYNOMIALS . . . . .	277
APPENDIX G	VECTOR MULTIPLICATION . . . . .	283
G.1	COUNTERCLOCKWISE FROM X-AXIS ANGLE CONVENTION . . . . .	284
G.2	CLOCKWISE FROM Y-AXIS ANGLE CONVENTION . . . . .	287
G.3	SWITCHING THE ANGLE CONVENTION FROM CW FROM THE Y-AXIS TO CCW FROM THE X-AXIS . . . . .	289
G.3.1	USEFUL VECTOR DOT PRODUCTS AND VECTOR MULTIPLICATION PRODUCTS . . . . .	291

TABLE OF CONTENTS – *Continued*

APPENDIX H ZEMAX MACRO FOR THE NEW SOLAR TELESCOPE ANALYSIS . . . . .	293
APPENDIX I MATLAB CODE . . . . .	297
I.1 FITTINGPOLYNOMIALS.M MATLAB CODE . . . . .	297
I.2 ASTIGMATISM.M MATLAB CODE . . . . .	315
I.3 COMA.M MATLAB CODE . . . . .	315
I.4 ZERNIKEPOLYSONGRID.M MATLAB CODE . . . . .	315
I.5 DOUBLEZERNIKEFIT.M MATLAB CODE . . . . .	316
APPENDIX J DERIVATIONS OF ORTHOGONAL FUNCTIONS . . . . .	318
J.1 ZERNIKE POLYNOMIALS TO DESCRIBE FIELD DEPENDENCE	319
J.2 ZERNIKE POLYNOMIALS TO DESCRIBE PUPIL DEPENDENCE	322
J.3 THIRD ORDER ABERRATION DERIVATIONS . . . . .	322
J.3.1 THIRD ORDER FIELD CURVATURE . . . . .	323
J.3.2 ASTIGMATISM . . . . .	324
J.3.3 COMA . . . . .	325
J.3.4 SPHERICAL ABERRATION . . . . .	326
J.4 FIFTH ORDER ABERRATION DERIVATIONS . . . . .	326
J.4.1 QUARTIC FIELD CURVATURE . . . . .	326
J.4.2 FIFTH ORDER ASTIGMATISM . . . . .	328
J.4.3 FIELD CUBED COMA . . . . .	334
J.4.4 TREFOIL . . . . .	338
J.4.5 FIELD CURVATURE FOR OBLIQUE SPHERICAL . . . . .	339
J.4.6 OBLIQUE SPHERICAL ABERRATION . . . . .	341
J.4.7 FIFTH ORDER (SECONDARY) COMA . . . . .	343
APPENDIX K ORTHOGONAL ABERRATION FUNCTIONS . . . . .	344
APPENDIX L DOUBLE ZERNIKE EXPANSIONS . . . . .	360
L.1 KWEE AND BRAAT'S DOUBLE EXPANSION . . . . .	360
L.2 AGUROK'S DOUBLE EXPANSION . . . . .	363
APPENDIX M ORTHOGONALITY OF DOUBLE ZERNIKE FUNCTIONS	367
APPENDIX N SHAPE ERRORS . . . . .	369
APPENDIX O HET CONTROL MODES . . . . .	371
REFERENCES . . . . .	399

## LIST OF FIGURES

1.1	The optical alignment problem. . . . .	26
2.1	The optical layout of the Wide Field Corrector for the Hobby-Eberly Telescope. The focus of the primary mirror lies near the central hole of M3. . . . .	36
2.2	Results from the Monte Carlo trials including perturbations to the manufacturing, alignment, and window degrees of freedom. . . . .	41
2.3	Comparison of correlated misalignment errors across the field. For predicting the performance, the largest value from any tilt configuration was chosen for the correlated misalignment error for each field in Table 2.9. . . . .	49
2.4	Contributions to HET WFC encircled energy. . . . .	51
2.5	Net effect of nominal design; alignment, manufacturing, and window tolerances; and 15 $\mu\text{m}$ compensator perturbations. . . . .	52
2.6	Expected performance of the Wide Field Corrector. . . . .	53
2.7	The encircled energy diameters are smaller when eight compensator degrees of freedom are included during optimization than when five compensator degrees of freedom are used. . . . .	54
3.1	Pupil and image conventions in a right-handed coordinate system. The components along the $\vec{x}$ and $\vec{y}$ axes respectively are $\rho \cos \phi$ and $\rho \sin \phi$ for $\vec{\rho}$ and $H_x$ and $H_y$ for $\vec{H}$ . . . . .	60
3.2	Graphical representations of spherical aberration through focus. (a) aberrated rays coming to focus; (b) spot diagrams through focus showing what an image of a point would look like (Burge, 1993). . . . .	66
3.3	Spot diagram for spherical aberration constant throughout the field (from Shack). . . . .	67
3.4	Ray diagram for coma. . . . .	69
3.5	Spot diagram for coma linear throughout the field (from Shack). . . . .	69
3.6	Ray diagram for astigmatism. . . . .	70
3.7	Spot diagram for astigmatism quadratic throughout the field. . . . .	70
3.8	Astigmatism field curves. The line astigmatic line images are formed on these surfaces. The medial focal surface (not shown) is located halfway in between the sagittal and tangential focal surfaces. . . . .	71
3.9	Field curvature (quadratic with field). . . . .	71
3.10	Distortion (cubic with field). . . . .	72

LIST OF FIGURES – *Continued*

3.11	Astigmatism nodes in a perturbed optical system. There are two nodes in the field which are determined by the vectors $\vec{a}_{222}$ and $\vec{b}_{222}$ (Thompson, 1980). . . . .	80
4.1	New Solar Telescope optical layout. . . . .	93
4.2	M2 rotation points shown in the nominal aligned system. . . . .	96
4.3	M2 degrees of freedom. (The rotations are shown larger than typical misalignments to clearly show the location of the rotation point.) . . .	97
4.4	Effect of M2 degrees of freedom on spot diagrams. The maximum field angle spots are shown for a full field of view of 3 arcminutes. Spot sizes are exaggerated by $4\times$ for clarity. (These spot diagrams are what you would see looking toward the focal plane, and not from behind it.) . . . . .	98
4.5	NST axisymmetric parent spot diagrams with linear astigmatism. . .	106
4.6	NST parent spot diagrams with constant coma. . . . .	107
4.7	Constant coma and linear astigmatism coefficients in NST for different rotation points. The scale for the coma coefficient $y$ -axis is $10\times$ larger than the astigmatism $y$ -axis. (All rotations were $0.1^\circ$ .) . . . .	111
4.8	Transformation of one unit of coma from NST parent telescope pupil to actual pupil. . . . .	113
4.9	Transformation of one unit of astigmatism from NST parent telescope pupil to actual pupil. . . . .	114
4.10	Transformation of one unit of spherical aberration from NST parent telescope pupil to actual pupil. . . . .	115
4.11	Spot diagrams when secondary mirror is rotated about the focus of the primary mirror $5.5^\circ$ . The spot diagrams shows a plate scale change for $\vec{x}$ rotations and image rotation for $\vec{y}$ rotations. The full field of view is 3 arcminutes. The spot size exaggeration is $\times 4$ . The rotations about $\vec{x}$ show plate scale change, while the rotations about $\vec{y}$ show image rotation. . . . .	119
4.12	Spot diagrams when secondary mirror is rotated about its center of curvature by $0.1^\circ$ . The full field of view is 3 arcminutes. The spot size exaggeration is $\times 4$ . . . . .	120
4.13	Spot diagrams when secondary mirror is rotated about its center of curvature by $0.1^\circ$ and images are refocused. The full field of view is 3 arcminutes. The spot size exaggeration is $\times 4$ . . . . .	120
4.14	Wavefront maps for the on axis field point when the secondary mirror is rotated in about its center of curvature ( $CoCRotY = 0.1^\circ$ ). . . . .	122
5.1	Hubble Space Telescope optical model. . . . .	127

LIST OF FIGURES – *Continued*

5.2	Hubble Space Telescope spot diagrams. The size of the Airy disk is shown in the center of the field of the ZEMAX spot diagram. . . . .	127
5.3	2-D Quadratic astigmatism in Hubble Space Telescope design (Example 1: no misalignment) in CODE V. The astigmatism node on-axis is shown with a red dot. . . . .	130
5.4	1-D Quadratic astigmatism in Hubble Space Telescope design (Example 1: no misalignment) in CODE V. . . . .	130
5.5	Fringe Zernike coefficients for Hubble Space Telescope (Example 1: no misalignment) in CODE V. . . . .	131
5.6	Astigmatism coefficients throughout the field in the Hubble Space Telescope Example 1. Quadratic astigmatism dominates and one node is seen on-axis. (The plots show an infinitesimal amount of linear astigmatism, which come from limitations in the numerical calculations.) . . . . .	132
5.7	Astigmatism and coma in Example 1, separated by field dependence, calculated from coefficients. The astigmatism node on-axis is shown with a red dot. The $H_x$ and $H_y$ fields are normalized to the half field of view of $0.233^\circ$ . . . . .	133
5.8	Spot diagram showing constant astigmatism for the Hubble Space Telescope with an astigmatism error on the primary mirror (Example 2). (The spots are shown at a defocused image plane 20 mm behind focus and are exaggerated by $50\times$ .) . . . . .	135
5.9	2-D Constant astigmatism in Hubble Space Telescope, with simulated primary mirror astigmatism shape error in CODE V. There are no nodes within the field of view. . . . .	136
5.10	1-D Constant astigmatism in Hubble Space Telescope, with simulated primary mirror astigmatism shape error in CODE V. . . . .	136
5.11	Astigmatism coefficients from the Hubble Space Telescope Example 2. Constant astigmatism (from the $Z_5$ term) dominates. (The plots show an infinitesimal amount of linear astigmatism, which come from limitations in the numerical calculations.) . . . . .	137
5.12	Astigmatism and coma in Example 2, separated by field dependence, calculated from coefficients. The $H_x$ and $H_y$ fields are normalized to the half field of view of $0.233^\circ$ . . . . .	138
5.13	Astigmatism coefficients from the Hubble Space Telescope Example 3. Binodal astigmatism (seen in the top left plot) is produced by a combination of quadratic and constant astigmatism. (The plots show an infinitesimal amount of linear astigmatism, which come from limitations in the numerical calculations.) . . . . .	139

LIST OF FIGURES – *Continued*

5.14	Astigmatism and coma, separated by field dependence, calculated from coefficients. Binodal astigmatism may be seen in the top left plot. The astigmatism nodes are shown with red dots. The $H_x$ and $H_y$ fields are normalized to the half field of view of $0.233^\circ$ . . . . .	140
5.15	2-D constant, linear and quadratic astigmatism in Hubble Space Telescope, with simulated secondary mirror misalignment in CODE V. . .	143
5.16	Astigmatism coefficients from the Hubble Space Telescope Example 4. Binodal astigmatism (top left plot) is produced by a combination of constant, linear and quadratic astigmatism. The nodes are no longer symmetric about the origin due to the linear component. . . . .	145
5.17	Astigmatism and coma, separated by field dependence, calculated from coefficients. Binodal astigmatism may be seen in the top left plot. (The astigmatism nodes are shown with red dots.) The $H_x$ and $H_y$ fields are normalized to the half field of view of $0.233^\circ$ . . . . .	145
5.18	2-D constant, linear and quadratic astigmatism in Hubble Space Telescope, with small simulated secondary mirror misalignment in CODE V. The astigmatism nodes are shown with red dots. . . . .	147
5.19	Astigmatism coefficients from the Hubble Space Telescope Example 5. Although the astigmatism is binodal here, the resolution is not high enough to show the two separate nodes in the top left plot The binodal nature of the astigmatism can be inferred by observing that the astigmatism has both quadratic and linear field dependent contributions. . . . .	148
5.20	Astigmatism and coma, separated by field dependence, calculated from coefficients. The two nodes for the astigmatism in the top left plot are shown with red dots. The $H_x$ and $H_y$ fields are normalized to the half field of view of $0.233^\circ$ . . . . .	148
6.1	Schematic representation for an off-axis beam. The large circle represents a mirror away from the stop of the system. The small circle in gray represents an off-axis beam. Figure courtesy of Tessieres (2004).	172
6.2	Astigmatism line images throughout the field when there is trefoil shape error on a mirror not at the stop. (This particular image, courtesy of Tessieres (2004) shows the effect for trefoil on the tertiary mirror in a 3-mirror telescope.) . . . . .	176

LIST OF FIGURES – *Continued*

6.3	The orthogonal field-dependent aberration coefficients that exist in a rotationally symmetric system (due to $W$ terms) for the HST with a secondary mirror misalignment. These coefficients have primarily a constant dependence on the value of the perturbation. As the perturbation becomes large, the nonlinearity is viewable for some of the terms. . . . .	184
6.4	The orthogonal field-dependent aberration coefficients that depend linearly on the misalignment (due to $\vec{A}$ terms), shown for the HST with a secondary mirror misalignment. These coefficients have primarily a linear dependence on the perturbation. As the perturbation becomes large, the nonlinearity is viewable for some of the terms. . . . .	184
6.5	The orthogonal field-dependent aberration coefficients that depend quadratically on the misalignment (due to $\vec{B}$ terms), shown for the HST with a secondary mirror misalignment. These coefficients have primarily a quadratic dependence on the perturbation. As the perturbation becomes large, the nonlinearity is viewable for some of the terms. . . . .	185
7.1	Rod and spring system example . . . . .	194
7.2	First configuration of forces examined for the rod and spring system example. Force one pushes on the rod directly over point $A$ and force two pushes on the rod directly over point $B$ . . . . .	196
7.3	Resulting control modes from SVD of sensitivity matrix for the first force configuration of the rod/spring system. The nominal position of the rod is shown in black. The first mode ( $u_1$ ) from the SVD causes the bar to move as shown in red. This is caused by the forces in the vector $v_1$ , shown by the vertical maroon lines ending in “×.” The second mode ( $u_2$ ) from the SVD causes the bar to move as shown in blue. This is caused by the forces in the vector $v_2$ , shown by the dark blue vertical lines, ending in circles. . . . .	199
7.4	Block diagram describing the process of recording the aberrations in a misaligned optical system to find the sensitivity matrix and to analyze the resulting modes from an SVD of this matrix. . . . .	206
7.5	Singular values of the axisymmetric Gregorian telescope. The vertical scales depends on the normalization of the $\mathbf{U}$ and $\mathbf{V}$ matrices. . . . .	208
7.6	Aberration modes from the $\mathbf{U}$ matrix. . . . .	210
7.7	Configuration modes from the $\mathbf{V}$ matrix. . . . .	211
8.1	Process of calculating the alignment sensitivity matrix. . . . .	216
8.2	Process of calculating the alignment state. . . . .	218

LIST OF FIGURES – *Continued*

8.3	Wavefront maps across the field showing constant coma from the first and second control modes. . . . .	223
8.4	Coefficients showing constant coma from the first and second control modes. . . . .	224
8.5	Defocus from third control mode. . . . .	226
8.6	Wavefront maps across the field showing focal plane tilt from fourth and fifth control modes. . . . .	227
8.7	Coefficients showing focal plane tilt from fourth and fifth control modes.	228
8.8	HET SVD singular values. . . . .	230
8.9	Required degrees of freedom to produce HET control modes. . . . .	232
8.10	Aberrations when degrees of freedom are perturbed and the system is corrected for constant coma with the WFC rigid body hexapod tilt.	237
F.1	Zernike polynomial wavefront maps. . . . .	281
G.1	Vector multiplication displayed in two different coordinate systems. The vectors $\vec{A}$ and $\vec{B}$ are each unit length, so the vector product is also unit length. . . . .	284
G.2	The angle of the vector is defined counterclockwise from the $x$ -axis. . . . .	285
G.3	The angle of the vector is defined clockwise from the $y$ -axis. . . . .	288
K.1	Orthogonal focus polynomials. . . . .	345
K.2	Orthogonal astigmatism polynomials. . . . .	347
K.3	Orthogonal coma polynomials. . . . .	351
K.4	Orthogonal trefoil polynomials. . . . .	354
K.5	Orthogonal spherical aberration polynomials. . . . .	356
K.6	Orthogonal oblique spherical aberration polynomials. . . . .	357
K.7	Orthogonal secondary coma polynomials. . . . .	359
O.1	HET SVD mode 1: constant coma. . . . .	374
O.2	HET SVD mode 2: constant coma. . . . .	375
O.3	HET SVD mode 3: defocus. . . . .	376
O.4	HET SVD mode 4: focal plane tilt (linear defocus). . . . .	377
O.5	HET SVD mode 5: focal plane tilt (linear defocus). . . . .	378
O.6	HET SVD mode 6: linear astigmatism. . . . .	379
O.7	HET SVD mode 7: linear astigmatism. . . . .	380
O.8	HET SVD mode 8: linear coma, quadratic astigmatism (axisymmetric errors). . . . .	381
O.9	HET SVD mode 9: constant secondary coma, linear secondary astigmatism. . . . .	382



LIST OF FIGURES – *Continued*

O.10 HET SVD mode 10: constant secondary coma, linear secondary astigmatism. . . . .	383
O.11 HET SVD mode 11: constant spherical and other axisymmetric error. . . . .	384
O.12 HET SVD mode 12. . . . .	385
O.13 HET SVD mode 13. . . . .	386
O.14 HET SVD mode 14. . . . .	387
O.15 HET SVD mode 15. . . . .	388
O.16 HET SVD mode 16. . . . .	389
O.17 HET SVD mode 17. . . . .	390
O.18 HET SVD mode 18. . . . .	391
O.19 HET SVD mode 19. . . . .	392
O.20 HET SVD mode 20. . . . .	393
O.21 HET SVD mode 21. . . . .	394
O.22 HET SVD mode 22. . . . .	395
O.23 HET SVD mode 23. . . . .	396
O.24 HET SVD mode 24. . . . .	397
O.25 HET SVD mode 25. . . . .	398

## LIST OF TABLES

2.1	HET WFC system prescription (not including the entrance and exit windows). . . . .	35
2.2	Mirror diameters used to convert tilt in degrees to microns. . . . .	36
2.3	HET WFC mirror aspheric terms. . . . .	37
2.4	Manufacturing tolerances for M2, M3, M4 M5 and focal plane. . . . .	42
2.5	Alignment tolerances in microns. . . . .	44
2.6	Operational tolerances in microns. . . . .	45
2.7	Compensator resolution tolerances. . . . .	46
2.8	Tolerances on the entrance and exit windows. . . . .	50
2.9	Contributions to 90% encircled energy (in arcseconds) for the HET WFC (90% confidence level). . . . .	51
2.10	Change in effective focal length (after optimization) due to correlated mirror motion and result on maximum relative image motion for two points on opposite edges of the field of view. . . . .	55
2.11	Compensator adjustments used during the Monte Carlo simulations and required by the hexapod on the telescope. . . . .	57
3.1	Third-order aberrations of a misaligned system developed by Thompson. . . . .	81
3.2	Fifth-order aberrations of a misaligned system developed by Thompson. . . . .	82
3.3	Perturbation vectors and scalars defined by Thompson (1980). . . . .	83
3.4	Field-dependent functions to describe the Zernike polynomial coefficients for third-order aberrations in a misaligned system. . . . .	84
3.5	Conversion from Thompson's notation into Tessieres's notation for third-order aberrations. . . . .	86
3.6	Field-dependent functions to describe the Zernike polynomial coefficients for fifth-order aberrations in a misaligned system. . . . .	87
3.7	Conversion from Thompson's notation into Tessieres's notation for fifth-order aberrations. . . . .	88
4.1	NST system prescription. . . . .	92
4.2	NST f/number calculations. . . . .	94
4.3	Secondary mirror degrees of freedom and effects on aberrations. . . . .	99

LIST OF TABLES – *Continued*

4.4	$\alpha$ and $\beta$ coefficients in units of microns for rotations of $0.1^\circ$ . Quadratic astigmatism is described by $\alpha_0$ . Linear astigmatism is described by $\alpha_1$ and $\alpha_2$ . Constant astigmatism is described by $\alpha_3$ and $\alpha_4$ . Linear coma is described by $\beta_0$ . Constant coma is described by $\beta_1$ and $\beta_2$ . . . . .	108
4.5	Conversion of Zernike polynomials to lower order for NST pupil. . . . .	113
4.6	Calculations for scaling aberrations in the off-axis NST. The aberrations that exist in the nominal (rotationally symmetric) system do not change with misalignment. The aberrations that scale with the misalignment are marked “/degree” in the calculation column, while the aberrations that depend on the square of the misalignment are given for case when the misalignment is $0.1^\circ$ . . . . .	116
4.7	Comparison of aberrations for on-axis and off-axis telescopes. . . . .	118
4.8	Pointing error (in mm) for NST. . . . .	123
5.1	Hubble Space Telescope (Good design). . . . .	126
5.2	Astigmatism examples studied in this chapter. . . . .	128
5.3	Astigmatism and coma coefficients in waves ( $\lambda = 586.56$ nm) from Tessier’s equations (Table 3.4) for Example 1. . . . .	133
5.4	Coefficients in waves ( $\lambda = 586.56$ nm) for three waves of constant astigmatism added to Hubble primary mirror (Example 2). . . . .	137
5.5	Coefficients in waves ( $\lambda = 586.56$ nm) for constant astigmatism added to Hubble primary mirror in Example 3. . . . .	139
5.6	Coefficients in waves ( $\lambda = 586.56$ nm) for Hubble Space Telescope Example 4 with a 10 mm secondary mirror misalignment in the $\vec{x}$ -direction. . . . .	144
5.7	Coefficients in waves ( $\lambda = 586.56$ nm) for Hubble Space Telescope Example 5 with a 0.1 mm secondary mirror misalignment in the $\vec{x}$ -direction. . . . .	147
6.1	Polynomials orthogonal over field and pupil space for focus terms $Z_4(\rho, \phi)$ . . . . .	164
6.2	Polynomials orthogonal over field and pupil space for astigmatism terms $Z_5(\rho, \phi)$ and $Z_6(\rho, \phi)$ . . . . .	165
6.3	Polynomials orthogonal over field and pupil space for coma terms $Z_7(\rho, \phi)$ and $Z_8(\rho, \phi)$ . . . . .	166
6.4	Polynomials orthogonal over field and pupil space for trefoil terms $Z_9(\rho, \phi)$ and $Z_{10}(\rho, \phi)$ . . . . .	166
6.5	Polynomials orthogonal over field and pupil space for spherical aberration terms $Z_{11}(\rho, \phi)$ . . . . .	167

LIST OF TABLES – *Continued*

6.6	Polynomials orthogonal over field and pupil space for oblique spherical terms $Z_{12}(\rho, \phi)$ and $Z_{13}(\rho, \phi)$ . . . . .	167
6.7	Polynomials orthogonal over field and pupil space for secondary coma terms $Z_{16}(\rho, \phi)$ and $Z_{17}(\rho, \phi)$ . . . . .	167
6.8	Orthogonal polynomials induced by mirror bending. The coefficients $a$ and $c$ describe the size and position of the beam on a mirror away from the stop. . . . .	174
6.9	Orthogonal field dependent aberrations that depend linearly on the misalignment. The aberrations with check marks in the column labeled “New?” have not been discussed in prior work. . . . .	179
6.10	Orthogonal field dependent aberration coefficients for the Hubble Telescope with misaligned secondary mirror in the $\vec{x}$ -direction. The coefficient units are waves of 587.6 nm. . . . .	182
7.1	SVD matrix definitions. . . . .	198
7.2	Table of rod/spring examples—first set. . . . .	201
7.3	Table of rod/spring examples—second set. . . . .	203
7.4	Coefficients for the orthogonal functions from the SVD analysis of the NST sensitivity matrix. . . . .	212
8.1	Values describing the sizes of the influence, control and sensitivity matrices. . . . .	215
8.2	Degrees of freedom for the HET WFC alignment sensitivity matrix. . . . .	219
8.3	Summary of HET control modes. . . . .	235
8.4	RSS of coefficients across modes 1–14. . . . .	243
E.1	Mirror displacement at $35^\circ$ in plane tilt and $0^\circ$ out of plane tilt (Nominal design). . . . .	274
E.2	Mirror displacement at $26.5^\circ$ in plane tilt and $0^\circ$ out of plane tilt. . . . .	274
E.3	Mirror displacement at $43.5^\circ$ in plane tilt and $0^\circ$ out of plane tilt. . . . .	275
E.4	Mirror displacement at $35^\circ$ in plane tilt and $8.5^\circ$ out of plane tilt. . . . .	275
E.5	Mirror displacement at $29^\circ$ in plane tilt and $6^\circ$ out of plane tilt. . . . .	275
E.6	Mirror displacement at $41^\circ$ in plane tilt and $6^\circ$ out of plane tilt. . . . .	275
E.7	Positive temperature change of $\Delta T = 20^\circ\text{C}$ , WFC at nominal orientation ( $35^\circ$ in plane tilt, $0^\circ$ out of plane tilt). . . . .	276
F.1	Orthonormal Zernike Circle Polynomials in polar form. . . . .	279
F.2	Orthonormal Zernike Circle Polynomials in Cartesian form . . . . .	280
L.1	Aberrations predicted by Agurok in a perturbed optical system. . . . .	365

LIST OF TABLES – *Continued*

N.1	Zernike standard polynomials in a scalar and vectorial form. The odd number after the name of each aberration (3 or 5) refers to the order of the aberration. . . . .	369
N.2	Table of the expressions to describe the field dependence of aberrations induced by the first 10 Zernike polynomial bending modes. . . .	370
N.3	Expression of the third order spherical aberration introduced as a bending mode. . . . .	370
O.1	HET SVD singular values. . . . .	372
O.2	HET ZEMAX verification of SVD mode 1. If mode 1 is approximated only as constant coma (even though there are small amounts of higher order coma), then the expected values are $Z_7 = -13.5$ waves and $Z_8 = 13.5$ waves for any field angle. The fields are normalized such that $H_x = H_y = 11$ arcmin and the Zernike coefficients are in waves ( $\lambda = 1 \mu\text{m}$ ). . . . .	373

## ABSTRACT

The performance of optical imaging systems relies on control of aberrations that can arise from limitations in the design, manufacture, or alignment. This dissertation addresses the form of aberrations that occur for misaligned reflective systems, such as telescopes. The relationship between a characteristic set of field-dependent aberrations and the misalignments that cause them is systematically explored. A comprehensive technique that quantifies field performance for a 5-mirror system is given, using Monte Carlo analysis to provide confidence levels of image quality as functions of manufacturing and alignment errors. This analysis is an example of the “forward problem”—determining optical performance of a system if the errors are assumed. The inverse problem—determining the state of alignment based on measurements of performance—is more difficult. The solution to the inverse problem for a multiple mirror system requires an understanding of the complex coupling between many degrees of freedom (tilt, decenter, despace, shape error) of the optical elements and field-dependent aberrations.

This work builds on previous treatment of field dependent optical aberrations from Tessieres, Thompson, Shack, Buchroeder and others. A basis set of field-dependent aberrations orthogonal over both field and pupil are developed here and used to describe systems with misaligned and misshapen optics. This description allows complete representation of high order and non-linear effects. The functional

form of aberrations that are characteristic of mirror tilt, shift, and deformation show some useful patterns that provide insight to the fundamental effects of misalignment.

The use of singular value decomposition to create orthogonal combinations of the field dependent aberrations provides a powerful tool for evaluating a system and for estimating the state of alignment using wavefront measurements. The following optical systems are evaluated to investigate the linear coupling between misalignment and the resulting field dependent aberrations:

- Two-mirror telescopes, evaluating well-understood effects for an axisymmetric system and developing the relationships for an unobscured system.
- Four-mirror correctors for a spherical primary telescope.

The tools and methods are applied to reflective optical systems for astronomical telescopes, but the methods are general and can be useful for any optical imaging system.

## CHAPTER 1

### INTRODUCTION, MOTIVATION AND BACKGROUND

This chapter contains a brief review of the historical background of the analysis of aberrations in optical systems for alignment and outlines the contents of the entire dissertation.

#### 1.1 INTRODUCTION - ABERRATIONS DEGRADE IMAGE QUALITY

The image quality of an optical system depends on the errors, called optical aberrations in the system. The aberrations from a variety of sources sum together to determine the total aberrations in the system. Optical components that are poorly mounted, incorrectly placed, or incorrectly made, and the residual design aberrations can all contribute to the total optical aberration in the system. Atmospheric turbulence and diffraction are other effects that can also add to the aberrations which degrade the image quality.

Optical engineers strive to minimize all sources of aberrations in order to achieve the highest optical quality possible. Minimizing aberrations in a computer model while designing a new optical system is just the first step. Often (unless the design is revolutionary), it is a routine procedure for a lens designer to modify an old design to meet new specifications. After an optical system is designed, before it is built, the engineer must specify all the fabrication and alignment tolerances for the optical components and related systems so that the final system can meet its



performance specifications. For example in a telescope, the allowed deviation in the radius of curvature and conic constant of the mirrors must be specified in addition to the mirror positions in order to achieve a given spot size at the detector. In addition, opto-mechanical engineers design mounts to prevent the optical elements from warping or sagging (especially important for large optics), and for telescopes, adaptive optics systems may be designed to minimize the effects of atmospheric turbulence.

## 1.2 MOTIVATION - MINIMIZING ALIGNMENT ABERRATIONS

For some optical designs, it is straightforward to assemble a system that meets the performance specifications. Other systems with tight specifications require careful planning to ensure that the built system is satisfactory. For these systems, tolerances must be chosen for all of the individual elements for fabrication and alignment. Inevitably, tighter tolerances increase the cost to build a system, so tolerances must be chosen sensibly, specifying tolerances only as tight as necessary. For example, mid-spatial frequency figure errors on the mirrors in a telescope may cause the same sort of image degradation as atmospheric turbulence. If the mirror figure tolerance is set to be far better than the best atmosphere that the telescope will see, then the fabrication costs will be unnecessarily high because the final image would be no better than if lower quality optics were used.

Predicting the optical performance of a system before it is built from a list of expected errors with tolerances can be done by a perturbation analysis. Monte Carlo simulations can help predict the performance of a unknown system where

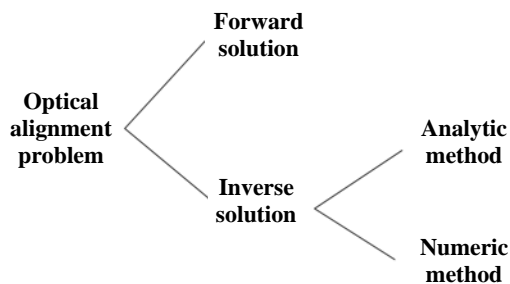


Figure 1.1: The optical alignment problem.

many errors can couple together in a complex way. This is considered a “forward problem.” The confidence level for achieving a given performance can be found, which helps with the process of assigning reasonable tolerances. Only statistics on what might exist before a system is built.

When a system is built, the performance will never be as good as the optical model. Some of the causes of image degradation might not be able to be further improved, such as the quality of the optics. However, the alignment errors can be adjusted if one can somehow infer what is wrong. Inferring the details of a particular system from measured data is known as an “inverse problem,” which is much harder to solve than a forward problem. Although the sensitivities to errors in the system are the same for both the forward problem and inverse problem, it is close to impossible to infer exactly everything that is wrong in an optical system from a fuzzy image. Measured wavefront data across the field gives more complete information about what is wrong in the system. This dissertation is concerned with the relation between alignment errors and optical aberrations, and how to determine one given the other. A diagram showing the forward vs. inverse problem is shown in Figure 1.1.

Alignment of an optical system begins when an optical element is first placed in the system. Usually the element is placed in approximately the correct place using measurements of physical dimensions. This “rough alignment” involves measuring the distances prescribed in the optical design carefully in order to put everything together as closely as possible. In a system with only a few elements or loose tolerances, methodically making small adjustments to the positions of the optical elements will eventually sufficiently align the system. In a system with more elements and/or tighter tolerances, more precise tools for physical measurements, such as alignment telescopes or metering rods, can be used for alignment.

If the previous technique does not sufficiently align the system to meet the performance specifications, then optical measurements of wavefront aberrations may need to be made and the inverse problem must be solved to calculate the misalignments. Since the misalignment determines the specific aberrations present and their amount, then measurements of the wavefront can be used to determine the misalignment to do the “fine alignment.” For example, in a two-element system, there are analytic solutions that describe the mirror adjustments necessary to minimize particular aberrations measured. In a typical Cassegrain (two-mirror reflecting) telescope, moving the secondary mirror in certain ways can correct aberrations, including constant coma or linear astigmatism. For a more complicated system with many degrees of freedom, the effect of each of the degrees of freedom of the optical elements on the aberrations has previously not been well-understood analytically. In such cases, numerical methods such as Singular Value Decomposition (SVD) have been used instead to make a maximum likelihood estimate (MLE) of the correction,

based on a model of the aberrations. This dissertation will expand on the work that has been done previously in both analytic solutions and numerical solutions. Numerical solutions will be especially useful for reflective systems with three or more optical elements.

### 1.3 BACKGROUND

Misaligned multi-element optical systems usually have no axis of symmetry and suffer from imaging aberrations that vary in a complex (but deterministic) way over the field of view, even if a system is designed to be aberration-free. An optical system designed to be aberration-free is a careful balance of elements that each add positive or negative amounts of different aberrations that are minimized over the image plane when the system is aligned. But when the system is misaligned, the perturbations move the aberration field centers associated with the various surfaces, and the net aberrations are no longer minimized throughout the field.

The aberrations resulting from tilting and decentering elements in a rotationally symmetric optical system have been studied by a variety of authors, including Buchroeder (1976), Shack & Thompson (1980); Shack (2005), Thompson (1980, 2005, 2009) and others (Rimmer, 1970; Turner Jr., 1992; Moore *et al.*, 2008) and are well-described. Aberration fields for these kinds of axisymmetric systems have been derived by modifying the well known wave aberrations of Hopkins (1950) for axisymmetric systems using polynomial expansions in pupil and field space. The result is a set of general equations describing the field dependence of various aberrations that can be used to evaluate the image degradation due to misalignments.

Unfortunately, these equations are limited for use in aligning a telescope or other optical system because they do not express the aberrations in terms of an orthogonal set of functions, such as Zernike polynomials, which causes coupling between low and high order effects.

On the other hand, many other publications describe the effect of low order aberrations on image quality for a generic two-mirror telescope (such as a Cassegrain or Ritchey-Chrétien), and quantify the amount of aberrations due to misalignments. These publications on astronomical optics (Fehniger, 1980; Bhatia, 1995; McLeod, 1996; Wilson, 1996; Wilson & Delabre, 1997; Mahajan, 1998; Schroeder, 1999; Noethe & Guisard, 2000) present a way to align the telescope by relating aberration coefficients to the tilts and decenters in the system. It is well-known that a decenter of the secondary will introduce coma which is constant over the field; and a tilt about the “coma-free point” will generate mostly astigmatism that varies linearly with the field. (A rotation about the coma-free point in a two-mirror system does not introduce coma. See Wilson (1996) for more explanations.) Some publications (Bhatia, 1995; Wilson & Delabre, 1997; Mahajan, 1998; Schroeder, 1999) are for specific two-mirror telescopes and relate aberration coefficients to tilts and decenters in the system, including some with useful relations between Zernike coefficients and field dependencies that can be used to retrieve the misalignments in a two-mirror telescope (McLeod, 1996; Schroeder, 1999). Unfortunately, the equations are for the specific systems in these publications, and do not offer general relations valid for any system.

The general polynomial wavefront expansions were transformed by Tessieres

(2003) into Zernike polynomial form. Knowing these field dependences allows a system to be modeled with optical software to find the relationships between the Zernike coefficients and the tilts and decenters of each element in the system. However, the field dependence of these functions developed by Tessieres are not orthogonal, meaning that the magnitude of each term depends on which other terms are included in the analysis. Fitting nonorthogonal functions to data is not ideal when there are lots of variables (in our case misalignment degrees of freedom) to be fit. Using an SVD to make MLE of the current alignment state of the system is one method that has been used to align a complicated optical system. Chapman & Sweeney (1998) used an SVD technique to align a microlithography system with many degrees of freedom. After this success, CODE V implemented an SVD alignment algorithm into the code (CODE V, 2009). This numerical method works well, as long as the measurements are accurate, but it is difficult to use without an understanding of the system. Without such insight, it is difficult to determine the number of measurements or the data quality required.

There are a variety of large telescopes newly constructed or under construction with more than two mirrors that will need to be aligned. These include the Large Synoptic Survey Telescope (LSST) (Claver *et al.*, 2004) with three mirrors (and three corrective lenses), the James Webb Space Telescope (JWST) (Clampin, 2008), also with three mirrors, and the Southern African Large Telescope (SALT) (O'Donoghue & Swat, 2002) and Hobby-Eberly Telescope (HET) (Booth *et al.*, 2006), each with a spherical primary mirror and a four-mirror corrector.

Each of these telescopes has been analyzed using SVD techniques to understand

the aberration modes which may result from misalignment or to simulate an alignment process:

1. LSST: Phillion *et al.* (2006) (includes mirror bending modes)
2. JWST: Shiri *et al.* (2007) (includes primary mirror segment misalignments)
3. SALT: Hvisc & Burge (2008)
4. HET: this dissertation

#### 1.4 OVERVIEW OF DISSERTATION

This dissertation continues in Chapter 2 with an example of looking at the effect of mirror misalignments (among other expected errors) in an optical system to predict system performance. This chapter is an example of the forward problem, while the rest of the dissertation concentrates on the inverse problem: how knowledge of the aberrations due to incorrectly placed components can aid the optical alignment process. The dissertation is split into two major sections after Chapter 2. Section I (Chapters 3–6) covers analytical alignment methods. Section II (Chapters 7–8) covers numeric alignment methods. The final chapter discusses how to choose which approach is more appropriate to solve the inverse problem, depending on the details of the particular situation.

Section I comprises Chapters 3–6. Chapter 3 provides a review of the basic principles of aberrations. It emphasizes the differences between aberrations of rotationally symmetric systems, and those of nonrotationally symmetric systems. Chapter 4 discusses the aberrations that occur in a model of a misaligned two-mirror Gregorian telescope (the New Solar Telescope). Chapter 5 compares Nodal Aberration Theory

and the functional description of the field-dependence of aberrations for the example of a Ritchey-Chretien telescope (the Hubble Space Telescope). Chapter 6 presents a new set of orthogonal field-dependent aberrations that are formed by combinations of Zernike polynomials in pupil and field space. The last part of Chapter 6 continues with the Hubble Telescope example to provide coefficients for a misaligned optical system.

Section II comprises Chapters 7–8. Chapter 7 discusses a general background for alignment of optical systems using a numeric method and provides two examples of SVD of a system. Chapter 8 discusses the specific case of SVD alignment of the Hobby-Eberly Telescope Wide Field Corrector (HET WFC). The orthogonal control modes are found by SVD of the influence matrix of the system. These control modes are examined in the orthogonal two-term double Zernike functions presented in Chapter 6.

The final chapter (Chapter 9) summarizes the entire dissertation and discusses how to approach a complicated alignment problem. Additionally, Chapter 9 proposes some ideas for future work in this area.



## CHAPTER 2

EXAMPLE OF AN EXPECTED OPTICAL PERFORMANCE CALCULATION  
BASED ON THE FORWARD ANALYSIS OF SYSTEM TOLERANCES

When building optical systems, error analysis is essential for planning and making decisions that will enable the system to be built to meet the performance specifications. In this chapter, Monte Carlo simulations are used to predict the performance of the Hobby-Eberly Telescope (HET) Wide Field Corrector (WFC) for a variety of sources of error, such as manufacturing errors, misalignment, or temperature changes.

In an initial analysis (Hvisc & Burge, 2008), the sensitivities to spot size were found for the alignment degrees of freedom of each of the mirrors. For an individual degree of freedom, the product of the sensitivity and the tolerance predicts the degradation to the root mean square (rms) spot size, averaged over a number of field points. Tolerances were chosen for each of the degrees of freedom, such that when the total rms spot size was found by root sum square, the resulting performance was acceptable. Finding the predicted performance using a root sum square assumes that the individual effects are uncorrelated; however, this assumption may not be completely valid, especially in a complex system, where there may be many factors that couple together in unknown ways that affect the performance.

To verify that the tolerances chosen would result in acceptable performance for the actual specification—the 80% encircled energy diameters throughout the field,

Monte Carlo simulations were done. In a Monte Carlo simulation, a large number of systems that meet all tolerances are generated. The first advantage of a Monte Carlo analysis is that it fully captures any coupling between the degrees of freedom that are toleranced. In addition, another advantage is that it does not require any assumptions about the linearity of the sensitivities to the degrees of freedom. (In later chapters, the nonlinear dependence of aberrations to misalignment is discussed. Although the nonlinearity is usually small over the chosen range tolerances, a Monte Carlo analysis saves the hassle of verifying for each degree of freedom and accounting for it when necessary.) The final advantage of Monte Carlo simulations is that since a large number of possible systems are generated, statistics may be found for any number of resulting quantities of interest, beyond the average spot size. In this case, the Monte Carlo analysis provides statistics on the encircled energy spot size diameters at different locations throughout the field, the required range of motion of the compensators and the resulting plate scale of the telescope. Before anything is made, there are no exact numbers for anything—only probabilities on what might be. A Monte Carlo analysis is the best way to give the project the confidence it needs to determine that the chosen tolerances give a high likelihood that an acceptable system will be built.

The chapter shows the results of an analysis for a newer design of the HET WFC than in the previous paper (Hvisc & Burge, 2008). The tolerances for the alignment degrees of freedom have been refined and tolerances on many additional degrees of freedom have been added. This chapter includes analysis that predicts the expected performance for the HET WFC that accounts for many sources of

errors, including manufacturing, alignment, assembly and operation, compensator positioning, polishing and support figure errors, correlated misalignment errors and window fabrication errors. Most of the effects for these individual error sources were found by Monte Carlo simulations.

## 2.1 INTRODUCTION TO THE HOBBY-EBERLY TELESCOPE

Some telescopes use primary mirrors with spherical shape to reduce the cost of the mirror fabrication and to allow the mirror to operate at fixed elevation. These advantages become significant as the size of the telescope grows. The HET is one such telescope that uses a spherical primary mirror (Ramsey *et al.*, 1998; Krabbendam *et al.*, 1998). However, the disadvantage of the spherical primary is a large amount of spherical aberration which needs to be corrected. The 11-m HET has a corrector for the spherical aberration, but a major upgrade, including a new Wide Field Corrector (WFC), is planned. The new corrector for the HET, shown in Figure 2.1, is currently under construction at the University of Arizona. The prescription for the corrector is listed in Table 2.1. The diameters of the mirrors are listed in Table 2.2.

Table 2.1: HET WFC system prescription (not including the entrance and exit windows).

Surface Name	Radius of curvature (mm)	Conic constant	Thickness (mm)
M1	-26163.9	0	-14014.6
M2	2620.8	0.663	982.6
M3	-2032.5	-7.711	-1559.8
M4	-376.7	-2.098	337.1
M5	-742.1	-0.2675	-2410.8
Focal plane	970.0		

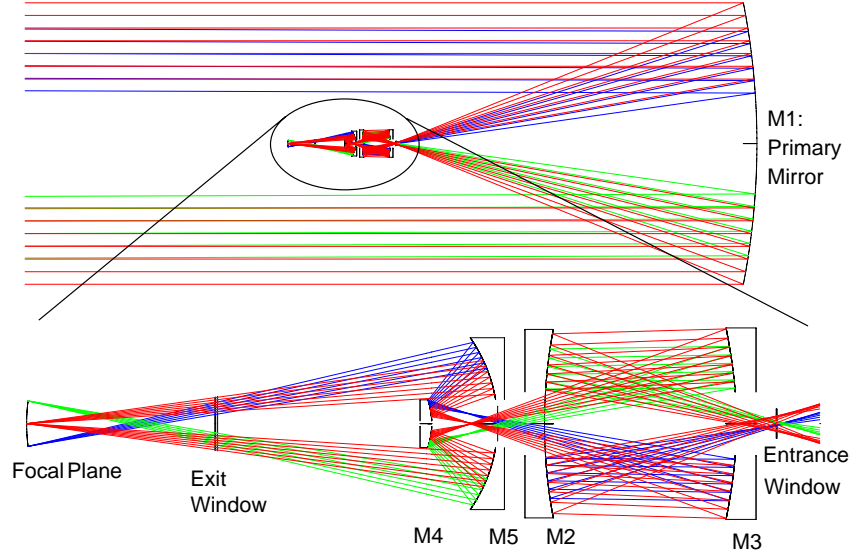


Figure 2.1: The optical layout of the Wide Field Corrector for the Hobby-Eberly Telescope. The focus of the primary mirror lies near the central hole of M3.

Table 2.2: Mirror diameters used to convert tilt in degrees to microns.

Degree of freedom	Diameter
M2 tilt	1020 mm
M3 tilt	1020 mm
M4 tilt	249 mm
M4 M5 tilt	900 mm
Focal plane tilt	232.56 mm
M2-M5 together tilt	1020 mm

Three of the four mirrors are concave and one (M4) is convex. All four mirrors are aspheric, with surfaces described by conic constants. In addition to the conic constants, M3 and M5 require higher order aspheric terms to describe their surfaces. The sag of a mirror  $z$  as a function of the radial position  $r$  is given by the following equation when there are higher order aspheric terms  $\alpha$ :

$$z(r) = \frac{cr^2}{1 + \sqrt{1 - (1+k)c^2r^2}} + \alpha_1 r^2 + \alpha_2 r^4 + \alpha_3 r^6 + \alpha_4 r^8 + \alpha_5 r^{10} \quad (2.1)$$

Table 2.3: HET WFC mirror aspheric terms.

Aspheric term	M2	M3	M4	M5
$\alpha_3$ 6th order term ( $1/\text{mm}^5$ )	-	$-8.263 \times 10^{-17}$	-	$5.762 \times 10^{-19}$
$\alpha_4$ 8th order term ( $1/\text{mm}^7$ )	-	$8.482 \times 10^{-23}$	-	$2.137 \times 10^{-25}$
$\alpha_5$ 10th order term ( $1/\text{mm}^9$ )	-	$-3.595 \times 10^{-29}$	-	-

where  $c$  is the curvature (or  $1/R$ , the inverse of the radius of curvature), and  $k$  is the conic constant of the mirror. The terms  $\alpha_1$  and  $\alpha_2$  are not needed because the radius of curvature determines the second order dependence on the radial pupil position and the conic constant describes the fourth order dependence. In addition, since the surfaces are radially symmetric, only even order terms are needed. The three aspheric terms for M3 and two aspheric terms for M5 are listed in Table 2.3.

This chapter describes the expected performance in terms of encircled energy (EE), based on the tolerances for manufacturing, alignment and operation of the WFC design. The process of finding the sensitivities, choosing tolerances and predicting the performance for an earlier design of the corrector was discussed in a previous paper (Hvisc & Burge, 2008). The tolerances and design have both gone through some slight modifications and this chapter uses the current design for the mirrors as they will be manufactured. The contributions from a variety of error sources are assumed to be uncorrelated, which allows their effects to be added together by root sum square to find the estimated performance. For example, the resulting 90% encircled energy diameter  $\Phi$  is given by:

$$\Phi = \sqrt{\Phi_0^2 + (\Delta\Phi_1)^2 + (\Delta\Phi_2)^2 + \dots} \quad (2.2)$$

where  $\Phi_0$  is the encircled energy for the nominal design and  $\Delta\Phi_i$  is the effect from each parameter  $i$ . The different parameters or sources of error ( $i = 1-8$ ), which will be discussed in Sections 2.3.1–2.3.8, are:

1. Manufacturing tolerances
2. Alignment tolerances
3. Assembly and operation tolerances
4. Compensator tolerances
5. Polishing and support figure errors
6. Temperature effects
7. Correlated misalignment errors
8. Entrance and exit windows

## 2.2 INTRODUCTION TO MONTE CARLO SIMULATIONS

Monte Carlo (MC) simulations are used to predict the statistical effect of tolerances on the degrees of freedom in a system by simulating a large number of random systems that meet all tolerances. For this system, the MC trials (described in Appendix A) are performed in ZEMAX using the tolerancing functions with a custom tolerancing script (included in Appendix B). In one trial, each degree of freedom is randomly perturbed to some number, within a uniform distribution of values allowed by the tolerance. The compensators (which are the five degrees of freedom

of the entire WFC with focal plane) are optimized to minimize the root mean square (rms) spot size at a collection of points in the field. (This collection includes five field positions: one on axis and at four at  $(\pm 5 \text{ arcmin}, \pm 5 \text{ arcmin})$ ). This is repeated for some number of trials, typically around 200 or more in this analysis. Confidence levels for the encircled energy diameters are calculated as the percent of systems that are better (smaller EE diameters) than the given number. While a 50% confidence level could be used to show what might be expected on average, there are many systems that could be much worse. In order to have a high confidence that the system will work, the 90% confidence levels are considered here. (It is entirely a coincidence that there are two different values that are both examined at 90% (confidence levels and encircled energies).)

The specification for this system is given in terms of the 80% encircled energy spot sizes through the field of view. However, this analysis uses 90% encircled energy sizes because it is assumed that 10% of the energy is scattered out of the image completely due to surface irregularities. The encircled energy diameters after many Monte Carlo trials are performed are evaluated in MATLAB at five points throughout the field of view (0, 2.7, 5.4, 8.1 and 11.0 arcminutes).

The ZEMAX file resulting from the Monte Carlo analysis lists the resulting encircled energy radii across in the field in units of microns, compensator motions and effective focal length for each trial. A MATLAB function (included in Appendix C.1) loads the data file with the ZEMAX tolerancing results for one set of Monte Carlo simulations. The MATLAB function searches for the relevant data in the data file and analyzes the results. The encircled energy radii in microns are converted into

diameters in arcseconds, and for each field radius, the encircled energy diameter is plotted as a cumulative probability, as shown in Figure 2.2b, based on the histograms of the resulting spot sizes from the Monte Carlo trials, as shown in Figure 2.2a.

The output of the function is a set of encircled energy diameters for different confidence levels across the field. These diameters are usually larger than the ones for the nominal system. The effect of the set of perturbations tested (e.g. the alignment tolerances or the manufacturing tolerances) is found by a root difference squared (Equation 2.3) of the set of perturbed encircled energy diameters and the ones without the perturbations. That is, the effect for just one perturbation (e.g.  $i = 1$ ) in Equation. 2.2 is

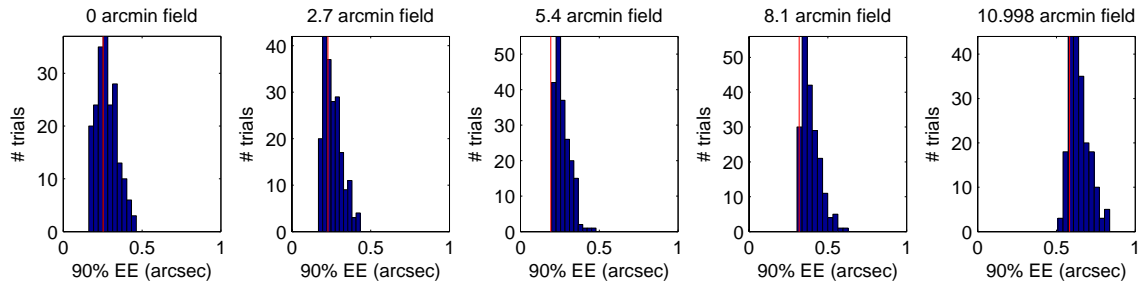
$$\Delta\Phi_1 = \sqrt{\Phi^2 - \Phi_0^2}. \quad (2.3)$$

This step is performed for each set of perturbations ( $i$ ) at each different field radius and confidence level in a second MATLAB file (included in Appendix C.3). The purpose of calculating the root difference squared is to be able to separate the contributions so they can be examined individually (in a table or a plot). In the end, the effect all the contributions together will be found by a root sum square. However, the assumption does not need to be made that the effects are not correlated in this case because the data are actually the statistical results from a Monte Carlo simulation to begin with and they were just arbitrarily separated for the table.

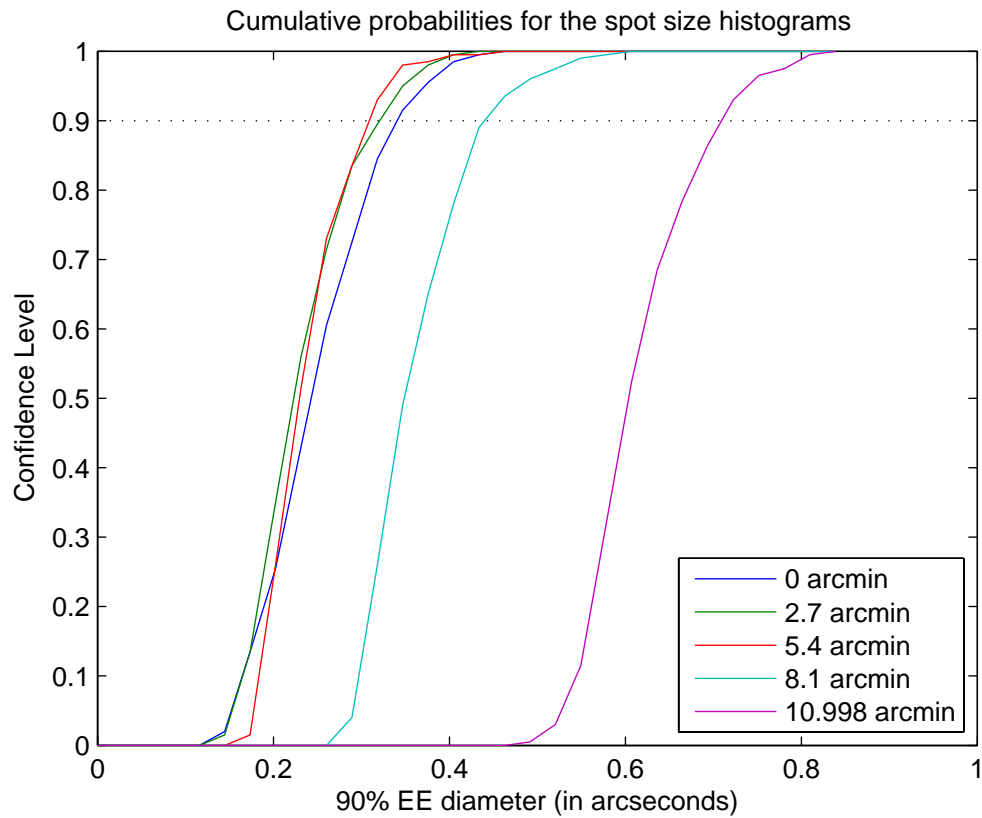
### 2.3 CONTRIBUTIONS

In this section, each of the contributions and their meanings are presented. Section 2.4 lists the encircled energy contributions and shows a plot of the results. All





(a) Histograms of the spot sizes. The red line indicates the nominal value of the model before the perturbations.



(b) Cumulative probability distribution for spot sizes. The intersection of each of the curves with the dotted line at 0.9 gives the the encircled energy diameters at the 90% confidence level.

Figure 2.2: Results from the Monte Carlo trials including perturbations to the manufacturing, alignment, and window degrees of freedom.

tolerances are listed as positive numbers, but are assumed the same in the positive and negative directions. For example, a tolerance listed as 50  $\mu\text{m}$  should be assumed to be  $\pm 50 \mu\text{m}$  (peak to valley).

### 2.3.1 MIRROR FABRICATION TOLERANCES

Mirror fabrication tolerances in this section include the error in the knowledge of the radii and conic constants. All other errors (such as those in the higher order aspheric terms) are included in the polishing and support figure errors section. The focal plane in the system has a radius of curvature because the light will be fed into optical fibers. The ends of the fibers will be on a spherical focal surface. (This is possible, while in general, a curved detector is quite hard to make.) This spherical focal surface (for the fiber ends) has an allowable radius of curvature tolerance variation chosen (3.44 mm) such that the change in sag at the edge of the surface at a radius of 116.28 mm (which is the edge of the field) is no more than  $\pm 25 \mu\text{m}$ . Table 2.4 lists the measurement tolerances used in the Monte Carlo simulation.

Table 2.4: Manufacturing tolerances for M2, M3, M4 M5 and focal plane.

	M2	M3	M4	M5	Focal Plane
$\Delta R$ uncertainty (mm)	0.075	0.075	0.045	0.04	3.44
$\Delta k$ uncertainty	0.0011	0.0008	0.0006	0.0001	NA

These uncertainties in Table 2.4 are different from the allowable  $\Delta R$  and  $\Delta k$  for actual variations in the final radius of curvature and conic constant. Errors in the radius of curvature and the conic constant result in radially symmetric aberrations, such as focus and spherical aberration. Small errors in the specified  $R$  and  $k$  values

are acceptable to some level because the spacings between the mirrors can be slightly adjusted during reoptimization in the optical model after the mirrors are fabricated, but before the system is put together, to remove these errors completely. It is the amount of uncertainty in the radius of curvature and conic constant that cause errors in the final system, because the resulting aberrations can not be corrected by deliberately changing the spacing between the elements.

A Monte Carlo simulation was performed where each of the mirrors were simultaneously perturbed by some amount that satisfied the tolerances in Table 2.4 and the total effect (listed in row  $\Phi_1$  of Table 2.9) was found using the technique described in Section 2.2.

### 2.3.2 ALIGNMENT TOLERANCES

Alignment tolerances include those tolerances for alignment of the individual mirrors and groups of mirrors, as listed below in Table 2.5. The decenters and tilts were perturbed in each of the  $\vec{x}$  and the  $\vec{y}$  directions with the same tolerances. During a previous analysis, it was determined that M4 and M5 should move together. Instead of specifying tight tolerances for both M4 and M5 positions in the system (which would be hard), one mirror could be held precisely (M4) as long as the other mirror (M5) moves with it. This concept affected the design of the mechanical structure holding the mirrors. All of the tilts in Table 2.5 are given in microns of displacement across the diameter. The diameters assumed were listed in Table 2.2. The tolerances on decenter, for example are given in the  $\vec{x}$  and  $\vec{y}$  directions. For an arbitrary direction, the total decenter can be  $\sqrt{2}$  times the tolerance. A Monte Carlo

simulation was performed where each of the mirrors were simultaneously perturbed by some amount that satisfied the tolerances in Table 2.5 and the total effect (listed in row  $\Phi_2$  of Table 2.9) was found using the technique described in Section 2.2.

Table 2.5: Alignment tolerances in microns.

Degree of Freedom	Tolerance
M2 to M3 axial	100
M2 to M5 axial	100
M4 to M5 axial	20
M2 decenter ( $\vec{x}$ or $\vec{y}$ )	50
M3 decenter ( $\vec{x}$ or $\vec{y}$ )	50
M4 decenter ( $\vec{x}$ or $\vec{y}$ )	20
M4 M5 decenter ( $\vec{x}$ or $\vec{y}$ )	50
M2 tilt ( $\vec{x}$ or $\vec{y}$ )	50
M3 tilt ( $\vec{x}$ or $\vec{y}$ )	50
M4 tilt ( $\vec{x}$ or $\vec{y}$ )	20
M4 M5 tilt ( $\vec{x}$ or $\vec{y}$ )	50
Focal plane axial position	1000
Focal plane tilt ( $\vec{x}$ or $\vec{y}$ )	50
M2-M5 together decenter ( $\vec{x}$ or $\vec{y}$ )	250
M2-M5 together tilt ( $\vec{x}$ or $\vec{y}$ )	250

### 2.3.3 ASSEMBLY AND OPERATION TOLERANCES

The assembly and operational tolerances are for the same degrees of freedom as the alignment tolerances in the previous section. The values of these tolerances, which account for operational changes, are 1/4 of the tolerances above and are listed in Table 2.6. This amount was chosen as a reasonable approximation for the tolerances. The total effect (which will be listed in row  $\Phi_3$  of Table 2.9) is 1/4 that for the alignment errors in row  $\Phi_2$ .

Table 2.6: Operational tolerances in microns.

Degree of freedom	Tolerance
M2 to M3 axial	25
M2 to M5 axial	25
M4 to M5 axial	5
M2 decenter ( $\vec{x}$ or $\vec{y}$ )	12.5
M3 decenter ( $\vec{x}$ or $\vec{y}$ )	12.5
M4 decenter ( $\vec{x}$ or $\vec{y}$ )	5
M4 M5 decenter ( $\vec{x}$ or $\vec{y}$ )	12.5
M2 tilt ( $\vec{x}$ or $\vec{y}$ )	12.5
M3 tilt ( $\vec{x}$ or $\vec{y}$ )	12.5
M4 tilt ( $\vec{x}$ or $\vec{y}$ )	5
M4 M5 tilt ( $\vec{x}$ or $\vec{y}$ )	12.5
Focal plane axial position	250
Focal plane tilt ( $\vec{x}$ or $\vec{y}$ )	12.5
M2-M5 together decenter ( $\vec{x}$ or $\vec{y}$ )	62.5
M2-M5 together tilt ( $\vec{x}$ or $\vec{y}$ )	62.5

### 2.3.4 COMPENSATOR TOLERANCES

There are five compensators used to correct manufacturing and alignment errors and other errors in the system. The compensators are the axial position,  $\vec{x}$  and  $\vec{y}$  tilt, and  $\vec{x}$  and  $\vec{y}$  decenter of the entire WFC, including the focal plane, which is mounted on a hexapod. However, these degrees can only be controlled with finite precision. The tolerance for each of these degrees of freedom is  $\pm 15 \mu\text{m}$ . For the system tilt,  $15 \mu\text{m}$  is across the 0.7 meter radius. These compensator resolution tolerances are given in Table 2.7.

To find the effect of the compensator tolerances, a Monte Carlo simulation was performed where each of the four mirrors were simultaneously perturbed by some amount that satisfied all of the manufacturing and alignment tolerances. Then another Monte Carlo simulation was performed where again each of the four mirrors

are perturbed according to the manufacturing and alignment tolerances, but this time, after the compensators are adjusted to their optimal positions according to the merit function, they are perturbed up to 15  $\mu\text{m}$ . The total effect of just perturbation of the compensators ( $\Phi_4$  in Table 2.9) is found by calculating the the root difference squared of the set of Monte Carlo trials without the compensator perturbations and the Monte Carlo trials with the compensator perturbations. The effect of perturbing the compensators is mostly constant across the field because these degrees of freedom (like tilting or decentering a secondary mirror of a two mirror telescope) cause mostly constant coma throughout the field.

### 2.3.5 POLISHING AND SUPPORT FIGURE ERRORS (INCLUDING TRANSMITTED WAVEFRONT ERROR FOR WINDOWS)

The tolerance on the effect on the polishing and support figure errors (including errors in the higher order aspheric coefficients) is 0.25 arcseconds for 90% EE, constant across the field. These numbers (not found by Monte Carlo analysis) can be considered to be at the 100% confidence level, since the mirror will be done when it meets this specification. For example, there would not be a 50% confidence of meeting 0.20 arcseconds polishing/figuring error across the field or better. The surface figure specification for the effect on the encircled energy diameter is in the units of

Table 2.7: Compensator resolution tolerances.

Degree of freedom	Tolerance
Axial Position	15 $\mu\text{m}$
$\vec{x}$ or $\vec{y}$ decenter	15 $\mu\text{m}$
$\vec{x}$ or $\vec{y}$ tilt	21.4 $\mu\text{rad}$ or 4.42 arcsec

the encircled energy diameter, which is arcseconds. An internal University of Arizona report (included in Appendix D) breaks down the 0.25 arcsecond contributions into the surface slope specifications for both polishing and support errors for each of the mirrors that are necessary to achieve this overall specification.

### 2.3.6 TEMPERATURE EFFECTS

The effect of uniform temperature changes in the system is examined. (Temperature gradients, or differential temperatures between M2 and M3, for example, are not considered.) The effect of the temperature change on each of the mirror positions was given by finite element analysis (FEA) (Appendix E), performed by Dr. Rob Stone. The temperature changes affect the mirror positions linearly for both positive and negative temperature changes. However, the optical effect on the encircled energy diameter is not linear.

To find the effect of the temperature, the design is put in the perturbed state as found by FEA for worst-case (largest  $\Delta T$ ). Next, all the fabrication, alignment and compensator degrees of freedom are perturbed, within their tolerances, through the normal MC routine. The statistics on encircled energy spot size are found for the resulting systems and compared to the statistics resulting from another Monte Carlo analysis using the same perturbed degrees of freedom and tolerances, but starting from the nominal system (at 10 °C). (In particular, the effect for just the temperature change is found by a root difference squared of the one set of Monte Carlo trials that started with a temperature change and another set of Monte Carlo trials that started at the nominal temperature.) The effects for each a positive and a negative

temperature change were found. However, negative temperature changes cause a larger effect on the encircled energy than positive temperature changes throughout the field. All the numbers listed in Table 2.9 are for a temperature of  $-10\text{ }^{\circ}\text{C}$ , which is a change of  $\Delta T = -20\text{ }^{\circ}\text{C}$  from the nominal temperature of  $10\text{ }^{\circ}\text{C}$ .

### 2.3.7 CORRELATED MISALIGNMENTS THAT DEPEND ON THE TILT OF THE ENTIRE WFC

The tracker on the telescope moves the position of the entire WFC and focal plane assembly to keep the position of the stars on the sky constant at the focal plane for long exposures. During this process, the WFC is tilted from its nominal elevation angle of  $35^{\circ}$  to keep its optical axis perpendicular to the surface of the primary mirror. Tilting the WFC with respect to gravity causes the four WFC mirrors to move out of position. Since the mirrors move together in a certain, repeatable way, this section is called “correlated misalignments.” The tracker may tilt the corrector up to  $8.5^{\circ}$  in any direction. The mechanical changes to the mirror positions due to tilting were found by FEA (Appendix E), performed by Dr. Rob Stone, for a number of configurations, representing the possible directions and ranges of tilt angles. These configurations include in plane for both  $26.5^{\circ}$  and  $43.5^{\circ}$ , and out of plane for  $35^{\circ}/8.5^{\circ}$ ,  $29^{\circ}/6^{\circ}$ ,  $41^{\circ}/6^{\circ}$ .

Next, the optical effect of these mirror changes due to tilting the corrector for each of these different pointing configurations was found. Similar to the Monte Carlo analysis done for the temperature changes, the mirrors are moved to perturbed positions, as calculated by the FEA. Then a Monte Carlo simulation is run, where



the manufacturing and alignment degrees of freedom are perturbed to new values within the allowed range defined by the tolerances. The effect for the correlated misalignments is found by a root difference squared between those results and those from another Monte Carlo analysis, with the same manufacturing and alignment perturbations, but starting with the mirrors in their positions defined by the nominal design. Not surprisingly, the different tracker positions affect the mirrors in the WFC and the resulting optical performance differently. In order to account for the worst-possible scenario for predicting the performance of the WFC, the largest effect at each field for any of the different tracker positions tested was used for the results in Table 2.9. This is shown graphically in Figure 2.3.

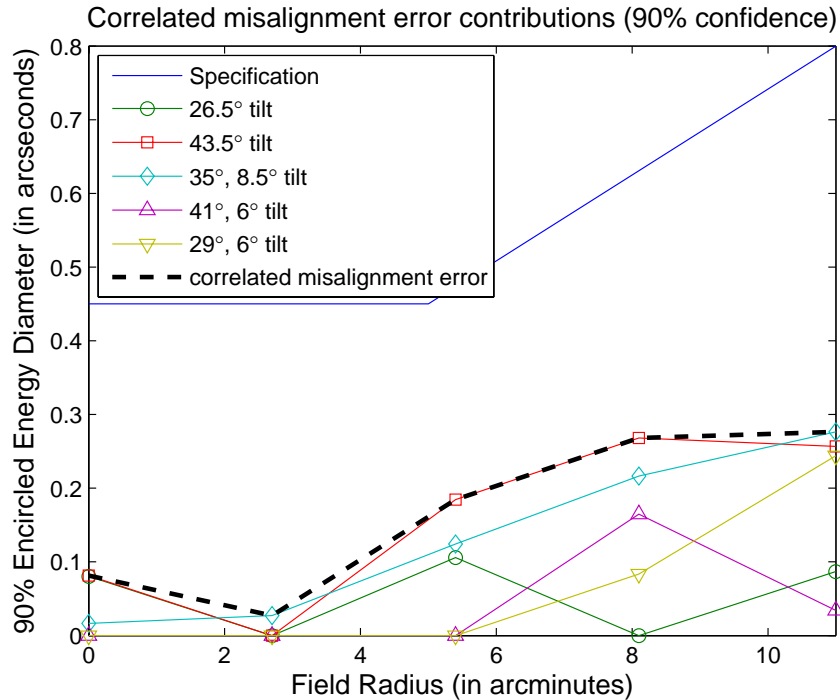


Figure 2.3: Comparison of correlated misalignment errors across the field. For predicting the performance, the largest value from any tilt configuration was chosen for the correlated misalignment error for each field in Table 2.9.

### 2.3.8 WEDGE AND THICKNESS VARIATION OF THE ENTRANCE AND EXIT WINDOWS

The entrance and exit windows will not significantly affect the performance, but are included here for completeness. The effect of wedge and thickness variations in the windows is modeled by Monte Carlo analysis. The tolerances for these effects are listed in Table 2.8. The windows may also have transmitted wavefront errors (from surface irregularities or index variations), but this effect is included in the 0.25 arcsec figure specification (in Section 2.3.5).

Table 2.8: Tolerances on the entrance and exit windows.

Degree of freedom	Tolerance
Wedge (amount allowed)	1 arcminute
Thickness (measurement uncertainty)	0.1 mm

## 2.4 RESULTS

The tolerances were chosen so that the 90% confidence level (almost) meets the specification. The net result of all contributions added by root sum square is shown in Figure 2.4. There is a 90% confidence that the system will perform this well or better. The numbers in this graph are provided in Table 2.9. Since many of the tolerances have a similar magnitude effect, it would be hard to improve performance significantly without tightening all of the tolerances.

As a final check, one Monte Carlo simulation that included perturbations to the alignment, fabrication, window and compensator degrees of freedom was performed and the net result was basically the same as the root sum square of the individual

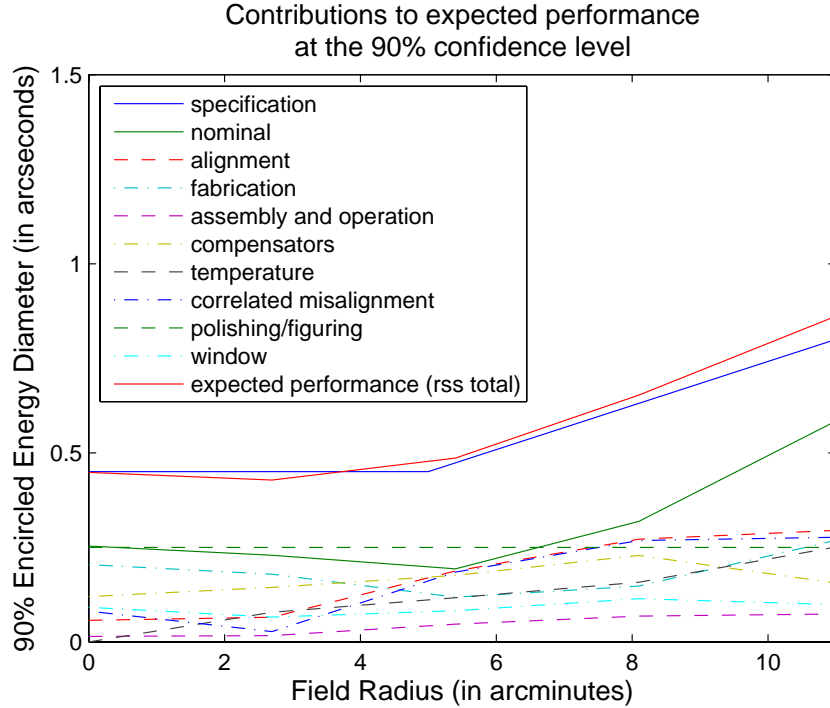


Figure 2.4: Contributions to HET WFC encircled energy.

Table 2.9: Contributions to 90% encircled energy (in arcseconds) for the HET WFC (90% confidence level).

Field Radius (arcminutes)		0	2.7	5.4	8.1	11.0
$\Phi_0$	Nominal design	0.253	0.229	0.193	0.318	0.584
$\Phi_1$	Mirror fabrication	0.204	0.179	0.118	0.148	0.269
$\Phi_2$	Alignment	0.057	0.065	0.187	0.272	0.295
$\Phi_3$	Assembly and operation	0.014	0.016	0.047	0.068	0.074
$\Phi_4$	Compensators–15 $\mu\text{m}$ accuracy	0.119	0.144	0.176	0.229	0.156
$\Phi_5$	Polishing and support figure	0.25	0.25	0.25	0.25	0.25
$\Phi_6$	Temperature effects	0.000	0.078	0.117	0.158	0.251
$\Phi_7$	Correlated misalignment	0.081	0.027	0.184	0.268	0.276
$\Phi_8$	Entrance and exit windows	0.091	0.065	0.082	0.114	0.099
$\Phi$	Root Sum Square Total	0.448	0.428	0.486	0.653	0.861

results (rows  $\Phi_0$ ,  $\Phi_1$ ,  $\Phi_2$ ,  $\Phi_4$  and  $\Phi_8$ ), as shown in Figure 2.5, as one would expect.

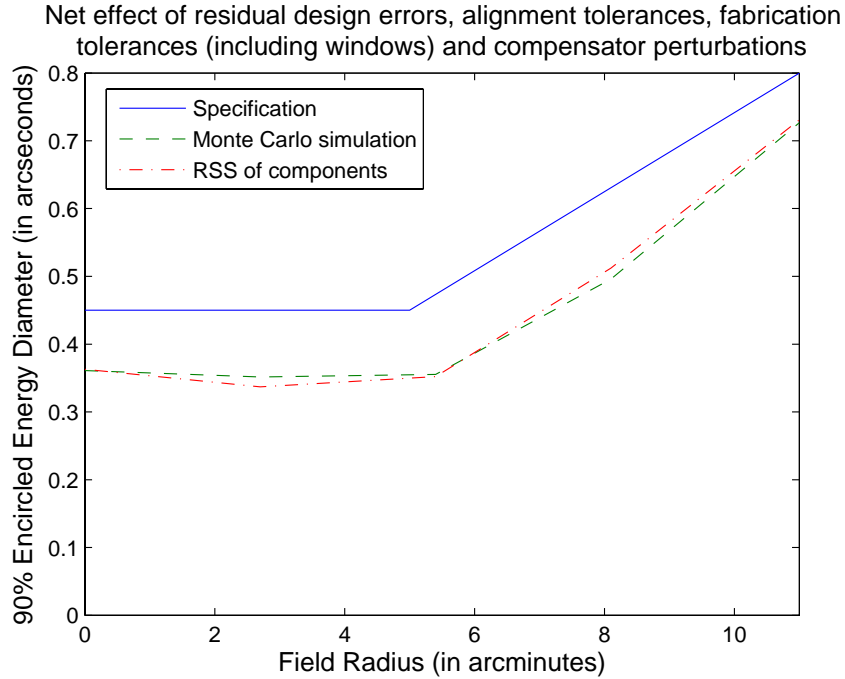


Figure 2.5: Net effect of nominal design; alignment, manufacturing, and window tolerances; and 15  $\mu\text{m}$  compensator perturbations.

## 2.5 OTHER CONFIDENCE LEVEL CURVES

For comparison, Figure 2.6 shows curves for different confidence levels. 90% of the systems generated randomly that met all tolerances (given throughout the report) will have encircled energy diameters equal to or smaller than the yellow line, which was shown in red in Figure 2.4. Other confidence levels, including 25%, 50%, 75%, 95% and 98% are shown now as well.

With such a complex system, the tolerance stack-up is necessarily statistical. It is unlikely, but possible, that many effects will be within tolerance, but will add in such a way that the system performance specification would not be met. This situation must be addressed by extra alignment steps during the system test. In the system test, the wavefront aberrations may be measured throughout the field

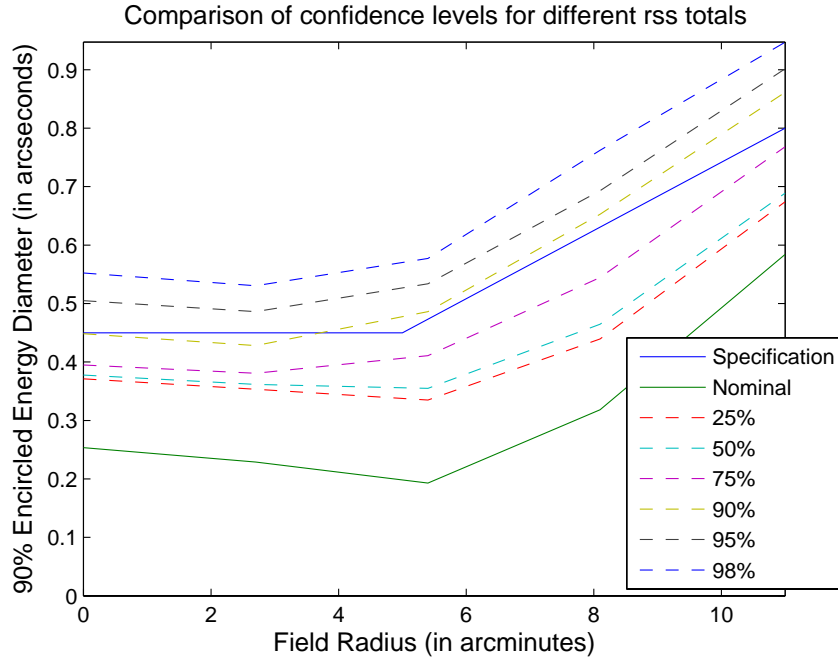


Figure 2.6: Expected performance of the Wide Field Corrector.

and a maximum likelihood estimation of the optimum mirror alignment motions to compensate can be made. This is the type of inverse problem that is considered in the rest of the dissertation. The field-dependent aberrations that are possible in the misaligned HET WFC system are considered in Chapter 8, although the system test itself will not be described in this dissertation.

## 2.6 IMPROVEMENT WHEN MORE COMPENSATORS ARE USED

During any optimization, adding extra degrees of freedom allows the system to have a wider solution space. If the additional degrees of freedom chosen are helpful in correcting the existing errors in the system, then the resulting performance improves. Including focal plane tilt (in two directions) and axial motion during optimization in the Monte Carlo analyses, in addition to the previous five rigid body motions

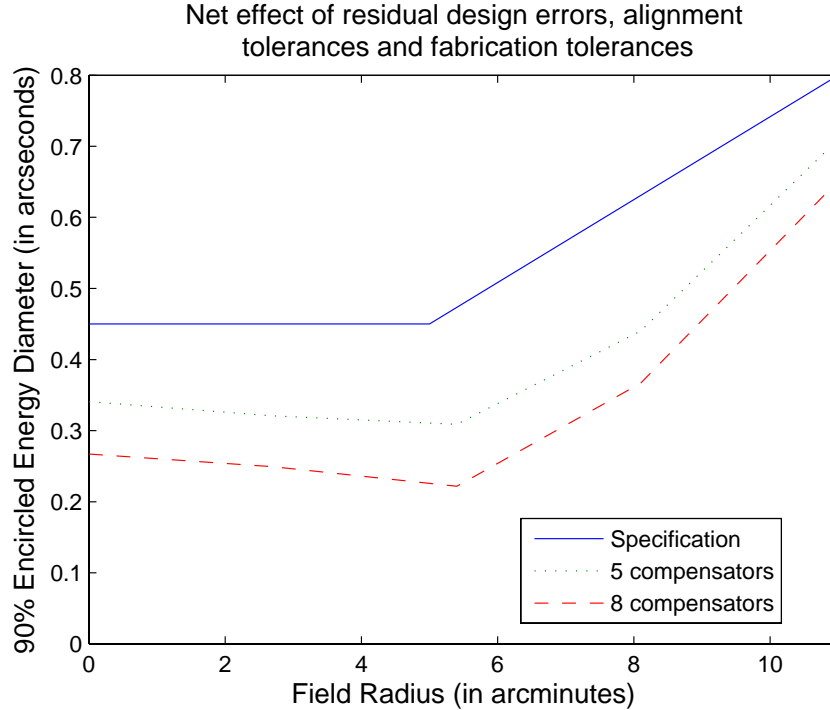


Figure 2.7: The encircled energy diameters are smaller when eight compensator degrees of freedom are included during optimization than when five compensator degrees of freedom are used.

of the corrector allowed by the hexapod, improves the performance of the system. While these additional degrees of freedom are not currently planned to be actively controlled during operation, they may be adjusted in the initial alignment. Figure 2.7 compares the 90% confidence level of the 90% encircled energy diameters from the Monte Carlo simulations when perturbations to the alignment, fabrications and compensator resolution degrees of freedom are included for the case of eight compensators to that of five. The root difference squared between these two curves is approximately 0.2 arcseconds and fairly uniform through the field.

## 2.7 EFFECTIVE FOCAL LENGTH CHANGE DUE TO CORRELATED MIRROR MISALIGNMENTS

When the mirrors move because of changes in orientation or temperature, there is a change in the plate scale, which is equivalent to a change in effective focal length. Table 2.10 shows the change in focal length of the system due to correlated mirror motions for different cases. The system is perturbed according to the values from the FEA analysis (Appendix E) and then the rigid body tilt of the entire corrector and focal plane is adjusted for compensation to minimize the same merit function for rms spot size throughout the field as was used during the Monte Carlo trials.

Table 2.10: Change in effective focal length (after optimization) due to correlated mirror motion and result on maximum relative image motion for two points on opposite edges of the field of view.

Configuration	EFL in mm	$\Delta$ EFL in mm	Change in separation in $\mu$ m
35° in plane, 0° out of plane tilt	36497.4		
26.5° in plane, 0° out of plane tilt	36498.7	1.3	8.32
43.5° in plane, 0° out of plane tilt	36496.2	-1.2	-7.68
35.0° in plane, 8.5° out of plane tilt	36497.4	0.0	0
29.0° in plane, 6.0° out of plane tilt	36498.3	0.90	5.76
41.0° in plane, 6.0° out of plane tilt	36496.5	-0.90	-5.76
Positive temperature change (+20°C)	36498.3	0.90	5.76
Negative temperature change (-20°C)	36496.4	-1.0	6.40

The effective focal lengths resulting from the Monte Carlo analyses that perturb the alignment and fabrication degrees of freedom have much greater variations. The standard deviation of the distribution of effective focal lengths for all the different Monte Carlo simulations was similar. For the set of Monte Carlo trials with the

widest distribution of resulting focal lengths, the standard deviation was 24.5 mm. The most extreme change in focal length from nominal out of all the Monte Carlo trials was 69.9 mm. This means that the nominal EFL once the system is built may be anything within a large range. The overall plate scale will be measured, and offsets from the nominal effective focal length are not important.

The main problem for HET will be plate scale changes that occur during an exposure due to correlated misalignment errors or temperature changes, which can not be corrected. For long exposures, the change in effective focal length can be approximated by determining how much the pointing of the WFC and temperature will change during the time of the exposure. The plate scale of a telescope can be expressed as

$$\text{Plate scale} = \frac{\text{EFL}}{206265 \text{ arcseconds/radian}}. \quad (2.4)$$

The specification for plate scale change is that no two points in the field can move relative to each other by more than 35  $\mu\text{m}$  during an exposure. When the plate scale changes during an exposure, the two points at opposite sides of the field of view (located 22 arcminutes apart) will have more relative motion than any other pair of two points. During an exposure, the relative motion of two points at the opposite sides of the field of view in microns can be calculated by the following expression:

$$\text{Motion} = (22 \text{ arcmin}) \left( \frac{60 \text{ arcsec}}{\text{arcmin}} \right) \left( \frac{1000 \mu\text{m}}{\text{mm}} \right) \left( \frac{\Delta\text{EFL (in mm)}}{206265 \text{ arcsec/rad}} \right). \quad (2.5)$$

This equation was used to calculate the “Change in separation” values in Table 2.10.



In the longest possible exposure, an object would be tracked over the entire range of motion of the WFC from  $26.5^\circ$  to  $43.5^\circ$  elevation angle and two points could change in position by  $8.32 \mu\text{m} - 7.68 \mu\text{m} = 16 \mu\text{m}$ . Since this change in relative position between two points is less than  $35 \mu\text{m}$ , the specification will be met.

## 2.8 RESULTING COMPENSATOR RANGES OF MOTION

One of the important results from the Monte Carlo simulations, in addition to the expected performance, is the expected range of values required of the compensators to correct the system when other errors are present. The compensators are the rigid body motions of the entire WFC and focal plane, which are controlled by the hexapod. These values, listed in Table 2.11, show the range of motion from the hexapod required during the Monte Carlo simulations (when five compensator degrees of freedom were used), and that it should be able to achieve in the system. Table 2.11 lists the maximum change recorded change out of all Monte Carlo trials performed, a typical value for the standard deviation for a set of Monte Carlo trials and two times the standard deviation, which represents the maximum range for 90% of the time.

Table 2.11: Compensator adjustments used during the Monte Carlo simulations and required by the hexapod on the telescope.

Compensator degree of freedom	maximum	$\sigma$	$2\sigma$
axial position (mm)	0.45	0.13	0.26
$x$ or $y$ decenter (mm)	10.6	1.8	3.6
$x$ or $y$ tilt (arcmin)	2.9	0.69	1.4

## 2.9 CONCLUSION

This chapter showed an example of the forward problem using the example of the Hobby-Eberly Telescope Wide Field Corrector, currently under construction at the University of Arizona. Many degrees of freedom were investigated individually to find their expected contribution to the degradation of the system performance. This type of tolerance analysis must always be done before building an optical system in order to ensure both that the final specifications will be met and that the tolerances are not so tight that they are detrimental to the project budget. A careful tolerance analysis is especially important for systems that have challenging specifications, such as this example.

This example addressed the degradation to system performance for many degrees of freedom, including both from manufacturing and alignment. The remainder of this dissertation will concentrate primarily on the errors in the optical wavefront from alignment. This information is useful for inverse-type problems for alignment. That is, given a measurement of the wavefront aberrations, how can one infer the alignment degrees of freedom that are incorrect?

## CHAPTER 3

## BASIC PRINCIPLES OF ABERRATIONS

In an ideal geometrical optical system, there is one to one correspondence between a point object and its point image. This type of system is completely described by first order optics (the focal length of the optical system, the object and image heights and positions, etc.). However in real life, there is no such thing as a perfect optical system and systems fail to make perfect images, due to diffraction and geometrical aberrations. The best possible images, limited only by diffraction when geometrical aberrations are not present, are referred to as diffraction-limited and for point objects appear as the well-known Airy pattern. The wavefront aberrations due to geometric considerations are discussed in this chapter. As a wavefront propagates, however, diffraction does causes the wavefront to change, and the effect is that different aberrations become coupled (Sasian, 2009; Zhou, 2009).

The aberrations in an optical system can be represented as optical path differences (OPD) between the light rays. The path a ray takes depends on how it is refracted or reflected by each element of the system. This path can be uniquely described by knowing the starting location and direction of the ray in any plane, or alternatively by knowing the coordinates of a ray in two different planes. Typically, for the purpose of aberrations, the ray is described by two vectors:  $\vec{H}$ , the field vector in the object plane, and  $\vec{\rho}$ , the pupil vector. These are shown in the right-hand convention in Figure 3.1.

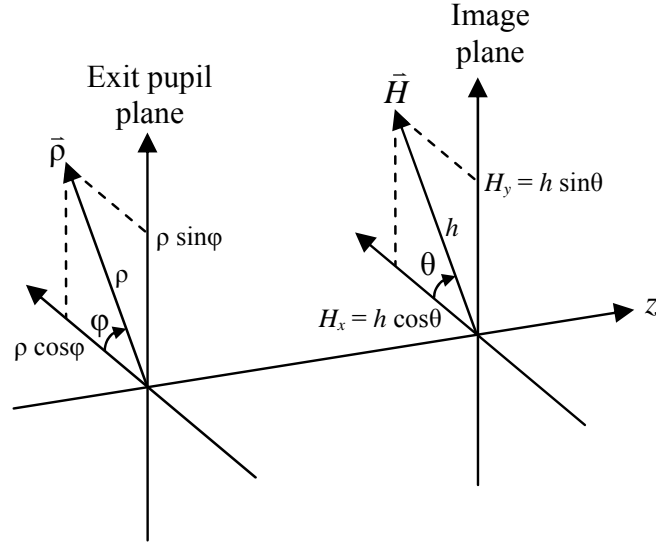


Figure 3.1: Pupil and image conventions in a right-handed coordinate system. The components along the  $\vec{x}$  and  $\vec{y}$  axes respectively are  $\rho \cos \phi$  and  $\rho \sin \phi$  for  $\vec{\rho}$  and  $H_x$  and  $H_y$  for  $\vec{H}$ .

Aberrations can be divided into two broad categories: chromatic aberrations and monochromatic aberrations. Chromatic aberrations occur in refractive optical systems due to the variation of index of refraction with wavelength, and are not considered in this dissertation. Monochromatic aberrations occur in both refractive and reflective optical systems and do not depend on the wavelength. Only monochromatic aberrations occur reflective optical systems, such as telescopes, and this are the type of aberration considered in this dissertation.

An optical wavefront is the surface connecting the ends of all the rays with equal optical path length. As a spherical wavefront emanating from a point source travels through an optical system, it departs from the original spherical shape. If there existed a perfectly spherical converging wavefront departing from the optical system, it would converge back into a point image. The wavefront aberrations in

an optical system are defined relative to this spherical reference wavefront. Many references, including Mahajan (1998, 2001), and Born & Wolf (1999) cover the basics of aberrations in detail.

### 3.1 ABERRATIONS OF ROTATIONALLY SYMMETRIC OPTICAL SYSTEMS

The aberrations in axially symmetric optical systems were first studied in the 1800's by Seidel (1865). Hopkins (1950) created the wave aberration expansion to describe the aberration contributions in a rotationally symmetric system. Welford (1974, 1986) is a classic text which describes these aberrations.

If the coordinate system of a rotationally symmetric optical system is rotated, there are some quantities that remain invariant. Simply stated, these are the length of the pupil and field vectors and the angle between the two vectors. Mathematically, these terms come from combinations of the dot products of the field and pupil vectors. The length of the field vector comes from a dot product of the field vector with itself:  $h = |\vec{H}| = \sqrt{\vec{H} \cdot \vec{H}}$ . The length of the pupil vector comes from a dot product of the pupil vector with itself:  $\rho = |\vec{\rho}| = \sqrt{\vec{\rho} \cdot \vec{\rho}}$ . Finally, the angle between the two vectors is determined by the dot product:  $\vec{H} \cdot \vec{\rho} = h \rho \cos(\theta - \phi)$ . If one assumes (in a rotationally symmetric system) that the object is on the  $\vec{x}$  axis, then  $\theta = 0$  and there is no loss of generality. The dot product between the field and pupil vectors becomes  $\vec{H} \cdot \vec{\rho} = h \rho \cos(\phi)$ .

One can describe the wavefront aberration function  $W(\vec{H}, \vec{\rho})$  as a power series sum of combinations of these vector dot products that remain invariant under

rotation:

$$W(\vec{H}, \vec{\rho}) = \sum_{p=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} W_{klm} (\vec{H} \cdot \vec{H})^p (\vec{\rho} \cdot \vec{\rho})^n (\vec{H} \cdot \vec{\rho})^m. \quad (3.1)$$

The aberration function can also be expressed in the standard scalar form  $W(h, \rho, \phi)$  as originally published by Hopkins:

$$\begin{aligned} W(h, \rho, \phi) &= \sum_{p=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} W_{klm} (h^2)^p (\rho^2)^n (h \rho \cos \phi)^m \\ &= \sum_{p=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} W_{klm} h^{2p+m} \rho^{2n+m} \cos^m \phi \end{aligned} \quad (3.2)$$

where the power of the field is  $k = 2p + m$  and the power of the pupil is  $l = 2n + m$ . The constant  $W_{klm}$  is the weighting coefficient for the aberration term, where the dependence of the aberration on the field  $\vec{H}$  is given by the first subscript (or index), the dependence on the pupil  $\vec{\rho}$  is given by the second subscript, and the dependence on the angle between the field and pupil (given by the dot product  $\vec{H} \cdot \vec{\rho}$ ) is given by the third subscript. This constant includes contributions from all of the optical surfaces  $j$  in the system, such that

$$W_{klm} = \sum_j (W_{klm})_j. \quad (3.3)$$

The order of the wave aberration is found by summing the powers of  $h$  and  $\rho$ . The order is equal to  $2(p + n + m)$  and thus is always even. Terms of order four, six and eight are called the primary (or Seidel), secondary (or Schwarzschild) and tertiary aberrations respectively. Terms of order two in the wave aberration correspond to changes in the first order properties of the optical system and are usually not

included as aberrations, since in those cases, points still image to points. These terms include:

- $W_{200}$ : piston (no effect on the image),
- $W_{020}$ : focus (affects the image plane axial location), and
- $W_{111}$ : tilt (affects the lateral position of image in the image plane).

Including all the fourth order terms in the wave aberration function leads to the equation:

$$\begin{aligned}
 W(h, \rho, \phi) = & W_{040} \rho^4 + W_{131} h \rho^3 \cos \phi + W_{222} h^2 \rho^2 \cos^2 \phi \\
 & + W_{220} h^2 \rho^2 + W_{311} h^3 \rho \cos \phi,
 \end{aligned} \tag{3.4}$$

where the  $W$ 's with subscripts on the right hand side of the equation are the scaling factors that describe the amount of each aberration. An alternate way of describing the aberration function for any one field  $h$  is:

$$W(\rho, \phi) = a_{40} \rho^4 + a_{31} \rho^3 \cos \phi + a_{22} \rho^2 \cos^2 \phi + a_{20} \rho^2 + a_{11} \rho \cos \phi, \tag{3.5}$$

where now the coefficients only have two subscripts, for the powers of  $\rho$  and  $\cos \phi$  respectively. This equation is called the Seidel aberration function and also the primary aberration function (Mahajan, 1998).

The primary aberrations of a rotationally symmetric optical system are familiar to most optical engineers as the five Seidel aberrations. The primary aberrations will describe the low-order aberrations that occur in a rotationally symmetric optical

system. The Seidel aberrations are:

- $W_{040}$ : spherical aberration ( $a_{40}$ ),
- $W_{131}$ : coma ( $a_{31}$ ),
- $W_{222}$ : astigmatism ( $a_{22}$ ),
- $W_{220}$ : field curvature ( $a_{22}$ ), and
- $W_{311}$ : distortion ( $a_{31}$ ).

Ray aberrations, an alternate way of looking at aberrations, describe where the rays intersect the image plane. This is useful for producing spot diagrams, which are visualizations of how images might actually appear. The ray aberrations are described by the coordinates of a ray in the image plane:

$$(H_x, H_y) = \frac{R}{n} \vec{\nabla} W \quad (3.6)$$

where  $R$  is the radius of curvature of the reference sphere,  $n$  is the index of refraction in image space and  $\vec{\nabla}$  is the gradient operator.

Thus, because the ray aberrations are derivatives of the wave aberrations, the order of the ray aberration is one power lower than the corresponding wave aberration. Fourth-order wave aberrations correspond to third-order ray aberrations and sixth-order wave aberrations correspond to fifth-order ray aberrations. This explains why some people use “fourth-order aberrations” while others use “third-order aberrations” to refer to the same thing. In this dissertation, I refer to the order of the



ray aberrations (e.g. third-order and fifth-order aberrations) because it appears to be a little more common.

### 3.1.1 ABERRATION DEFINITIONS

Aberrations are always named according to their dependence on the pupil coordinates,  $\rho$  and  $\phi$ . The particular field dependence that each aberration has should be considered a characteristic of the aberration and not a defining term for the aberration. In this section, each of the Seidel (or primary) aberrations, which are the ones where the combined powers of the field and pupil sum to four, are described. Extra emphasis on the field dependence in a rotationally symmetric system is given because, as shown later, these dependencies change when the system is no longer rotationally symmetric.

#### **$W_{040}$ : Spherical aberration**

Spherical aberration occurs when the rays from a point object focus at different axial positions. The location of the focus depends on which zone of the pupil the rays pass through: the paraxial rays (those that are close to the axis) focus at a different axial position than the marginal rays (those that pass through the edge of the pupil). Thus, there are a variety of focal positions, including those that depend on where the rays pass through the pupil (such as paraxial focus or marginal focus) and those that can be considered “best focus” (under different merits, such as minimum wavefront error, minimum rms spot size, or circle of least confusion), as shown in Figure 3.2. For more information on these, see Mahajan (1998).

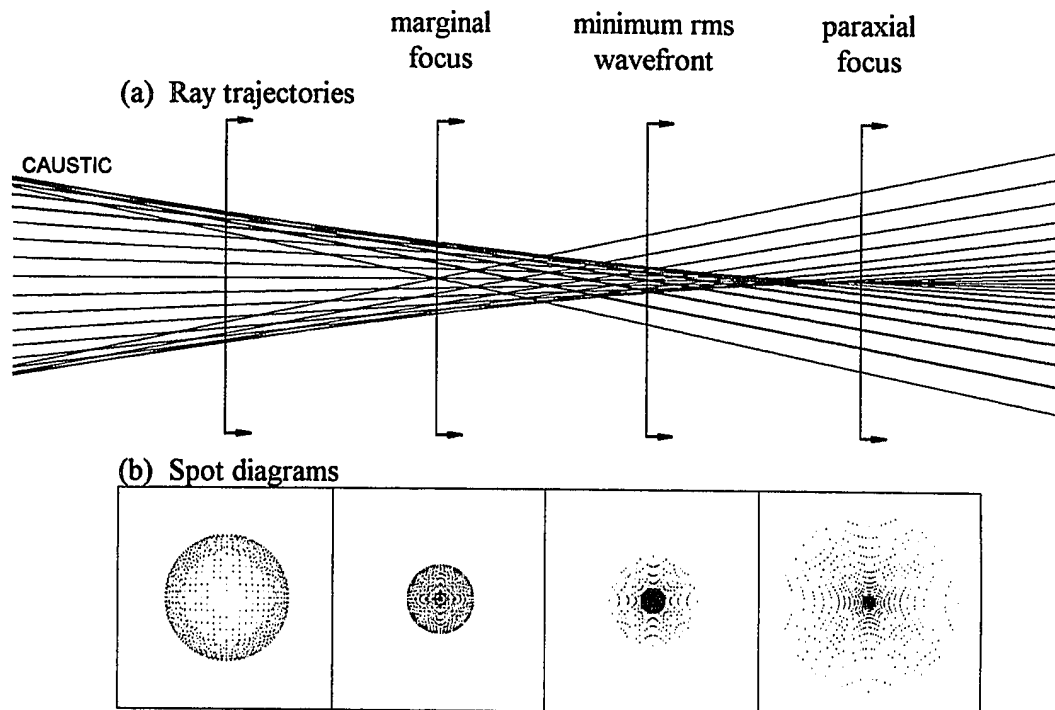


Figure 3.2: Graphical representations of spherical aberration through focus. (a) aberrated rays coming to focus; (b) spot diagrams through focus showing what an image of a point would look like (Burge, 1993).

Spherical aberration has no dependence on the field and since it occurs equally everywhere in the field (see Figure 3.3), including on-axis, designers strive hard to minimize this aberration in the optical system. The term “spherical aberration” comes from the fact that it occurs when a spherical mirror images a source at infinity. To eliminate this aberration from the optical design, mirrors can be made conic (paraboloids etc), or complex correctors, as for the Hobby-Eberly Telescope, can be used. In practice it is difficult to completely eliminate spherical aberration in a system with large mirrors because it is hard to fabricate the mirrors exactly to meet the optical specification. In order to polish the mirror’s conic constant exactly, it must be able to be measured exactly, but even this is hard to do. The Hubble Space Telescope was accidentally made with an especially bad case of spherical aberration when it was original built due to an error in the optical test, but this was subsequently corrected by additional optics. Spherical aberration is not limited to reflective systems. Refractive systems also may have this aberration that depends on the pupil location as  $\rho^4$ .

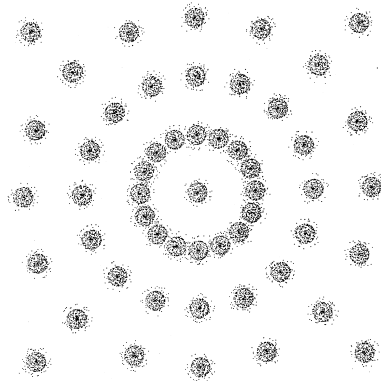


Figure 3.3: Spot diagram for spherical aberration constant throughout the field (from Shack).

### $W_{131}$ : Coma

Coma is an aberration that causes a point object to image into a comet-shaped image. Many think that this is the origin of the word “coma.” However, the name “coma” comes from the Greek word for hair (Wilson, 1996). One online dictionary (foreignword.com) translates “hair” into  $\kappa\omicron\mu\eta$ . Coma occurs when the rays from the different zones of the pupil image to incorrect field positions in the image plane, as shown in Figure 3.4. For pupil positions that are larger, the location of the position of the circles move out. In particular, coma depends on the pupil position as  $\rho^3 \cos \phi$ . In a rotationally symmetric system, coma increases linearly with field angle, as shown in Figure 3.5, and so is larger for small field angles than the other aberrations that have a quadratic or cubic dependence on the field (given the same coefficient values for the aberrations which are measured at the edge of the field). It is the biggest (or most limiting type of) aberration in Cassegrain telescopes. Coma occurs when there is a violation of the sine condition. Satisfying the sine condition actually corrects all linear field-dependent aberrations, not just coma (Zhao & Burge, 2002; Born & Wolf, 1999).

### $W_{222}$ : Astigmatism

Astigmatism is an aberration that occurs when the tangential rays focus at a different axial distance than the sagittal rays, as shown in Figure 3.6. When the object lies on the  $y$ -axis, the tangential rays are those in the  $y - z$  plane, while the sagittal rays are those in the  $x - z$  plane. The particular wavefront aberration that causes

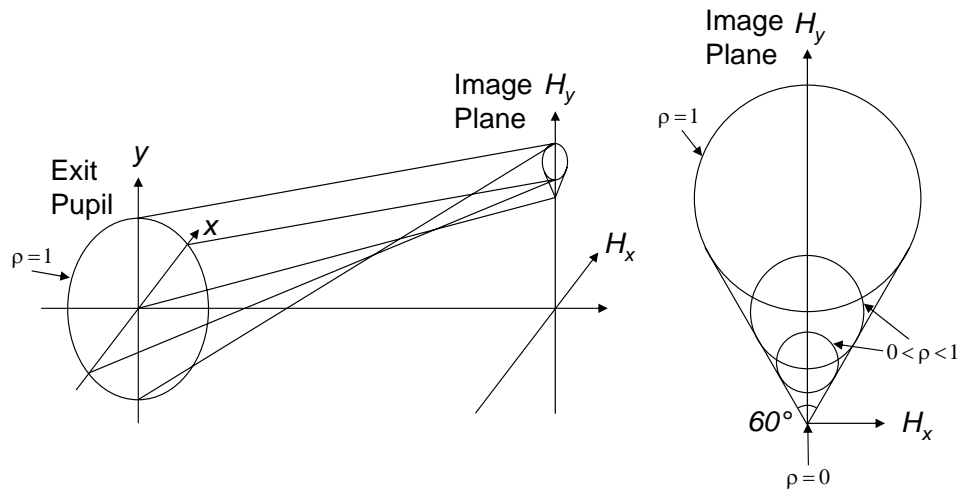


Figure 3.4: Ray diagram for coma.

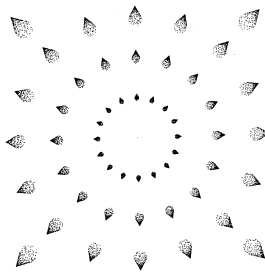


Figure 3.5: Spot diagram for coma linear throughout the field (from Shack).

astigmatism is  $\rho^2\phi^2$ . Astigmatism is quadratic in field in a rotationally symmetric optical system, as shown in Figure 3.7. Line images are formed on the sagittal and tangential focal surfaces, while circular images are formed on the medial focal surface. The sagittal and tangential focal surfaces are shown in Figure 3.8. The tangential and sagittal line images are always at right angles to each other.

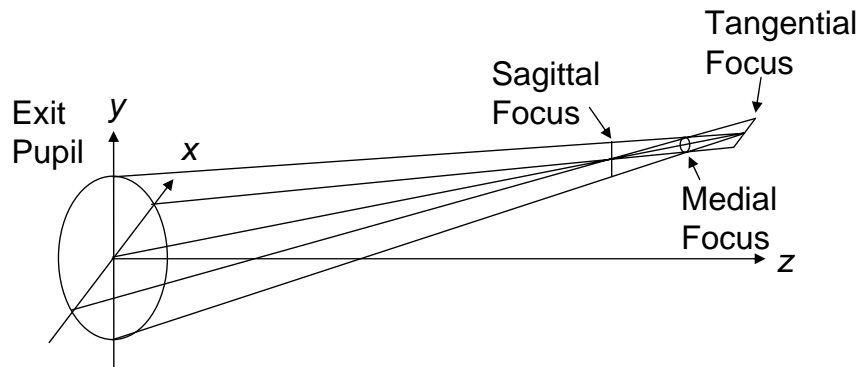


Figure 3.6: Ray diagram for astigmatism.

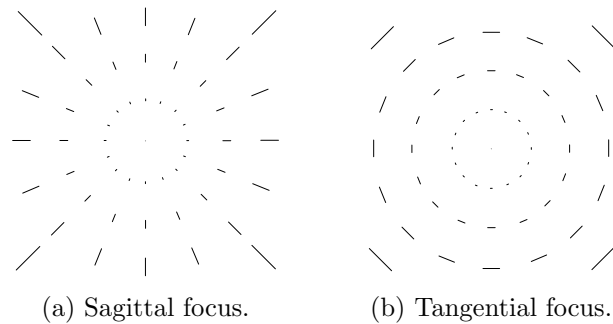


Figure 3.7: Spot diagram for astigmatism quadratic throughout the field.

### $W_{220}$ : Field Curvature

Field curvature occurs because the surface of sharp images is not flat. Field curvature is quadratic in field, as shown in Figure 3.9a. Because of the  $\rho^2$  pupil de-

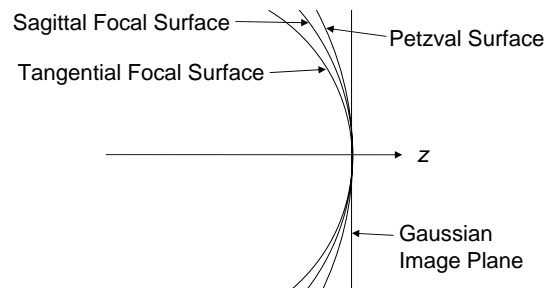
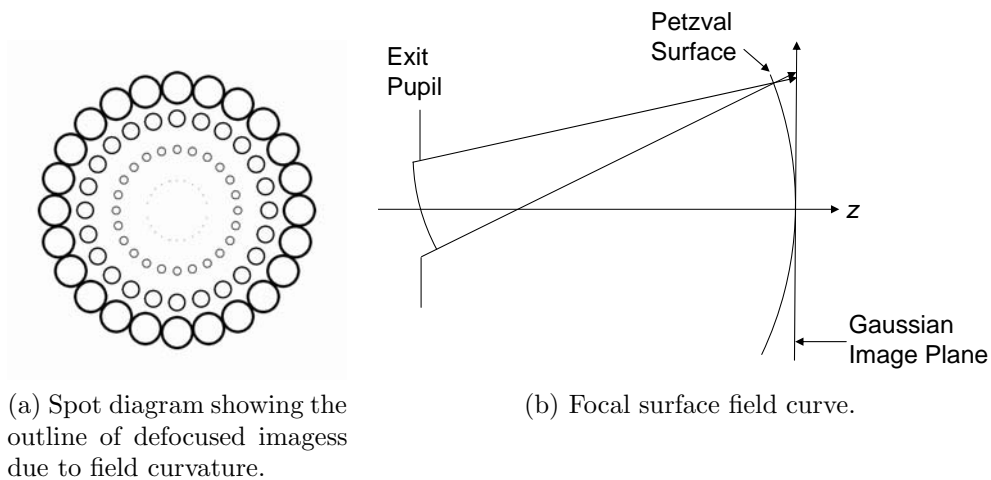


Figure 3.8: Astigmatism field curves. The line astigmatic line images are formed on these surfaces. The medial focal surface (not shown) is located halfway in between the sagittal and tangential focal surfaces.

pendence, field curvature is equivalent to defocus at a single field point. However, field curvature refers to a type of defocus with a specific field dependence. In a rotationally symmetric system, field curvature has a quadratic field dependence. If an image plane could be made spherically curved with the Petzval curvature, this aberration could be eliminated from the optical system.



(a) Spot diagram showing the outline of defocused images due to field curvature.

(b) Focal surface field curve.

Figure 3.9: Field curvature (quadratic with field).

### $W_{311}$ : Distortion

Distortion is a different sort of aberration than the others because with it a point source may still be imaged to a point image. This imaged point is, however, in the wrong position. Since it is a mapping error, distortion is best visualized by showing how a square grid maps in the image plane. A mapping error can also be considered as magnification that is a function of field or a plate scale (scale in the image plane) that is not constant but varies in both field angle and direction. For most applications, distortion is “not so bad” as an aberration because it still allows a point source to image to a perfect point and the mapping error can be calibrated out. In a rotationally symmetric system, distortion is cubic in field, as shown in Figure 3.10.

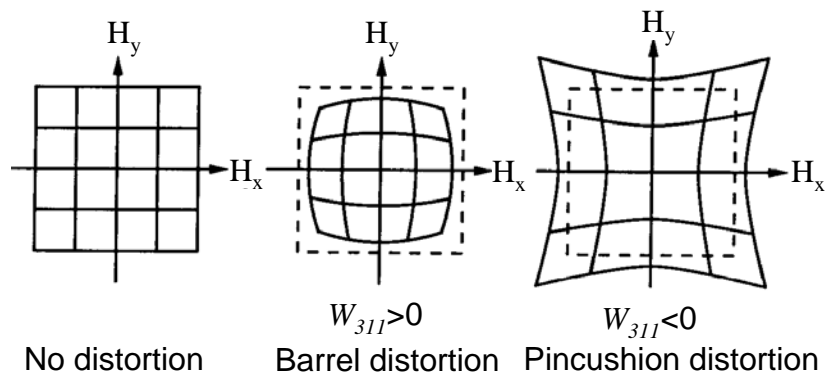


Figure 3.10: Distortion (cubic with field).

### Higher order aberrations

One can continue expanding the wave aberration function to describe even higher order aberrations. Fifth order aberrations include:



- $W_{060}$ : fifth-order spherical aberration,
- $W_{151}$ : fifth-order coma,
- $W_{420}$ : field curvature for fifth-order astigmatism,
- $W_{422}$ : fifth-order astigmatism,
- $W_{511}$ : fifth-order distortion,
- $W_{240}$ : field curvature for oblique spherical aberration,
- $W_{242}$ : oblique spherical aberration,
- $W_{331}$ : field cubed coma, and
- $W_{333}$ : elliptical coma (trefoil).

For more details about these aberrations, please see Buchdahl (1968), Sasian (2009), and Tessieres (2003).

### 3.1.2 DESCRIPTION OF THIRD-ORDER ABERRATIONS USING ZERNIKE POLYNOMIALS

Zernike polynomials (describing in detail in Appendix F) offer an alternate, orthogonal basis for describing the aberration function of an optical wavefront. A wavefront (or an optical surface) can be described as a sum of Zernike polynomials terms such as

$$W(\rho, \phi) = \sum_j C_j Z_j(\rho, \phi), \quad (3.7)$$

where  $C_j$  are the expansion coefficients that represent the scale of the corresponding Zernike polynomials  $Z_j$ , with pupil coordinates  $(\rho, \phi)$ .

Since Equation 3.7 has no explicit field dependence, if one wants to use Zernike polynomials to describe the field dependence of the aberrations, then the field dependence needs to be added to Equation 3.7. One way to do this is to make the expansion coefficients  $C_j$  a function of the field. This can be done in an equation such as

$$W(\rho, \phi, H_x, H_y) = \sum_j C_j(H_x, H_y) Z_j(\rho, \phi), \quad (3.8)$$

where now  $C_j(H_x, H_y)$  is a function that represents the scale and field dependence of the corresponding Zernike polynomials. These functions were found by Tessieres (2003).

For example, the following equation shows the quadratic field dependence that primary astigmatism (Zernike terms 5 and 6) has in a rotationally symmetric optical system:

$$\begin{aligned} W_{\text{astigmatism}}(\rho, \phi, H_x, H_y) &= C_5(H_x, H_y) Z_5(\rho, \phi) + C_6(H_x, H_y) Z_6(\rho, \phi) \\ &= 2 \alpha_0 H_x H_y Z_5(\rho, \phi) + \alpha_0 (H_x^2 - H_y^2) Z_6(\rho, \phi), \end{aligned} \quad (3.9)$$

where  $\alpha_0 = \frac{1}{2\sqrt{6}} W_{222}$ .

As seen in Section 3.1.1, primary coma (Zernike polynomial terms 7 and 8) has a linear field dependence in a rotationally symmetric optical system. The particular linear dependence on  $H_x$  and  $H_y$  for each of the primary coma Zernike polynomial

terms is given by:

$$\begin{aligned} W_{\text{coma}}(\rho, \phi, H_x, H_y) &= C_7(H_x, H_y) Z_7(\rho, \phi) + C_8(H_x, H_y) Z_8(\rho, \phi) \\ &= \beta_0 H_y Z_7(\rho, \phi) + \beta_0 H_x Z_8(\rho, \phi), \end{aligned} \quad (3.10)$$

where  $\beta_0$  relates to  $W_{131}$ , the original Seidel coefficient for coma.

Field curvature has the same pupil dependence as defocus, as mentioned in Section 3.1.1, so it is described by the Zernike focus term  $Z_4(\rho, \phi)$ . Equation 3.11 shows the quadratic field dependence that field curvature has in a rotationally symmetric optical system:

$$\begin{aligned} W_{\text{field curvature}}(\rho, \phi, H_x, H_y) &= C_4(H_x, H_y) Z_4(\rho, \phi) \\ &= \gamma_0 (H_x^2 + H_y^2) Z_4(\rho, \phi), \end{aligned} \quad (3.11)$$

where  $\gamma_0$  relates to the original Seidel field curvature coefficient  $W_{220}$ .

Spherical aberration does not have any dependence on the field. Thus the standard coefficient for the spherical aberration Zernike term  $Z_{11}$  is valid for any field angle:

$$W_{\text{spherical}} = C_{11} Z_{11}(\rho, \phi) = \nu_0 Z_{11}(\rho, \phi) \quad (3.12)$$

where  $\nu_0$  relates to  $W_{040}$ , the original Seidel coefficient for spherical aberration.

If fifth order aberrations are considered, then there exist higher order types of spherical aberration, such as  $W_{240}$ , that have field dependence. This contribution from  $W_{040}$  must be added to the total spherical aberration.

### 3.2 ABERRATIONS OF NON-ROTATIONALLY SYMMETRIC OPTICAL SYSTEMS

The aberrations in non-rotationally symmetric optical systems have been studied at the University of Arizona's College of Optical Sciences over the years. An optical system becomes non-rotationally symmetric when one of the components in the systems becomes tilted or decentered. An optical element is tilted when the optical axis of the element is not parallel to the optical axis of the system. An optical element is decentered when its vertex is displaced from the optical axis. An optical element can also be despaced along the optical axis, but this does not violate the rotational symmetry of the optical system and is not considered here. This degree of freedom may cause rotationally symmetric aberrations (such as defocus or spherical aberration), but the effect is decoupled from that of the non-rotationally symmetric degrees of freedom. The analysis included here for non-rotationally symmetric systems is valid for both small and large deviations from axial symmetry. The equations could be used to aid in the design of non-axial systems. However, the concentration of this dissertation is the study of the aberrations resulting from unintentional misalignments.

One of the first students at the University of Arizona to study aberrations in non-rotationally symmetric optical systems was Buchroeder (1976), a student of Roland Shack. The main result of his work is the hypothesis that the aberration field at the image plane of a misaligned or non-symmetric optical system is still a sum of the individual surface contributions, but that the contributions no longer have a common center on-axis in the image plane (Buchroeder, 1976).

Kevin Thompson, another student of Roland Shack, described the wavefront aberrations (expanded up to 5th order) that can occur in non-symmetric optical systems using the notion proposed by Buchroeder that the field centers become displaced from the optical axis. One important conclusion of Thompson’s dissertation is that no new aberrations are created in non-rotationally symmetric optical systems. The only thing that happens is that new field dependencies occur for those aberrations that are already known, and the new field dependence goes down in order from the power that exists in the rotationally symmetric system. For example, astigmatism which grows quadratically with field in a rotationally symmetric system also has linear and constant contributions (in field) in a non-symmetric system. Thompson also introduced Nodal Aberration Theory (Thompson, 1980, 2005, 2009) to describe the locations in the field where aberrations sum to zero.

### 3.2.1 NODAL ABERRATION THEORY

This section shows the theory of the field centers (or nodes) and then expands the equations for the case of astigmatism. Critical to Thompson’s dissertation is the operation of vector multiplication (established by Roland Shack). This vector operation, described in Appendix G, is neither a dot product nor a cross product and is similar to multiplication of complex numbers when the vectors are written as phasors. In this section, the symbol “.” corresponds to the dot product and no sign between two vectors denotes a vector multiplication.

For the case of a perturbed optical system, the effective field height  $\vec{H}_{Aj}$  replaces

$\vec{H}$  in the wave aberration expansion (Equation 3.1):

$$W(\vec{H}, \vec{\rho}, \vec{\sigma}_j) = \sum_j \sum_{p=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} (W_{klm})_j \left( \vec{H}_{Aj} \cdot \vec{H}_{Aj} \right)^p (\vec{\rho} \cdot \vec{\rho})^n \left( \vec{H}_{Aj} \cdot \vec{\rho} \right)^m. \quad (3.13)$$

This is Equation (3-9) in Thompson (1980). The effective field height for the  $j$ -th surface contribution  $\vec{H}_{Aj}$  may be expressed as:

$$\vec{H}_{Aj} = \vec{H} - \vec{\sigma}_j, \quad (3.14)$$

where the vector  $\vec{\sigma}_j$  in the image plane, first introduced by Buchroeder, describes the location of the center of rotational symmetry for a perturbed surface  $j$  with respect to the optical axis. The vector  $\vec{\sigma}_j$  can be calculated using real-ray-based methods by an optical design program (Thompson *et al.*, 2009). The wave aberration in Equation 3.13 above was modified from Equation 3.1 to also explicitly include the contributions from each of the surfaces  $j$  in the system. The coefficients  $(W_{klm})_j$  are not affected by tilts and decentrations because they are functions of paraxially determined quantities (Thompson, 1980).

In this section, the astigmatism will be defined from the medial surface. The medial surface in the case of astigmatism is the focal surface where point objects image to circles. This surface is inbetween the two surfaces where the astigmatic line images are formed. The medial surface is obtained from the coefficient  $W_{220_M} = W_{220} + \frac{1}{2} W_{222}$ . Thus, the astigmatism relative to the medial surface is given by  $\frac{1}{2} W_{222} \left( \vec{H}^2 \cdot \vec{\rho}^2 \right)$  (instead of  $W_{222} \left( \vec{H}^2 \cdot \vec{\rho}^2 \right)$  which is to the sagittal focal surface).

The astigmatism contribution in a perturbed optical system with respect to the

medial surface is given by Equation (3-33) in Thompson (1980):

$$\begin{aligned}
W &= \frac{1}{2} \sum_j W_{222j} \left[ \left( \vec{H} - \vec{\sigma}_j \right)^2 \cdot \vec{\rho}^2 \right] \\
&= \frac{1}{2} \left[ \sum_j W_{222j} \vec{H}^2 - 2\vec{H} \left( \sum_j W_{222j} \vec{\sigma}_j \right) + \sum_j W_{222j} \vec{\sigma}_j^2 \right] \cdot \vec{\rho}^2 \quad (3.15) \\
&= \frac{1}{2} \left( W_{222} \vec{H}^2 - 2\vec{H} \vec{A}_{222} + \vec{B}_{222}^2 \right) \cdot \vec{\rho}^2,
\end{aligned}$$

where the perturbation vectors  $\vec{A}_{222}$  and  $\vec{B}_{222}^2$  are

$$\vec{A}_{222} \equiv \sum_j W_{222j} \vec{\sigma}_j, \text{ and} \quad (3.16)$$

$$\vec{B}_{222}^2 \equiv \sum_j W_{222j} \vec{\sigma}_j^2. \quad (3.17)$$

The perturbation vectors in Equations 3.16 and 3.17 can be normalized such that:

$$\vec{a}_{222} \equiv \vec{A}_{222}/W_{222}, \text{ and} \quad (3.18)$$

$$\vec{b}_{222}^2 \equiv \vec{B}_{222}^2/W_{222} - \vec{a}_{222}^2. \quad (3.19)$$

Now Equation 3.15 may be written using the normalized vectors as:

$$W = \frac{1}{2} W_{222} \left[ \left( \vec{H} - \vec{a}_{222} \right)^2 + \vec{b}_{222}^2 \right] \cdot \vec{\rho}^2. \quad (3.20)$$

The astigmatism nodes in the image plane may be found by solving Equation 3.20

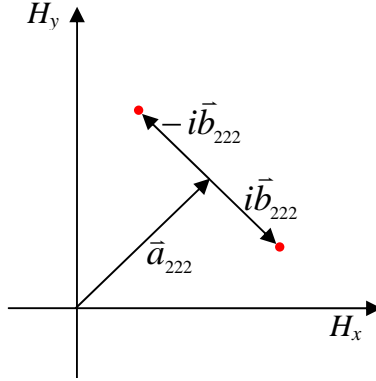


Figure 3.11: Astigmatism nodes in a perturbed optical system. There are two nodes in the field which are determined by the vectors  $\vec{a}_{222}$  and  $\vec{b}_{222}$  (Thompson, 1980).

for  $W = 0$ :

$$0 = \left( \vec{H} - \vec{a}_{222} \right)^2 + \vec{b}_{222}^2 \quad (3.21)$$

$$\vec{H} = \vec{a}_{222} \pm \sqrt{-\vec{b}_{222}^2} \quad (3.22)$$

$$\vec{H} = \vec{a}_{222} \pm i \vec{b}_{222}. \quad (3.23)$$

These two locations where the astigmatism is zero are the nodes, and they are shown conceptually in Figure 3.11. This behavior was dubbed “binodal astigmatism” by Roland Shack.

Thompson (1980) includes figures showing the nodes, such as Figure 3.11, for all of the different aberrations in his dissertation. Tables 3.1 and 3.2 list the rest of the third and fifth-order Seidel aberrations that are possible in a misaligned system developed by Thompson. These equations were used by Tessieres to find the  $H_x$  and  $H_y$  field dependencies of the Zernike polynomials, as described in the next section.

Note that in Table 3.2, the terms for  $W_{242}$ ,  $W_{151}$  and  $W_{240M}$  look very similar (in terms of field-dependence) to the third-order aberration terms in Table 3.1 for



Table 3.1: Third-order aberrations in a misaligned system developed by Thompson (1980). (The symbol “ $\cdot$ ” corresponds to the dot product and no sign between two vectors denotes a vector multiplication.)

<p><b>Spherical aberration:</b> <math>W_{040}</math></p> $W_{040} (\vec{\rho} \cdot \vec{\rho})^2$
<p><b>Third order astigmatism:</b> <math>W_{222}</math></p> $\frac{1}{2} \left( W_{222} \vec{H}^2 - 2\vec{H} \vec{A}_{222} + \vec{B}_{222}^2 \right) \cdot \vec{\rho}^2$
<p><b>Third order coma:</b> <math>W_{131}</math></p> $\left[ \left( W_{131} \vec{H} - \vec{A}_{131} \right) \cdot \vec{\rho} \right] (\vec{\rho} \cdot \vec{\rho})$
<p><b>Field curvature:</b> <math>W_{220_M}</math></p> $\left[ W_{220_M} \left( \vec{H} \cdot \vec{H} \right) - 2 \left( \vec{H} \cdot \vec{A}_{220_M} \right) + B_{220_M} \right] (\vec{\rho} \cdot \vec{\rho})$
<p><b>Third order distortion:</b> <math>W_{311}</math></p> $\left[ W_{311} (\vec{H} \cdot \vec{H}) \vec{H} - 2(\vec{H} \cdot \vec{A}_{311}) \vec{H} + 2B_{311} \vec{H} - (\vec{H} \cdot \vec{H}) \vec{A}_{311} + (\vec{B}_{311}^2 \vec{H}^*) - \vec{C}_{311} \right] \cdot \vec{\rho}$

$W_{222}$ ,  $W_{131}$  and  $W_{220_M}$ , respectively. The only difference is the addition of a factor  $\rho \cdot \rho$ .

The perturbation vectors and scalars derived by Thompson (1980, 2005) and used in Tables 3.1 and 3.2 are summarized in Table 3.3.

Table 3.2: Fifth-order aberrations in a misaligned system developed by Thompson (1980). (The symbol “.” corresponds to the dot product and no sign between two vectors denotes a vector multiplication.)

<p><b>Spherical aberration:</b> <math>W_{060}</math></p> $W_{060} (\vec{\rho} \cdot \vec{\rho})^3$
<p><b>Field cubed coma:</b> <math>W_{331}</math></p> $\left[ W_{331M} (\vec{H} \cdot \vec{H}) \vec{H} - 2 (\vec{H} \cdot \vec{A}_{331M}) \vec{H} + 2B_{331M} \vec{H} - (\vec{H} \cdot \vec{H}) \vec{A}_{331M} + \vec{B}_{331M}^2 \vec{H}^* - \vec{C}_{331M} \right] \cdot \vec{\rho} (\vec{\rho} \cdot \vec{\rho})$
<p><b>Elliptical coma (Trefoil):</b> <math>W_{333}</math></p> $\left[ \frac{1}{4} W_{333} \vec{H}^3 - \frac{3}{4} (\vec{H}^2 \vec{A}_{333}) + \frac{3}{4} (\vec{H} \vec{B}_{333}^2) - \frac{1}{4} \vec{C}_{333}^3 \right] \cdot \vec{\rho}^3$
<p><b>Quartic field curvature:</b> <math>W_{420M}</math></p> $\left[ W_{420M} (\vec{H} \cdot \vec{H})^2 - 4 (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{A}_{420M}) + B_{420M} (\vec{H} \cdot \vec{H}) + 2 (\vec{H}^2 \cdot \vec{B}_{420M}^2) - 4 (\vec{H} \cdot \vec{C}_{420M}) + D_{420M} \right] (\vec{\rho} \cdot \vec{\rho})$
<p><b>Quartic astigmatism:</b> <math>W_{422}</math></p> $\left[ \frac{1}{2} W_{422} (\vec{H} \cdot \vec{H}) \vec{H}^2 - (\vec{H} \cdot \vec{H}) (\vec{H} \vec{A}_{422}) + \frac{3}{2} (\vec{H} \cdot \vec{H}) \vec{B}_{422}^2 - (\vec{H} \cdot \vec{A}_{422}) \vec{H}^2 - \frac{1}{2} \vec{C}_{422}^3 \vec{H}^* + \frac{3}{2} B_{422} \vec{H}^2 - \frac{3}{2} (\vec{H} \vec{C}_{422}) + \frac{1}{2} \vec{D}_{422}^2 \right] \cdot \vec{\rho}^2$
<p><b>Fifth order coma:</b> <math>W_{151}</math></p> $\left[ (W_{151} \vec{H} - \vec{A}_{151}) \cdot \vec{\rho} \right] (\vec{\rho} \cdot \vec{\rho})^2$
<p><b>Oblique spherical aberration:</b> <math>W_{242}</math></p> $\frac{1}{2} \left[ (W_{242} \vec{H}^2 - 2\vec{H} \vec{A}_{242} + \vec{B}_{242}^2) \cdot \vec{\rho}^2 \right] (\vec{\rho} \cdot \vec{\rho})$
<p><b>Field curvature for oblique spherical:</b> <math>W_{240M}</math></p> $\left[ W_{240M} (\vec{H} \cdot \vec{H}) - 2 (\vec{H} \cdot \vec{A}_{240M}) + B_{240M} \right] (\vec{\rho} \cdot \vec{\rho})^2$
<p><b>Fifth order distortion:</b> <math>W_{511}</math></p> $\left[ W_{511} (\vec{H} \cdot \vec{H})^2 \vec{H} - 4 (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{A}_{511}) \vec{H} + 6B_{511} (\vec{H} \cdot \vec{H}) \vec{H} + 2(\vec{H}^2 \cdot \vec{B}_{511}^2) \vec{H} - 4 (\vec{H} \cdot \vec{C}_{511}) \vec{H} + 3D_{511} \vec{H} - (\vec{H} \cdot \vec{H})^2 \vec{A}_{511} + 2 (\vec{H} \cdot \vec{H}) \vec{B}_{511}^2 \vec{H}^* - 4 (\vec{H} \cdot \vec{H}) \vec{C}_{511} - \vec{H}^2 \vec{C}_{511}^* - \vec{C}_{511}^3 \vec{H}^{2*} + 2\vec{D}_{511}^2 \vec{H}^* - \vec{E}_{511} \right] \cdot \vec{\rho}$

Table 3.3: Perturbation vectors and scalars defined by Thompson (1980).

$$\begin{aligned}
W_{klm} &= \sum_j W_{klm_j} \\
\vec{A}_{klm} &= \sum_j W_{klm_j} \vec{\sigma}_j \\
B_{klm} &= \sum_j W_{klm_j} (\vec{\sigma}_j \cdot \vec{\sigma}_j) \\
\vec{B}_{klm}^2 &= \sum_j W_{klm_j} \vec{\sigma}_j^2 \\
\vec{C}_{klm} &= \sum_j W_{klm_j} (\vec{\sigma}_j \cdot \vec{\sigma}_j) \vec{\sigma}_j \\
\vec{C}_{klm}^3 &= \sum_j W_{klm_j} \vec{\sigma}_j^3 \\
D_{klm} &= \sum_j W_{klm_j} (\vec{\sigma}_j \cdot \vec{\sigma}_j)^2 \\
\vec{D}_{klm}^2 &= \sum_j W_{klm_j} (\vec{\sigma}_j \cdot \vec{\sigma}_j) \vec{\sigma}_j^2 \\
\vec{E}_{klm} &= \sum_j W_{klm_j} (\vec{\sigma}_j \cdot \vec{\sigma}_j)^2 \vec{\sigma}_j
\end{aligned}$$

### 3.2.2 DESCRIPTION OF THIRD-ORDER ABERRATIONS USING ZERNIKE POLYNOMIALS

In his thesis, Tessieres expanded on Thompson's work, transferring the aberrations generated from vector multiplication that are possible in a misaligned system into Zernike polynomials for both third and fifth order aberrations (Tessieres, 2003). The derivations for all the 3rd order and 5th order equations are included in his thesis using the convention where the angle is measured clockwise from the  $\vec{y}$ -axis (Tessieres, 2003), which is different from the convention used in this dissertation. These derivations also require vector multiplication as described in Appendix G. Appendix G also includes a discussion about how the equations are modified, depending on whether the angles are measured clockwise from the  $\vec{y}$ -axis or counterclockwise

Table 3.4: Field-dependent functions to describe the Zernike polynomial coefficients for third-order aberrations in a misaligned system.

<b>Third order field curvature:</b> $W_{220M}$	
$C_4(H_x, H_y) = \gamma_0 (H_x^2 + H_y^2) + \gamma_1 H_y + \gamma_2 H_x + \gamma_3$	
<b>Third order astigmatism:</b> $W_{222}$	
$C_5(H_x, H_y) = 2 \alpha_0 H_x H_y + \alpha_1 H_x + \alpha_2 H_y + \alpha_3$	
$C_6(H_x, H_y) = \alpha_0 (H_x^2 - H_y^2) - \alpha_1 H_y + \alpha_2 H_x + \alpha_4$	
<b>Third order coma:</b> $W_{131}$	
$C_7(H_x, H_y) = \beta_0 H_y + \beta_1$	
$C_8(H_x, H_y) = \beta_0 H_x + \beta_2$	
<b>Spherical aberration:</b> $W_{040}$	
$C_{11} = \nu_0$	

from the  $\vec{x}$  axis.

Table 3.4 lists the field-dependent Zernike coefficients in a misaligned optical system for astigmatism, coma, and focus (field curvature) using the standard Zernike polynomial ordering (Noll, 1976). These equations will use the same Greek letter variables (e.g.  $\alpha$  for astigmatism,  $\beta$  for coma) for the aberration coefficients as in Section 3.1.2, however now there will be nonzero values for the subscripts for field dependencies that did not exist in an aligned system. Spherical aberration is included here for completeness, although it is unchanged by misalignments. Distortion is not included because it is not easily measured with wavefront sensors in a misaligned system. Also, the Zernike polynomial equations for distortion can not be directly derived from Thompson's equation for  $W_{311}$  because they would also require additional terms to cancel the tilt terms in  $Z_7(\rho, \phi)$  and  $Z_8(\rho, \phi)$  terms which are not part of third-order coma.

Table 3.5 lists a conversion between the Greek letter notation and Thompson’s notation. The terms are listed in order of increasing subscript value for each aberration. It is interesting to note that the terms from Thompson that are scalars correspond to one Greek letter coefficient according to Tessieres, while the terms from Thompson that are vectors correspond to two Greek letter coefficients, which makes sense. In Table 3.4, the aberration corresponding to  $W_{220_M}$  is labeled “third-order field curvature,” however the name corresponding to the  $W_{220_M}$  is changed to “focus” in Table 3.5. This makes more sense because, when one observes the field dependence of the terms in the equation, there are terms that are linear and constant in the field. This linear term is often called “focal plane” tilt. The constant term is simply defocus. The terms in Table 3.5 that are directly proportional to the misalignment are the ones that are often discussed as being induced by misalignments in the system. These terms include focal plane tilt (linear focus), linear astigmatism, and constant coma.

Table 3.6 lists the *additional* field-dependent Zernike coefficients in a misaligned optical system for the 5th order aberrations (still using the standard Zernike polynomial ordering (Noll, 1976)). As before, the terms with subscript zero occur in a rotationally symmetric system and the higher subscripts indicate the terms that can be induced due to misalignments. Once again, distortion is not included because it is not easily measured with wavefront sensors in a misaligned system.

Table 3.5: Conversion from Thompson's notation into Tessieres's notation for third-order aberrations.

Zernike Term	Aberration Name	Field dependence	Tessieres	Thompson	Dependence on misalignment
4	Focus	quadratic	$\gamma_0$	$W_{220_M}$	none
		linear	$\gamma_1, \gamma_2$	$\vec{A}_{220_M}$	$\propto$ perturbation
		constant	$\gamma_3$	$B_{220_M}$	$\propto$ perturbation <sup>2</sup>
5, 6	Astig.	quadratic	$\alpha_0$	$W_{222}$	none
		linear	$\alpha_1, \alpha_2$	$\vec{A}_{222}$	$\propto$ perturbation
		constant	$\alpha_3, \alpha_4$	$\vec{B}_{222}$	$\propto$ perturbation <sup>2</sup>
7, 8	Coma	linear	$\beta_0$	$W_{131}$	none
		constant	$\beta_1, \beta_2$	$\vec{A}_{131}$	$\propto$ perturbation
11	Spherical	constant	$\nu_0$	$W_{040}$	none

Table 3.6: Field-dependent functions to describe the Zernike polynomial coefficients for fifth-order aberrations in a misaligned system.

<b>Quartic field curvature: <math>W_{420}</math></b>	
$C_4(H_x, H_y) = \psi_0 (H_x^4 + H_y^4 + 2 H_x^2 H_y^2) + \psi_1 (H_y^3 + H_x^2 H_y)$	
$\quad + \psi_2 (H_x^3 + H_x H_y^2) + \psi_3 (H_x^2 + H_y^2) + \psi_4 H_x H_y$	
$\quad + \psi_5 (H_x^2 - H_y^2) + \psi_6 H_y + \psi_7 H_x + \psi_8$	
<b>Quartic astigmatism: <math>W_{422}</math></b>	
$C_5(H_x, H_y) = 2 \chi_0 (H_x H_y^3 + H_x^3 H_y) + \chi_1 (3 H_x^2 H_y + H_y^3)$	
$\quad + \chi_2 (H_x^3 + 3 H_x H_y^2) + \chi_3 (H_x^2 + H_y^2) + \chi_5 2 H_x H_y - \chi_6 H_y$	
$\quad + \chi_7 H_x + \chi_8 H_y + \chi_9 H_x + \chi_{10}$	
$C_6(H_x, H_y) = \chi_0 (H_x^4 - H_y^4) + 2 \chi_1 H_x^3 - 2 \chi_2 H_y^3 + \chi_4 (H_x^2 + H_y^2)$	
$\quad + \chi_5 (H_x^2 - H_y^2) + \chi_6 H_x + \chi_7 H_y + \chi_8 H_x - \chi_9 H_y + \chi_{11}$	
<b>Field cubed coma: <math>W_{331}</math></b>	
$C_7(H_x, H_y) = \xi_0 (H_y^3 + H_x^2 H_y) + \xi_1 H_y^2 + \xi_2 H_x H_y + \xi_3 H_y + \xi_4 (H_x^2 + H_y^2)$	
$\quad + \xi_6 H_x + \xi_7 H_y + \xi_8$	
$C_8(H_x, H_y) = \xi_0 (H_x^3 + H_x H_y^2) + \xi_1 H_x H_y + \xi_2 H_x^2 + \xi_3 H_x + \xi_5 (H_x^2 + H_y^2)$	
$\quad + \xi_6 H_y - \xi_7 H_x + \xi_9$	
<b>Elliptical coma (trefoil): <math>W_{333}</math></b>	
$C_9(H_x, H_y) = \mu_0 (3 H_x^2 H_y - H_y^3) + \mu_1 (H_x^2 - H_y^2) + \mu_2 2 H_y H_x$	
$\quad + \mu_3 H_y + \mu_4 H_x + \mu_5$	
$C_{10}(H_x, H_y) = \mu_0 (H_x^3 - 3 H_x H_y^2) - 2 \mu_1 H_y H_x + \mu_2 (H_x^2 - H_y^2)$	
$\quad + \mu_3 H_x - \mu_4 H_y + \mu_6$	
<b>Fifth order field curvature: <math>W_{240M}</math></b>	
$C_{11}(H_x, H_y) = \delta_0 (H_x^2 + H_y^2) + \delta_1 H_y + \delta_2 H_x + \delta_3$	
<b>Oblique spherical aberration: <math>W_{242}</math></b>	
$C_{12}(H_x, H_y) = \eta_0 (H_x^2 - H_y^2) - \eta_1 H_y + \eta_2 H_x + \eta_3$	
$C_{13}(H_x, H_y) = \eta_0 2 H_x H_y + \eta_1 H_x + \eta_2 H_y + \eta_4$	
<b>Fifth order coma: <math>W_{151}</math></b>	
$C_{16}(H_x, H_y) = \kappa_0 H_x + \kappa_1$	
$C_{17}(H_x, H_y) = \kappa_0 H_y + \kappa_2$	

Table 3.7 lists a conversion between the Greek letter notation and Thompson's notation. As in Table 3.5, the rows are listed in the order of Zernike polynomial coefficients for increasing Greek letter subscript values. The terms that scale directly to the misalignment now include cubic focus, cubic astigmatism, quadratic coma, quadratic trefoil, linear oblique spherical aberration, and constant secondary ( $\rho^5$ ) coma. These terms are the aberrations that one should expect to find in a misaligned system with many degrees of freedom after constant coma, linear astigmatism and focal plane tilt have been corrected.

While before the scalars from Thompson corresponded to one Greek letter coefficient, and the vectors corresponded to two coefficients, this does not have to be the case. When there are multiple terms that use the same vectors (for example  $(-\frac{1}{2}\vec{C}_{422}^3 \vec{H}^*)$  and  $(-\frac{3}{2}\vec{H} \vec{C}_{422})$ ), two Greek letters could be chosen to correspond to  $C_{422x}$  and  $C_{422y}$ , but since these coefficients are coupled in a complex way to the field dependencies, it is simpler to use four coefficients.

Table 3.7: Conversion from Thompson's notation into Tessieres's notation for fifth-order aberrations.

Zernike Term	Aberration Name	Field dependence	Tessieres	Thompson	Dependence on misalignment
4	Focus	quartic	$\psi_0$	$W_{420M}$	none
		cubic	$\psi_1, \psi_2$	$\vec{A}_{420M}$	$\propto$ perturbation
		quadratic #1	$\psi_3$	$B_{420M}$	$\propto$ perturbation <sup>2</sup>
		quadratic #2	$\psi_4, \psi_5$	$\vec{B}_{420M}$	$\propto$ perturbation <sup>2</sup>
		linear	$\psi_6, \psi_7$	$\vec{C}_{420M}$	$\propto$ perturbation <sup>3</sup>
		constant	$\psi_8$	$D_{420M}$	$\propto$ perturbation <sup>4</sup>

*continued on next page*



Table 3.7: *continued*

Zernike Term	Aberration Name	Field dependence	Tessieres	Thompson	Dependence on misalignment
5, 6	Astig.	quartic	$\chi_0$	$W_{422}$	none
		cubic	$\chi_1, \chi_2$	$\vec{A}_{422}$	$\propto$ perturbation
		quadratic #1	$\chi_3, \chi_4$	$\vec{B}_{422}$	$\propto$ perturbation <sup>2</sup>
		quadratic #2	$\chi_5$	$B_{422}$	$\propto$ perturbation <sup>2</sup>
		linear #1	$\chi_6, \chi_7$	$\vec{C}_{422}$	$\propto$ perturbation <sup>3</sup>
		linear #2	$\chi_8, \chi_9$	$\vec{C}_{422}$	$\propto$ perturbation <sup>3</sup>
		constant	$\chi_{10}, \chi_{11}$	$\vec{D}_{422}$	$\propto$ perturbation <sup>4</sup>
7, 8	Coma	cubic	$\xi_0$	$W_{331_M}$	none
		quadratic #1	$\xi_1, \xi_2$	$\vec{A}_{331_M}$	$\propto$ perturbation
		quadratic #2	$\xi_4, \xi_5$	$\vec{A}_{331_M}$	$\propto$ perturbation
		linear #1	$\xi_3$	$B_{331_M}$	$\propto$ perturbation <sup>2</sup>
		linear #2	$\xi_6, \xi_7$	$\vec{B}_{331_M}$	$\propto$ perturbation <sup>2</sup>
		constant	$\xi_8, \xi_9$	$\vec{C}_{331_M}$	$\propto$ perturbation <sup>3</sup>
9,10	Trefoil	cubic	$\mu_0$	$W_{333}$	none
		quadratic	$\mu_1, \mu_2$	$\vec{A}_{333}$	$\propto$ perturbation
		linear	$\mu_3, \mu_4$	$\vec{B}_{333}$	$\propto$ perturbation <sup>2</sup>
		constant	$\mu_5, \mu_6$	$\vec{C}_{333}$	$\propto$ perturbation <sup>3</sup>
11	Oblique Sph.	quadratic	$\delta_0$	$W_{240_M}$	none
		linear	$\delta_1, \delta_2$	$\vec{A}_{240_M}$	$\propto$ perturbation
		constant	$\delta_3$	$B_{240_M}$	$\propto$ perturbation <sup>2</sup>
12, 13	Oblique Sph.	quadratic	$\eta_0$	$W_{242}$	none
		linear	$\eta_1, \eta_2$	$\vec{A}_{242}$	$\propto$ perturbation
		constant	$\eta_3, \eta_4$	$\vec{B}_{242}$	$\propto$ perturbation <sup>2</sup>
16, 17	Coma ( $\rho^5$ )	linear	$\kappa_0$	$W_{151}$	none
		constant	$\kappa_1, \kappa_2$	$\vec{A}_{151}$	$\propto$ perturbation

### 3.3 CONCLUSION

This chapter introduced the wavefront aberration function to describe the aberrations in optical systems. First, rotationally symmetric systems were considered and the different types of aberrations (including spherical aberration, coma, astigma-

tism, field curvature and distortion) were introduced. Each of these aberrations has a characteristic field dependence in an axially-symmetric system. However, in non-rotationally symmetric systems, the aberrations develop different field dependencies. Equations (developed by Tessieres) were introduced to describe the functional dependence of the aberrations when they are measured as Zernike polynomials in a misaligned system. Chapter 4 shows an application using these equations for studying the alignment of the New Solar Telescope (NST). Chapter 4 also examines in more detail and shows plots of two of the most common aberrations found by Tessieres (and others) that are possible in a misaligned two mirror system including constant astigmatism and linear astigmatism.

## CHAPTER 4

DEMONSTRATION OF COUPLING OF LOW ORDER FIELD-DEPENDENT  
ABERRATIONS TO ALIGNMENT OF AN OFF-AXIS TWO MIRROR  
TELESCOPE

The aberrations due to misalignment of the secondary mirror in the New Solar Telescope (NST) are reviewed in this chapter. The majority of this chapter was previously published as an SPIE proceeding (Manuel & Burge, 2009) and appears here with permission of SPIE. The NST is an off-axis Gregorian telescope at Big Bear Solar Observatory (BBSO) (Didkovsky *et al.*, 2004). The telescope will operate in the wavelength range of 390 nm into the far infrared. The design of the telescope is presented in Section 4.1. Section 4.2 includes a discussion of the aberrations due to misalignments of the secondary mirror for a general on-axis Gregorian telescope with the same first order properties as NST. The observed aberrations of constant coma and linear astigmatism for this “parent” telescope from the ZEMAX model match what is expected in a misaligned telescope. Section 4.3 includes a discussion of the changes to the observed aberrations due to the off-axis nature of the NST. Finally, Section 4.5 discusses the effect of the misalignments on the pointing of the telescope. A summary of the chapter is given in Section 4.6

#### 4.1 OPTICAL DESIGN OF THE NEW SOLAR TELESCOPE

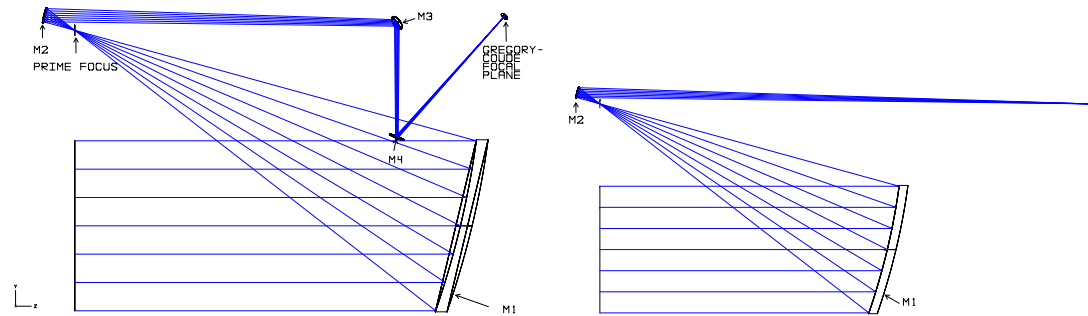
The New Solar Telescope is an off-axis telescope with the same design parameters (e.g. radii, conic constants) as a theoretical on-axis parent telescope, but only certain off-axis sections of the theoretical on-axis telescope are fabricated and used. The reasons for the off-axis design, discussed in detail by Didkovsky *et al.* (2004), are to avoid having a central obscuration and spiders. The primary and secondary mirrors (M1 and M2) of NST together compose a Gregorian telescope and then either one or two flat fold mirrors (M3 and M4) are used to send the light to either the Nasmyth or Gregory-Coudé focal plane. The two mirror system prescription is listed in Table 4.1.

Table 4.1: NST system prescription.

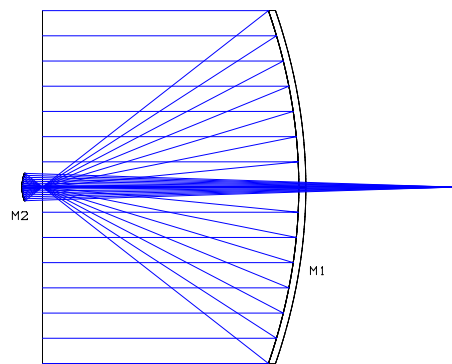
Surface Name	Radius of curvature (mm)	Conic constant	Thickness (mm)
M1	-7700	-1	3850
Prime Focus			300.05
M2	573.5828	-0.83087	6490.259
Gregorian Focus			

The (off-axis) primary mirror (f/2.4) has a radius of curvature of 7700 mm and a clear aperture diameter of 1600 mm, as shown in the telescope design in Figure 4.1a. This mirror was figured and tested at the University of Arizona as a 1/5 scale demonstration (Martin *et al.*, 2006) for the 8.4 m diameter off-axis mirror segments of the 25 m f/0.7 Giant Magellan Telescope (GMT) (Johns, 2008) primary mirror. For simplification during the analysis of the off-axis telescope, the two fold mirrors were removed, as in Figure 4.1b. To make the axisymmetric parent telescope of NST in ZEMAX, the coordinate breaks and off-axis aperture definitions were removed, as

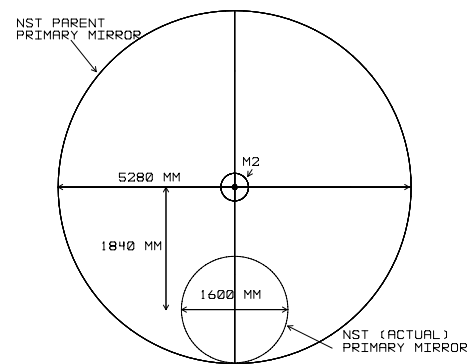
shown in Figure 4.1c. The aperture stop is the primary mirror and a comparison of the two pupils is shown in Figure 4.1d—the end view of the optical design.



(a) NST with all four mirrors ( $f/52.05$  system) (b) NST after fold mirrors removed ( $f/52.05$  system)



(c) Side view of NST on-axis parent telescope ( $f/15.77$  system)



(d) End view of the two pupils

Figure 4.1: New Solar Telescope optical layout. The  $f$ /number calculations are provided in Table 4.2.

Table 4.2 provides a comparison of the  $f$ /numbers for the NST system and primary mirror for both the actual and parent pupils.

Table 4.2: NST f/number calculations.

Label	$f/\#$	Calculation
NST primary mirror	$f/2.406$	3850/1600
NST system	$f/52.05$	83277.8/1600
Parent primary mirror	$f/0.729$	3850/5280
Parent system	$f/15.77$	83277.8/5280

## 4.2 ABERRATIONS OF THE PARENT ON-AXIS GREGORIAN TELESCOPE

### 4.2.1 DEFINITION OF THE DEGREES OF FREEDOM STUDIED FOR NST

The analysis in this chapter assumes that all the optics are fabricated and supported perfectly and that the only aberrations beyond the residual aberrations are introduced by rigid body misalignments of those optics. (The residual design aberrations are those that exist in the optical model, since it is impossible to design a system to be absolutely perfect.) The image quality of a two mirror telescope is dependent on the position of the secondary mirror relative to the primary mirror. The position of the primary mirror, which serves as the reference, only affects the pointing of the telescope. As for any object, the secondary mirror has six rigid body degrees of freedom as defined from the surface vertex:  $\vec{x}$ ,  $\vec{y}$ , and  $\vec{z}$  rotations; and  $\vec{x}$ ,  $\vec{y}$ , and  $\vec{z}$  translations. The first two degrees of freedom for the secondary mirror may be defined as: 1. rotation about the optical axis, which has no effect in a rotationally symmetric system and 2. translation of the secondary along the optical axis, which only causes defocus and a small amount of spherical aberration. These degrees of freedom will not be further discussed in this chapter because their effect on system performance is decoupled from the remaining non-axisymmetric degrees of freedom,

which are the emphasis of the chapter. The other four degrees of freedom (two tilts and two decenters) of the secondary mirror will each individually cause both coma and astigmatism where the sign of the aberration depends on the direction of the misalignment and the amount depends on the magnitude of the perturbation. Other higher order aberrations may occur in addition to third order coma and astigmatism, but are not significant in a system like this without too many degrees of freedom.

However, it is possible to define the degrees of freedom differently such that the  $\vec{x}$  and  $\vec{y}$  translations and rotations about the vertex are combined into rotations about two different points along the optical axis. Different linear combinations of rotations about these points can achieve any of the same vertex rotations plus translations as before. In a two-mirror telescope, there is a point on the optical axis where rotations about this point do not introduce any coma (Wilson, 1996; Schroeder, 1999; Mahajan, 1998). This “coma-free point” or “neutral point” exists because different misalignment degrees of freedom produce the same types of aberrations with different magnitude ratios. Thus, it is possible to cancel the coma introduced by a decenter with the coma that is introduced by a tilt (if properly chosen). This linear combination of tilt and decenter is equivalent to a rotation about the coma-free point and since the effect of a perturbation on the induced coma is linear, it is always balanced for any rotation angle of the secondary mirror about the coma-free point. However, rotations about this coma-free point still introduce field astigmatism. The coma-free point for a classical telescope such as this is at the focus of the primary mirror, as explained the next paragraph. For more complex systems, the position of the coma-free point can be found using the equations in the references (Schroeder,

1999; Wilson, 1996) or a lens design program. (This point will be abbreviated PF for prime focus in the figures and tables below.) In NST, the prime focus is located 300.1 mm away from the vertex of the secondary mirror, as shown in Figures 4.2 and 4.3a.

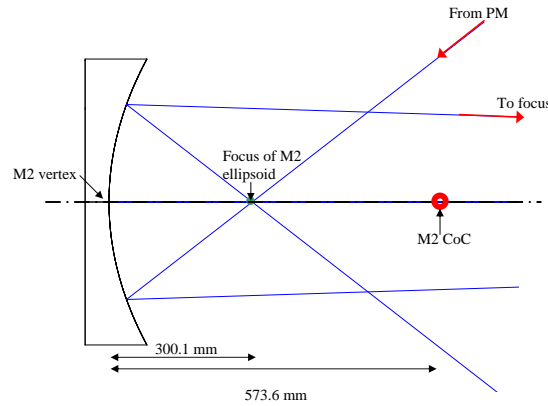


Figure 4.2: M2 rotation points shown in the nominal aligned system.

In a classical telescope, the primary mirror is a paraboloid (conic constant  $k = -1$ ). The type of telescope is defined by whether the secondary mirror is concave (Gregorian) or convex (Cassegrain). In a Gregorian telescope, the secondary mirror is an oblate ellipsoid ( $-1 < k < 0$ ) such that one focus of the ellipse is at the prime focus and the other is at the system focal plane. The two foci are conjugate points and the light rays can transfer perfectly from one focus to the other focus according to Fermat's Principle. As the secondary mirror ellipsoid rotates about the prime focus (one of its focal points), the rays for the on-axis field that go through the prime focus on the optical axis still can form a perfect image at the other focal point (the system focus) which is displaced from nominal by the rotation. However, as the telescope field angle changes, aberrations from the primary and secondary



mirrors are created.

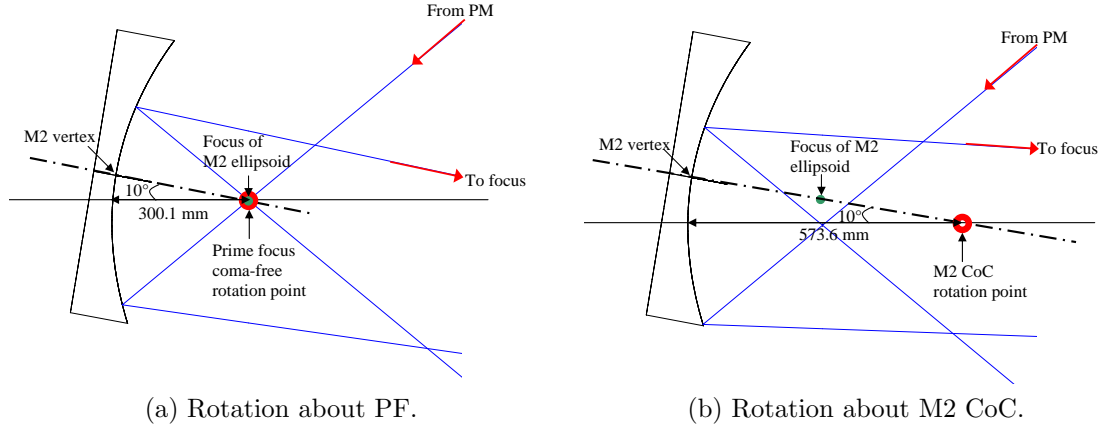


Figure 4.3: M2 degrees of freedom. (The rotations are shown larger than typical misalignments to clearly show the location of the rotation point.)

A spot diagram for NST with M2 having some rotation about the coma-free point is shown in Figure 4.4a. This spot diagram shows astigmatism, which causes the focused spots to be elongated. At the sagittal and tangential image surfaces, the images would form perfect line images. (The tangential image surface is that where line images from objects in the  $y - z$  plane are formed, while the sagittal image surface is that where line images from objects in the  $x - z$  plane are formed.) In particular, the astigmatism shown here has a linear field dependence, which can be seen by the length of the spots throughout the field and the particular orientation pattern of the elongated spots throughout the field. For the large rotation angle used in Figure 4.4a, the pointing error would be very large, so it is not included in the spot diagram.

Specifying  $\vec{x}$  and  $\vec{y}$  rotations about the coma-free point uses two degrees of freedom. The remaining two degrees of freedom may either be regular vertex tilts or decenters or a combination that results in rotation about another point on the axis.

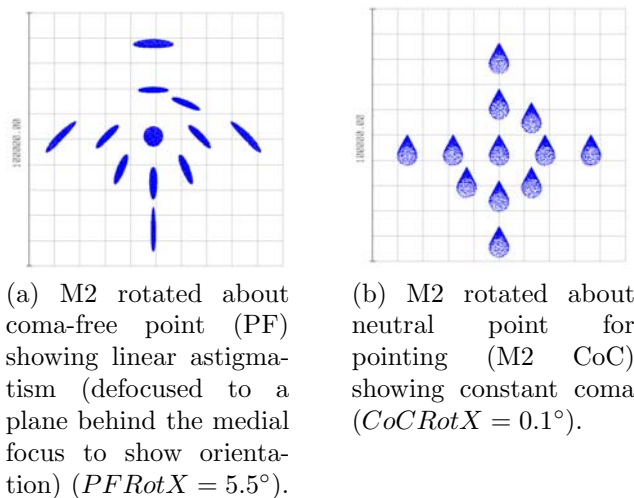


Figure 4.4: Effect of M2 degrees of freedom on spot diagrams. The maximum field angle spots are shown for a full field of view of 3 arcminutes. Spot sizes are exaggerated by  $4\times$  for clarity. (These spot diagrams are what you would see looking toward the focal plane, and not from behind it.)

For this chapter, the final two degrees of freedom are defined as  $\vec{x}$  and  $\vec{y}$  rotations of the secondary mirror about its center of curvature (CoC). Perturbing this degree of freedom results in the aberrations of constant coma and linear astigmatism; however it does not result in a pointing error for the paraxial rays. In NST, the CoC of the secondary mirror is located 573.6 mm away from the vertex of the secondary mirror, as shown in Figure 4.3b. The center of curvature of the secondary mirror is always a pointing-free neutral point for paraxial rays (independent of the secondary mirror shape) (Wilson, 1996). This can be visualized by considering the case of a spherical secondary mirror. For this case, rotations about the center of curvature have absolutely no effect on the shape of the optical surface. The paraxial rays which do not have a pointing error in Figure 4.4b are those at the pointed tip of the coma pattern. (Comatic images are identified by their teardrop or comet shape.) Since the coma aberration causes a nonsymmetric image blur, the centroid pointing does indeed

change for center of curvature rotations. If desired, a lens design program could be used to find an alternate rotation point that does not cause centroid pointing errors. This point would be near the M2 center of curvature, but slightly displaced along the optical axis.

The definition of these degrees of freedom for the secondary mirror provides a convenient choice because it decouples the effects of coma and pointing. Table 4.3 summarizes the results that a rotation of the secondary mirror about the coma-free point will only introduce astigmatism and a pointing change, while a rotation of the secondary about its center of curvature causes no pointing error, but introduces both constant coma and linear astigmatism. The name of the aberration (i.e. coma or astigmatism) describes the shape of the wavefront in the exit pupil while the preceding modifier (i.e. constant or linear) describes the magnitude of the aberration as a function of the field of view.

Table 4.3: Secondary mirror degrees of freedom and effects on aberrations.

<b>Rotation point for secondary mirror</b>	<b>Prime Focus</b> (neutral point for coma or coma-free point)	<b>M2 Center of Curvature</b> (neutral point for pointing)
Distance from M2 vertex	300.1 mm	573.6 mm
Aberrations:		
Constant coma	No	Yes (dominating effect)
Linear astigmatism	Yes	Yes (but small effect)
Pointing change	Yes	No

#### 4.2.2 USING ABERRATIONS FOR ALIGNING THE SYSTEM

Knowledge of the aberrations and their sensitivities to misalignment can be used to align an optical system. In a two mirror telescope, one can rotate the secondary mirror about its center of curvature until the constant coma is eliminated, then rotate the secondary about the prime focus (coma-free point) until the linear astigmatism is gone. This rotation about prime focus will also simultaneously eliminate pointing errors, which alternatively may be easier to measure and correct. Pointing errors are discussed in Section 4.5.

It is interesting to note that if the primary mirror surface is incorrectly figured with coma, the telescope images during operation will appear with the same constant coma as that caused by a secondary mirror misalignment. This means that some amount of coma figure error on the primary can be accommodated by purposely misaligning the telescope without any degradation in performance. Similarly, if the primary mirror is incorrectly figured with astigmatism, this astigmatism will show up equally for all field angles. However, this constant astigmatism can not be corrected by purposely misaligning the secondary mirror because the astigmatism due to misalignment has linear field dependence. Thus, the astigmatism error in the mirror shape must be corrected by other means, such as additional polishing or by bending the surface with active supports during operation. It is also possible to correct the astigmatism in one mirror by introducing astigmatism on a different mirror. This is discussed more in Section 6.7.

The total aberrations measured in a system are a sum of the aberrations that exist in the rotationally symmetric system, the aberrations that are induced by

misalignments and scale linearly with them, the aberrations that are induced by misalignments that scale with the misalignment squared, or higher, and the aberrations due to bending. This can be expressed conceptually in the following equation:

$$\begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix} = \begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix}_{\text{nominal}} + \begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix}_{\text{misalignment}} + \begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix}_{\text{misalignment}^2} + \begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix}_{\text{bending}} . \quad (4.1)$$

The telescope alignment adjustments can be calculated if the sensitivities to misalignments of the secondary mirror are known from the optical model. The sensitivities of the aberrations that arise with a tilted or decentered secondary mirror are described using Zernike polynomial coefficients in this section. This description is based on the work of Tessieres (2003) and Thompson (1980, 2005). A similar alignment problem was described for the prime focus corrector of the Bok 90-inch telescope on Kitt Peak by Tessieres (2003). In this chapter, the information is stored in a matrix that relates the Zernike coefficients to the perturbations. In a system with only two elements, tilt and decenter perturbations only cause primary coma and astigmatism, which are represented by the coefficients  $C_5$ – $C_8$  for the standard Zernike polynomials in  $Z_5$ – $Z_8$ . The standard Zernike definitions described by Noll (1976), where subscripts 5 and 6 correspond to astigmatism and 7 and 8 correspond to coma, are used. (It is assumed that a separate procedure will be used to focus the telescope and set the secondary mirror spacing to eliminate focus ( $Z_4$ ) and spherical aberration ( $Z_{11}$ ).) Therefore, a  $4 \times 4$  interaction matrix  $\mathbf{M}$  is sufficient to describe

the alignment aberrations:

$$\begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix}_{\text{misalignment}} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \cdot \begin{bmatrix} PFRotX \\ PFRotY \\ CCRotX \\ CCRotY \end{bmatrix}. \quad (4.2)$$

In a system that has more elements that may be perturbed, there should be an additional column in  $\mathbf{M}$  for each degree of freedom. These extra degrees of freedom cause additional aberrations to become significant and, therefore, more rows would be required. Equation 4.2 can also be written in shortened form as  $Z = \mathbf{M}A$  where  $Z$  is the vector of Zernike coefficients  $C_5$ – $C_8$  and  $A$  is a vector storing the alignment perturbations. When the interaction matrix  $\mathbf{M}$  and the specific Zernike aberration coefficients  $C_5$ – $C_8$  are known, the current misalignments can be determined using the equation  $A = \mathbf{M}^{-1}Z$ . The system can then be aligned once the misalignments are calculated. Chapter 7 includes more discussion on solving linear systems in the presence of noise.

The aberrations that can occur due to misalignments may only have certain field dependencies and these must be known in order to fill the  $\mathbf{M}$  matrix. The field dependencies have been described by Thompson for third and fifth order Seidel aberrations (Thompson, 1980, 2005) and by Tessieres for Zernike polynomials (Tessieres, 2003). In this chapter, the coefficients are fit to the equations describing the field dependencies of the Zernike polynomials as found by Tessieres.

### 4.2.3 FIELD-DEPENDENT ZERNIKE ABERRATIONS IN A MISALIGNED SYSTEM

Typically, a single wavefront  $W(\rho, \phi)$  can be described by a sum of scaled Zernike polynomials, as in Equation 3.7. However one might also describe the wavefront for any field angle in a system by making the weighting coefficients field-dependent functions  $C_i(H_x, H_y)$ , as in Equation 3.8. A complete set of equations have been developed to describe the Zernike polynomials in a misaligned system (Tessieres, 2003), as described in Section 3.2.2. In Section 4.2, it was shown that constant coma and linear astigmatism are the dominant aberrations caused by misalignments in a two-mirror optical system. The following equations (introduced in Chapter 3) express the field-dependent Zernike coefficients of astigmatism ( $Z_5$  and  $Z_6$ ) and coma ( $Z_7$  and  $Z_8$ ) in a misaligned system:

$$C_5(H_x, H_y) = 2 \alpha_0 H_x H_y + \alpha_1 H_x + \alpha_2 H_y + \alpha_3 \quad (4.3)$$

$$C_6(H_x, H_y) = \alpha_0 (H_x^2 - H_y^2) - \alpha_1 H_y + \alpha_2 H_x + \alpha_4 \quad (4.4)$$

$$C_7(H_x, H_y) = \beta_0 H_y + \beta_1 \quad (4.5)$$

$$C_8(H_x, H_y) = \beta_0 H_x + \beta_2. \quad (4.6)$$

The Greek letter coefficients with subscript zero correspond to aberrations that occur in a rotationally symmetric system (Tessieres, 2003). Equations 4.3 and 4.4 show that the total astigmatism can be a combination of astigmatism components that are quadratic, linear and constant in field, while Equations 4.5 and 4.6 show that coma can be linear or constant in field. The combination of quadratic and linear

astigmatism results in two locations in the image field where the total astigmatism is zero. This is referred to as “binodal astigmatism” and is discussed more in Chapter 5. By itself, the linear coma that exists in a rotationally-symmetric system has one location in the field where the total coma is zero, which is the on-axis field point. When a system has both linear and constant coma, the location of the coma node shifts away from the on-axis field. If the amount of constant coma is large, as in the examples in this chapter, the coma node will be far outside of the field of view.

The optical model used in this section is the NST axisymmetric parent telescope, described in Section 4.1. The specific aberrations in the actual off-axis NST will be different, but it is important to first show the technique that will be used. Each column of the matrix  $\mathbf{M}$  shown in Equation 4.2 corresponds to a perturbation of one alignment degree of freedom for the secondary mirror. The columns are calculated one at a time (left to right), starting with a coma-free (prime focus) rotation about  $\vec{x}$ , then about  $\vec{y}$ , then pointing-free (CoC) rotation about  $\vec{x}$  and then about  $\vec{y}$ . The process of finding the aberrations in ZEMAX consists of the following steps for each misalignment perturbation.

1. Record the Zernike coefficients at a grid of field angles for some amount of perturbation.
2. Perform a least-squares fit on the recorded Zernike coefficients to find the coefficients  $\alpha_0$ - $\alpha_4$  and  $\beta_0$ - $\beta_2$  that describe each degree of freedom.



## **Record the Zernike coefficients throughout the field for perturbations of M2**

The first step in the process of finding the equations that relate the Zernike coefficients to the perturbations is to perturb the secondary mirror with a significant amount of rotation. For each perturbation, the Zernike coefficients are recorded for a grid of 49 different field positions using a ZEMAX macro, provided in Appendix H. The Zernike coefficients are then stored in five separate data files for: 1. the nominal system, 2. the secondary mirror PF  $\vec{x}$  rotation, 3. the secondary mirror PF  $\vec{y}$  rotation, 4. the secondary mirror  $\vec{x}$  CoC rotation and 5. the secondary mirror  $\vec{y}$  CoC rotation. The residual design aberrations of the nominal system are recorded, so the effect of the misalignment can be found by taking a difference between the coefficients before and after the misalignment is applied.

### **Rotation of the secondary mirror about the coma-free point (prime focus)**

Rotating the secondary mirror about the coma-free point (the prime focus) results in linear astigmatism, as shown by the spot diagrams in Figure 4.5. As mentioned earlier, linear astigmatism always has this pattern of line images throughout the field when out of focus. (In the medial plane, astigmatism still results in circular images.) However the pattern for the entire field rotates by  $90^\circ$ , depending on the direction of the rotation of M2. (The orientation angle of the line image for any field angle may be calculated and is  $1/2$  the arctangent of the ratio of the Zernike coefficients  $C_5$  and  $C_6$  (Tessieres, 2003).) Errors in this degree of freedom are tolerable because the astigmatism grows linearly from the on-axis field and therefore small field angles

have small aberrations. In Figure 4.5, a rather large rotation angle of  $5.5^\circ$  is used, so that the induced linear astigmatism is clearly shown over the residual field curvature in the design. This large misalignment causes a large amount of pointing error, so the spots in Figure 4.5 are shown as referenced to the centroid.

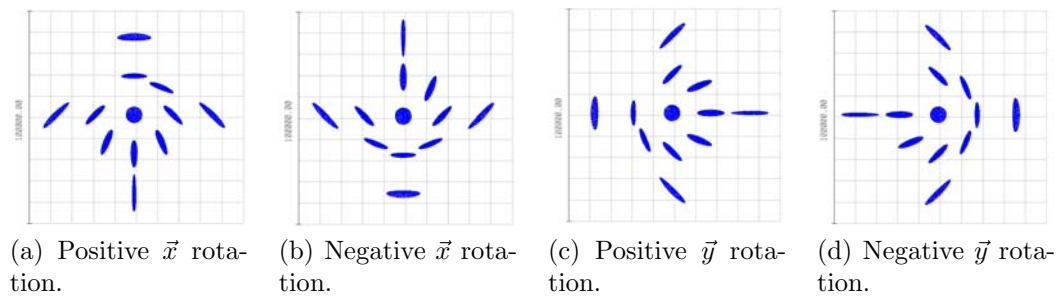


Figure 4.5: Spot diagrams of NST axisymmetric parent with M2 rotated about prime focus by  $5.5^\circ$  showing linear astigmatism in an image plane behind the medial focus. The full field of view is 3 arcminutes. Spot sizes are exaggerated by  $4\times$  for clarity.

### Rotation of the secondary mirror about its center of curvature

Rotating the secondary mirror about its CoC causes mostly constant coma, as seen in Figure 4.6. The direction of the coma pattern depends on the axis and direction of rotation. This secondary mirror perturbation also causes a small amount of astigmatism compared to the coma, but it is not visible in the figure. Notice that the tip of the coma pattern for the on-axis field is centered in the field, so there is no pointing error. The rotation angle for Figure 4.6 is much smaller than the rotation about prime focus in the Figure 4.5, but the spot size exaggeration is the same, so the amount of coma is very large. Since this misalignment causes so much coma, it is an important degree of freedom to correct, especially since it affects all field angles equally.

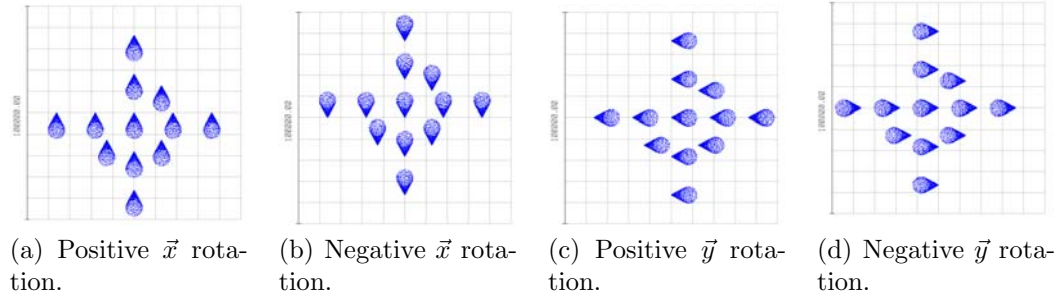


Figure 4.6: Spot diagrams of NST axisymmetric parent showing constant coma caused by rotating the secondary mirror about its center of curvature by  $0.1^\circ$ . The full field of view is 3 arcminutes. Spot sizes are exaggerated by  $4\times$  for clarity.

### Perform a least-squares fit to find the values of $\alpha_n$ and $\beta_n$

To find the equations that relate the Zernike coefficients to the perturbations and the field positions, a least-squares fit is performed on the recorded data using a MATLAB script, provided in Appendix I. For each perturbation, the script finds the values of all the  $\alpha_n$  and  $\beta_n$  terms to describe the field dependent Zernike coefficients. The results are listed in Table 4.4. One of the linear astigmatism coefficients ( $\alpha_1$  or  $\alpha_2$ ) is significant for all the degrees of freedom, as expected. The center of curvature rotations have a significant amount of constant coma ( $\beta_1$  or  $\beta_2$ ) and some constant astigmatism ( $\alpha_4$ ). The prime focus rotations, which were intended to be coma-free degrees of freedom, result in a very small amount of constant coma after all. All the coefficients from the least-squares fit with magnitudes less than  $1 \times 10^{-10}$  microns were set to zero in Table 4.4. A large number of significant digits are shown so that the nonlinearity can be seen in the linear coma coefficient ( $\beta_0$ ). The nonlinearity of the field-dependent aberration functions is discussed in Chapter 6. The constant coma sensitivity to center of curvature rotations ( $3.951 \mu\text{m}/0.1^\circ$ ) is much larger than

the linear astigmatism sensitivity to prime focus rotations ( $0.140 \mu\text{m}/0.1^\circ$ ), which shows that the system is much more sensitive to aberrations caused by center of curvature rotations.

Table 4.4:  $\alpha$  and  $\beta$  coefficients in units of microns for rotations of  $0.1^\circ$ . Quadratic astigmatism is described by  $\alpha_0$ . Linear astigmatism is described by  $\alpha_1$  and  $\alpha_2$ . Constant astigmatism is described by  $\alpha_3$  and  $\alpha_4$ . Linear coma is described by  $\beta_0$ . Constant coma is described by  $\beta_1$  and  $\beta_2$ .

	<i>Nominal</i>	<i>PFRotX</i>	<i>PFRotY</i>	<i>CoCRotX</i>	<i>CoCRotY</i>
$\alpha_0$	0.022033	0.022033	0.022033	0.022033	0.022033
$\alpha_1$	0	0.14033	0	0.1293	0
$\alpha_2$	0	0	-0.14033	0	-0.1293
$\alpha_3$	0	0	0	0	0
$\alpha_4$	0	$-2.50e-6$	$2.50e-6$	0.37234	-0.37234
$\beta_0$	-0.035714	-0.035715	-0.035715	-0.035691	-0.035691
$\beta_1$	0	0	$-4.2584e-4$	0	3.9531
$\beta_2$	0	$4.2584e-4$	0	-3.9531	0

The quadratic astigmatism  $\alpha_0$  and the linear coma term  $\beta_0$  were found to be the same (or near numerically the same) for all misalignment degrees of freedom and for different magnitudes of misalignment. This makes sense because these are the terms that occur in the rotationally symmetric system and are not affected by alignment.

The net field-dependent astigmatism has both quadratic and linear field-dependent terms, which have a combined effect of creating two nodes, which are locations where the net astigmatism is zero. One node is on axis because both functions are zero on axis. The other node occurs at the single point where the quadratic astigmatism compensates the linear astigmatism.

The misalignment of  $0.1^\circ$  used in Table 4.4 was found to be in the range of rotation angles that have a linear effect on the resulting aberrations for linear astig-

matism and constant coma. For example, rotations of  $0.01^\circ$  result in coefficients ten times smaller, while rotations of  $1^\circ$  result in coefficients ten times larger. Constant-field astigmatism terms depend on the square of the misalignment. For example, rotations of  $0.01^\circ$  result in coefficients one hundred times smaller, while rotations of  $1^\circ$  result in coefficients one hundred times larger. Since the  $\alpha_3$  and  $\alpha_4$  coefficients for constant astigmatism do not depend linearly on the misalignment, they should not be used in the LS fit.

#### 4.2.4 SUMMARY OF LEAST-SQUARES FIT COEFFICIENTS

Next, the Zernike coefficients can be expressed in terms of the perturbations and the field positions in the matrix form. The significant terms that are induced by misalignment and included in the sensitivity matrix in the following equation are  $\alpha_1$ ,  $\alpha_2$ ,  $\beta_1$  and  $\beta_2$ . These terms depend linearly on the misalignment degree of freedom. In order to specify the rotation angle in degrees, these values will be multiplied by ten (because the original misalignment was  $0.1^\circ$ ). The equation describing the aberrations resulting in the system due to different misalignment degrees of freedom is:

$$\begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix} = \begin{bmatrix} 1.403 H_x & -1.403 H_y & 1.293 H_x & -1.293 H_y \\ -1.403 H_y & -1.403 H_x & -1.293 H_y & -1.293 H_x \\ 0 & 0 & 0 & 39.53 \\ 0 & 0 & -39.53 & 0 \end{bmatrix} \cdot \begin{bmatrix} PFRotX \\ PFRotY \\ CoCRotX \\ CoCRotY \end{bmatrix} \quad (4.7)$$

where  $H_x$  and  $H_y$  are the normalized field angles, the Zernike coefficients are in microns, and the misalignment rotations are in degrees. If one desires the field angles to be specified in degrees, then the coefficients of  $H_x$  and  $H_y$  must be divided

by the half field of view  $F_r = 0.025^\circ$ . When the field coordinates are in degrees, then the equation becomes:

$$\begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix} = \begin{bmatrix} 56.13 f_x & -56.13 f_y & 51.72 f_x & -51.72 f_y \\ -56.13 f_y & -56.13 f_x & -51.72 f_y & -51.72 f_x \\ 0 & 0 & 0 & 39.53 \\ 0 & 0 & -39.53 & 0 \end{bmatrix} \cdot \begin{bmatrix} PFRotX \\ PFRotY \\ CoCRotX \\ CoCRotY \end{bmatrix} \quad (4.8)$$

where now the field coordinates  $f_x = H_x F_r$  and  $f_y = H_y F_r$  (using the notation of ZEMAX) are in degrees, the Zernike coefficients are in microns, and the misalignment rotations are in degrees.

#### 4.2.5 ALTERNATE DEGREES OF FREEDOM

In Section 4.2.1, two different rotation points for the secondary mirror were defined that do not cause constant coma or pointing to be induced in the optical system when misaligned. Misalignments of both of these degrees of freedom induce linear astigmatism in the optical system. This may cause one to wonder if there is a rotation point about which the secondary mirror can be misaligned that does not cause linear astigmatism. This was tested through a simulation in the optical design program. The constant coma and linear astigmatism coefficients were found for a constant rotation angle of  $0.1^\circ$  for rotation points at varying distances from the secondary mirror. A linear astigmatism-free rotation point was found 3548 mm away from the secondary mirror toward the primary mirror. A plot of the constant coma and linear astigmatism coefficients is provided in Figure 4.7.

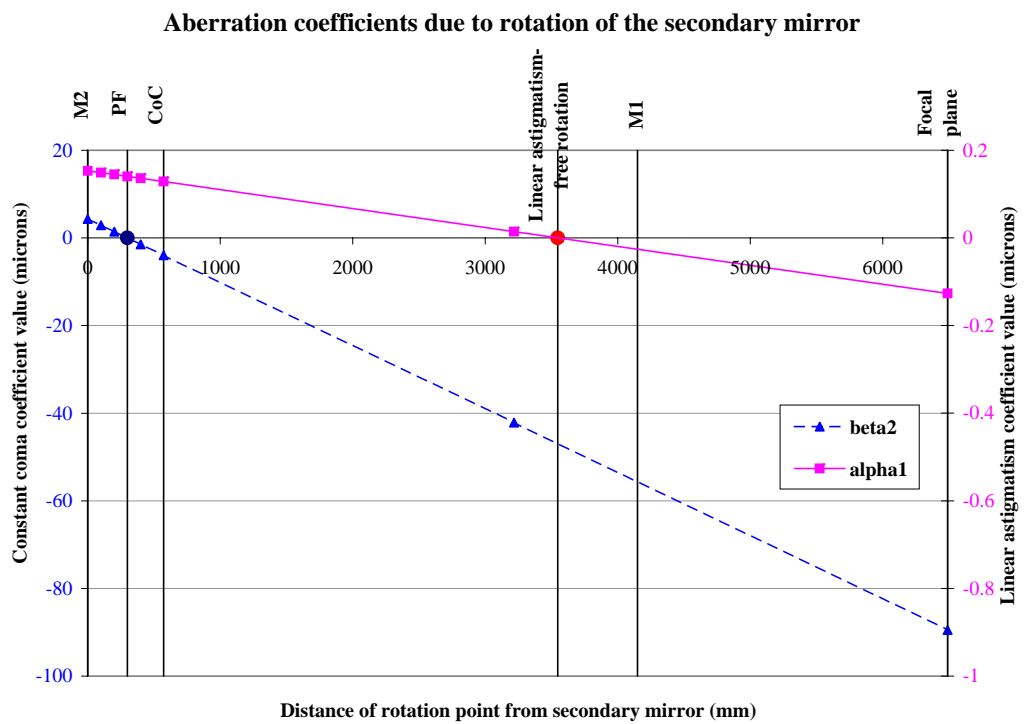


Figure 4.7: Constant coma and linear astigmatism coefficients in NST for different rotation points. The scale for the coma coefficient  $y$ -axis is  $10\times$  larger than the astigmatism  $y$ -axis. (All rotations were  $0.1^\circ$ .)

### 4.3 ABERRATIONS OF NST (THE ACTUAL OFF-AXIS TELESCOPE)

The analysis done in Section 4.2 was performed for the full NST parent telescope. The actual NST will just be an off-axis portion of the entire telescope. The NST pupil has a 1600 mm diameter which is decentered 1840 mm from the optical axis of the parent telescope, as shown in Figure 4.1d. The aberrations in the parent telescope must be multiplied by the scale factors found in the following section to find the aberrations caused by misalignments in the actual NST pupil.

#### 4.3.1 ABERRATION SCALE FACTORS FOR THE OFF-AXIS NST PUPIL

Using convenient MATLAB code written by Lundström & Unsbo (2007), the Zernike coefficients of one circular (or elliptical) pupil can be converted into coefficients of another scaled, translated, or rotated pupil. The coefficients of the Zernike polynomials in the subpupil only contain orders equal to or lower than the Zernike coefficients originally placed on the pupil. Table 4.5 shows how one unit of an aberration in the NST parent pupil transforms into lower order aberrations for the actual NST pupil. Since the spherical aberration residual in the nominal system design also transforms into lower aberrations, such as the coma and astigmatism, its contribution must be considered as well. Spherical aberration ( $Z_{11}$ ) in the parent pupil does not transform into trefoil in the NST pupil, so rows for Zernike polynomials  $Z_9$  and  $Z_{10}$  were omitted from the table.

As seen earlier, constant coma is caused by rotating the secondary mirror about any point other than the prime focus. The degree of freedom used in this chapter is the M2 center of curvature. The wavefronts in Figure 4.8 shows conceptually how



Table 4.5: Conversion of Zernike polynomials to lower order for NST pupil.

Aberration in NST	Aberration in parent telescope					
		Astigmatism		Coma		Spherical
		$Z_5 = 1$	$Z_6 = 1$	$Z_7 = 1$	$Z_8 = 1$	$Z_{11} = 1$
Piston	$Z_1$	0	-1.1899	0.5268	0	-0.4967
Tilt	$Z_2$	-0.5173	0	0	-0.1539	0
Tilt	$Z_3$	0	0.5173	1.0952	0	-0.2663
Focus	$Z_4$	0	0	-0.3135	0	0.3681
Astigmatism	$Z_5$	0.0918	0	0	-0.2217	0
Astigmatism	$Z_6$		0.0918	0.2217	0	-0.4886
Coma	$Z_7$			0.0278	0	-0.1227
Coma	$Z_8$				0.0278	0
Spherical	$Z_{11}$					0.0084

coma over the parent telescope pupil transforms into astigmatism and coma (once piston, tip, tilt and focus are removed) for the NST pupil (a subaperture of the larger pupil) according to the values in Table 4.5.

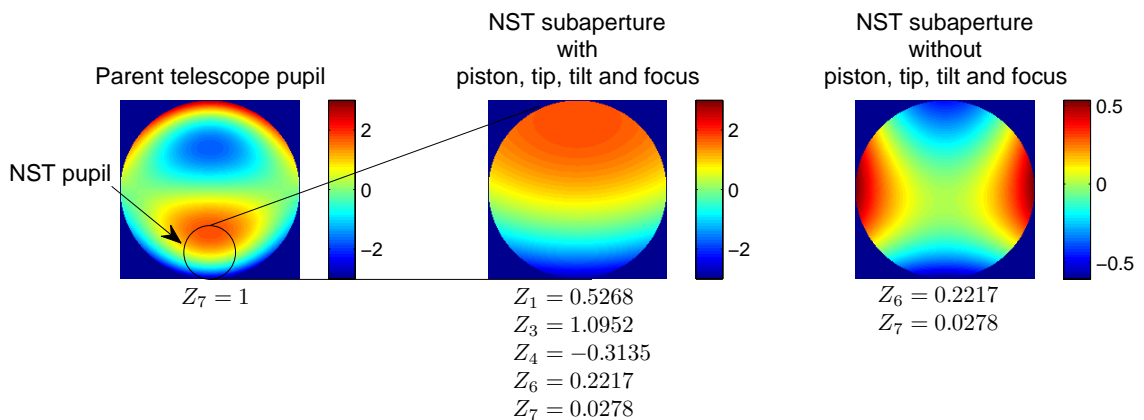


Figure 4.8: Transformation of one unit of coma from NST parent telescope pupil to actual pupil.

Since aberrations on one larger pupil only transform to aberrations of an equal or

lower order on a subaperture, astigmatism ( $Z_5$  or  $Z_6$ ) transforms into Zernike terms 1–6. However, if piston and tilt are once again removed, then only astigmatism remains (similar to the case of coma). As another example, shown in Figure 4.9, one unit of  $Z_5$  astigmatism on the full aperture transforms to  $-0.5173$  units of  $Z_2$  tilt and  $0.0918$  units of  $Z_5$  astigmatism on the subaperture.

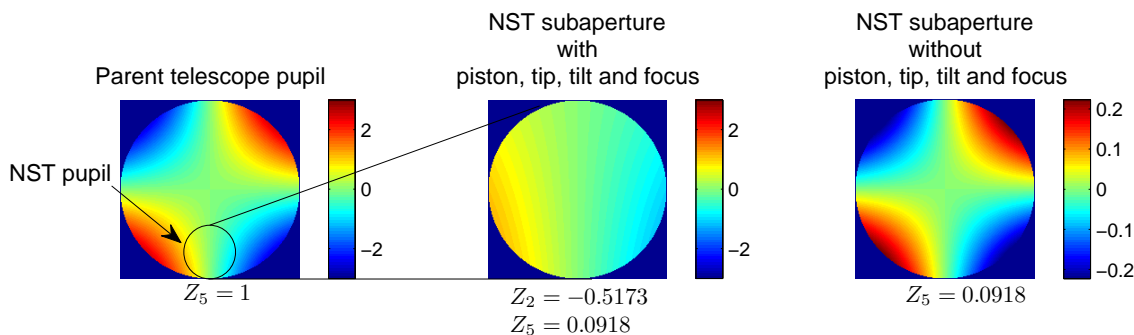


Figure 4.9: Transformation of one unit of astigmatism from NST parent telescope pupil to actual pupil.

It is possible to completely eliminate spherical aberration in a two mirror telescope by appropriate choice of conic constants and axial positions of each of the mirrors. If there are any errors in the conic constants of the mirrors in the telescope, there will be spherical aberration (constant in field) in the final system. Spherical aberration adds a contribution to the astigmatism and coma terms in a misaligned system and may be considered. As a final example, shown in Figure 4.10, one unit of  $Z_{11}$  on the full aperture looks like  $-0.1227$  units of  $Z_7$  coma and  $-0.4886$  units of  $Z_6$  astigmatism on the subaperture. These scale factors will be multiplied by the value for the spherical aberration (constant in field) for the NST parent telescope system. In the particular computer model analyzed, the small amount of spherical aberration was  $0.028 \mu\text{m}$ . (This spherical aberration would be zero if the conic

constant of the secondary mirror was decreased by 0.00019 to  $k = -0.83106$ .)

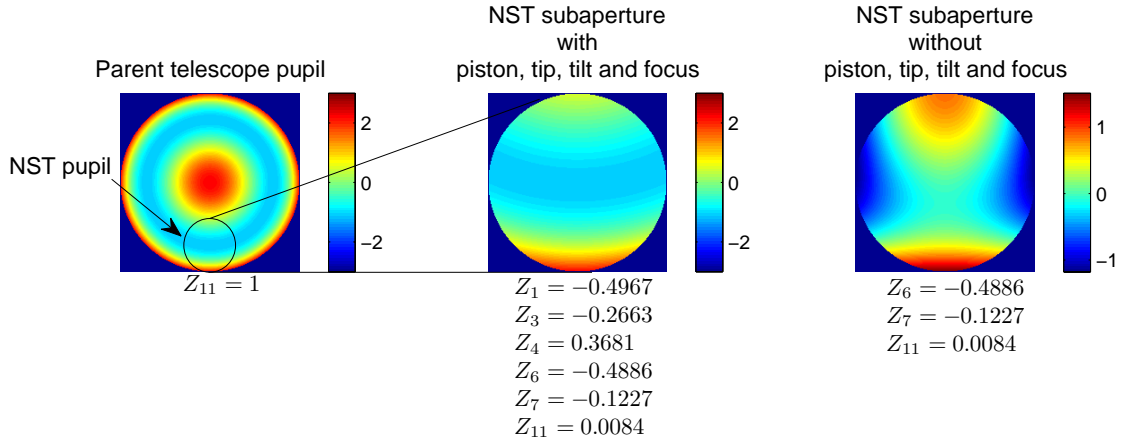


Figure 4.10: Transformation of one unit of spherical aberration from NST parent telescope pupil to actual pupil.

### 4.3.2 COMBINING THE RESULTS

This section will combine the least-squares coefficients for the on-axis system and the scaling factors for the off-axis pupil to determine the aberrations of the misaligned off-axis NST system. Table 4.6 shows the calculations.

The total astigmatism in the misaligned off-axis NST system includes contributions from a variety of source that are quadratic, linear and constant in field. The quadratic astigmatism comes from the astigmatism in the rotationally symmetric parent telescope. There are two sources of the linear astigmatism: the linear astigmatism from misalignment of the secondary and the transformed linear coma from the rotationally symmetric parent telescope pupil. There are three sources of constant astigmatism: the constant astigmatism from the misalignment of the secondary (which depends on the square of the misalignment), the transformed constant

Table 4.6: Calculations for scaling aberrations in the off-axis NST. The aberrations that exist in the nominal (rotationally symmetric) system do not change with misalignment. The aberrations that scale with the misalignment are marked “/degree” in the calculation column, while the aberrations that depend on the square of the misalignment are given for case when the misalignment is  $0.1^\circ$ .

<b>Aberration</b>				
Field dependence	Source	Parent Aberration	Calculation (units are microns)	
<b>Astigmatism</b>				
Quadratic	Nominal	Astig.	$0.0918 * 0.022033$	$= 0.0020$
Linear	<i>PFRot</i>	Astig.	$0.0918 * 1.403$	$= 0.1288$ /degree
Linear	<i>CoCRot</i>	Astig.	$0.0918 * 1.293$	$= 0.1187$ /degree
Constant	<i>PFRot</i>	Astig.	$0.0918 * -2.5 \times 10^{-6}$	$= 0.0000$ for $0.1^\circ$
Constant	<i>CoCRot</i>	Astig.	$0.0918 * 0.37234$	$= 0.0342$ for $0.1^\circ$
Linear	Nominal	Coma	$0.2217 * -0.03571$	$= -0.0079$
Constant	<i>CoCRot</i>	Coma	$0.2217 * 39.53$	$= 8.7638$ /degree
Constant	Nominal	Spherical	$-0.4886 * 0.028$	$= -0.0137$
<b>Coma</b>				
Linear	Nominal	Coma	$0.0278 * -0.3571$	$= -0.0099$
Constant	<i>CoCRot</i>	Coma	$0.0278 * 39.53$	$= 1.0989$ /degree
Constant	Nominal	Spherical	$-0.1227 * 0.028$	$= -0.0034$

coma from the parent telescope pupil in the misaligned system and the transformed constant spherical aberration of the rotationally symmetric parent telescope system.

The total coma is the sum of the linear coma in the parent telescope design and the constant coma. This constant coma comes from two sources: the constant coma in the parent telescope induced by the misalignment and that transformed from the spherical aberration in the axially symmetric parent telescope.

The following equation shows the sensitivity matrix for the off-axis NST (com-

pare to Equation 4.7):

$$\begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix} = \begin{bmatrix} 0.129 H_x & -0.129 H_y & 0.119 H_x + 8.8 & -0.119 H_y \\ -0.129 H_y & -0.129 H_x & -0.119 H_y & -0.119 H_x + 8.8 \\ 0 & 0 & 0 & 1.10 \\ 0 & 0 & -1.10 & 0 \end{bmatrix} \cdot \begin{bmatrix} PFRotX \\ PFRotY \\ CoCRotX \\ CoCRotY \end{bmatrix}. \quad (4.9)$$

If the field angles are specified in degrees, then the coefficients of  $H_x$  and  $H_y$  must be divided by the half field of view  $F_r = 0.025^\circ$ , as was done for the parent telescope equations. In this case, the result is

$$\begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix} = \begin{bmatrix} 5.15 H_x & -5.15 H_y & 4.75 H_x + 8.8 & -4.75 H_y \\ -5.15 H_y & -5.15 H_x & -4.75 H_y & -4.75 H_x + 8.8 \\ 0 & 0 & 0 & 1.10 \\ 0 & 0 & -1.10 & 0 \end{bmatrix} \cdot \begin{bmatrix} PFRotX \\ PFRotY \\ CoCRotX \\ CoCRotY \end{bmatrix}, \quad (4.10)$$

where the field coordinates  $f_x = H_x F_r$  and  $f_y = H_y F_r$  are in degrees, the Zernike coefficients are in microns, and the misalignment rotations are in degrees.

Previously the field dependencies from the nominal design (quadratic astigmatism and linear coma) were not included because they had different field dependencies than the ones induced by the misalignment. However the linear coma from the nominal design also contributes to linear astigmatism. Also, the constant spherical aberration from the nominal design contributes to the constant astigmatism and constant coma in the off-axis system. These aberrations are small, but could be calculated and subtracted from the measured aberrations for increased accuracy. The aberrations are:

$$\begin{bmatrix} C_5 \\ C_6 \\ C_7 \\ C_8 \end{bmatrix}_{\text{nominal}} = \begin{bmatrix} 0.0079 H_x \\ -0.0079 H_y - 0.01368 \\ -0.00344 \\ 0 \end{bmatrix}. \quad (4.11)$$

#### 4.4 SPOT DIAGRAMS FOR THE MISALIGNED OFF-AXIS NST

Table 4.7 shows conceptually how the dominant aberrations that occur for the misaligned parent telescope transform into different aberrations for the off-axis NST, following the results in Table 4.5.

Table 4.7: Comparison of aberrations for on-axis and off-axis telescopes.

On-axis parent telescope	Off-axis telescope
Constant coma (39.5 $\mu\text{m RMS}/^\circ$ )	Constant coma (1.1 $\mu\text{m RMS}/^\circ$ )
	Constant astigmatism (8.8 $\mu\text{m RMS}/^\circ$ )
	Constant tilt (pointing)
	Constant focus (defocus)
Linear astigmatism (56.1 or 51.2 $\mu\text{m RMS}/^\circ/^\circ$ ) <sup>a</sup>	Linear astigmatism (5.12 or 4.76 $\mu\text{m RMS}/^\circ/^\circ$ )
	Linear tilt (plate scale change or rotation)
Pointing	Pointing

<sup>a</sup>The two different linear astigmatism sensitivities are for the two different types of degrees of freedom (prime focus rotations and center of curvature rotations). See Equation 4.10.

Next, spot diagrams are observed for the off-axis NST. The degrees of freedom used are exactly the same as those used for the parent telescope. Figure 4.11 shows the spot diagrams for rotations about the prime focus for the off-axis NST. Here, the secondary mirror is rotated by the same angle and the spot size exaggeration is the same, so these figures may be directly compared to those in Figure 4.5. The rays in this figure are exactly a subset of the rays in the Figure 4.5. Since normally prime focus rotations result in mostly linear astigmatism, one should expect that

in the off-axis system, the aberrations will include a linear astigmatism and a plate scale change, according to Table 4.7. This is indeed what is shown in Figure 4.11.

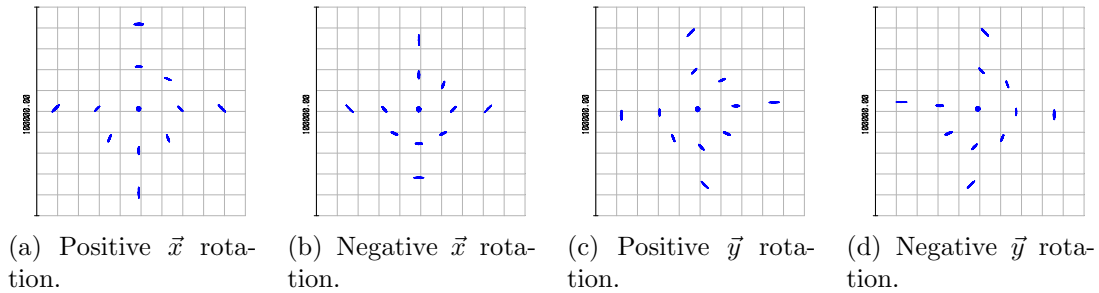


Figure 4.11: Spot diagrams when secondary mirror is rotated about the focus of the primary mirror  $5.5^\circ$ . The spot diagrams shows a plate scale change for  $\vec{x}$  rotations and image rotation for  $\vec{y}$  rotations. The full field of view is 3 arcminutes. The spot size exaggeration is  $\times 4$ . The rotations about  $\vec{x}$  show plate scale change, while the rotations about  $\vec{y}$  show image rotation.

Figures 4.12 and 4.13 show the spot diagrams for rotations about the center of curvature for the off-axis NST. Once again, the secondary mirror is rotated by the same angle and the spot size exaggeration is the same, so these two figures may be directly compared to those in Figure 4.6. Since in the parent telescope, center of curvature rotations result in mostly constant coma, one should expect that in the off-axis system, the aberrations will include a little bit of constant coma, a lot of constant astigmatism and some defocus error, according to Table 4.7. Indeed, Figure 4.12a and Figure 4.12b show elongated images throughout the field that look like constant astigmatism with a defocus error while Figure 4.12c and Figure 4.12d show some constant coma. If the two spot diagrams in Figure 4.12a and Figure 4.12b are refocused, then they also show some constant coma (Figure 4.13). The constant astigmatism is no longer visible when the system is at the medial focus for astigmatism.

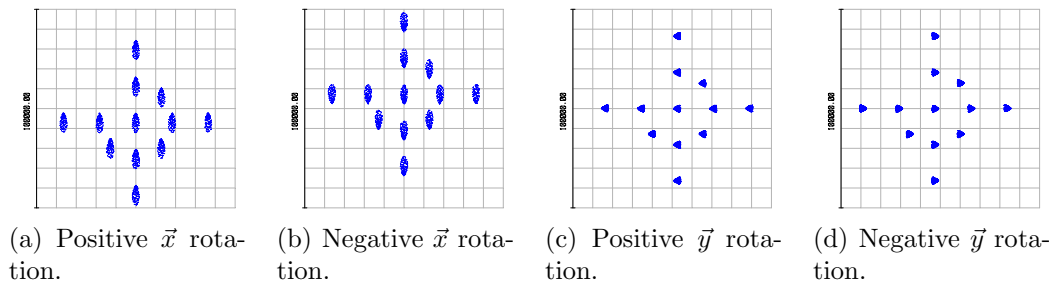


Figure 4.12: Spot diagrams when secondary mirror is rotated about its center of curvature by  $0.1^\circ$ . The full field of view is 3 arcminutes. The spot size exaggeration is  $\times 4$ .

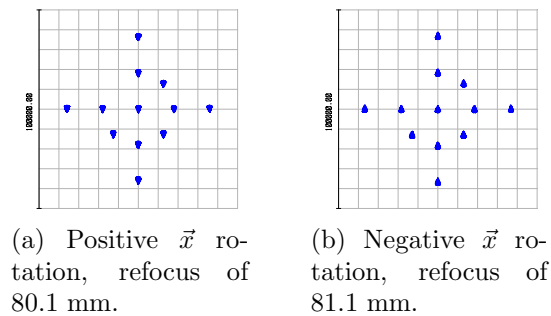


Figure 4.13: Spot diagrams when secondary mirror is rotated about its center of curvature by  $0.1^\circ$  and images are refocused. The full field of view is 3 arcminutes. The spot size exaggeration is  $\times 4$ .

For a general misaligned state of the off-axis NST, there is constant coma and binodal astigmatism. The constant coma is derived from the off-axis constant coma. The binodal astigmatism is formed by the combination of constant astigmatism, derived from the off-axis constant coma, linear astigmatism, derived from off-axis linear astigmatism and quadratic astigmatism. If the nodes are out of the field of view, the binodal nature of the astigmatism might not be obvious from a spot diagram.

As a final check, Figure 4.14a shows the wavefront map for the on-axis field



point in the parent system when the secondary mirror is rotated about the center of curvature by  $0.1^\circ$ . As expected, this map is dominated by coma. Since center of curvature rotations of the secondary mirror result in constant coma, all of the field points in the system would have a similar wavefront map. ZEMAX reports the RMS after tilt is removed is 3.9543 waves (where  $\lambda = 1 \mu\text{m}$ ), which is very close to the constant coma value in Table 4.4 ( $\beta_1 = 3.9531 \mu\text{m}$ ). Figure 4.14b shows the wavefront map for the same misalignment with an aperture set on the primary mirror to define the off-axis telescope. Figure 4.14b has the same  $0.1^\circ$  center of curvature rotation, so the resulting wavefront is expected to show astigmatism and a very small amount of coma. ZEMAX reports the RMS after tilt is removed is 0.8915 waves (where  $\lambda = 1 \mu\text{m}$ ). The approximate RMS from this misalignment can be found by calculating the root sum square of the two largest expected aberrations in the off-axis NST from Table 4.6 for the on-axis field for  $0.1^\circ$  rotations about the center of curvature. A root sum square of the astigmatism that comes from the constant coma on the parent ( $0.88 \mu\text{m}$ ) and the coma from the constant coma on the parent ( $0.11 \mu\text{m}$ ) results in  $0.89 \mu\text{m}$ . (There are other contributions to the wavefront for the off-axis NST from other effects, like the astigmatism and coma from the spherical aberration on the parent and the constant astigmatism resulting from the rotation, but these are small and do not significantly affect the RSS.)

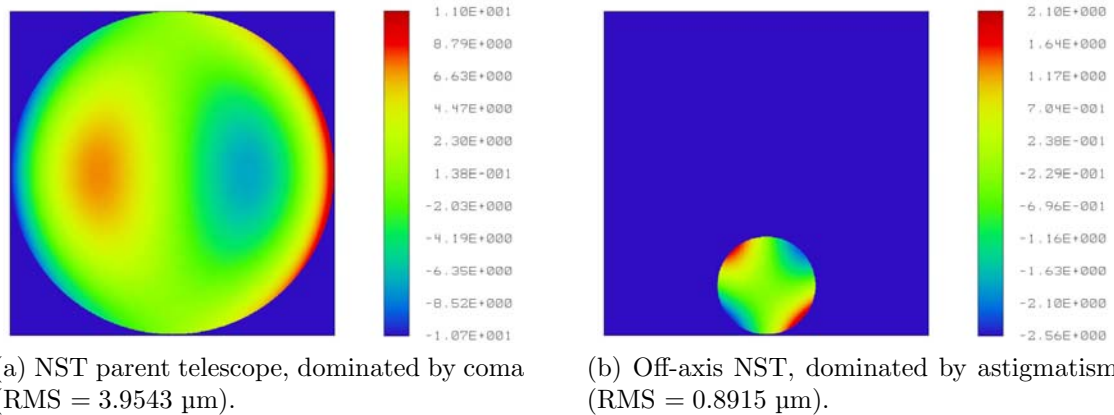


Figure 4.14: Wavefront maps for the on axis field point when the secondary mirror is rotated in about its center of curvature ( $CoCRotY = 0.1^\circ$ ). The wavefront tilt was removed from each map. The units of the color bars are microns.

#### 4.5 EFFECT OF MISALIGNMENTS ON POINTING

In some cases, it might be easier to use knowledge of the pointing than to measure the quantity of specific aberrations. To implement this, a laser tracker can be used to align the focal plane to the parent axis of the primary mirror (Burge *et al.*, 2007). It is therefore useful to know the amount of the image shift caused by rotation of the secondary mirror. The rotation of M2 about the optical axis and axial translation of M2 in the  $\vec{z}$ -direction are not considered here because they do not affect the pointing. For this section, the centroid position (for the on axis field point) will be used as the measure of the pointing error in the system. In ZEMAX, the operands CENX and CENY give the  $x$  and  $y$  coordinates of the centroid in the image plane. Table 4.8 compares the pointing errors for rotation in both the parent telescope and

the off-axis NST. The plate scale of the telescope (Equation 2.4) is

$$\text{Plate scale} = \frac{83277.8 \text{ mm}}{206265 \text{ arcsec/rad}} = 403.7 \text{ } \mu\text{m/arcsec}. \quad (4.12)$$

Thus for every millimeter of pointing error in Table 4.8, the pointing changes by about 2.5 arcseconds.

Table 4.8: Pointing error (in mm) for NST.

Degree of freedom	Angle= 1°	Parent telescope		Off-axis telescope	
		CENX	CENY	CENX	CENY
Residual (no rotation)		0	0	0	-0.02
Prime focus rotation	<i>PFRotX</i>	0	-108.0	0	-108.1
	<i>PFRotY</i>	108.0	0	108.0	-0.04
Center of curvature rotation	<i>CoCRotX</i>	0	-11.23	0	-16.02
	<i>CoCRotY</i>	11.23	0	6.53	0.31

For the parent telescope, rotation of M2 about the coma-free point (prime focus) results in a large pointing error (108 mm), as expected. The other effect of rotation about prime focus in the parent telescope is linear astigmatism (Table 4.3). Since linear astigmatism in the parent telescope corresponds to linear astigmatism, and a plate scale change or image rotation in the off-axis telescope, no additional pointing errors are expected in the off-axis telescope, and the pointing errors match for the on-axis and off-axis cases.

For rotation about the center of curvature, a small pointing error (11 mm) is observed for the parent telescope. (The pointing error for the CoC rotation in Table 4.8 is nonzero because the table is recording the centroid pointing, and not the paraxial pointing.) For the off-axis telescope, one should expect the pointing

errors to be modified from the parent telescope. Since rotation about the CoC results in primarily constant coma (Table 4.3) for the parent telescope, and this effect translates into a constant tilt (Table 4.7) for the off-axis telescope, one should expect the pointing error to be significantly changed. This is confirmed in Table 4.8.

#### 4.6 CONCLUSION

In this chapter, the aberrations due to misalignments of the secondary mirror of a rotationally symmetric Gregorian telescope (the “parent telescope”) are compared to those of an off-axis version of the same telescope. The degrees of freedom studied were the rotations about the coma-free point (the focus of the primary mirror) and rotations about the center of curvature for the secondary mirror. Rotations about the coma-free point primarily result in both linear astigmatism and pointing errors for the parent telescope while the off-axis telescope includes the same errors, plus a change in plate scale or image rotation. Rotations about the center of curvature of the secondary mirror result in primarily constant coma and a small amount of linear astigmatism, but no pointing error for the parent telescope. For the off-axis telescope, the constant coma is converted into constant astigmatism, constant tilt (pointing error), and defocus.

## CHAPTER 5

COMPARISON OF NODAL ABERRATION THEORY AND THE  
FUNCTIONAL DESCRIPTION OF ABERRATIONS IN A MISALIGNED  
RITCHEY-CHRETIEN TELESCOPE

The purpose of this chapter is to compare the Zernike polynomial description of aberrations in a misaligned optical system with Nodal Aberration Theory (NAT) (Thompson, 1980, 2005, 2009). The systematic treatment using wavefront coefficients is useful for quantifying the causes of aberrations for some applications and does not conflict with the nodal theory.

The effect of misalignment or shape (figure) errors is a displacement of the aberration nodes. The misalignment or shape error may be calculated from the location of the nodes using Nodal Aberration Theory. However, another way to determine the misalignment or shape error is to calculate mathematically the field dependence of the aberration coefficients (shown here in this chapter for the case of astigmatism). In a real system, these coefficients can be measured by a wavefront sensor. This is convenient because in practice, it is common to take measurements of the wavefront throughout the field. One major advantage of this method, which will be shown in this chapter, is that the changes in the coefficients are decoupled for different causes. Wavefront measurements can be taken anywhere in the field and the misalignments can be calculated. (This requires knowing the sensitivities of the Zernike coefficients to each of the degrees of freedom expected to have errors, which

can be found by perturbing the optical model on the computer.)

## 5.1 INTRODUCTION TO THE RITCHEY-CHRETIEN TELESCOPE ASTIGMATISM EXAMPLES

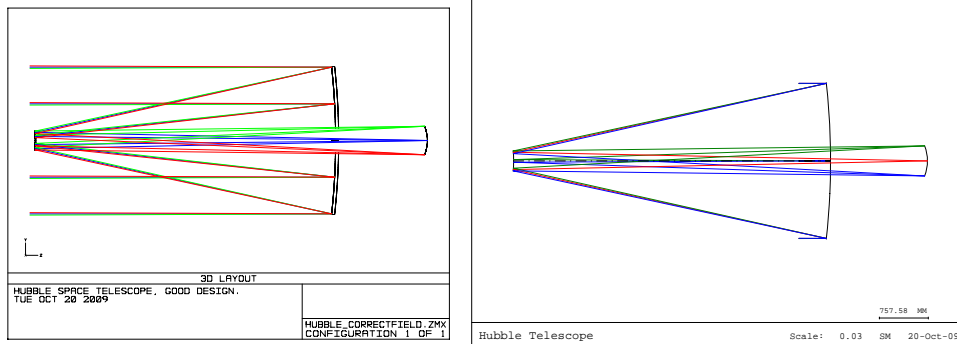
Some examples of a Ritchey-Chretien telescope with different errors causing astigmatism with different field dependencies will be shown. The optical model used is the Hubble telescope (a “good design” as described by the ZEMAX file). The optical model for the telescope is listed in Table 5.1 and shown in Figure 5.1. (Note that the image formed after the secondary is observed on a curved focal plane. At first, this seems impractical because curved detectors are hard to make, but in practice, there are multiple instruments packages with additional relay optics. Curving the image plane alters the Zernike coefficient for focus, which is not studied here.) The half field of view of the telescope is  $0.233^\circ$  or 14 arcminutes and the wavelength used is  $\lambda = 586.56$  nm.

Table 5.1: Hubble Space Telescope (Good design).

Surface Name	Radius	Thickness	Semi-Diameter	Conic
Primary (stop)	-11040	-4906.071	1200.0	-1.002299
Secondary	-1358	6406.200	155.0	-1.49686
Image	-631.079			

Figure 5.2 shows spot diagrams for the aligned telescope. The ZEMAX spot diagram is on the left and the CODE V spot diagram is on the right. The spot sizes are exaggerated in each, but by different amounts.

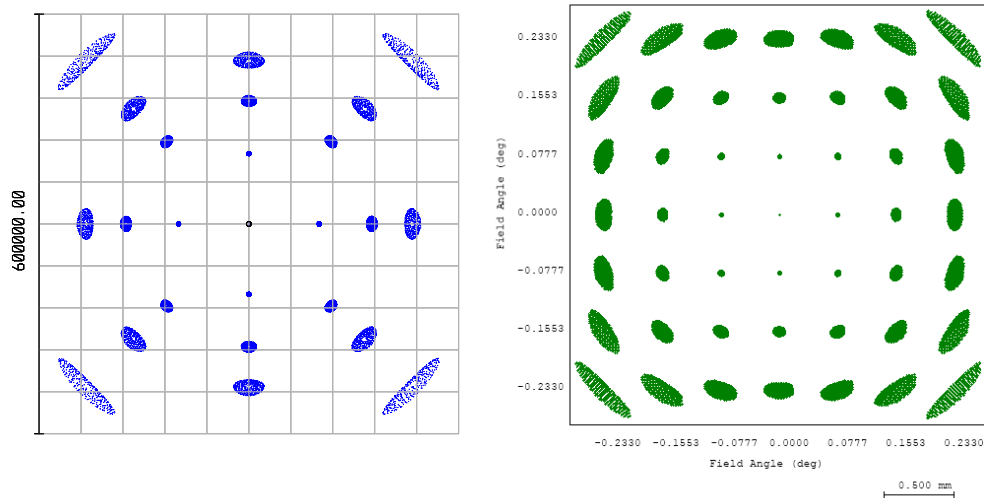
Examples will show variations of this system that are dominated by different astigmatism field dependences and some that have approximately balanced amounts



(a) ZEMAX model.

(b) CODE V model.

Figure 5.1: Hubble Space Telescope optical model.



(a) ZEMAX model (200× exaggeration).

(b) CODE V model.

Figure 5.2: Hubble Space Telescope spot diagrams. The size of the Airy disk is shown in the center of the field of the ZEMAX spot diagram.

of astigmatism. The astigmatism throughout the field will be shown as a representation of the coefficients, as recorded in ZEMAX and analyzed by MATLAB and as full field displays from CODE V. The different examples studied are listed in Table 5.2.

Table 5.2: Astigmatism examples studied in this chapter.

Description	Astigmatism (dominating field dependence)	Nodal behavior	Orientation of lines
1 Nominal system	Quadratic	One (in center)	Radially- outward
2 Large shape error (astigmatism) on primary	Constant	None observed	Parallel
3 Small shape error (astigmatism) on primary	Constant + Quadratic	Binodal (nodes symmetric about origin)	Field- asymmetric
4 Large misalignment of secondary	Quadratic + Linear + Constant	Binodal (nodes not symmetric about origin)	Field- asymmetric
5 Smaller misalignment of secondary	Quadratic + Linear	Binodal (nodes not symmetric about origin)	Field- asymmetric

The nodal behavior of each system will be discussed. Astigmatism in general has a binodal form. For any location in the field, the value of the astigmatism coefficient is proportional to the product of the two distances from that field point to each of the two nodes. The locations of the nodes depend on the details of the system design and alignment.



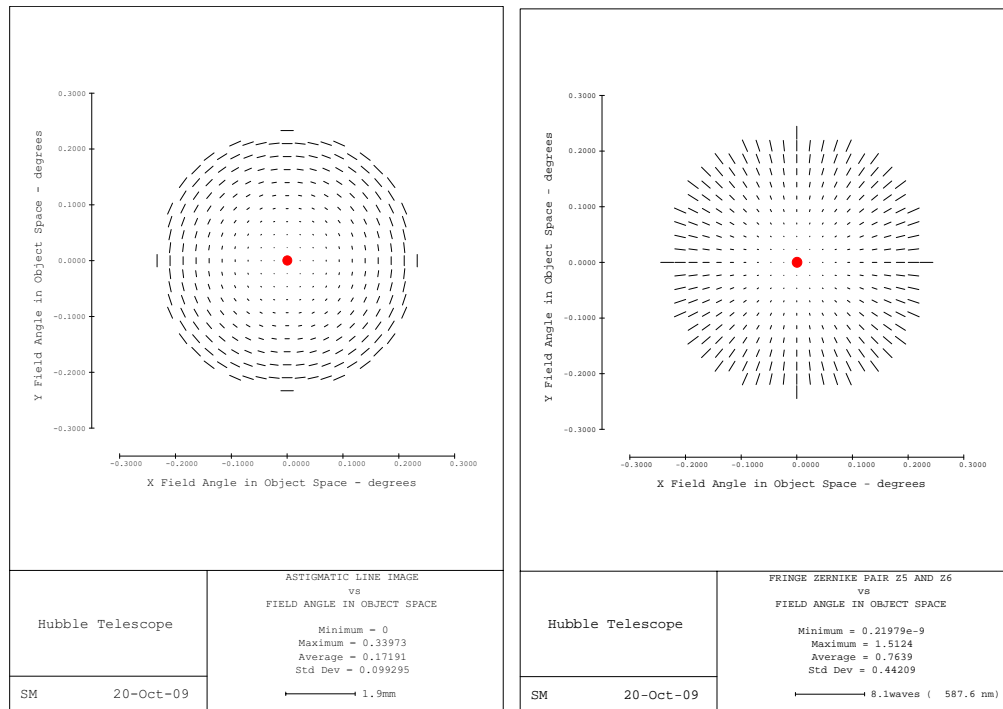
### 5.1.1 QUADRATIC ASTIGMATISM IN THE NOMINAL OPTICAL DESIGN

The first example shows the quadratic astigmatism that is native to the axially-symmetric optical system.

#### **Example 1: Native astigmatism in the nominal optical design**

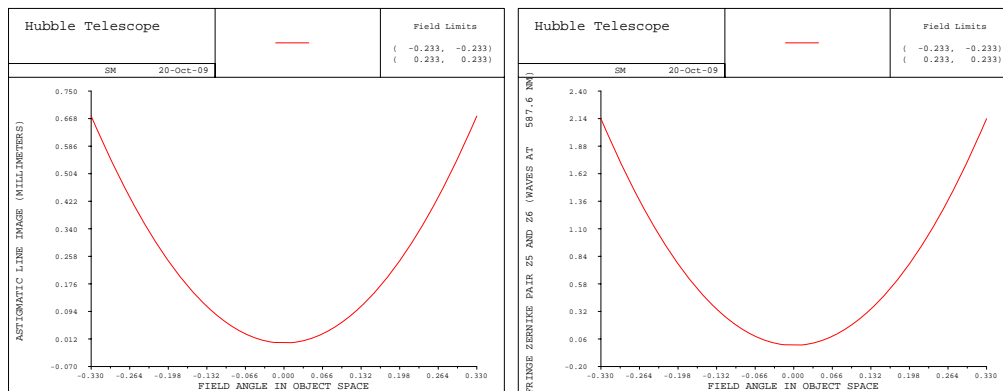
The nominal system is a well-corrected, rotationally-symmetric optical system. One can say that the two nodes both overlap at the center of the field and as a result, the small amount of astigmatism that exists in the system has a quadratic dependence on the field. This can be seen through the length of the line images in the 2-D plots in Figure 5.3. In Figure 5.3, the astigmatism is plotted as an “Astigmatic Line Image” from CODE V and as the coupled coefficients for the fringe Zernike coefficients ( $Z_5$  and  $Z_6$ ). The relative length of the lines through the field in each of these two figures is equivalent, but the orientation of each individual line rotates by  $90^\circ$ . The “Astigmatic line image” in CODE V is calculated through a Coddington ray trace (which depends on paraxial quantities) while the “Fringe Zernike Pair” is calculated through the coefficients that are fit in the pupil. Since the orientations of the lines rotate by  $90^\circ$ , the two figures appear like the orientations of the line images on either side of the medial focus. (These line images are formed on curved images planes, such as the ones labeled “sagittal” and “tangential” in Figure 3.8, that come together for the on-axis field.) Figure 5.4 shows 1-D slices of Figure 5.3 that are clearly quadratic.

The uncoupled plots for the Zernike coefficients are shown in Figure 5.5. The 2-D plots show the magnitude of the coefficients, but their signs are not clear. In



(a) 2-D “Astigmatism line image.”      (b) 2-D “Fringe Zernike pair  $Z_5$  and  $Z_6$ .”

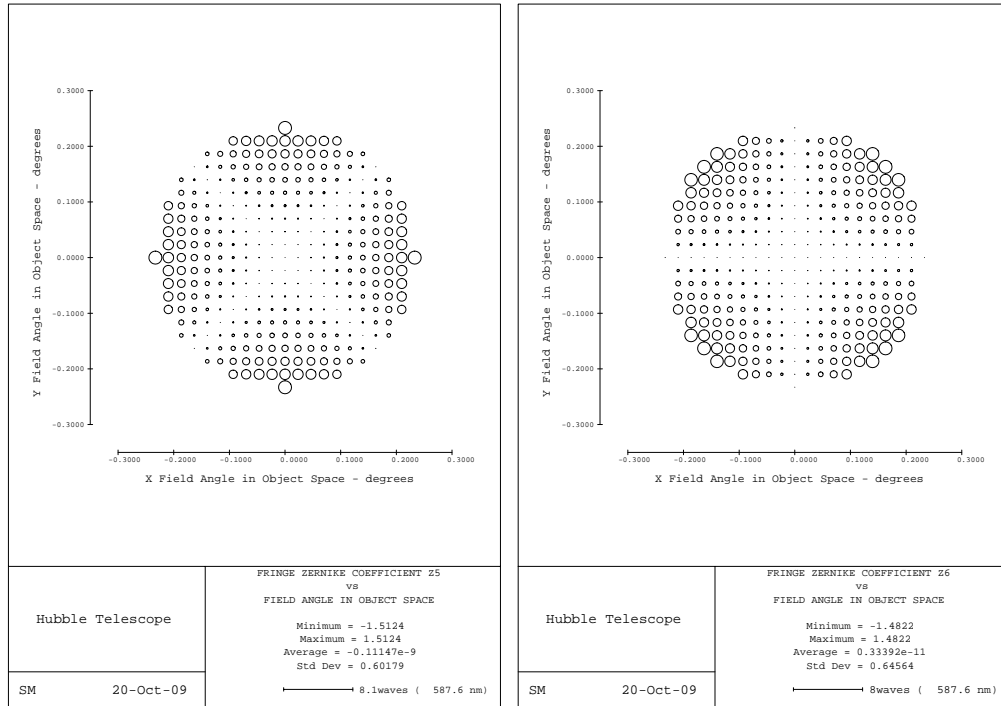
Figure 5.3: 2-D Quadratic astigmatism in Hubble Space Telescope design (Example 1: no misalignment) in CODE V. The astigmatism node on-axis is shown with a red dot.



(a) 1-D “Astigmatism line image.”      (b) 1-D “Fringe Zernike pair  $Z_5$  and  $Z_6$ .”

Figure 5.4: 1-D Quadratic astigmatism in Hubble Space Telescope design (Example 1: no misalignment) in CODE V.

CODE V, these fringe Zernike terms are  $Z_5 = \rho^2 \cos(2\phi)$  and  $Z_6 = \rho^2 \sin(2\phi)$ .



(a) 2-D “Fringe Zernike Coefficient  $Z_5$ .”      (b) 2-D “Fringe Zernike Coefficient  $Z_6$ .”

Figure 5.5: Fringe Zernike coefficients for Hubble Space Telescope (Example 1: no misalignment) in CODE V.

The MATLAB representation of the Zernike standard coefficients for astigmatism, as recorded by ZEMAX, is shown in Figure 5.6. The standard Zernike coefficients in ZEMAX are given by  $Z_5 = \rho^2 \sin(2\phi)$  and  $Z_6 = \rho^2 \cos(2\phi)$ , so  $Z_5$  and  $Z_6$  are switched from the fringe Zernike polynomials in CODE V. The “Total C5” in Figure 5.6 has the same field dependence as shown in the CODE V fringe Zernike  $Z_6$  (Figure 5.5b). Similarly, the “Total C6” in Figure 5.6 has the same field dependence as shown in the CODE V fringe Zernike  $Z_5$  (Figure 5.5a). Figure 5.6 also shows how the magnitude of these recorded coefficients are split into constant, linear and quadratic contributions throughout the field by a least-squares fit to the equations of Tessieres

in Table 3.4.

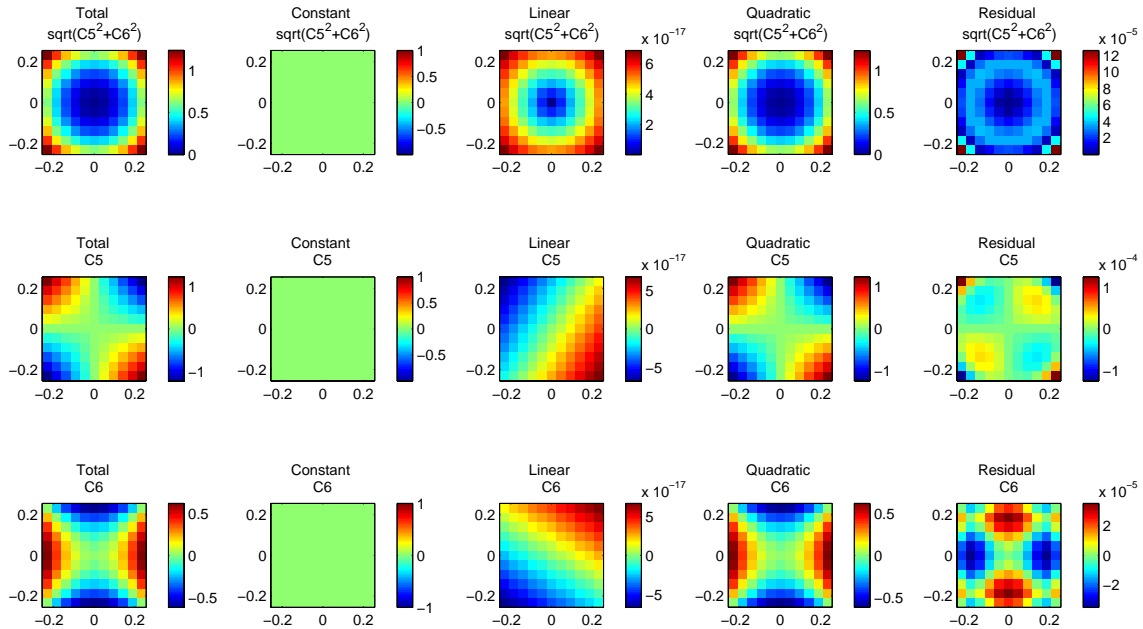


Figure 5.6: Astigmatism coefficients throughout the field in the Hubble Space Telescope Example 1. Quadratic astigmatism dominates and one node is seen on-axis. (The plots show an infinitesimal amount of linear astigmatism, which come from limitations in the numerical calculations.)

The coefficients (from a least-squares fit to equations in Table 3.4) for astigmatism and coma in the nominal Hubble Space Telescope design are listed in Table 5.3. These coefficients are for the case when the field is normalized to unity in the equations. As expected, only the terms with subscript “0” (meaning for the rotationally symmetric system) are nonzero. (The other terms are essentially zero.) Since this telescope has a Ritchey-Chretien design, which is aplanatic, it is corrected for spherical aberration and coma. The limiting aberration expected is quadratic astigmatism, which is the largest coefficient value listed in Table 5.3.

Figure 5.6 shows the magnitude of the astigmatism coefficients, but these values can also be plotted as line images where the length of the line is proportional to

Table 5.3: Astigmatism and coma coefficients in waves ( $\lambda = 586.56$  nm) from Tessier's equations (Table 3.4) for Example 1.

Coefficient	Aberration name	Value
$\alpha_0$	Quadratic astigmatism	0.61747
$\alpha_1$	Linear astigmatism 1	0
$\alpha_2$	Linear astigmatism 2	0
$\alpha_3$	Constant astigmatism 1	0
$\alpha_4$	Constant astigmatism 2	0
$\beta_0$	Linear coma	-0.012668
$\beta_1$	Constant coma 1	0
$\beta_2$	Constant coma 2	0

the total astigmatism (root sum square of the  $Z_5$  and  $Z_6$  components) and the orientation of the line is given by one half the arctangent of the ratio of the two components. Figure 5.7 shows how the coefficients in Table 5.3 turn into line images. Here, only quadratic astigmatism is visible.

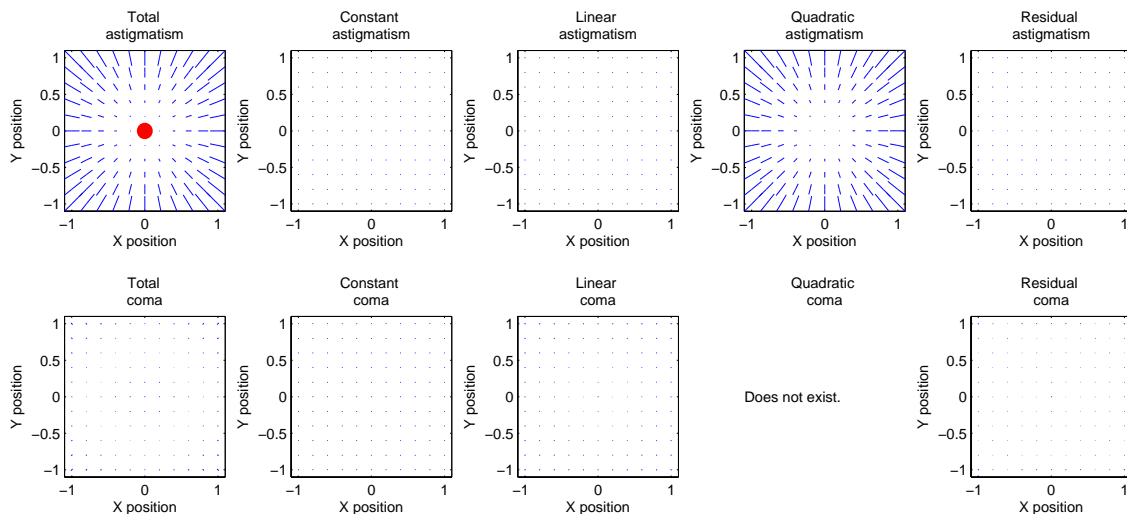


Figure 5.7: Astigmatism and coma in Example 1, separated by field dependence, calculated from coefficients. The astigmatism node on-axis is shown with a red dot. The  $H_x$  and  $H_y$  fields are normalized to the half field of view of  $0.233^\circ$ .

### 5.1.2 CONSTANT ASTIGMATISM FROM PRIMARY MIRROR BENDING

If the primary mirror (entrance pupil) has an astigmatism shape error, then this astigmatism from the surface is added to all of the points in the field. Since all of the fields are equally affected, the induced astigmatism is constant in the field. Examples 2 and 3 show the effect of adding an astigmatism shape error to the primary mirror is constant astigmatism. Example 2 shows a large shape error so the constant astigmatism dominates, while Example 3 shows a more realistic shape error where the total astigmatism is a sum of the constant and native quadratic components.

#### **Example 2: Constant astigmatism from large primary mirror bending error**

In this example, the astigmatism shape error on the primary mirror was chosen to be sufficiently large, so that the constant astigmatism will dominate over the residual quadratic astigmatism of the aligned system. In ZEMAX, three waves of error in  $Z_5 = \sqrt{6}\rho^2 \sin(2\phi)$  using a “Zernike Standard Phase” surface on the primary mirror results in the spot diagram shown in Figure 5.8. The normalizing radius for the Zernike polynomial term is the same radius as that of the primary mirror (1200 mm). (An alternate way to create this surface is to use a “Zernike Standard Sag” surface. Since the wavefront error doubles upon reflection, the astigmatism coefficient for the sag to create the equivalent error should be 1.5 waves = 0.008813 mm.) The spot diagram is for an image plane 20 mm behind focus to show the line images. There are two nodes (as there will always be for astigmatism), but both nodes are out of

the field of view and not seen.

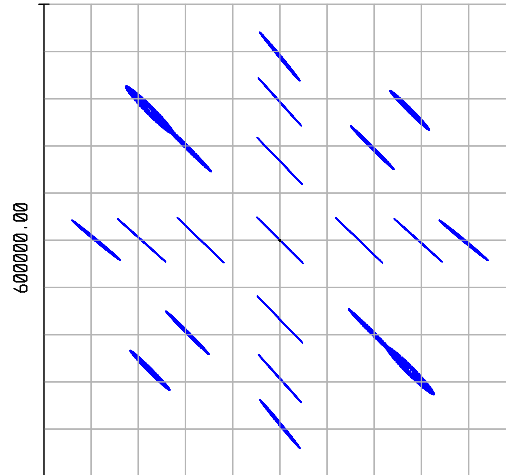
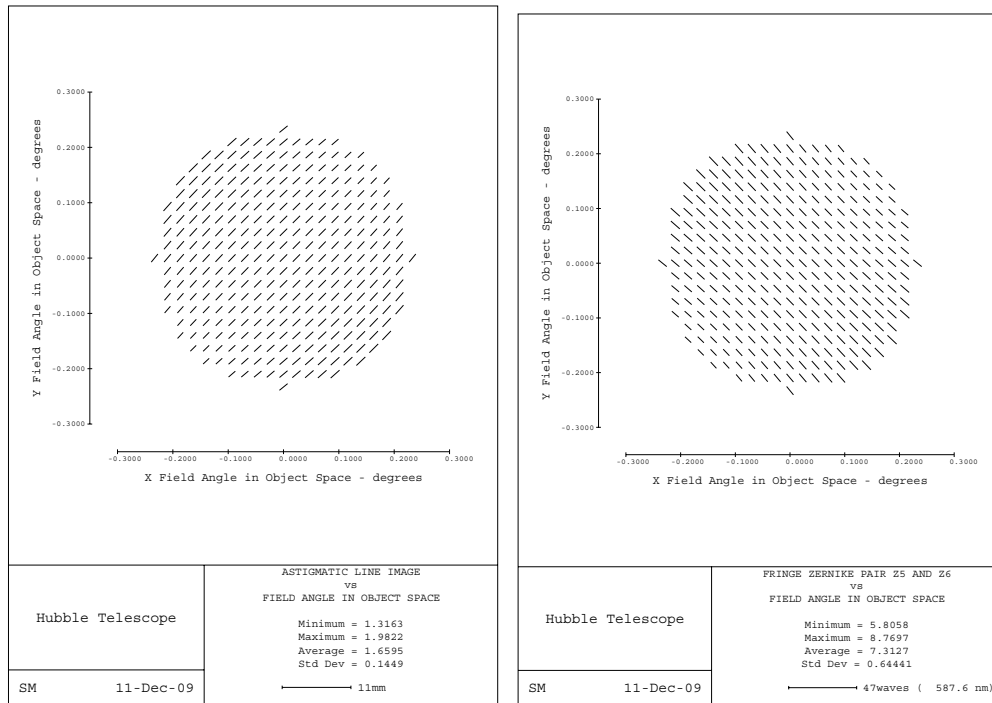


Figure 5.8: Spot diagram showing constant astigmatism for the Hubble Space Telescope with an astigmatism error on the primary mirror (Example 2). (The spots are shown at a defocused image plane 20 mm behind focus and are exaggerated by  $50\times$ .)

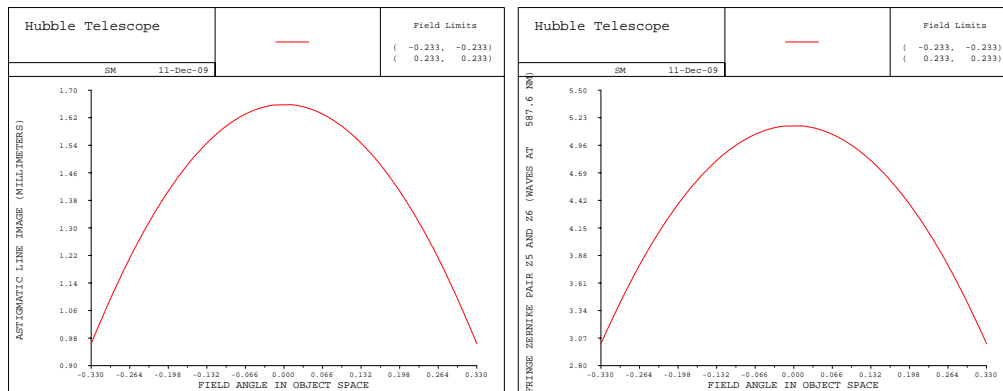
In CODE V, 0.00216 mm of  $Z_6$  astigmatism are placed on the primary mirror with a Zernike polynomial surface. (This includes the normalization factor of  $\sqrt{6}$  which is not included in CODE V for the Zernike polynomial surface.) Figures 5.9 and 5.10 show the results. Constant astigmatism results in line images of the same length and orientation throughout the field. The magnitude of the astigmatism is shown in Figure 5.10. Here, a parabolic shape is seen (from the quadratic native astigmatism), but the  $y$ -axis scales show that the overall effect is mostly constant through the field.

The coefficients for constant, linear and quadratic astigmatism were found from the ZEMAX model using MATLAB and are listed in Table 5.4. All of the numbers are the same as for the nominal design, except for the constant astigmatism term for  $Z_5$ . The MATLAB plots are shown in Figure 5.11 and Figure 5.12.



(a) 2-D “Astigmatism line image.” (b) 2-D “Fringe Zernike pair  $Z_5$  and  $Z_6$ .”

Figure 5.9: 2-D Constant astigmatism in Hubble Space Telescope, with simulated primary mirror astigmatism shape error in CODE V. There are no nodes within the field of view.



(a) 1-D “Astigmatism line image.” (b) 1-D “Fringe Zernike pair  $Z_5$  and  $Z_6$ .”

Figure 5.10: 1-D Constant astigmatism in Hubble Space Telescope, with simulated primary mirror astigmatism shape error in CODE V.



Table 5.4: Coefficients in waves ( $\lambda = 586.56$  nm) for three waves of constant astigmatism added to Hubble primary mirror (Example 2).

Coefficient	Aberration name	Value
$\alpha_0$	Quadratic astigmatism	0.61747
$\alpha_1$	Linear astigmatism 1	0
$\alpha_2$	Linear astigmatism 2	0
$\alpha_3$	Constant astigmatism 1	-3
$\alpha_4$	Constant astigmatism 2	0
$\beta_0$	Linear coma	-0.012668
$\beta_1$	Constant coma 1	0
$\beta_2$	Constant coma 2	0

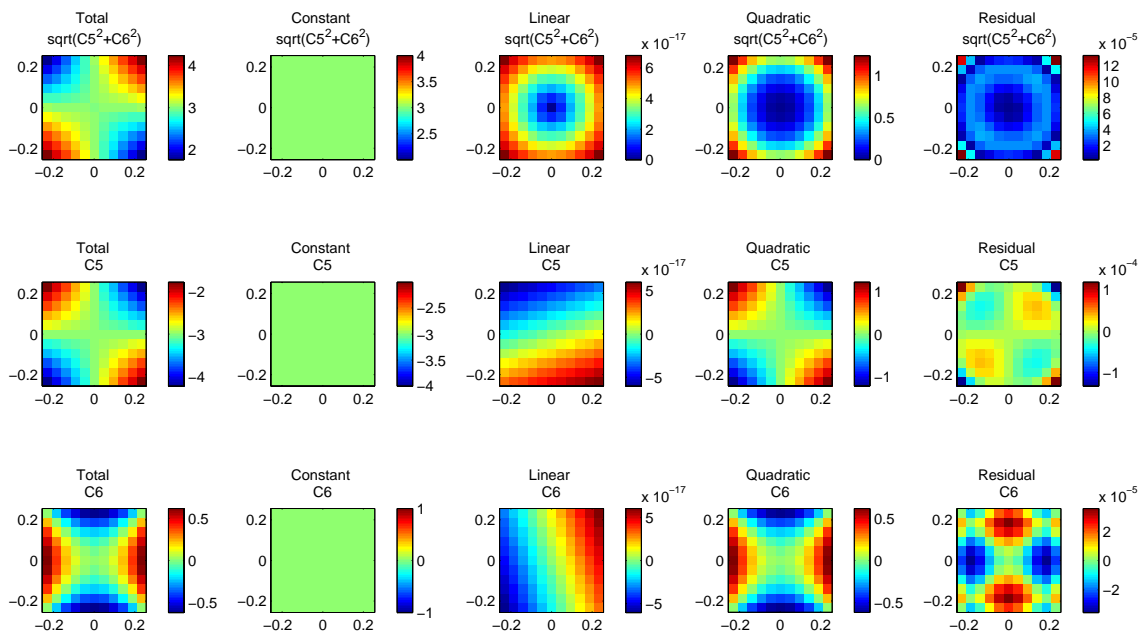


Figure 5.11: Astigmatism coefficients from the Hubble Space Telescope Example 2. Constant astigmatism (from the  $Z_5$  term) dominates. (The plots show an infinitesimal amount of linear astigmatism, which come from limitations in the numerical calculations.)

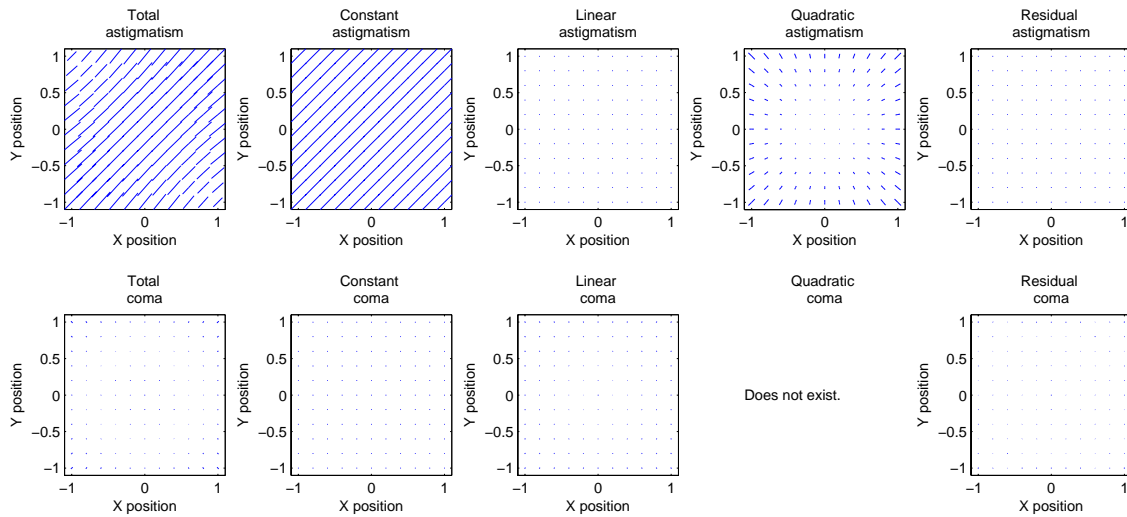


Figure 5.12: Astigmatism and coma in Example 2, separated by field dependence, calculated from coefficients. The  $H_x$  and  $H_y$  fields are normalized to the half field of view of  $0.233^\circ$ .

### Example 3: Native astigmatism in the optical design with a modest amount of astigmatism from the primary mirror bending

In this example, the primary shape error ( $-0.61747$  waves) is chosen so that the constant astigmatism equals the quadratic astigmatism at the edge of the field ( $0.233^\circ$ ). This results in two positions at the edge of the field where the astigmatism cancels. Thus, two nodes are observed at positions that are symmetric about the origin (where the field angle is zero). The astigmatism coefficients found using ZEMAX and MATLAB are listed in Table 5.5. All of the values are the same as for the nominal design, except for the constant astigmatism term for  $Z_5$ . The MATLAB plots are shown in Figure 5.13 and Figure 5.14.

It is important to note that the new field distribution can be calculated as the sum of the two effects:

1. Quadratic astigmatism, native to the design. This is fully described by one

Table 5.5: Coefficients in waves ( $\lambda = 586.56$  nm) for constant astigmatism added to Hubble primary mirror in Example 3.

Coefficient	Aberration name	Value
$\alpha_0$	Quadratic astigmatism	0.61747
$\alpha_1$	Linear astigmatism 1	0
$\alpha_2$	Linear astigmatism 2	0
$\alpha_3$	Constant astigmatism 1	-0.61747
$\alpha_4$	Constant astigmatism 2	0
$\beta_0$	Linear coma	-0.01266
$\beta_1$	Constant coma 1	0
$\beta_2$	Constant coma 2	0

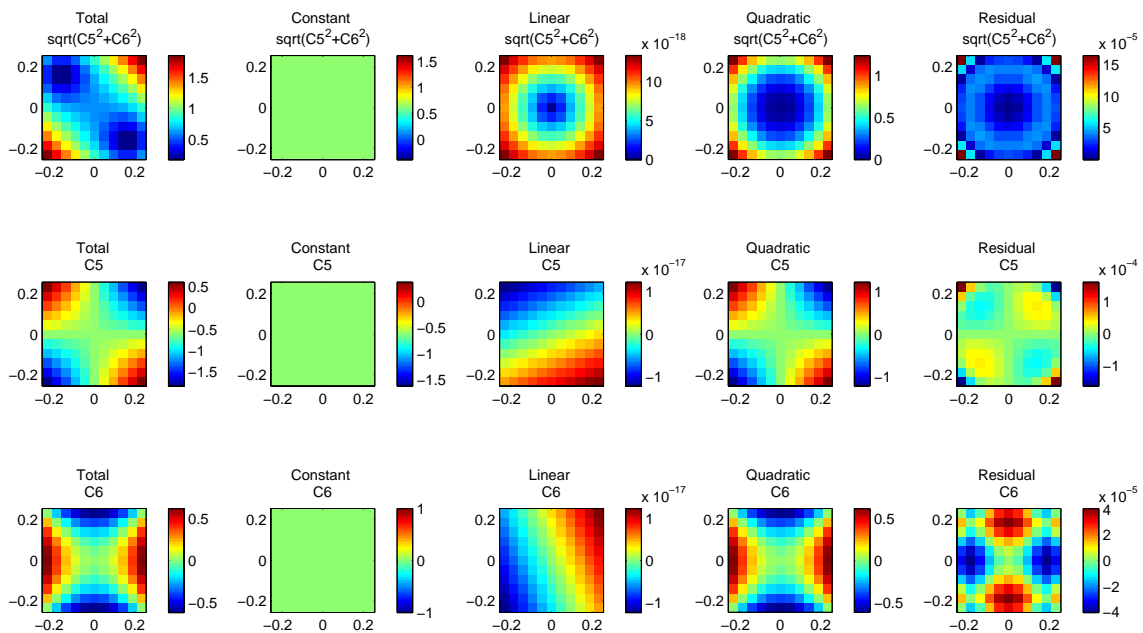


Figure 5.13: Astigmatism coefficients from the Hubble Space Telescope Example 3. Binodal astigmatism (seen in the top left plot) is produced by a combination of quadratic and constant astigmatism. (The plots show an infinitesimal amount of linear astigmatism, which come from limitations in the numerical calculations.)

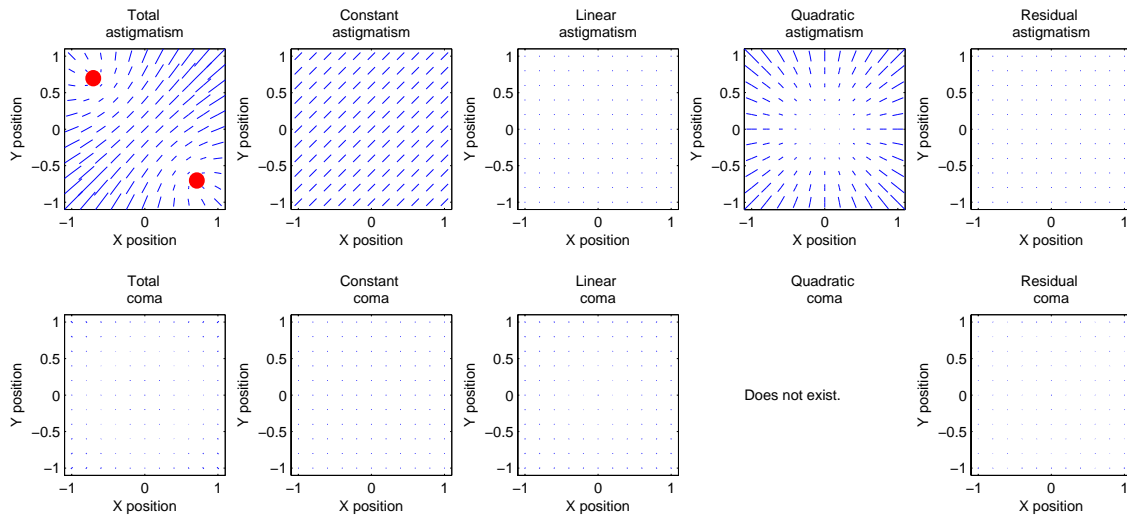


Figure 5.14: Astigmatism and coma, separated by field dependence, calculated from coefficients. Binodal astigmatism may be seen in the top left plot. The astigmatism nodes are shown with red dots. The  $H_x$  and  $H_y$  fields are normalized to the half field of view of  $0.233^\circ$ .

coefficient, and it does not change as the constant astigmatism is added.

2. Constant astigmatism. This is in general described with two coefficients (one each for the  $0^\circ$  and  $45^\circ$  astigmatism contributions).

The binodal effect is created by a linear sum of these.

### 5.1.3 LINEAR ASTIGMATISM FROM SECONDARY MIRROR MISALIGNMENT

In Example 2, the magnitude of the primary mirror shape error was chosen so that the resulting constant astigmatism dominates over the quadratic astigmatism, which is native to the original design. In order to produce linear astigmatism, the secondary mirror must be misaligned. However, this perturbation also produces constant astigmatism. These contributions are predicted by Thompson's equation:

$$W(\vec{H}, \vec{\rho}) = \frac{1}{2} \left( W_{222} \vec{H}^2 - 2\vec{H} \vec{A}_{222} + \vec{B}_{222}^2 \right) \cdot \vec{\rho}^2. \quad (5.1)$$

For this design, it is not possible to choose a large misalignment such that the linear astigmatism is the dominating astigmatism field dependence in the system. For small misalignments, the added linear astigmatism (from the  $\vec{A}_{222}$  term) is small compared to the native quadratic astigmatism. As the amount of misalignment increases, the amount of linear astigmatism also increases linearly. This is because the vector  $\vec{A}_{222}$  scaling the astigmatism component that is linear in field depends linearly on the misalignment. In particular, the misalignment causes the field center to move off-axis, according to the perturbation vector  $\vec{\sigma}$  in the equation  $\vec{A}_{222} = \sum_j W_{222j} \vec{\sigma}_j$ . The constant astigmatism from the misalignment grows with the square of the misalignment ( $\vec{B}_{222}^2 = \sum_j W_{222j} \vec{\sigma}_j^2$ ). As a result, there is no misalignment that can be chosen to show linear misalignment which is large compared to both the quadratic and the constant astigmatism. The following two examples show how the coefficients depend on the magnitude of the misalignment. Of course, constant coma will also be generated by this degree of freedom, but that will not be emphasized here.

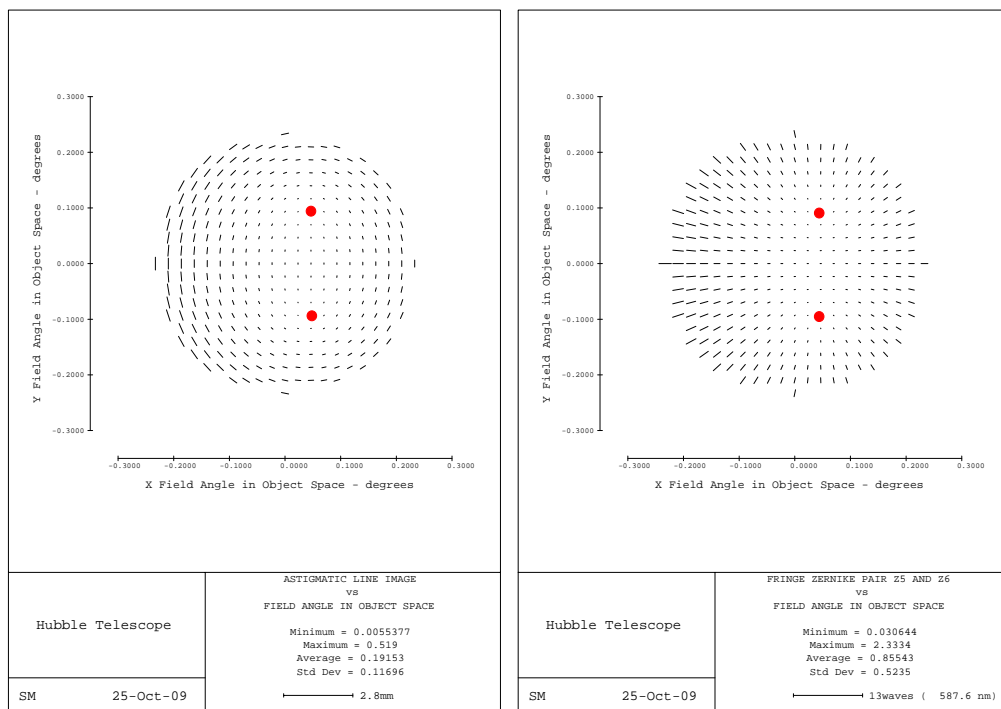
As shown in the previous two examples, constant astigmatism causes the astigmatic line images to have the same orientation and length throughout the field. The new field dependence introduced in the next two examples is linear astigmatism. The system can not be perturbed such that the total astigmatism has only linear field dependence, but the linear contribution can be seen when the total astigmatism is divided into its components.

For linear astigmatism, the length of the line images grows linearly as a function of the field radius, and the orientation varies throughout the field. Linear astigmatism line images always have the same field-asymmetric shape, but the overall rotation of the lines throughout the field may change (i.e. see Figure 4.5) depending on the direction of the misalignment that caused the linear astigmatism and whether the focus is at the sagittal or tangential image surface. One may call this “field-asymmetric, field-linear astigmatism” (Thompson *et al.*, 2008), to be more descriptive. This long name is not needed to distinguish it from the case of “field-symmetric, field-linear astigmatism” because that aberration can not exist. A symmetric shape is impossible (for linear field dependence), based on the argument that the magnitude of the astigmatism at some field point is the product of the two distances from the field point to each of the nodes.

**Example 4: Constant and linear astigmatism from relatively large secondary mirror misalignment combined with native quadratic astigmatism**

In this example, the secondary mirror is misaligned by a relatively large amount (decenter in  $\vec{x}$  by 10 mm). This produces both constant and linear astigmatism contributions and causes the astigmatism line images to become binodal, as shown in the CODE V images in Figure 5.15. However, now the nodes are not symmetric about the origin anymore due to the addition of the linear astigmatism component.

The coefficients for this system, found using ZEMAX and MATLAB, are listed in Table 5.6. The linear coma ( $\beta_0$ ) and quadratic astigmatism ( $\alpha_0$ ) terms slightly



(a) 2-D “Astigmatism line image.”      (b) 2-D “Fringe Zernike pair  $Z_5$  and  $Z_6$ .”

Figure 5.15: 2-D constant, linear and quadratic astigmatism in Hubble Space Telescope, with simulated secondary mirror misalignment in CODE V.

changed from the nominal amounts. This is the nonlinear effect of the square of the misalignment becoming significant. The MATLAB plots are shown in Figure 5.16 and Figure 5.17.

Table 5.6: Coefficients in waves ( $\lambda = 586.56$  nm) for Hubble Space Telescope Example 4 with a 10 mm secondary mirror misalignment in the  $\vec{x}$ -direction.

Coefficient	Aberration name	Value
$\alpha_0$	Quadratic astigmatism	0.61795
$\alpha_1$	Linear astigmatism 1	-0.21655
$\alpha_2$	Linear astigmatism 2	0
$\alpha_3$	Constant astigmatism 1	0
$\alpha_4$	Constant astigmatism 2	0.11973
$\beta_0$	Linear coma	-0.015939
$\beta_1$	Constant coma 1	-5.077
$\beta_2$	Constant coma 2	0



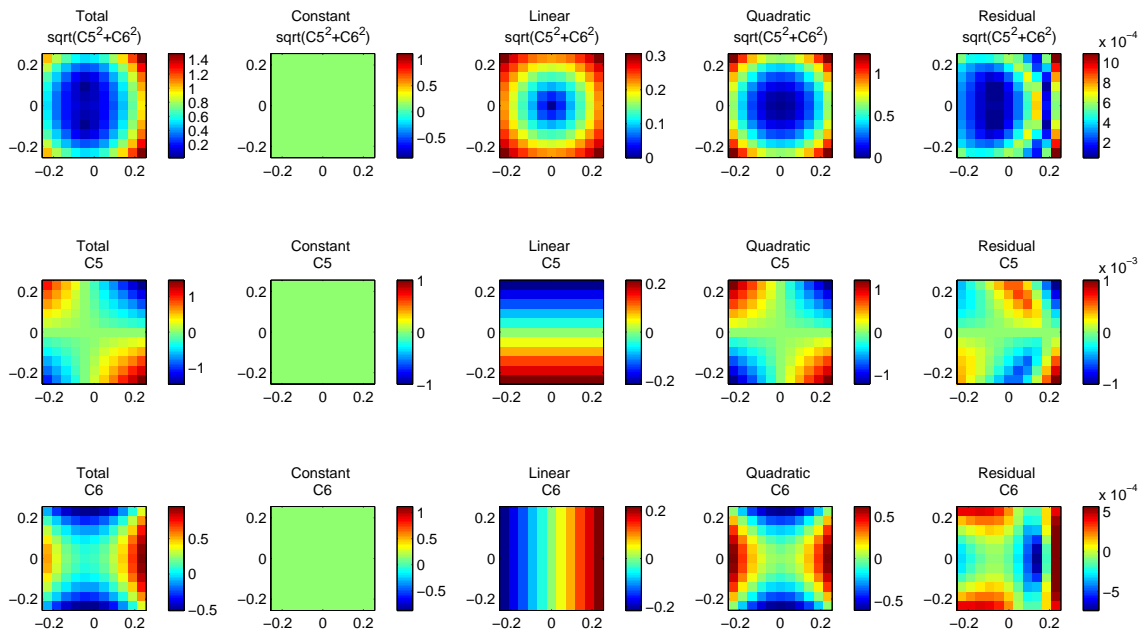


Figure 5.16: Astigmatism coefficients from the Hubble Space Telescope Example 4. Binodal astigmatism (top left plot) is produced by a combination of constant, linear and quadratic astigmatism. The nodes are no longer symmetric about the origin due to the linear component.

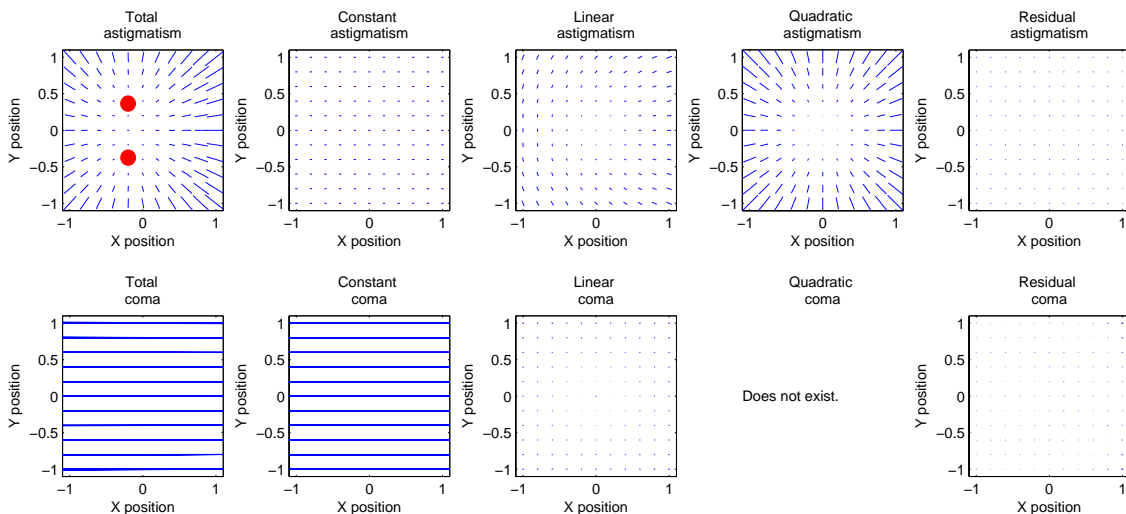
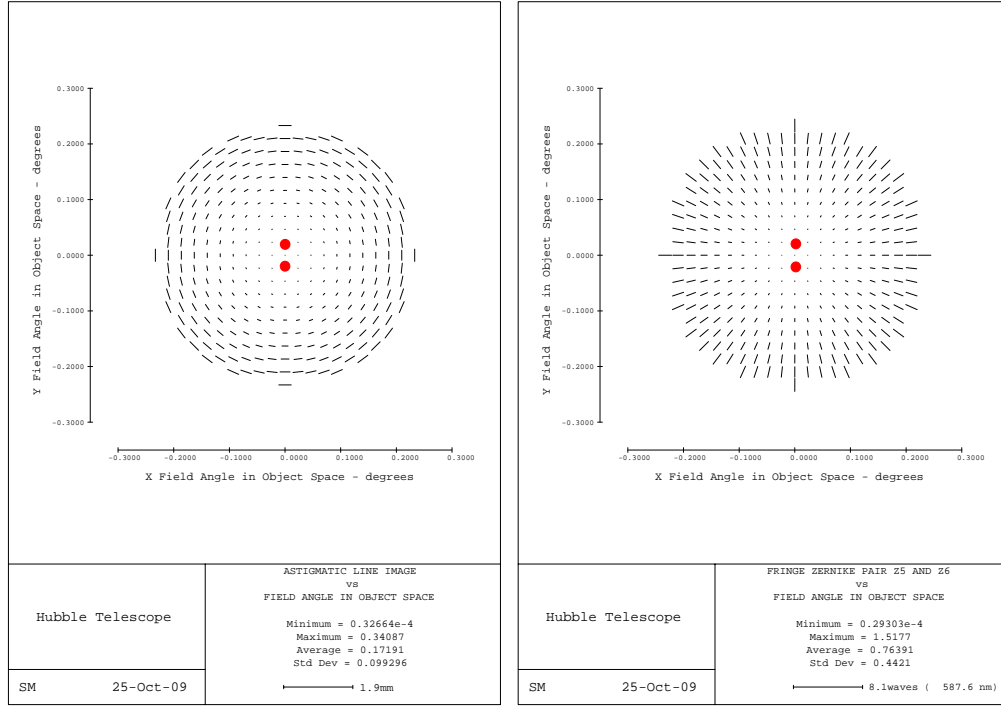


Figure 5.17: Astigmatism and coma, separated by field dependence, calculated from coefficients. Binodal astigmatism may be seen in the top left plot. (The astigmatism nodes are shown with red dots.) The  $H_x$  and  $H_y$  fields are normalized to the half field of view of  $0.233^\circ$ .

**Example 5: Native astigmatism in the optical design with a modest amount of linear astigmatism from secondary mirror misalignment**

Now the secondary mirror is misaligned in the  $\vec{x}$ -direction by 0.1 mm. This is a smaller amount to show the change in the fit coefficients. In terms of nodes, there are two again (as there will always be for astigmatism), but since the misalignment is so small, the nodes are both very close to the origin, as shown by the red dots in Figure 5.18. Since the nodes are close together, they are hard to distinguish with the plots with 21 points across the pupil diameter in Figure 5.18 and practically impossible to see with the 7 points across the pupil diameter in Figure 5.19. Since the combination of linear astigmatism and quadratic astigmatism causes binodal behavior, the knowledge that there is a linear astigmatism contribution (from either Table 5.7 or Figure 5.19) tells us that there must be two nodes for the astigmatism.

The coefficients for this system, found using ZEMAX and MATLAB, are listed in Table 5.7 and plotted in Figures 5.19 and 5.20. The linear coma ( $\beta_0$ ) and quadratic astigmatism ( $\alpha_0$ ) terms now are much closer to the nominal amounts (different from Example 4). The perturbation amount in this example is  $100\times$  smaller than in Example 4. As a result, one would expect the linear astigmatism coefficient ( $\alpha_1$ ) and constant coma 1 ( $\beta_1$ ) terms to be  $100\times$  smaller. The ratios were for  $\alpha_1$   $-0.21655 / -0.0021613 = 100.1943$  and for  $\beta_1$   $-5.077 / -0.050775 = 99.9902$ . One should expect the aberration that depends on the square of the perturbation, constant astigmatism 2 ( $\alpha_4$ ) to be  $100^2\times$  or  $10000\times$  smaller than in Example 4. The actual ratio was  $0.11973 / 1.1876 \times 10^{-5} = 10082$ . Thus, the system behaves as expected.



(a) 2-D “Astigmatism line image.”      (b) 2-D “Fringe Zernike pair  $Z_5$  and  $Z_6$ .”

Figure 5.18: 2-D constant, linear and quadratic astigmatism in Hubble Space Telescope, with small simulated secondary mirror misalignment in CODE V. The astigmatism nodes are shown with red dots.

Table 5.7: Coefficients in waves ( $\lambda = 586.56$  nm) for Hubble Space Telescope Example 5 with a 0.1 mm secondary mirror misalignment in the  $\vec{x}$ -direction.

Coefficient	Aberration name	Value
$\alpha_0$	Quadratic astigmatism	0.61748
$\alpha_1$	Linear astigmatism 1	-0.0021613
$\alpha_2$	Linear astigmatism 2	0
$\alpha_3$	Constant astigmatism 1	0
$\alpha_4$	Constant astigmatism 2	$1.1876e-5$
$\beta_0$	Linear coma	-0.01266
$\beta_1$	Constant coma 1	-0.050775
$\beta_2$	Constant coma 2	0

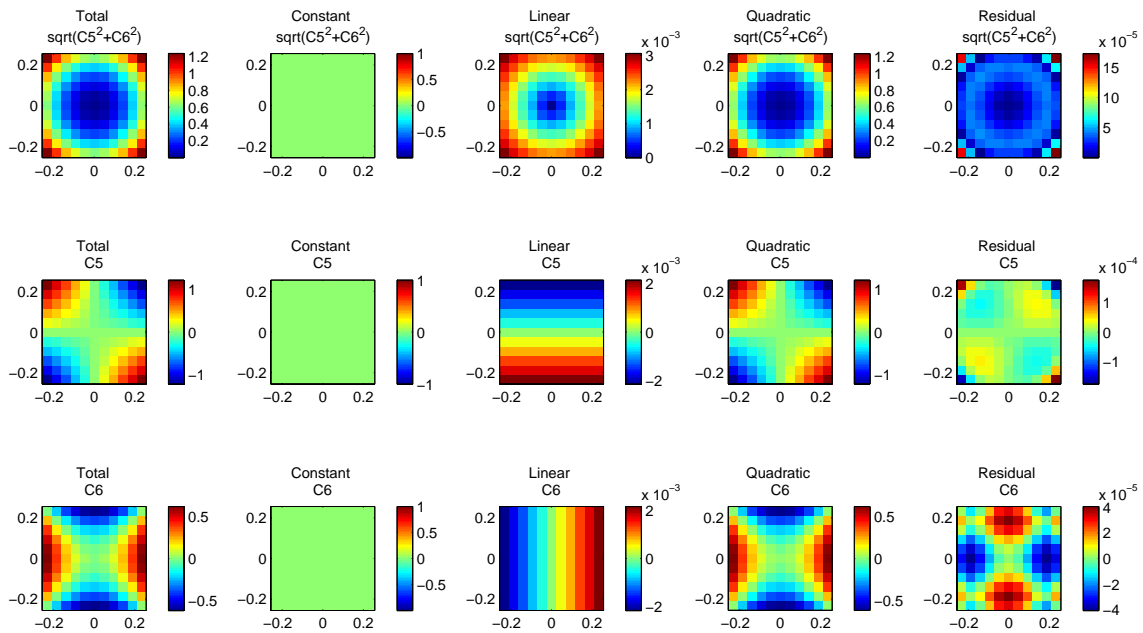


Figure 5.19: Astigmatism coefficients from the Hubble Space Telescope Example 5. Although the astigmatism is binodal here, the resolution is not high enough to show the two separate nodes in the top left plot. The binodal nature of the astigmatism can be inferred by observing that the astigmatism has both quadratic and linear field dependent contributions.

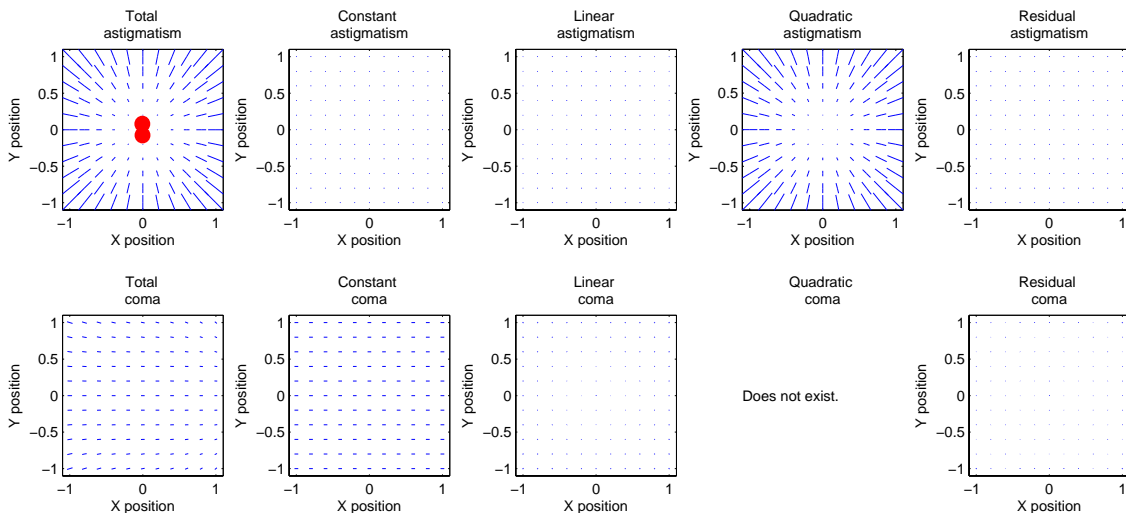


Figure 5.20: Astigmatism and coma, separated by field dependence, calculated from coefficients. The two nodes for the astigmatism in the top left plot are shown with red dots. The  $H_x$  and  $H_y$  fields are normalized to the half field of view of  $0.233^\circ$ .

## 5.2 DISCUSSION OF NODAL ABERRATION THEORY

The effects of misalignments are aberration fields that depend in a repeatable, understandable way based on the misalignment. This field can have some locations where the value of the aberration is zero. These are the nodes, which depend on the misalignment. The number of nodes expected depends on the aberration under consideration. This chapter considered the case of astigmatism, which may have two nodes, as it is more interesting than the case of coma which will have only one node.

Because astigmatism has two nodes, it is treated as “binodal.” However, sometimes this is apparent and sometimes it is not. If both nodes occur together for the on-axis field or if the nodes are not very far apart, then they appear as only one node. Some systems can be found where one node is in the field of view and the other is out of the field. And sometimes, neither of the nodes may be in the field of view. Using the node locations for calculating misalignments is possible, but not the only method.

For a two mirror system, the secondary mirror can be aligned. If only tilts and decenters are considered, there are four degrees of freedom (considering both  $\vec{x}$  and  $\vec{y}$  misalignments). These four misalignments result mainly in four effects: constant coma (two directions) and linear astigmatism (two directions). The secondary mirror can be corrected by first fixing the coma (two variables:  $Z_7$  and  $Z_8$  at one field). Then the secondary mirror can be rotated about its coma-free point to correct the remaining astigmatism ( $Z_5$  and  $Z_6$ ). Nodal Aberration Theory requires that the astigmatism be measured at three fields to determine the misalignments or shape

errors causing astigmatism (Schmid *et al.*, 2008). The astigmatism is determined by the shape and orientation of a defocused image. The nodes are not required to be in the field of view for this method to work.

Alternatively, these four degrees of freedom can be corrected in one step by a least-squares fit to the field dependent functions. There are five astigmatism coefficients ( $\alpha_0$ – $\alpha_4$ ) according to Tessieres’s notation, or:

- Constant astigmatism, mostly from the mirror bending (but also affected by misalignment), two components
- Linear field dependent astigmatism from the secondary mirror alignment, two components, and
- Quadratic astigmatism from the native design, described by a single parameter.

Two Zernike polynomials ( $Z_5$  and  $Z_6$ ) measured at three field point provide six measurements to solve for the five coefficients. This number of required field points to determine the astigmatism is in agreement with Nodal Aberration Theory. If the native quadratic astigmatism is known, and if the primary is known to be perfect, then the wavefront at a single field point, described with two Zernike coefficients, is needed to quantify the linear field dependent astigmatism, thus the secondary mirror misalignment. The three coma coefficients ( $\beta_0$ – $\beta_2$ ) can be found by four measurements ( $Z_7$  and  $Z_8$  each at two fields). In practice, multiple points are used with a least-squares fit to reduce the coupling from noise.

This Zernike polynomial wavefront decomposition method is powerful because it can be extended to systems with more degrees of freedom. If there are more mirrors, then there will be more aberrations that become important, so other measurements need to be made to calculate focal plane tilt ( $Z_4$ ) and/or higher order Zernike polynomials ( $Z_9, Z_{10}, Z_{11}$  etc.) The correct number of wavefront measurements need to be made to account for the number of unknowns in the equations. Once the Zernike polynomials that need to be considered are determined, only the number of variables in the equation need to be counted and then the appropriate number of measurements can be taken to solve for them. Additional measurements will help reduce the noise by  $1/\sqrt{n}$ , where  $n$  is the number of measurements.

It has been argued that the Cassegrain telescope has binodal astigmatism and that only the three-mirror anastigmat (TMA) design will suffer linear field dependent astigmatism. It should be clear from the above analysis that the two nodes in the two mirror design are described by the two components: quadratic and linear field dependent astigmatism. These have different causes and different effects. The usefulness of the method presented here is the ability to decouple these. By fitting the coefficients to data, the linear field dependent astigmatism contribution is quantified in a simple way. There exists a one-to-one relationship between this linear field dependent astigmatism and secondary mirror misalignment. This relationship is independent of the correction of the quadratic field dependence, thus the treatment of this form is identical for a TMA as a Cassegrain design. The net combination of terms and the nodal appearance may be different, but the cause and effect of the linear field dependent astigmatism is the same. (A similar relationship exists for the

combined effect of constant and quadratic or linear astigmatism, which also creates nodal behavior.)

The systematic Zernike representation of the field dependent aberrations is correct, complete, and useful. It is fully consistent with the nodal treatment. The systematic description of the field dependence predicts the nodes, but does not give them special significance. This in no way detracts from other work that focuses on the nodes. It simply provides a different way to describe the field dependence that may be useful.

In addition, no significance is placed on the orientation of the astigmatic lines in a defocused image. For systems with many degrees of freedom (more than three mirrors), the images may be so aberrated that it is not possible to determine the orientation of the astigmatism in the presence of other aberrations. The orientation of the line image may be calculated from the Zernike coefficients measured by the wavefront sensor, but there is no reason to do this.

Astigmatism was treated completely in this chapter to provide a clear example. The Zernike polynomial decomposition is general. All higher order terms are treated in the same way. Likewise, it would be possible to find the zeros or nodes for all cases. Like the case of astigmatism, it may be possible to describe the field dependence using nodes. But like the case of astigmatism, a complete description that does not give special significance to the nodes is both correct and useful.



## CHAPTER 6

## AN ORTHOGONAL SET OF FIELD-DEPENDENT ABERRATIONS

This chapter begins in Section 6.1 with a review of the concept of orthogonality, as applied to polynomials. Zernike polynomials, described in Section 6.2, are an example of orthogonal polynomials commonly used in optics. For this reason, Tessieres (2003) modified the equations describing the possible field-dependent aberrations in a misaligned system from Thompson's wave expansion form into a form using Zernike polynomials over the pupil, as seen in Chapter 3. However, the field-dependent aberrations developed by Tessieres, as reviewed in Section 6.3, are not orthogonal over the field, which limits their usefulness. Therefore, an alternate set of polynomials, orthogonal over both the field and pupil, is introduced in Section 6.4.

These orthogonal polynomials are based on double Zernike polynomial expansions using products of Zernike polynomials in field space and Zernike polynomials in pupil space. The polynomials introduced are an efficient basis for describing systems with tilted and decentered elements because they match the field-dependent aberrations that are induced by misalignment. In order to do this, two double Zernike polynomial terms will be required in the function when the Zernike polynomial in either the field or the pupil has azimuthal dependence. The functions in the new orthogonal basis are compared to the field-dependent aberrations predicted by Thompson in Section 6.5. Section 6.6 describes completing this basis set of functions. Next, the field-dependent aberrations caused by bending or shape errors on

any mirror in the optical system are discussed in Section 6.7. The orthogonal double Zernike polynomials are shown to be a useful basis for describing these errors, as well.

Section 6.8 includes a summary of the orthogonal functions that depend linearly on the misalignments. These are the terms that will be most useful when doing a least-squares fit of the misalignments to the wavefront sensor measurements to determine the current misalignment state of the system. Chapter 6 concludes in Section 6.9 with the results of a numerical simulation listing coefficients for the proposed orthogonal functions in a misaligned system. It is shown that the aberrations scale with the misalignments in the system as expected.

## 6.1 INTRODUCTION TO ORTHOGONAL POLYNOMIALS

Orthogonal polynomials are of practical use in many applications. Many books have been written about orthogonal polynomials, but often these are very mathematical and hard for an engineer or scientist to digest. Fortunately, at least one comprehensible book has been written by Beckmann (1973).

Oftentimes, one wishes to expand a function (perhaps representing measured data) into other functions. An example of this is a Taylor series that expands the function  $g(x)$  into a system of expansion functions, such as

$$g(x) = C_0 + C_1 f_1(x) + C_2 f_2(x) + \dots, \quad (6.1)$$

where the  $C_n$  are the coefficients weighting the expansion functions  $f_n(x) = \{1, x, x^2, \dots\}$ . Expansion functions can also be dependent on any number of vari-

ables. An example of this in optics is a expansion that depends on the two pupil coordinates,  $\rho$  and  $\phi$ . The Seidel aberration function (Equation 3.5) expands a wavefront into aberration terms for spherical aberration, coma, astigmatism, focus and distortion:

$$W(\rho, \phi) = a_{40} \rho^4 + a_{31} \rho^3 \cos \phi + a_{22} \rho^2 \cos^2 \phi + a_{20} \rho^2 + a_{11} \rho \cos \phi, \quad (6.2)$$

recalling that the individual terms were discussed in Chapter 3. Neither of the sets of functions in these two examples ( $f_n = \{1, x, x^2, \dots\}$  or  $\{\rho^4, \rho^3 \cos \phi, \rho^2 \cos^2 \phi, \rho^2, \rho \cos \phi, \dots\}$ ) is orthogonal, so their use is limited because the coefficient values depend on the number of terms in the fit.

Orthogonal functions are often preferable because the coefficient of any term can be calculated directly from the following inner product without requiring a least-squares fit (Beckmann, 1973):

$$C_n = \frac{1}{h_n^2} \int_a^b g(x) f_n(x) dx. \quad (6.3)$$

In Equation 6.3,  $C_n$  is again the coefficient of the  $n^{\text{th}}$  function  $f_n(x)$ . The functions are from an orthogonal basis set, each with norm  $h_n$ , and  $g(x)$  is the function to be fit (e.g. the measured data). The functions are orthogonal over the line segment  $a \leq x \leq b$ . For a basis set of one-dimensional functions  $\{f_1(x), f_2(x), \dots\}$  that are orthogonal in the interval from  $a$  to  $b$ , the orthogonality relation is:

$$\int_a^b f_m(x) f_n(x) dx = h_n^2 \delta_{mn}, \quad (6.4)$$

where  $\delta_{mn}$  is the Kronecker delta:

$$\delta_{mn} = \begin{cases} 0, & \text{for } m \neq n; \\ 1, & \text{for } m = n. \end{cases} \quad (6.5)$$

The functions are considered orthogonal based on the analogy with vector algebra. Two vectors are orthogonal if and only if their scalar product vanishes. There are infinite different sets of functions that are orthogonal. For example the system  $\{1, \cos x, \cos 2x, \dots\}$  is orthogonal because the system  $\{6, 12 \cos x, 1970 \cos 2x, \dots\}$  is orthogonal. If the functions  $f_n(x)$  with norms  $h_n$  are orthogonal, then the functions  $f_n(x)/h_n$  are orthonormal. The act of specifying the multiplicative constants (or the norms  $h_n$ ) is called standardizing the system. Normalization is a special case of standardization when all the norms are set equal to unity. This type of orthogonal system is called orthonormal.

Another important consideration of orthogonal systems is completeness. To be able to represent any physical system using a sum of orthogonal functions, the set of orthogonal functions needs to be complete. For example, the set of functions  $\{\cos x, \cos 2x, \cos 3x\}$  is incomplete because there are physical functions that no combination of terms can possibly represent, such as  $\sin x$ . A complete set of functions, such as  $\{1, \cos x, \cos 2x, \cos 3x, \dots, \sin x, \sin 2x, \sin 3x, \dots\}$  in this example, is needed to represent a function by its Fourier expansion.

For a two dimensional function, an integral, similar to Equation 6.4, defines the orthogonality over some area. In optics, usually this area is a unit circle defining

the pupil, which results in the following integral:

$$\int_0^{2\pi} \int_0^1 f_m(\rho, \phi) f_n(\rho, \phi) \rho d\rho d\phi = \delta_{nm}, \quad (6.6)$$

where the functions  $f_n(\rho, \phi)$  are orthonormal. However, orthogonality over other areas, like rectangular regions, annular circles (Mahajan, 1981), or polygons (such as hexagons (Upton & Ellerbroek, 2004; Mahajan & Dai, 2006)), may be chosen depending on the application. This usually occurs when the pupil or optical surface (if using Zernike polynomials to describe a surface measured during the optical test) is not circular.

## 6.2 ORTHOGONALITY OF ZERNIKE POLYNOMIALS

Similarly to Equation 6.1, the wavefront describing the aberrations in an optical system can be written as a sum of polynomials:

$$W(\rho, \phi) = \sum_{j=1}^n C_j Z_j(\rho, \phi), \quad (6.7)$$

where the functions  $Z_j(\rho, \phi)$  are Zernike polynomials. (This is the same equation as described in Section 3.7.) When the pupil is a complete circle, the coefficients  $C_j$  in Equation 6.7 are independent of the number of terms  $n$  in the fit. Since the Zernike polynomials are orthogonal over a unit circle, the following equation holds:

$$\frac{1}{\pi} \int_0^{2\pi} \int_0^1 Z_i(\rho, \phi) Z_j(\rho, \phi) \rho d\rho d\phi = \delta_{ij}. \quad (6.8)$$

The  $\frac{1}{\pi}$  is included in the orthogonality integral to change the normalization such that

$$\frac{\int_0^{2\pi} \int_0^1 Z_i(\rho, \phi) Z_j(\rho, \phi) \rho d\rho d\phi}{\int_0^{2\pi} \int_0^1 \rho d\rho d\phi} = \delta_{ij}. \quad (6.9)$$

This is done so that the expansion coefficient  $C_j$  of a given Zernike polynomial term is equal to the rms value of the term. An equation similar to Equation 6.3 exists in 2D form, where  $W(\rho, \phi)$  replaces  $g(x)$  and  $Z_j(\rho, \phi)$  replaces  $f_n(x)$ :

$$C_j = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 W(\rho, \phi) Z_j(\rho, \phi) \rho d\rho d\phi. \quad (6.10)$$

There are many varieties of numbering schemes and normalizations used for Zernike polynomials, as mentioned in Appendix F, but all different groups of Zernike polynomials are composed of the same orthogonal functions.

### 6.3 ABERRATIONS FOR MISALIGNED SYSTEMS USING ZERNIKE POLYNOMIALS

Equation 6.7 describes a single optical surface or the wavefront aberration for just one chosen field angle in an optical system. However, functions can be written to describe the aberrations of an optical system at any field angle by including terms for the field dependence. In his thesis, Tessieres (2003) described the wavefront aberrations that are possible in a tilted and decentered system throughout the whole field by describing the  $H_x$  and  $H_y$  dependence of the (previously constant) coefficients of

the Zernike polynomials, in the form of

$$W(\rho, \phi, H_x, H_y) = \sum_{i=1}^n C_i(H_x, H_y) Z_i(\rho, \phi). \quad (6.11)$$

He compiled the functions for third-order aberrations and for fifth order aberrations (Tessieres, 2003). As a reminder, systems with third-order aberrations may be described using Zernike polynomial terms for field curvature ( $i = 4$ ), astigmatism ( $i = 5, 6$ ) and coma ( $i = 7, 8$ ). The indices in parentheses are for the standard Zernike polynomial ordering (Noll, 1976). Some of the Zernike polynomials have the same radial dependence, but varying azimuthal dependence ( $\sin \phi$  or  $\cos \phi$ ). For example,  $Z_7(\rho, \phi) = \rho^3 \sin \phi$  and  $Z_8(\rho, \phi) = \rho^3 \cos \phi$  both have the same  $\rho^3$ -dependence, but different azimuthal dependence ( $\sin \phi$  or  $\cos \phi$ ). There are some Zernike polynomials that are radially symmetric and have no dependence on angle, but the rest of the terms come in pairs with sine and cosine dependence. For the Zernike polynomials that do come in pairs, the Greek letter coefficients of Tessieres are always coupled. For example, the linear coma coefficient  $\beta_0$  describes both the  $H_y$  field dependence of  $Z_7(\rho, \phi)$  and the  $H_x$  dependence of  $Z_8(\rho, \phi)$ . Or for the case of astigmatism, recalling the example in Chapter 3 for a misaligned system, the coefficients  $C_5$  and  $C_6$  are described by the functions:

$$C_5(H_x, H_y) = 2 \alpha_0 H_x H_y + \alpha_1 H_y + \alpha_2 H_x + \alpha_3 \quad (6.12)$$

$$C_6(H_x, H_y) = \alpha_0 (H_x^2 - H_y^2) + \alpha_1 H_x - \alpha_2 H_y + \alpha_4 \quad (6.13)$$

when considering third-order aberrations. The term  $\alpha_0$  describes the quadratic astigmatism that is present in a rotationally symmetric system, while the  $\alpha_1$  and  $\alpha_2$  terms describe the two components of linear astigmatism and the  $\alpha_3$  and  $\alpha_4$  terms describe the two components of constant astigmatism in the misaligned system. When considering fifth-order aberrations (from  $W_{422}$ ), the equations for the coefficients  $C_5$  and  $C_6$  of the Zernike terms  $Z_5(\rho, \phi)$  and  $Z_6(\rho, \phi)$  have a modified form and more higher power field terms (up to 4th order in field) must be included:

$$\begin{aligned}
C_5(H_x, H_y) = & 2\chi_0(H_x H_y^3 + H_x^3 H_y) + \chi_1(3H_x^2 H_y + H_y^3) \\
& + \chi_2(H_x^3 + 3H_x H_y^2) + \chi_3(H_x^2 + H_y^2) + \chi_5 2H_x H_y \\
& - \chi_6 H_y + \chi_7 H_x + \chi_8 H_y + \chi_9 H_x + \chi_{10},
\end{aligned} \tag{6.14}$$

$$\begin{aligned}
C_6(H_x, H_y) = & \chi_0(H_x^4 - H_y^4) + 2\chi_1 H_x^3 - 2\chi_2 H_y^3 \\
& + \chi_4(H_x^2 + H_y^2) + \chi_5(H_x^2 - H_y^2) + \chi_6 H_x \\
& + \chi_7 H_y + \chi_8 H_x - \chi_9 H_y + \chi_{11}.
\end{aligned} \tag{6.15}$$

The coefficients  $\alpha_0$ - $\alpha_4$  calculated previously are no longer meaningful. Additionally, when fifth order aberrations are considered, more functions must be added to describe the coefficients for trefoil ( $i = 9, 10$ ), fifth-order field curvature ( $i = 11$ ), oblique spherical aberration ( $i = 12, 13$ ) and fifth-order coma ( $i = 16, 17$ ).



#### 6.4 EXAMPLE DERIVATIONS OF ORTHOGONAL FIELD-DEPENDENT ABERRATIONS

Using functions  $S_j(\rho, \phi, h, \theta)$  that are orthogonal over the pupil *and the field* as in the equation

$$\int_0^{2\pi} \int_0^1 \int_0^{2\pi} \int_0^1 S_i(\rho, \phi, h, \theta) S_j(\rho, \phi, h, \theta) \rho d\rho d\phi h dh d\theta = \delta_{ij} \quad (6.16)$$

converts the wavefront sum in Equation 6.11 into the following form:

$$W(\rho, \phi, h, \theta) = \sum_{j=1}^n C_j S_j(\rho, \phi, h, \theta). \quad (6.17)$$

Functions such as this, which allow the coefficients  $C_j$  to be independent of the number of terms included, will be derived in this section.

In order to have an efficient basis for describing the field-dependent aberrations, the low order orthogonal functions should correspond directly to the aberration coefficients, given as Greek letters by Tessieres (2003). The first step is to write the function that describes each Greek letter coefficient in Tessieres' equations (listed in Table 3.4). The next step is to change the  $(H_x, H_y)$ -dependence into  $(h, \theta)$ -dependence and convert the resulting equations into Zernike notation. To demonstrate this, the five orthogonal functions for astigmatism — quadratic ( $1\times$ ), linear ( $2\times$ ) and constant ( $2\times$ ) with field — are derived in this section. The functions for the other aberrations are derived in Appendix J.

For the case of quadratic astigmatism, collect all the terms from Equations 6.12

and 6.13 with contributions for  $\alpha_0$ :

$$\begin{aligned}
W_{\alpha_0 \text{ terms only}} &= 2 H_x H_y Z_5(\rho, \phi) + (H_x^2 - H_y^2) Z_6(\rho, \phi) \\
&= 2h^2 \sin \theta \cos \theta Z_5(\rho, \phi) + (h^2 \cos^2 \theta - h^2 \sin^2 \theta) Z_6(\rho, \phi) \\
&= h^2 \sin 2\theta Z_5(\rho, \phi) + h^2 \cos 2\theta Z_6(\rho, \phi)
\end{aligned} \tag{6.18}$$

$$S(\rho, \phi, h, \theta) = Z_5(h, \theta) Z_5(\rho, \phi) + Z_6(h, \theta) Z_6(\rho, \phi).$$

For linear astigmatism in one direction, the terms with contributions from  $\alpha_1$  are:

$$\begin{aligned}
W_{\alpha_1 \text{ terms only}} &= H_y Z_5(\rho, \phi) + H_x Z_6(\rho, \phi) \\
&= h \sin \theta Z_5(\rho, \phi) + h \cos \theta Z_6(\rho, \phi)
\end{aligned} \tag{6.19}$$

$$S(\rho, \phi, h, \theta) = Z_3(h, \theta) Z_5(\rho, \phi) + Z_2(h, \theta) Z_6(\rho, \phi),$$

and the linear astigmatism in the other direction has contributions from  $\alpha_2$ :

$$\begin{aligned}
W_{\alpha_2 \text{ terms only}} &= H_x Z_5(\rho, \phi) - H_y Z_6(\rho, \phi) \\
&= h \cos \theta Z_5(\rho, \phi) - h \sin \theta Z_6(\rho, \phi)
\end{aligned} \tag{6.20}$$

$$S(\rho, \phi, h, \theta) = Z_2(h, \theta) Z_5(\rho, \phi) - Z_3(h, \theta) Z_6(\rho, \phi).$$

Constant astigmatism, which has no field dependence, is described by Zernike piston terms in field, such that:

$$W_{\alpha_3 \text{ terms only}} = Z_5(\rho, \phi) \tag{6.21}$$

$$S(\rho, \phi, h, \theta) = Z_1(h, \theta) Z_5(\rho, \phi)$$

$$\begin{aligned}
W_{\alpha_4 \text{ terms only}} &= Z_6(\rho, \phi) \\
S(\rho, \phi, h, \theta) &= Z_1(h, \theta) Z_6(\rho, \phi).
\end{aligned}
\tag{6.22}$$

The functions  $Z_5(h, \theta)Z_5(\rho, \phi) + Z_6(h, \theta)Z_6(\rho, \phi)$  (for quadratic astigmatism),  $Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi)$  (for linear astigmatism),  $Z_2(h, \theta)Z_5(\rho, \phi) - Z_3(h, \theta)Z_6(\rho, \phi)$  (for linear astigmatism),  $Z_1(h, \theta)Z_5(\rho, \phi)$  (for constant astigmatism), and  $Z_1(h, \theta)Z_6(\rho, \phi)$  (for constant astigmatism) are the orthogonal functions  $S(\rho, \phi, h, \theta)$  for third-order astigmatism in a misaligned system. The remaining functions are derived in Appendix J and summarized in Section 6.5.

## 6.5 SUMMARY OF ORTHOGONAL FIELD-DEPENDENT ABERRATIONS

The final results, compiled in Tables 6.1–6.7, list the functions describing combinations of Zernike polynomials in pupil and field space that are possible in misaligned systems. Each aberration (in pupil space) is given its own table. The functions in Tables 6.1–6.7 are ordered according to increased field dependence. Deciding how to order these polynomials is not a trivial task. As was seen in Appendix F, there are many different ways to order Zernike polynomials, which are polynomials of only two indices; it is that much more unnatural to order polynomials with four index terms.

The aberrations that come in pairs in pupil space (coma, astigmatism, trefoil, etc.) are combined into one table for each pair. In addition, these functions are combined with the Zernike polynomials in field space that come in pairs. The

rotationally symmetric aberrations that do not come in pairs, such as piston ( $i = 1$ ), focus/field curvature ( $i = 4$ ), and spherical aberration ( $i = 11$ ), contain just one term for the orthogonal function, as shown in Tables 6.1 and 6.5 and in the rows containing  $Z_1(h, \theta)$  or  $Z_4(h, \theta)$  in the other tables.

Table 6.1: Polynomials orthogonal over field and pupil space for focus terms  $Z_4(\rho, \phi)$ .

Function	Tessieres	Thompson
$Z_1(h, \theta)Z_4(\rho, \phi)$ (cont.)	$\gamma_0, \gamma_3, \psi_0, \psi_3$ $\psi_8, \delta_0, \nu_0$	$W_{220_M}, B_{220_M}, W_{420_M}, B_{420_M},$ $D_{420_M}, W_{240_M}, W_{040}$
$Z_2(h, \theta)Z_4(\rho, \phi)$	$\gamma_1, \psi_1, \psi_6, \delta_2$	$\vec{A}_{220_M}, \vec{A}_{420_M}, \vec{C}_{420_M}, \vec{A}_{240_M}$
$Z_3(h, \theta)Z_4(\rho, \phi)$	$\gamma_2, \psi_2, \psi_7, \delta_1$	$\vec{A}_{220_M}, \vec{A}_{420_M}, \vec{C}_{420_M}, \vec{A}_{240_M}$
$Z_4(h, \theta)Z_4(\rho, \phi)$	$\gamma_0, \psi_0, \psi_3, \delta_0$	$W_{220_M}, W_{420_M}, B_{420_M}, W_{240_M}$
$Z_5(h, \theta)Z_4(\rho, \phi)$	$\psi_4$	$\vec{B}_{420_M}$
$Z_6(h, \theta)Z_4(\rho, \phi)$	$\psi_5$	$\vec{B}_{420_M}$
$Z_7(h, \theta)Z_4(\rho, \phi)$	$\psi_2$	$\vec{A}_{420_M}$
$Z_8(h, \theta)Z_4(\rho, \phi)$	$\psi_1$	$\vec{A}_{420_M}$
$Z_{11}(h, \theta)Z_4(\rho, \phi)$	$\psi_0$	$W_{420_M}$

Table 6.2: Polynomials orthogonal over field and pupil space for astigmatism terms  $Z_5(\rho, \phi)$  and  $Z_6(\rho, \phi)$ .

Function	Tessieres	Thompson
$Z_1(h, \theta)Z_5(\rho, \phi)$	$\alpha_3, \chi_3, \chi_{10}, \eta_4$	$\vec{B}_{222}, \vec{B}_{422}, \vec{D}_{422}, \vec{B}_{242}$
$Z_1(h, \theta)Z_6(\rho, \phi)$	$\alpha_4, \chi_4, \chi_{11}, \eta_3$	$\vec{B}_{222}, \vec{B}_{422}, \vec{D}_{422}, \vec{B}_{242}$
$Z_2(h, \theta)Z_5(\rho, \phi) + Z_3(h, \theta)Z_6(\rho, \phi)$	$\chi_7$	$\vec{C}_{422}$
$Z_2(h, \theta)Z_5(\rho, \phi) - Z_3(h, \theta)Z_6(\rho, \phi)$	$\alpha_2, \chi_2, \chi_9, \eta_1$	$\vec{A}_{222}, \vec{A}_{422}, \vec{C}_{422}, \vec{A}_{242}$
$Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi)$	$\alpha_1, \chi_1, \chi_8, \eta_2$	$\vec{A}_{222}, \vec{A}_{422}, \vec{C}_{422}, \vec{A}_{242}$
$Z_3(h, \theta)Z_5(\rho, \phi) - Z_2(h, \theta)Z_6(\rho, \phi)$	$\chi_6$	$\vec{C}_{422}$
$Z_4(h, \theta)Z_5(\rho, \phi)$	$\chi_3$	$\vec{B}_{422}$
$Z_4(h, \theta)Z_6(\rho, \phi)$	$\chi_4$	$\vec{B}_{422}$
$Z_5(h, \theta)Z_5(\rho, \phi) + Z_6(h, \theta)Z_6(\rho, \phi)$	$\alpha_0, \chi_0, \chi_5, \eta_0$	$W_{222}, W_{422}, B_{422}, W_{242}$
$Z_7(h, \theta)Z_5(\rho, \phi) + Z_8(h, \theta)Z_6(\rho, \phi)$	$\chi_1$	$\vec{A}_{422}$
$Z_8(h, \theta)Z_5(\rho, \phi) - Z_7(h, \theta)Z_6(\rho, \phi)$	$\chi_2$	$\vec{A}_{422}$
$Z_9(h, \theta)Z_5(\rho, \phi) + Z_{10}(h, \theta)Z_6(\rho, \phi)$	$\chi_1$	$\vec{A}_{422}$
$Z_{10}(h, \theta)Z_5(\rho, \phi) - Z_9(h, \theta)Z_6(\rho, \phi)$	$\chi_2$	$\vec{A}_{422}$
$Z_{13}(h, \theta)Z_5(\rho, \phi) + Z_{12}(h, \theta)Z_6(\rho, \phi)$	$\chi_0$	$W_{422}$

Table 6.3: Polynomials orthogonal over field and pupil space for coma terms  $Z_7(\rho, \phi)$  and  $Z_8(\rho, \phi)$ .

Function	Tessieres	Thompson
$Z_1(h, \theta)Z_7(\rho, \phi)$	$\beta_1, \xi_1, \xi_4, \xi_9, \kappa_2$	$\vec{A}_{131}, \vec{A}_{331M}, \vec{C}_{331M}, \vec{A}_{151}$
$Z_1(h, \theta)Z_8(\rho, \phi)$	$\beta_2, \xi_2, \xi_5, \xi_{10}, \kappa_1$	$\vec{A}_{131}, \vec{A}_{331M}, \vec{C}_{331M}, \vec{A}_{151}$
$Z_2(h, \theta)Z_7(\rho, \phi) + Z_3(h, \theta)Z_8(\rho, \phi)$	$\xi_6$	$\vec{B}_{331M}$
$Z_3(h, \theta)Z_7(\rho, \phi) + Z_2(h, \theta)Z_8(\rho, \phi)$	$\beta_0, \xi_0, \xi_3, \kappa_0$	$W_{131}, W_{331M}, B_{331M}, W_{151}$
$Z_3(h, \theta)Z_7(\rho, \phi) - Z_2(h, \theta)Z_8(\rho, \phi)$	$\xi_7$	$\vec{B}_{331M}$
$Z_4(h, \theta)Z_7(\rho, \phi)$	$\xi_1, \xi_4$	$\vec{A}_{331M}$
$Z_4(h, \theta)Z_8(\rho, \phi)$	$\xi_2, \xi_5$	$\vec{A}_{331M}$
$Z_5(h, \theta)Z_7(\rho, \phi) + Z_6(h, \theta)Z_8(\rho, \phi)$	$\xi_2$	$\vec{A}_{331M}$
$Z_6(h, \theta)Z_7(\rho, \phi) - Z_5(h, \theta)Z_8(\rho, \phi)$	$\xi_1$	$\vec{A}_{331M}$
$Z_7(h, \theta)Z_7(\rho, \phi) + Z_8(h, \theta)Z_8(\rho, \phi)$	$\xi_0$	$W_{331M}$

Table 6.4: Polynomials orthogonal over field and pupil space for trefoil terms  $Z_9(\rho, \phi)$  and  $Z_{10}(\rho, \phi)$ .

Function	Tessieres	Thompson
$Z_1(h, \theta)Z_9(\rho, \phi)$	$\mu_5$	$\vec{C}_{333}$
$Z_1(h, \theta)Z_{10}(\rho, \phi)$	$\mu_6$	$\vec{C}_{333}$
$Z_2(h, \theta)Z_9(\rho, \phi) - Z_3(h, \theta)Z_{10}(\rho, \phi)$	$\mu_4$	$\vec{B}_{333}$
$Z_3(h, \theta)Z_9(\rho, \phi) + Z_2(h, \theta)Z_{10}(\rho, \phi)$	$\mu_3$	$\vec{B}_{333}$
$Z_5(h, \theta)Z_9(\rho, \phi) + Z_6(h, \theta)Z_{10}(\rho, \phi)$	$\mu_2$	$\vec{A}_{333}$
$Z_6(h, \theta)Z_9(\rho, \phi) - Z_5(h, \theta)Z_{10}(\rho, \phi)$	$\mu_1$	$\vec{A}_{333}$
$Z_9(h, \theta)Z_9(\rho, \phi) + Z_{10}(h, \theta)Z_{10}(\rho, \phi)$	$\mu_0$	$W_{333}$

Table 6.5: Polynomials orthogonal over field and pupil space for spherical aberration terms  $Z_{11}(\rho, \phi)$ .

Function	Tessieres	Thompson
$Z_1(h, \theta)Z_{11}(\rho, \phi)$	$\nu_0, \delta_0, \delta_3$	$W_{040}, W_{240M}, B_{240M}$
$Z_2(h, \theta)Z_{11}(\rho, \phi)$	$\delta_2$	$\vec{A}_{240M}$
$Z_3(h, \theta)Z_{11}(\rho, \phi)$	$\delta_1$	$\vec{A}_{240M}$
$Z_4(h, \theta)Z_{11}(\rho, \phi)$	$\nu_0, \delta_0$	$W_{040}, W_{240M}$

Table 6.6: Polynomials orthogonal over field and pupil space for oblique spherical terms  $Z_{12}(\rho, \phi)$  and  $Z_{13}(\rho, \phi)$ .

Function	Tessieres	Thompson
$Z_1(h, \theta)Z_{12}(\rho, \phi)$	$\eta_3$	$\vec{B}_{242}$
$Z_1(h, \theta)Z_{13}(\rho, \phi)$	$\eta_4$	$\vec{B}_{242}$
$Z_2(h, \theta)Z_{12}(\rho, \phi) + Z_3(h, \theta)Z_{13}(\rho, \phi)$	$\eta_2$	$\vec{A}_{242}$
$Z_3(h, \theta)Z_{12}(\rho, \phi) - Z_2(h, \theta)Z_{13}(\rho, \phi)$	$\eta_1$	$\vec{A}_{242}$
$Z_6(h, \theta)Z_{12}(\rho, \phi) + Z_5(h, \theta)Z_{13}(\rho, \phi)$	$\eta_0$	$W_{242}$

Table 6.7: Polynomials orthogonal over field and pupil space for secondary coma terms  $Z_{16}(\rho, \phi)$  and  $Z_{17}(\rho, \phi)$ .

Function	Tessieres	Thompson
$Z_1(h, \theta)Z_{16}(\rho, \phi)$	$\kappa_1$	$\vec{A}_{151}$
$Z_1(h, \theta)Z_{17}(\rho, \phi)$	$\kappa_2$	$\vec{A}_{151}$
$Z_2(h, \theta)Z_{16}(\rho, \phi) + Z_3(h, \theta)Z_{17}(\rho, \phi)$	$\kappa_0$	$W_{151}$

The previous tables only included the terms for the third and fifth order aberrations. If higher order aberrations are included, the list of functions becomes longer.

Double Zernike polynomial expansions in field and pupil space have previously been proposed in the literature. These publications did not propose grouping the double Zernike terms together when they have the same coefficients as done here, so these publications require more coefficients and more functions overall to describe the same wavefront aberrations. In addition, the papers did not contain such a thorough analysis of the aberrations possible in a misaligned system and their dependences on the amount of the misalignment. When the optical systems under consideration do not contain too many degrees of freedom, fewer functions, as listed in the papers, are sufficient. These papers are reviewed in Appendix L.

## 6.6 COMPLETING THE BASIS

One complete basis of orthogonal functions could be one that has each of the Zernike polynomials in field space multiplied by each of the polynomials in pupil space. However, the functions listed in Section 6.5 are expressed in pairs of terms that are combinations of the terms sharing the same radial dependence. This was done in order to match the low order aberrations generated by misalignment. This allows fewer coefficients to describe a misaligned system.

In general, if the field dependence is described by the Zernike polynomials  $Z_k(h, \theta)$  and  $Z_l(h, \theta)$  which have the same radial dependence, but different azimuthal dependence (varying by a dependence on either sine or cosine) and if the pupil dependence is described by the Zernike polynomials  $Z_i(\rho, \phi)$  and  $Z_j(\rho, \phi)$  which have



the same radial dependence, but different azimuthal dependence, then there are four possible combinations to describe the resulting double Zernike pairs. These are  $Z_k(h, \theta)Z_i(\rho, \phi)$ ,  $Z_k(h, \theta)Z_j(\rho, \phi)$ ,  $Z_l(h, \theta)Z_i(\rho, \phi)$  and  $Z_l(h, \theta)Z_j(\rho, \phi)$ . In a system misaligned in one direction, only two of these functions occur and they each have the same coefficient. Instead of having two functions where the coefficients have the same value, and two functions with coefficients of zero, there can be one function with a nonzero coefficient and three functions with coefficients of zero. These functions are:

$$Z_k(h, \theta)Z_i(\rho, \phi) + Z_l(h, \theta)Z_j(\rho, \phi),$$

$$Z_k(h, \theta)Z_i(\rho, \phi) - Z_l(h, \theta)Z_j(\rho, \phi),$$

$$Z_l(h, \theta)Z_i(\rho, \phi) + Z_k(h, \theta)Z_j(\rho, \phi), \text{ and}$$

$$Z_l(h, \theta)Z_i(\rho, \phi) - Z_k(h, \theta)Z_j(\rho, \phi),$$

where ( $i$  and  $j$ ) and ( $k$  and  $l$ ) are pairs of Zernike polynomials that only differ by their sine or cosine dependence. (The orthogonality of these functions is shown in Appendix M.)

All four functions are needed to form a mathematically complete basis. The functions, as they were listed in Tables 6.1–6.7, are not complete. For example, secondary astigmatism which depends quadratically in field in a rotationally symmetric system (from  $W_{242}$ ) is described by the function  $Z_6(h, \theta)Z_{12}(\rho, \phi) + Z_5(h, \theta)Z_{13}(\rho, \phi)$ , but in order to complete the basis of terms with this combination of field and pupil dependence,  $Z_6(h, \theta)Z_{12}(\rho, \phi) - Z_5(h, \theta)Z_{13}(\rho, \phi)$ ,  $Z_5(h, \theta)Z_{12}(\rho, \phi) + Z_6(h, \theta)Z_{13}(\rho, \phi)$  and

$Z_5(h, \theta)Z_{12}(\rho, \phi) - Z_6(h, \theta)Z_{13}(\rho, \phi)$  are needed also. (In reality, it is not possible to write in a list all of the functions to complete the basis because there are infinitely many of them when increasingly higher orders of field and pupil dependence are considered.)

The functions to complete the basis that are not listed in Section 6.5 may still be possible due to misalignments. However, the terms not listed are expected to be very small. By considering increasingly higher orders of field dependencies (more  $\vec{H} \cdot \vec{H}$  terms so the first subscript of the aberration increases by two:  $W_{222}$ ,  $W_{422}$ ,  $W_{622}$  etc. for the case of astigmatism), it is possible to create any of the functions throughout the field in the entire basis for any given aberration.

## 6.7 ORTHOGONAL ABERRATIONS INDUCED BY SHAPE ERRORS

It is not uncommon for a mirror in a telescope to have the wrong surface shape. Figure errors may be left in the optical surface after the fabrication process is completed. Or because large mirrors are often relatively thin (for weight or thermal reasons), they may be able to bend easily. The low order bending modes of a mirror resemble Zernike polynomials. For example, a mirror can easily bend in astigmatism or trefoil, especially if there are three mounts supporting the mirror. These surface shape errors usually increase as the size of the mirror increases because it is harder to make and support these large mirrors, which are usually made as thin as possible to reduce the overall weight and amount of glass necessary.

The largest mirror in a telescope is the primary mirror, which is also usually the entrance pupil of the optical system. This means that the shape error on the

primary mirror affects all of the field angles equally. For example, if there is an astigmatism shape error on the primary, there will be constant astigmatism measured throughout the field and if there is a coma shape error on the primary, there will be constant coma measured throughout the field. (Since constant coma is easily induced by mirror misalignments, this shape error can be accommodated by intentionally misaligning another mirror in the system, as mentioned previously in Chapter 4.)

The next generation of telescopes has increasingly larger secondary or tertiary mirrors. Thus, the bending errors for mirrors not at a pupil in the optical system will become more important. The effect that shape errors on these mirrors have on the aberrations in the optical system needs to be understood as well. Since these mirrors are not at a pupil, finding the effect is not as straightforward as primary mirror bending errors. The aberrations observed are different from the shape error itself because the footprint of the beam for an off-axis field point only covers part of the secondary or tertiary mirror. This is exactly the same consideration as in Chapter 4 for the case of an off-axis telescope viewed as part of a symmetric parent telescope. The aberration that exists in the parent telescope always turns into aberrations that are lower order, as shown in Table 4.5. However, in the case of a bending error on mirrors not at the stop, the footprint of the beam moves around the surface, depending on the field angle, and a simple table with conversion constants (like Table 4.5) can not be written. The scaling factors for the lower order Zernike polynomials terms induced by bending errors include dependence on the field angle ( $H_x$  and  $H_y$ ). These dependencies were derived by Tessieres (2004) in a report,

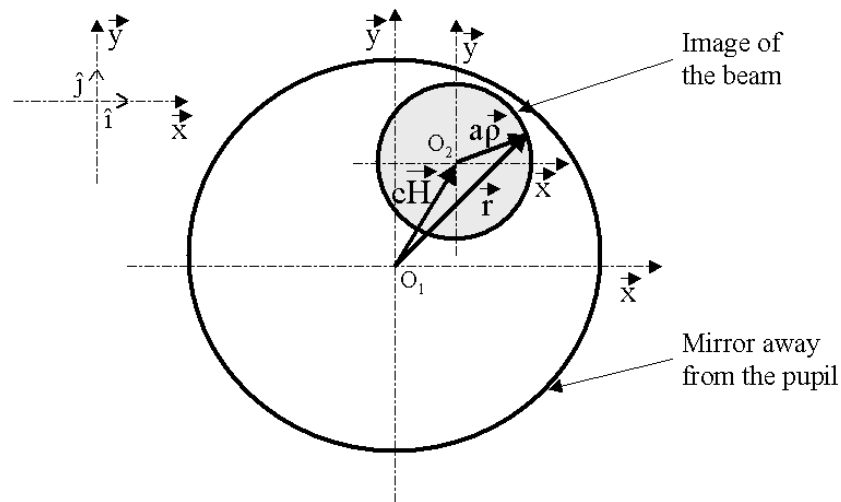


Figure 6.1: Schematic representation for an off-axis beam. The large circle represents a mirror away from the stop of the system. The small circle in gray represents an off-axis beam. Figure courtesy of Tessieres (2004).

where he explains the process in detail. Basically, the Zernike polynomials were written in a vector form  $Z(\vec{r}, \hat{i}, \hat{j})$  and the equation  $\vec{r} = a\vec{\rho} + c\vec{H}$ , which represents the location of a ray hitting a mirror away from the pupil, was substituted for  $\vec{r}$ . The functions were expanded with the new terms and terms containing each Zernike polynomial were combined together. The coefficient  $a$  represents the the ratio of the diameter of the beam footprint on the mirror to the diameter of the mirror while the coefficient  $c$  represents the ratio of displacement on the mirror to displacement in the field. The coefficients  $a$  and  $c$  are constants for each mirror within an optical system and can be calculated by tracing the chief and marginal rays (Tessieres, 2004). For a mirror at an entrance pupil (e.g. a primary mirror) in a telescope,  $a = 1$  and  $c = 0$ . For the other mirrors,  $0 < a < 1$  and  $0 < c < 1$ . This schematic representation for an off-axis beam is shown in Figure 6.1.

Tessieres (2004) lists the Zernike polynomials in vector form and the results for the effect of those bending errors for the first 37 terms. For convenience, the first 11 are repeated in Appendix N. By expressing the field dependence as Zernike polynomials in field space, the orthogonal double Zernike polynomials to describe the aberrations can be written. These are shown for the first 11 bending modes in Table 6.8. The effects of the shape errors on  $Z_2$  and  $Z_3$  were included by Tessieres, but they are not reproduced here because they correspond to distortion, or mapping errors, which are not easily measured by a wavefront sensor. However, bending errors do indeed create distortion. This is expected because the distortion terms ( $Z_2$  and  $Z_3$ ) are very low Zernike polynomials and shape errors for high order Zernike polynomials (for mirrors not at a pupil) produce lower order errors at the exit pupil. The table includes a factor of two due to the wavefront error doubling upon reflection.

Table 6.8 shows that the effect of a shape error of a mirror surface in an optical surface is described well by the double Zernike functions. Each Zernike polynomial shape error transforms into lower order observed Zernike polynomial aberrations that are described cleanly by a numerical constant and the values  $a$  and  $c$ , which are intrinsic to the mirror in the optical system. Since the basis of double Zernike functions is complete, it can describe any aberration resulting from any cause. However, these functions are very convenient because they match the effect of the shape error very well. (In other words, it does not take 10 or 100 functions or more to describe the result.)

Some interesting observations can be made from Table 6.8. The first observation

Table 6.8: Orthogonal polynomials induced by mirror bending. The coefficients  $a$  and  $c$  describe the size and position of the beam on a mirror away from the stop.

Bending mode	Orthogonal functions	Compensated by alignment?
$Z_4$	$2 a^2 Z_1(h, \theta) Z_4(\rho, \phi)$	✓ (Axial)
$Z_5$	$2 a^2 Z_1(h, \theta) Z_5(\rho, \phi)$	
$Z_6$	$2 a^2 Z_1(h, \theta) Z_6(\rho, \phi)$	
$Z_7$	$2 a^3 Z_1(h, \theta) Z_7(\rho, \phi)$ $2 \sqrt{3} a^2 c [Z_2(h, \theta) Z_5(\rho, \phi) - Z_3(h, \theta) Z_6(\rho, \phi)]$ $2 \sqrt{6} a^2 c Z_3(h, \theta) Z_4(\rho, \phi)$	✓ (Tilt/decenter)
$Z_8$	$2 a^3 Z_1(h, \theta) Z_8(\rho, \phi)$ $2 \sqrt{3} a^2 c [Z_3(h, \theta) Z_5(\rho, \phi) + Z_2(h, \theta) Z_6(\rho, \phi)]$ $2 \sqrt{6} a^2 c Z_2(h, \theta) Z_4(\rho, \phi)$	✓ (Tilt/decenter)
$Z_9$	$2 a^3 Z_1(h, \theta) Z_9(\rho, \phi)$ $2 \sqrt{3} a^2 c [Z_2(h, \theta) Z_5(\rho, \phi) + Z_3(h, \theta) Z_6(\rho, \phi)]$	
$Z_{10}$	$2 a^3 Z_1(h, \theta) Z_{10}(\rho, \phi)$ $2 \sqrt{3} a^2 c [Z_3(h, \theta) Z_5(\rho, \phi) - Z_2(h, \theta) Z_6(\rho, \phi)]$	
$Z_{11}$	$2 a^4 Z_1(h, \theta) Z_{11}(\rho, \phi)$ $2 \sqrt{10} a^3 c [Z_3(h, \theta) Z_7(\rho, \phi) + Z_2(h, \theta) Z_8(\rho, \phi)]$ $4 \sqrt{5} a^2 c^2 [Z_5(h, \theta) Z_5(\rho, \phi) + Z_6(h, \theta) Z_6(\rho, \phi)]$ $4 \sqrt{5} a^2 c^2 Z_4(h, \theta) Z_4(\rho, \phi)$ $2 \sqrt{5} a^2 [\sqrt{3}(a^2 - 1) + 2 c^2] Z_1(h, \theta) Z_4(\rho, \phi)$	✓ (Axial)

is one that should be expected: For the case of a mirror at the pupil ( $a = 1$  and  $c = 0$ ), the only effect of the shape error is the same aberration constant throughout the field (dependence on  $Z_1(h, \theta)$ ). The second observation is that a shape error always results in focus ( $Z_4(\rho, \phi)$ ) and/or astigmatism ( $Z_5(\rho, \phi)$  and  $Z_6(\rho, \phi)$ ) errors.

For a coma shape error ( $Z_7$  or  $Z_8$ ) on a mirror not at a pupil, the result is constant coma, linear astigmatism and focal plane tilt. The linear astigmatism has the same functional form ( $Z_2(h, \theta)Z_5(\rho, \phi) - Z_3(h, \theta)Z_6(\rho, \phi)$  or  $Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi)$ ) as the linear astigmatism which can be induced by and *depends linearly* on the misalignment of an element in the optical system. (See the lines with  $\vec{A}_{222}$  in Table 6.2.) In addition, constant coma and focal plane tilt also are field-dependent aberrations that match what is expected in a misaligned system (i.e. depend linearly on the misalignment according to  $\vec{A}_{131_M}$  or  $\vec{A}_{220_M}$ , respectively). Thus, a small coma shape error on a mirror really can be compensated by intentional misalignment. (Large shape errors would require large misalignments to compensate. If the misalignment required to compensate is too large, then the other aberrations that are induced and grow by the cube of the misalignment may become large as well.)

For a trefoil shape error ( $Z_9$  or  $Z_{10}$ ) on a mirror not at a pupil, the result is constant trefoil and linear astigmatism. According to the field center theory, it is possible to induce constant trefoil by misalignment, but it depends on the vector  $\vec{C}_{333}$  (as seen in Table 6.4, which depends on the cube of misalignment). This means only very minute amounts of constant trefoil can be induced by misalignment, while other aberrations come in much more quickly. If constant trefoil is observed in a

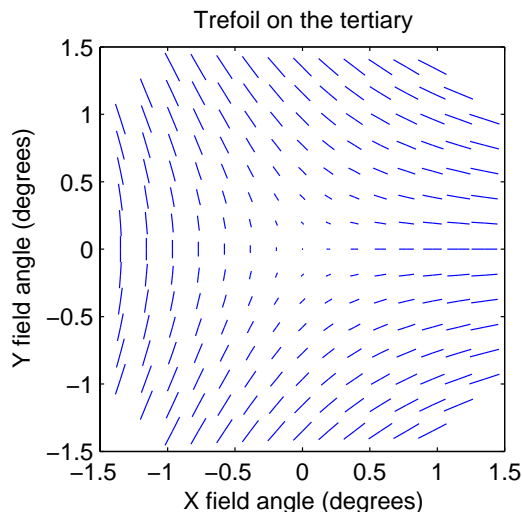


Figure 6.2: Astigmatism line images throughout the field when there is trefoil shape error on a mirror not at the stop. (This particular image, courtesy of Tessieres (2004) shows the effect for trefoil on the tertiary mirror in a 3-mirror telescope.)

system, then it is a pretty safe assumption that some mirror has a trefoil shape error. The linear astigmatism caused by a trefoil shape error ( $Z_2(h, \theta)Z_5(\rho, \phi) + Z_3(h, \theta)Z_6(\rho, \phi)$  or  $Z_3(h, \theta)Z_5(\rho, \phi) - Z_2(h, \theta)Z_6(\rho, \phi)$ ) has the same functional form as the linear astigmatism that can only be caused by the vector  $\vec{C}_{422}$ . Since this is a  $\vec{C}$  vector, it has a very small dependence on the misalignment. The form of the linear astigmatism induced by a trefoil shape error on a mirror has a different characteristic shape than the linear astigmatism caused by misalignment (e.g. Figure 4.5) when plotted as line images throughout the field, as shown in Figure 6.2.

For a spherical aberration shape error on a mirror in the optical system, the result is constant spherical aberration, linear coma, quadratic astigmatism and quadratic focus. The linear coma and the quadratic astigmatism have exactly the same form as those aberrations that exist in rotationally symmetric optical system. (These aberrations depend on the terms  $W_{040}$ ,  $W_{131}$ ,  $W_{222}$  and  $W_{220_M}$  in Tables 6.1–6.5.)



### 6.7.1 COMPENSATION OF SHAPE ERRORS ON ONE MIRROR BY SHAPE ERRORS ON OTHER MIRRORS

If there is astigmatism ( $Z_5$  or  $Z_6$ ) shape error on one of the mirror surfaces in the optical system, then constant astigmatism is the one and only aberration produced by that shape error. Thus, if the right amount of astigmatism shape error on another mirror in the system is induced (i.e. by the actuators supporting the mirror), the error can be completely compensated.

If there is a coma shape error ( $Z_7$  or  $Z_8$ ) on a mirror, then constant coma, linear astigmatism and focal plane tilt are produced in the system. If the constant coma from one mirror's shape error is corrected by inducing a coma shape error on one of the other mirrors or by misalignment, then the overall constant coma in the system will be corrected. However, the linear astigmatism and focal plane tilt induced by these shape errors will not automatically cancel. The form of the linear astigmatism induced by coma ( $Z_2(h, \theta)Z_5(\rho, \phi) - Z_3(h, \theta)Z_6(\rho, \phi)$  and  $Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi)$ ) is the same as that which is induced by misalignment. Therefore, the linear astigmatism can be corrected by rotating a mirror about its coma-free point, and the linear focus error (focal plane tilt) can be corrected by tilting the focal plane.

If there is a trefoil shape error on one mirror, then there is constant trefoil in the system (as measured in the exit pupil). This constant trefoil can be corrected by inducing trefoil on one of the other mirrors. Each trefoil shape error also contributes some amount of linear astigmatism. However, the form of the linear astigmatism induced ( $Z_2(h, \theta)Z_5(\rho, \phi) + Z_3(h, \theta)Z_6(\rho, \phi)$  and  $Z_3(h, \theta)Z_5(\rho, \phi) - Z_2(h, \theta)Z_6(\rho, \phi)$ )

is not possible to produce from misalignments. This error will remain uncorrected in the system once the trefoil is corrected. (It is still helpful to correct the constant trefoil, even if linear astigmatism is remaining because any constant aberration is detrimental for all field angles, while linear aberrations are small throughout the center of the field and are larger only towards the edge of the field.)

## 6.8 SUMMARY OF FIELD DEPENDENT ABERRATIONS THAT DEPEND LINEARLY ON MISALIGNMENT

This section includes a summary of the orthogonal field dependent aberrations that are induced by and *depend linearly* on the misalignment. These are the terms that depend on the  $\vec{A}$  vectors for each of the aberrations ( $\vec{A}_{220_M}$ ,  $\vec{A}_{222}$ ,  $\vec{A}_{131}$  etc.). Table 6.9 was hinted at near the end of Chapter 3. In Chapter 3, the particular field dependencies were not given, as they are in this table. (For example, in Chapter 3, there was no discussion of whether the quadratic coma that is expected in a misaligned system is radially symmetric or not. These are the functions that are useful to fit to the data while performing system alignment.

Table 6.9: Orthogonal field dependent aberrations that depend linearly on the misalignment. The aberrations with check marks in the column labeled “New?” have not been discussed in prior work.

Function	Description	New?
$Z_2(h, \theta)Z_4(\rho, \phi)$	Focal plane tilt	
$Z_3(h, \theta)Z_4(\rho, \phi)$	Focal plane tilt	
$Z_7(h, \theta)Z_4(\rho, \phi)$	Cubic focus	✓
$Z_8(h, \theta)Z_4(\rho, \phi)$	Cubic focus	✓
$Z_2(h, \theta)Z_5(\rho, \phi) - Z_3(h, \theta)Z_6(\rho, \phi)$	Linear astigmatism	
$Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi)$	Linear astigmatism	
$Z_7(h, \theta)Z_5(\rho, \phi) + Z_8(h, \theta)Z_6(\rho, \phi)$	Cubic astigmatism	✓
$Z_8(h, \theta)Z_5(\rho, \phi) - Z_7(h, \theta)Z_6(\rho, \phi)$	Cubic astigmatism	✓
$Z_9(h, \theta)Z_5(\rho, \phi) + Z_{10}(h, \theta)Z_6(\rho, \phi)$	Cubic astigmatism	✓
$Z_{10}(h, \theta)Z_5(\rho, \phi) - Z_9(h, \theta)Z_6(\rho, \phi)$	Cubic astigmatism	✓
$Z_1(h, \theta)Z_7(\rho, \phi)$	Constant coma	
$Z_1(h, \theta)Z_8(\rho, \phi)$	Constant coma	
$Z_4(h, \theta)Z_7(\rho, \phi)$	Quadratic coma	✓
$Z_4(h, \theta)Z_8(\rho, \phi)$	Quadratic coma	✓
$Z_5(h, \theta)Z_7(\rho, \phi) + Z_6(h, \theta)Z_8(\rho, \phi)$	Quadratic coma	✓
$Z_6(h, \theta)Z_7(\rho, \phi) - Z_5(h, \theta)Z_8(\rho, \phi)$	Quadratic coma	✓
$Z_5(h, \theta)Z_9(\rho, \phi) + Z_6(h, \theta)Z_{10}(\rho, \phi)$	Quadratic trefoil	✓
$Z_6(h, \theta)Z_9(\rho, \phi) - Z_5(h, \theta)Z_{10}(\rho, \phi)$	Quadratic trefoil	✓
$Z_2(h, \theta)Z_{11}(\rho, \phi)$	Linear spherical aberration	✓
$Z_3(h, \theta)Z_{11}(\rho, \phi)$	Linear spherical aberration	✓
$Z_2(h, \theta)Z_{12}(\rho, \phi) + Z_3(h, \theta)Z_{13}(\rho, \phi)$	Linear oblique spherical aberration	✓
$Z_3(h, \theta)Z_{12}(\rho, \phi) - Z_2(h, \theta)Z_{13}(\rho, \phi)$	Linear oblique spherical aberration	✓
$Z_1(h, \theta)Z_{16}(\rho, \phi)$	Constant secondary coma	✓
$Z_1(h, \theta)Z_{17}(\rho, \phi)$	Constant secondary coma	✓

## 6.9 NUMERICAL SIMULATIONS OF MISALIGNED SYSTEMS

In this section, some orthogonal field-dependent coefficients are listed for the Hubble Space Telescope with a misaligned secondary mirror. In this example, the secondary mirror is misaligned in the  $\vec{x}$ -direction by both 0.1 mm and 1 mm for comparison, in order to show linearity. Coefficients will be listed for terms that exist for the rotationally symmetric system and for terms that vary linearly with the misalignment. In Tables 6.1–6.7, there are 11 terms that occur in a rotationally symmetric system and these 11 coefficients are shown in Table 6.10. In Table 6.9, there are 24 terms that vary linearly in misalignment. However, for misalignments in only one direction, half of the terms will be zero. Thus, only 12 terms that depend linearly on the mirror misalignment are listed in Table 6.10 to conserve space.

The terms for constant focus, quadratic focus, quartic focus, quadratic astigmatism, linear coma, cubic coma, and constant spherical aberration are expected in a rotationally symmetric system. These terms can be found to be approximately the same across all of the columns. As the misalignment grows, these terms may change from their nominal value because these terms also have contributions that depend on the cube of the misalignment. One can see, by comparing the columns for 0.1 mm and 1 mm, that the values for 1 mm have a larger departure from the nominal value than the 0.1 mm terms. This is to be expected and not simply numerical errors in the calculation of the coefficients. The terms that are zero for the nominal system were those chosen to be included in the table because they depend linearly on the misalignment. These terms include focal plane tilt (linear focus), cubic focus, linear astigmatism, cubic astigmatism, constant coma, quadratic coma,

and linear spherical aberration. The values for these terms may be observed to change by a factor of ten from the 0.1 mm column to the 1 mm column. Whenever a function depends linearly on the misalignment, it also depends on the cube of the misalignment. For small misalignments, the nonlinearity is small, but it becomes larger for larger misalignments as expected. This nonlinearity explains why some of the terms do not differ by exactly a factor of ten. If the two misalignments had been chosen to be 1 mm and 10 mm, there would be even more variation between the columns.

Table 6.10: Orthogonal field dependent aberration coefficients for the Hubble Telescope with misaligned secondary mirror in the  $\vec{x}$ -direction. The coefficient units are waves of 587.6 nm.

Function	Nominal	$\vec{x}$ decenter 0.1 mm	$\vec{x}$ decenter 1 mm
$Z_1(h, \theta)Z_4(\rho, \phi)$	-0.04381	-0.04378	-0.04157
$Z_2(h, \theta)Z_4(\rho, \phi)$	0	-0.001274	-0.01271
$Z_4(h, \theta)Z_4(\rho, \phi)$	-0.04668	-0.04668	-0.04684
$Z_8(h, \theta)Z_4(\rho, \phi)$	0	0.0001717	0.001722
$Z_{10}(h, \theta)Z_4(\rho, \phi)$	0	0.0001121	0.001118
$Z_{11}(h, \theta)Z_4(\rho, \phi)$	-0.03020	-0.03020	-0.03000
$Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi)$	0	-0.002245	-0.02245
$Z_5(h, \theta)Z_5(\rho, \phi) + Z_6(h, \theta)Z_6(\rho, \phi)$	0.5481	0.5481	0.5481
$Z_7(h, \theta)Z_5(\rho, \phi) + Z_8(h, \theta)Z_6(\rho, \phi)$	0	-0.0002945	-0.002945
$Z_9(h, \theta)Z_5(\rho, \phi) + Z_{10}(h, \theta)Z_6(\rho, \phi)$	0	-0.0002727	-0.002727
$Z_{13}(h, \theta)Z_5(\rho, \phi) + Z_{12}(h, \theta)Z_6(\rho, \phi)$	0.1131	0.1131	0.1131
$Z_1(h, \theta)Z_8(\rho, \phi)$	0	-0.05080	-0.5080
$Z_3(h, \theta)Z_7(\rho, \phi) + Z_2(h, \theta)Z_8(\rho, \phi)$	-0.008121	-0.008121	-0.008160
$Z_4(h, \theta)Z_8(\rho, \phi)$	0	-0.003402	-0.03402
$Z_5(h, \theta)Z_7(\rho, \phi) + Z_6(h, \theta)Z_8(\rho, \phi)$	0	$3.946 \times 10^{-5}$	$3.945 \times 10^{-4}$
$Z_7(h, \theta)Z_7(\rho, \phi) + Z_8(h, \theta)Z_8(\rho, \phi)$	-0.004264	-0.004265	-0.004271
$Z_1(h, \theta)Z_{11}(\rho, \phi)$	-0.0005678	-0.0005679	-0.0005801
$Z_2(h, \theta)Z_{11}(\rho, \phi)$	0	$8.795 \times 10^{-6}$	$8.736 \times 10^{-5}$
$Z_4(h, \theta)Z_{11}(\rho, \phi)$	-0.0003897	-0.0003897	-0.0003897
$Z_2(h, \theta)Z_{12}(\rho, \phi) + Z_3(h, \theta)Z_{13}(\rho, \phi)$	0	$2.064 \times 10^{-6}$	$1.988 \times 10^{-5}$
$Z_6(h, \theta)Z_{12}(\rho, \phi) + Z_5(h, \theta)Z_{13}(\rho, \phi)$	-0.0002035	-0.0002032	-0.0002033
$Z_1(h, \theta)Z_{16}(\rho, \phi)$	0	0.0001468	0.001467
$Z_2(h, \theta)Z_{16}(\rho, \phi) + Z_3(h, \theta)Z_{17}(\rho, \phi)$	-0.004073	-0.004073	-0.004082

The aberration coefficients can be recorded for a range of values for the secondary mirror misalignment. Then the coefficients can be sorted according to whether they are expected to be constant, linear or quadratic with the value of the misalignment. The measured coefficients are plotted for the orthogonal double Zernike polynomial functions in Figures 6.3–6.5. Plots are included with both regular (linear) axes and either semi-log or log-log axes.

Figure 6.3 shows the aberration coefficients for a range of misalignments that depend on  $W$  in Tables 6.1–6.7. These coefficients are constant for a range of misalignments, however when the misalignment is large enough, the higher order effect (from the  $\vec{B}$  terms) becomes visible. These lines correspond to terms that have nonzero values in Table 6.10. The highest two lines, constant at 0.55 and at 0.11, correspond to the quadratic astigmatism ( $W_{222}$ ) and quartic astigmatism ( $W_{422}$ ) terms. The coefficients shown are the same on the linear and semi-log plots.

Figure 6.4 shows the aberration coefficients for a range of misalignments that depend on the  $\vec{A}$  vectors in Tables 6.1–6.7. These coefficients depend linearly on the misalignment. This can be seen by the linear plots on the left. The right side of Figure 6.4 shows the absolute value of the coefficients in a log-log plot. (The absolute value is taken because it is impossible to find the logarithm of negative values.) All of the curves have a slope of one on the log-log plot. The curve that is separate from the rest (purple line passing through (1, -0.5) in the linear graph on the left) corresponds to constant coma  $Z_1(h, \theta)Z_8(\rho, \phi)$  which depends on  $\vec{A}_{131}$ .

Figure 6.5 shows the aberration coefficients for a range of misalignments that depend on the  $\vec{B}$  vectors in Tables 6.1–6.7 without any  $W$  terms. These coefficients

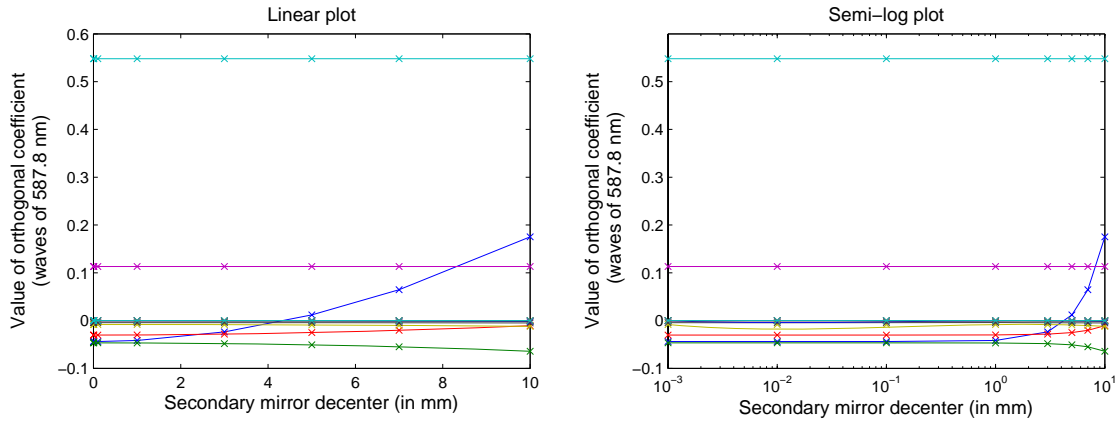


Figure 6.3: The orthogonal field-dependent aberration coefficients that exist in a rotationally symmetric system (due to  $W$  terms) for the HST with a secondary mirror misalignment. These coefficients have primarily a constant dependence on the value of the perturbation. As the perturbation becomes large, the nonlinearity is viewable for some of the terms.

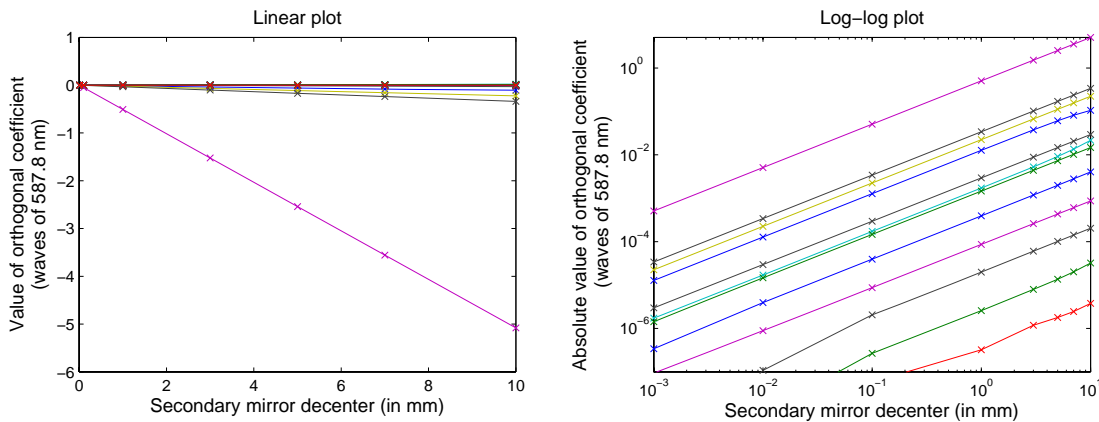


Figure 6.4: The orthogonal field-dependent aberration coefficients that depend linearly on the misalignment (due to  $\vec{A}$  terms), shown for the HST with a secondary mirror misalignment. These coefficients have primarily a linear dependence on the perturbation. As the perturbation becomes large, the nonlinearity is viewable for some of the terms.



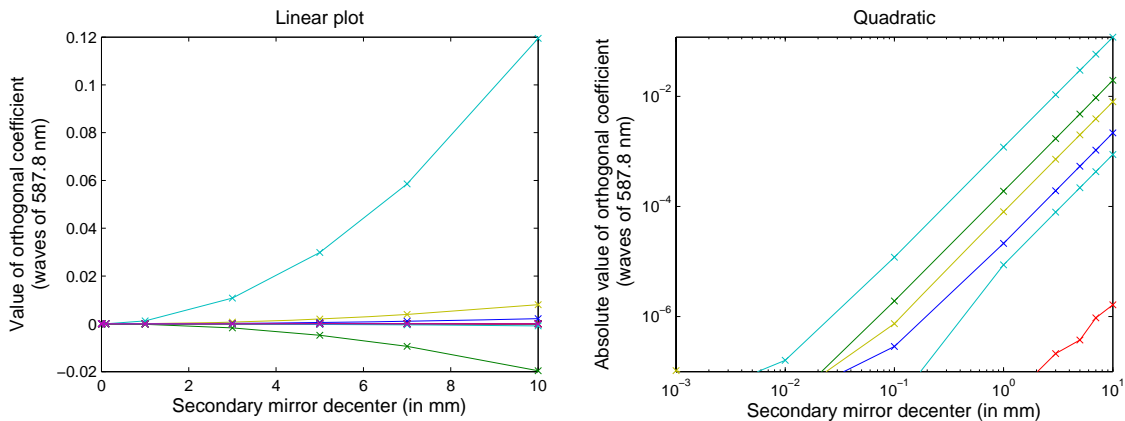


Figure 6.5: The orthogonal field-dependent aberration coefficients that depend quadratically on the misalignment (due to  $\vec{B}$  terms), shown for the HST with a secondary mirror misalignment. These coefficients have primarily a quadratic dependence on the perturbation. As the perturbation becomes large, the nonlinearity is viewable for some of the terms.

depend quadratically on the misalignment. All curves have a slope of two in the log-log plot. The curve with the largest values corresponds to constant astigmatism  $Z_1(h, \theta)Z_6(\rho, \phi)$  which depends on  $\vec{B}_{222}$ .

## 6.10 FINAL COMMENTS ON CHOICE OF BASIS FUNCTIONS

The pupil terms for the functions proposed in Chapter 6 could be annular Zernike polynomials for the case of many of the reflective telescopes presented here. However, since vignetting changes the shape of the pupil for large field angles, annular Zernike polynomials do not exactly describe the pupil shape either. Regular circular Zernike polynomials will work fine in many cases when the obscuration is small. In addition, when Zernike polynomials are sampled numerically, they are not completely orthogonal anyway (see Mahajan's chapter in Malacara (2007)). Also, many telescope systems use rectangular detectors, so circular Zernike polynomials are not

orthogonal over the field. Legendre polynomials are a better choice for a rectangular field.

## CHAPTER 7

LINEAR ANALYSIS OF SYSTEMS WITH MULTIPLE DEGREES OF  
FREEDOM

An optical system can be aligned by measuring wavefront data to infer the current misalignment state of the system. To solve this inverse problem, a least-squares fit of the misalignment to the measured data determines the misalignment. This type of numerical method may be used to align optical systems. If the sensitivity matrix is ill-conditioned, then small errors in the measured data can result in a poor solution of the alignment state of the system. A singular value decomposition (SVD) of the sensitivity matrix can be used to condition the matrix to get better results. This method is described mathematically in Section 7.1.

A singular value decomposition (SVD) of the sensitivity matrix gives an orthogonal set of aberration modes and the corresponding orthogonal combinations of misalignments that cause them. This information can be used to align the optical system. An SVD method was successfully used to align a complex microlithographic system (Chapman & Sweeney, 1998). The Large Synoptic Survey Telescope (LSST) and the James Webb Space Telescope (JWST) plan to use an SVD alignment approach. Not just large budget projects can use SVD for alignment: CODE V implemented a SVD alignment solution in the code for anyone to use. A numeric solution using SVD will always work, if done correctly. However, this method may be difficult to use if one does not understand how noise couples into the system. It

is important to use the SVD to gain physical insight into the system. This helps achieve an understanding on the important degrees of freedom in the system. This physical insight is introduced in Section 7.2, using two different example systems. The first example in Section 7.2.1 uses a simple mechanical rod and spring system to illustrate the meaning of each of the three matrices resulting from the SVD. The second example, in Section 7.2.2, uses a Gregorian telescope to discuss the SVD of a sensitivity matrix in an optical alignment problem. This example serves to bridge the gap between the very simple mechanical rod and spring example and the discussion of a complex alignment problem, the alignment of the four-mirror wide field corrector of the Hobby-Eberly Telescope, discussed in Chapter 8.

## 7.1 INTRODUCTION TO USING SINGULAR VALUE DECOMPOSITIONS IN LINEAR SYSTEMS ANALYSIS

A system of linear equations can be written as

$$\mathbf{A}x = b, \quad (7.1)$$

where  $\mathbf{A}$  has  $m$  rows and  $n$  columns ( $m \times n$  matrix),  $x$  is an  $n \times 1$  vector, and  $b$  is an  $m \times 1$  vector. If  $\mathbf{A}$  and  $b$  are known,  $x$  may be found using the following equation:

$$x = \mathbf{A}^{-1}b. \quad (7.2)$$

If the matrix  $\mathbf{A}$  has more rows than columns, then the system of linear equations is over-determined. (i.e.  $\mathbf{A}$  ( $m \times n$  matrix), where  $m > n$ .) When the system

is over-determined, a least-squares solution may be found for  $x$ . The least-squares solution is the solution for  $x$  that yields the smallest residual error. If the residual vector  $r$  for some  $x$  is

$$r = b - \mathbf{A}x, \quad (7.3)$$

then the vector  $\hat{x}$ , which is an approximation to  $x$  and gives the smallest possible residual, is the least-squares solution:

$$\|r\| = \|b - \mathbf{A}\hat{x}\| \leq \|b - \mathbf{A}x\| \quad \text{for all } x, \quad (7.4)$$

where  $\|r\|$  indicates the norm (or length) of the vector  $r$ . A least-squares solution for an over-constrained system always exists (i.e. some value of  $\|r\|$  has to be smallest), however there may be multiple solutions for  $\hat{x}$  with the same residual. (Thus,  $\hat{x}$  is not unique.) However, out of all of these different possible solutions  $\hat{x}$ , one of them will have the smallest norm  $\|\hat{x}\|$ . This solution is unique and is found by multiplying each side of Equation 7.1 by the transpose of  $\mathbf{A}$  and then solving for  $\hat{x}$ :

$$\mathbf{A}^T \mathbf{A} \hat{x} = \mathbf{A}^T b, \quad (7.5)$$

$$(\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{A}) \hat{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T b, \quad (7.6)$$

$$\hat{x} = \mathbf{A}^+ b, \quad (7.7)$$

where  $\mathbf{A}^+$  is the pseudoinverse of  $\mathbf{A}$ . A matrix  $\mathbf{A}$  is ill-conditioned, if small changes in  $b$  lead to relatively large changes in the solution  $\hat{x}$ . Ill-conditioned matrices are those that have zero or extremely small singular values, which will be explained in

more detail later in this section. SVD is a technique that can be used to find useful solutions for  $x$  when either the  $\mathbf{A}$  or the  $\mathbf{A}^T \mathbf{A}$  matrix is ill-conditioned.

In Chapter 4, Equation 7.1 was expressed as  $Z = \mathbf{M}A$  where  $Z$  is a vector of Zernike coefficients (used in place of  $b$ ) that are caused by a vector  $A$  describing the alignment state of the system (used in place of  $x$ ). The sensitivity of the resulting Zernike coefficients to the system alignment is described by the matrix  $\mathbf{M}$ . This notation will be used for the rest of the chapter.

Any real matrix  $\mathbf{M}$  can be decomposed uniquely into orthogonal matrices  $\mathbf{U}$  and  $\mathbf{V}$  using an SVD:

$$\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^T. \quad (7.8)$$

If  $\mathbf{M}$  is size  $m \times n$ , then  $\mathbf{U}$  is  $m \times m$ ,  $\mathbf{S}$  is  $m \times n$  and  $\mathbf{V}$  is  $n \times n$ .  $\mathbf{S}$  is a diagonal matrix composed of non-negative real values. The matrix is ordered such that the values decrease in size ( $\mathbf{S} = \text{diag}(s_1, s_2, \dots, s_n)$ ). Since there are only  $n$  singular values in  $\mathbf{S}$  (because there is one in each column), one can define an “economy” or “reduced” SVD such that  $\mathbf{S}$  is a square matrix of size  $n \times n$ :

$$\mathbf{S} = \begin{bmatrix} s_1 & 0 & \dots & 0 \\ 0 & s_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & s_n \end{bmatrix}, \quad (7.9)$$

where  $s_1 \geq s_2 \geq \dots \geq s_n$ . In the case of the economy SVD, then  $\mathbf{U}$  is now  $m \times n$  and  $\mathbf{V}$  is  $n \times n$ .

If any values  $s_i$  are zero, then the matrix  $\mathbf{M}$  is singular. If any values  $s_i$  are very small, then the matrix  $\mathbf{M}$  is ill-conditioned. If this is the case, then the inverse

of  $\mathbf{M}$  (or inverse of  $\mathbf{M}^T\mathbf{M}$  in an over-constrained system) does not give workable solutions. In an optical system, this would mean that if there is a small change in the measured Zernike coefficients (in vector  $Z$ ), then large changes in the solution of the required alignment changes in  $A$  may occur.

If  $\mathbf{M}$  is a  $n \times n$  nonsingular matrix, then its inverse may be expressed as

$$\mathbf{M}^{-1} = \mathbf{V}\mathbf{S}^{-1}\mathbf{U}^T, \quad (7.10)$$

where  $\mathbf{S}^{-1} = \text{diag}(\frac{1}{s_1}, \frac{1}{s_2}, \dots, \frac{1}{s_n})$ , or:

$$\mathbf{S}^{-1} = \begin{bmatrix} \frac{1}{s_1} & 0 & \dots & 0 \\ 0 & \frac{1}{s_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{s_n} \end{bmatrix}. \quad (7.11)$$

This matrix shows why small (or zero) singular values cause such problems in computing the inverse of a matrix: the inverse of a very small number is very large, or even infinite if the “small” number is precisely zero. These large numbers lead to problems in Equation 7.11. One way to handle this is to set all the small singular values below some small threshold  $t$  to have no effect in the system. The value of the threshold  $t$  needs to be carefully chosen for each system to achieve the desired result. When this is done, the SVD may be used to approximate the inverse of the matrix  $\mathbf{M}$ :

$$\mathbf{M}^+ = (\mathbf{U}\mathbf{S}\mathbf{V}^T)^{-1} = \mathbf{V}\mathbf{S}_0^{-1}\mathbf{U}^T, \quad (7.12)$$

where the values of the diagonal of  $\mathbf{S}_0^{-1}$  are  $1/s_i$  if  $s_i > t$  or 0 otherwise. This pseudoinverse is known as the Moore-Penrose Inverse (Barrett & Myers, 2004). The

singular value decomposition is an operation standard in every major mathematical software package.

## 7.2 PHYSICAL INSIGHT INTO THE SINGULAR VALUE DECOMPOSITION

The SVD in Section 7.1 is a useful technique for dealing with inverse problems when the system is ill-conditioned. However, using SVD is not just a mathematical trick to approximate the least-squares solution in an ill-conditioned system. It is a technique that can be used to gain insight into the physical system. In an optical system, a singular value decomposition of the sensitivity matrix reveals the orthogonal aberration modes of the system and how they are caused (i.e. what combination of misalignments). An SVD is useful for determining which modes of the system can be seen, measured and corrected. The actual singular values (or the ratios between them) give information about the dynamic range and signal to noise required for the measurement system to be able to see all the different modes. In practice, this insight is needed to carefully choose how many orthogonal control modes of the system should be corrected (or equivalently, how to choose the appropriate threshold  $t$ ).

Before the SVD of the system sensitivity matrix  $\mathbf{M}$  can be performed, the sensitivity matrix of the system itself must be found. The sensitivity matrix can be found by physically perturbing the system or by calculating the expected sensitivities using a model. Using a model to calculate the sensitivity matrix is a practical method because it does not involve actually touching the system and is the method used in this dissertation. Calculating the sensitivities from a model does not intro-



duce errors due to noise, but may not capture subtle differences between the actual system and the model. These differences become very important when controlling modes with small singular values.

Each column of the sensitivity matrix is a record of the effect of one individual degree of freedom. The aberrations due to misalignments of an optical system can be found by direct simulation and collected into a sensitivity matrix for the system. The sensitivity matrix describes the aberrations due to each degree of freedom being incorrect individually.

Each of the three matrices  $\mathbf{U}$ ,  $\mathbf{S}$  and  $\mathbf{V}$  gives useful information about the modes that can be controlled in the system. Overall, they describe the independent “effects” in the system that are most easily produced (largest effect for the smallest cause), and the “causes” of those effects. One way to remember the difference between  $\mathbf{U}$  and  $\mathbf{V}$  is to notice that  $\mathbf{U}$  and “effect” both start with vowels and  $\mathbf{V}$  and “cause” both start with consonants. The modes of the system are cause/effect pairs that are completely orthogonal, or decoupled from every other mode. These control modes of the system are decoupled from each other because the columns of the  $\mathbf{U}$  and  $\mathbf{V}$  matrices are orthogonal. Each column of the  $\mathbf{U}$  and  $\mathbf{V}$  matrices corresponds to a control mode that has some strength given by the appropriate singular value from  $\mathbf{S}$ . For example, the first column of  $\mathbf{U}$  is a combination of effects in the system that is caused by the combination of causes listed in the first column of  $\mathbf{V}$ . This is true in general for all of the column vectors in  $\mathbf{U}$  and  $\mathbf{V}$ :

$$\mathbf{M}v_i = s_i u_i. \quad (7.13)$$

These control modes will be described in the following subsections using first a mechanical system and then an optical system.

### 7.2.1 A MECHANICAL SYSTEM SVD EXAMPLE

In this section, a simple example is used to show the meaning of the SVD. Consider the rod and spring system shown in Figure 7.1 where the rod is supported by two springs (with spring constants  $k_A$  and  $k_B$ ) at its endpoints (positions  $A$  and  $B$ ). In this system, the rod is sitting on top of the springs, but not fixed to the springs. (Thus, if a force pushes down on the rod at point  $B$ , the position (height) of the rod at point  $A$  is unchanged.) For this model system, the sensitivity matrix and its singular value decomposition (SVD) are found for different arrangements of the positions of the two forces.

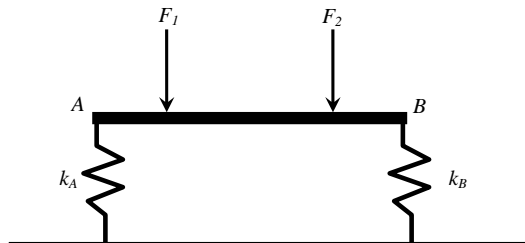


Figure 7.1: Rod and spring mechanical system example. Two forces  $F_1$  and  $F_2$  act on a rod and the overall effect is that points  $A$  and  $B$  on the rod move up and down by some amount depending on the magnitudes and locations of the forces.

A sensitivity matrix for this system can be determined by putting one unit of force into  $F_1$  and examining the effect on the height of the rod in positions  $A$  and  $B$  ( $x_A$  and  $x_B$ ) and then putting one unit of force into  $F_2$  and looking at the effect of  $x_A$  and  $x_B$ . The total motion of each point ( $A$  and  $B$ ) is the sum of the motion

due to each of the forces. The first part is linear algebra:

$$x_A = a F_1 + b F_2 \quad (7.14)$$

$$x_B = c F_1 + d F_2 \quad (7.15)$$

where  $a$ ,  $b$ ,  $c$ , and  $d$  are constants that depend on the spring constants and the locations of the forces. These two equations can be written in matrix form:

$$\begin{pmatrix} x_A \\ x_B \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}. \quad (7.16)$$

The  $abcd$  matrix is the sensitivity matrix, denoted by  $\mathbf{M}$ . Equation 7.16 can also be expressed as  $x = \mathbf{M}F$ , where  $x$  is a vector containing the two displacements  $x_A$  and  $x_B$  and  $F$  is a vector containing the magnitudes of the two forces ( $F_1$  and  $F_2$ ).  $\mathbf{M}$  may be determined from a model of this system, or in real life by doing one perturbation at a time and taking measurements. The units of the sensitivity matrix are determined by the other two vectors. For example, if the displacement  $x$  is in meters (m) and the force  $F$  is in Newtons (N), then the sensitivity matrix is in units of meters per Newton (m/N).

Different configurations of the positions of the two forces will be considered for this rod and spring system example.

### Force configuration I

In this first configuration, the two forces act on the opposite ends of the rod at positions  $A$  and  $B$ , as shown in Figure 7.2. The forces act on the spring that is

directly underneath and assume that the effect on the displacement of the rod at positions  $A$  and  $B$  is completely decoupled.

We have the normal equations for a spring  $F_1 = k_A x_A$  and  $F_2 = k_B x_B$ . Therefore:

$$\begin{pmatrix} x_A \\ x_B \end{pmatrix} = \begin{pmatrix} \frac{1}{k_A} & 0 \\ 0 & \frac{1}{k_B} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}. \quad (7.17)$$

The sensitivity matrix is in units of inverse spring constants. (The spring constant has units of [N/m], so the sensitivity matrix has units of [m/N].) Now for simplification, if  $k_A = k_B = 1$ , then the sensitivity matrix is:

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7.18)$$

In this example, the sensitivity matrix describes the position of the rod relative to its nominal position. That is, with zero force, the displacement is zero. Likewise, in the optical systems considered later, the sensitivity matrix describes the changes in aberrations that result from misalignments. The aberrations will be nonzero when the misalignments are zero due to residual design aberrations, but the sensitivity



Figure 7.2: First configuration of forces examined for the rod and spring system example. Force one pushes on the rod directly over point  $A$  and force two pushes on the rod directly over point  $B$ .

matrix describes the change in the aberration due to the misalignment. For example, here the sensitivity matrix values describe the change in the displacement over the change in the force:

$$m_{ij} = \frac{\Delta x_j}{\Delta F_i}. \quad (7.19)$$

An SVD of the sensitivity matrix in Equation 7.18 results in the following three matrices:

$$\mathbf{U} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7.20)$$

The matrices  $\mathbf{U}$  and  $\mathbf{V}$  are not necessarily unique, but the resulting modes are unique. The difference between the different possible  $\mathbf{U}$  and  $\mathbf{V}$  matrices is that the signs of the modes may be different. For example, another mathematical software package gives the decomposition of the identity matrix as:

$$\mathbf{U} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \quad \mathbf{V} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \quad \mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7.21)$$

These matrices describe the same control modes, with different signs. The signs of the vectors do not matter because they mean the bar moves in the other direction when you pull instead of push. The singular values describing the modes are unchanged, and the modes are listed in the same order, so the  $\mathbf{S}$  matrix is unique.

One should observe that the  $\mathbf{U}$  and  $\mathbf{V}$  matrices are both row and column orthonormal, as they should be. One useful check is that multiplying the matrices together gives you the original sensitivity matrix back. In this case (with all iden-

tity matrices), the result is obvious:

$$\mathbf{USV}^T = \mathbf{M} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7.22)$$

The inverse of the sensitivity matrix is:

$$\mathbf{M}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7.23)$$

The meanings of each of the matrices in the SVD are listed in Table 7.1.

Table 7.1: SVD matrix definitions.

Name	Description	Units
$\mathbf{U}$	This is the results or “effect matrix.” It contains the values of the displacement of the rod at points $A$ and $B$ . In the next optical example (Section 7.2.2), it contains the aberration modes.	The $\mathbf{U}$ matrix is unitless.
$\mathbf{V}$	This is the “cause matrix.” It contains the configuration vectors which describe what needs to be done in the system to get the results in the $\mathbf{U}$ matrix. In this example, it contains the magnitude of the forces, but it will be mirror misalignments in the next example.	The $\mathbf{V}$ matrix is unitless.
$\mathbf{S}$	This matrix contains the singular values, which give the strength of the control modes, in decreasing order along the diagonal of the matrix. The singular values may be extracted and listed in a vector.	The singular values have units of meters/Newton for the rod and spring example.

The example using the first configuration of forces is trivial, but it is still helpful to examine  $\mathbf{U}$ ,  $\mathbf{S}$  and  $\mathbf{V}$  anyway.

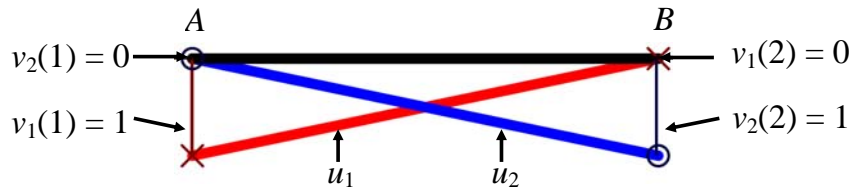


Figure 7.3: Resulting control modes from SVD of sensitivity matrix for the first force configuration of the rod/spring system. The nominal position of the rod is shown in black. The first mode ( $u_1$ ) from the SVD causes the bar to move as shown in red. This is caused by the forces in the vector  $v_1$ , shown by the vertical maroon lines ending in “×.” The second mode ( $u_2$ ) from the SVD causes the bar to move as shown in blue. This is caused by the forces in the vector  $v_2$ , shown by the dark blue vertical lines, ending in circles.

Column one of the  $\mathbf{U}$  matrix ( $u_1$ ) is the “effect” that corresponds to the “cause” in column one of the  $\mathbf{V}$  matrix ( $v_1$ ). The vector  $u_i$  gives a displacement vector containing two elements ( $x_A$  and  $x_B$ ) while  $v_i$  gives the forces ( $F_1$  and  $F_2$ ) that create that displacement.  $u_1$  and  $v_1$  show that pushing down at  $A$  with one unit of force in vector  $F_1$ , results in one unit of displacement at  $A$ . This is mode one:

$$u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (7.24)$$

Mode one is depicted in red in Figure 7.3. The vertical maroon lines ending with an “×” show the position and normalized size of the forces  $F_1$  and  $F_2$ . Since  $v_1(1) = 1$ , there is a one unit force downward at point  $A$  ( $v_1(1) = F_1 = 1$ ) and since  $v_1(2) = F_2 = 0$ , there is no force at point  $B$ .

The other columns  $u_2$  and  $v_2$  show that pushing down at  $B$  with one unit of force in vector  $F_2$ , results in one unit of displacement at  $B$ . This is mode two:

$$u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (7.25)$$

Mode two is depicted in blue in Figure 7.3. The blue bar moves down at  $B$  because of force two ( $v_2(2) = F_2 = 1$ ). The mode two forces are shown by the vertical blue lines ending in a circle.

The singular values in a vector form are

$$s = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (7.26)$$

and in matrix form are

$$\mathbf{S} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7.27)$$

Since the singular values are equal, both modes are equally easy to produce and control in the system, which is obvious from the symmetry in this case.

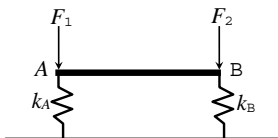
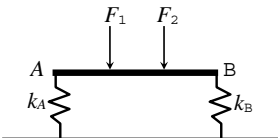
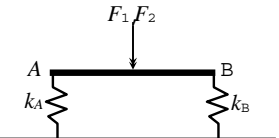
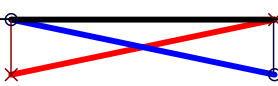
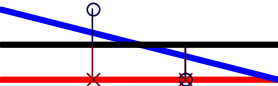
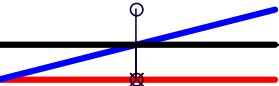
### Other configurations of force locations

Five other configurations of forces, along with the first configuration, are shown in Tables 7.2 and 7.3.

In the second configuration, the two forces act each  $1/3$  the way from the edges of the rod. Each force acts  $2/3$  on the spring that is closest and  $1/3$  on the spring that is farther away. An SVD of this sensitivity matrix results in one up/down mode (mode one shown in red) and one rotational mode (mode two shown in blue). The forces act equally in opposite directions for the rotational mode and equally in the same direction for the up/down mode. The singular value for the up and down mode is larger by a factor of three than the rotational mode, so it takes less total force to produce one normalized unit of effect (motion of the rod).



Table 7.2: Table of rod/spring examples—first set.

	Configuration I	Configuration II	Configuration III
			
$M$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.667 & 0.333 \\ 0.333 & 0.667 \end{pmatrix}$	$\begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$
$M^{-1}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$	Does not exist
$U$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -0.707 & 0.707 \\ -0.707 & -0.707 \end{pmatrix}$	$\begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix}$
$S$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0.333 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
$V$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -0.707 & 0.707 \\ -0.707 & -0.707 \end{pmatrix}$	$\begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix}$
			

In the third configuration, the two forces both act in the exact middle of the rod. Each force acts 1/2 on each spring. An SVD of this sensitivity matrix results again in one rotational mode and one up/down mode. The forces act equally in opposite directions for the rotational mode and equally in the same direction for the up/down mode. However, in this case the singular value for the second (rotational) mode is zero, which means that this mode is impossible to control. This makes sense considering the physical system. Pushing down in exactly the middle of the rod, will definitely move the rod downward (mode one), but it will be impossible to control the rotation of the rod (mode two). Since the singular value is zero, the inverse of the sensitivity matrix can not even be calculated. In this case, the Moore-Penrose inverse can be calculated. The important step in the Moore-Penrose inverse calculation is the calculation of the  $\mathbf{S}_0^{-1}$  matrix. Since  $s_2 = 0$  is below the threshold, and the value of  $1/s_2$  is infinite, the control of this mode should set to zero by setting  $1/s_2 = 0$  in  $\mathbf{S}_0^{-1}$ :

$$\mathbf{M}^+ = \mathbf{V}\mathbf{S}_0^{-1}\mathbf{U}^T \quad (7.28)$$

$$= \begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix}^T \quad (7.29)$$

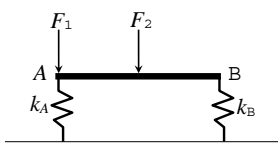
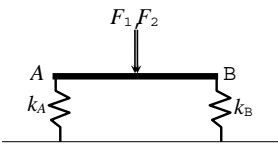
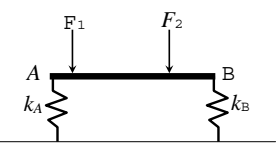
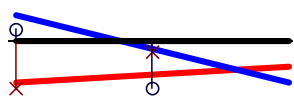
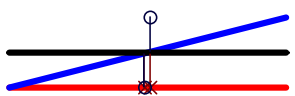
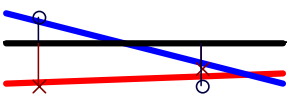
$$= \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}. \quad (7.30)$$

Since the inverse matrix does not have any minus signs, it means that there is no possibility of trying to control the rotational mode (which has  $F_1$  pushing down at the same time  $F_2$  is pulling up), as expected since that mode was set to zero in the  $\mathbf{S}_0^{-1}$  matrix. The total force applied is split equally between  $F_1$  and  $F_2$ . The

columns and the two modes of the pseudoinverse matrix are degenerate.

In the fourth configuration, the first force is on the end of the rod and completely acts on the spring underneath, while a second force pushes in the middle of the rod and acts equally on both springs. An SVD of this sensitivity matrix results in two modes that both rotate and move up and down. The singular values are both nonzero, so the inverse sensitivity matrix can be calculated and both modes can be controlled.

Table 7.3: Table of rod/spring examples—second set.

	Configuration IV	Configuration V	Configuration VI
			
$M$	$\begin{pmatrix} 1 & 0 \\ 0.5 & 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.51 & 0.49 \\ 0.49 & 0.51 \end{pmatrix}$	$\begin{pmatrix} 0.88 & 0.12 \\ 0.29 & 0.71 \end{pmatrix}$
$M^{-1}$	$\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 25.5 & -24.5 \\ 24.5 & 25.5 \end{pmatrix}$	$\begin{pmatrix} 1.203 & -0.203 \\ 0.492 & 1.492 \end{pmatrix}$
$U$	$\begin{pmatrix} -0.851 & 0.526 \\ -0.526 & -0.851 \end{pmatrix}$	$\begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix}$	$\begin{pmatrix} -0.801 & 0.599 \\ -0.599 & 0.801 \end{pmatrix}$
$S$	$\begin{pmatrix} 1.144 & 0 \\ 0 & 0.437 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0.02 \end{pmatrix}$	$\begin{pmatrix} 1.021 & 0 \\ 0 & 0.578 \end{pmatrix}$
$V$	$\begin{pmatrix} -0.973 & 0.23 \\ -0.23 & -0.973 \end{pmatrix}$	$\begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix}$	$\begin{pmatrix} -0.86 & 0.51 \\ -0.51 & -0.86 \end{pmatrix}$
			

In the fifth configuration, the two forces act almost in the middle of the rod, so it is similar to the third configuration. ( $F_1$  is located 49% of the way from point  $A$  to point  $B$ , while  $F_2$  is located 49% of the way from point  $B$  to point  $A$ .) As in the third configuration, an SVD of this sensitivity matrix results in one rotational mode and one up/down mode. The forces still act equally in opposite directions for the rotational mode and equally in the same direction for the up/down mode. However, this time, the second singular value is small, but nonzero. Although the inverse of the sensitivity matrix can be calculated, it could be hard to control the solution, since the second singular value is so small. The inverse of the sensitivity matrix contains fairly large numbers, compared to the other examples. This means that large forces are required, which essentially have a very small effect, to control movement of the rod. The singular value for the second mode for the  $50\times$  smaller than the first singular value. If both the measurement system and the forces have a dynamic range of at least 50, then both modes can be controlled. Without that dynamic range, errors could be amplified by a factor of 50.

If only the up and down mode is controlled, then the pseudoinverse matrix can be calculated, as for Configuration 3. In this case,  $1/s_2$  is set to zero in the  $\mathbf{S}_0^{-1}$  matrix and the inverse is calculated:

$$\mathbf{M}^+ = \mathbf{V}\mathbf{S}_0^{-1}\mathbf{U}^T \quad (7.31)$$

$$= \begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix}^T \quad (7.32)$$

$$= \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}. \quad (7.33)$$

The values in  $\mathbf{M}^+$  in Equation 7.33 are much smaller than those in Table 7.3 for  $\mathbf{M}^{-1}$  in this configuration. Practically, this means much smaller forces are required to create almost the same effect.

The sixth configuration shows a case for somewhat randomly chosen force positions. The modes can still be calculated and the numbers look random, but the modes can still be understood in exactly the same way as all of the other configurations studied. For example, the modes can still be drawn and the columns of  $\mathbf{U}$   $\mathbf{V}$  are still normalized to unity, even though it is not immediately obvious:

$$\sqrt{0.86^2 + 0.51^2} = 1 \quad (7.34)$$

$$\sqrt{0.80^2 + 0.60^2} = 1. \quad (7.35)$$

## 7.2.2 AN SVD EXAMPLE USING A GREGORIAN TELESCOPE

In this section, an SVD analysis is shown for a simple two-mirror Gregorian telescope system. The telescope used will be the axisymmetric parent version of the New Solar Telescope, presented in Table 4.1. The flowchart indicating the steps in this analysis is shown in Figure 7.4.

Finding the sensitivity matrix can be done by physically perturbing the system, or by using an optical model (preferred choice). In this method, the degrees of freedom of the system are perturbed one at a time and the resulting Zernike polynomial aberration coefficients are recorded throughout the field. In this example, the same degrees of freedom are used as in Chapter 4:  $x$  and  $y$  rotations of the secondary mirror about both its center of curvature and the prime focus. Because four degrees

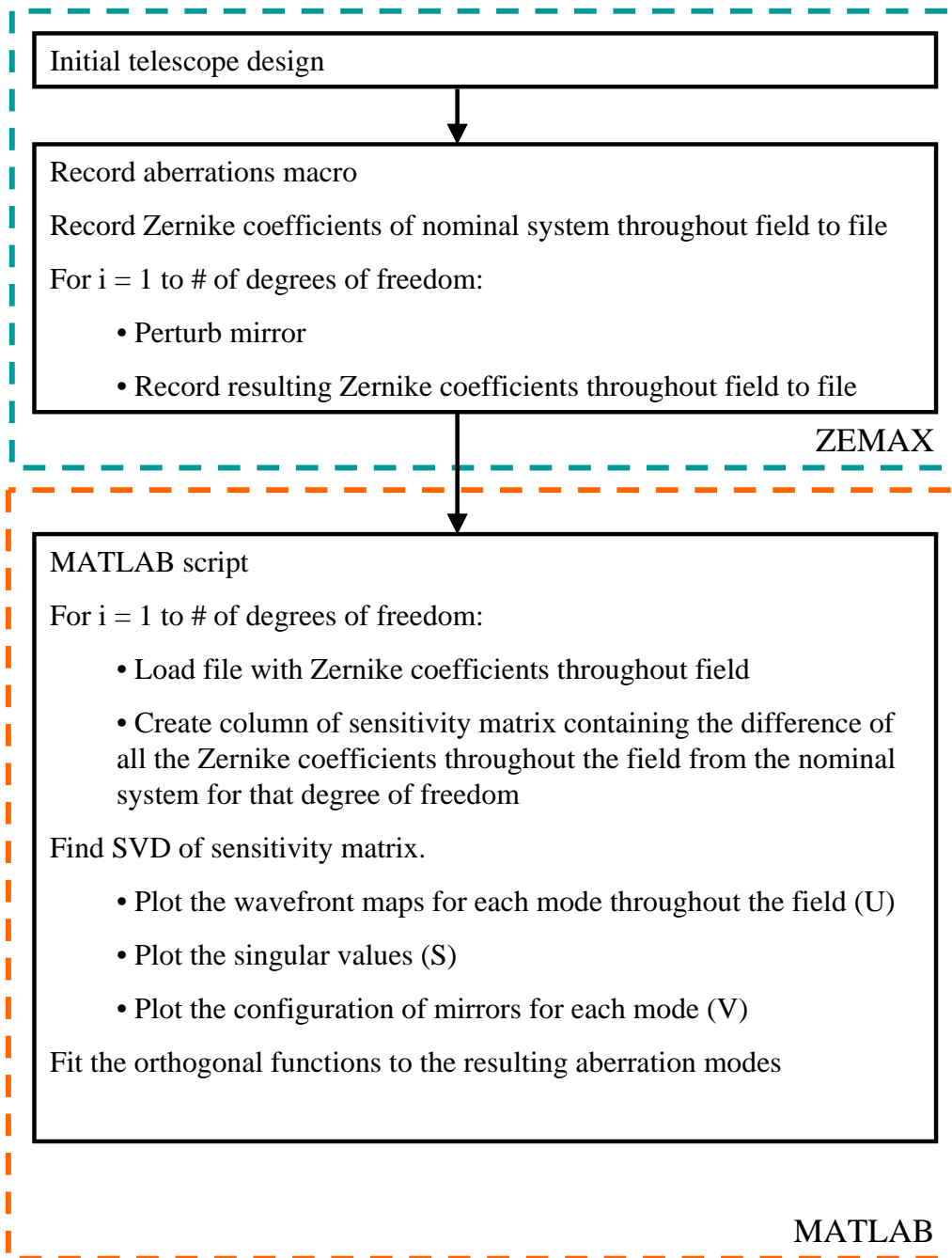


Figure 7.4: Block diagram describing the process of recording the aberrations in a misaligned optical system to find the sensitivity matrix and to analyze the resulting modes from an SVD of this matrix.

of freedom are perturbed, the sensitivity matrix has four columns. Eight Zernike coefficients ( $Z_4$ – $Z_{11}$ ) are recorded throughout the field at a grid of 121 field angles. Thus, the number of rows of the sensitivity matrix is 968.

The units of the singular value matrix  $\mathbf{S}$  are the same as those used for the sensitivity matrix. For example, in the previous example, the units were meters/Newton. In general, the matrix has units of aberration per misalignment. In this example, the unit of the misalignment is rotation angle, so the units of  $\mathbf{S}$  are microns/ $0.1^\circ$ . (The two different degrees of freedom are rotations about two different points: the center of curvature and the prime focus.)

Since the size of  $\mathbf{M}$  is  $968 \times 4$ ,  $\mathbf{U}$  is also  $968 \times 4$ . Mathematically,  $\mathbf{U}$  is size  $968 \times 968$ . However, since there are only four degrees of freedom, there can be at most four nonzero singular values and four resulting control modes. Since the modes that exist in columns 4–958 of the  $\mathbf{U}$  matrix cannot be controlled, there is no reason to record them and the “economy” SVD is used. Therefore,  $\mathbf{U}$  is  $968 \times 4$  and  $\mathbf{S}$  and  $\mathbf{V}$  are both  $4 \times 4$  matrices.

The four singular values correspond to the amount of aberration that can be induced by a normalized amount of each of the four different modes. In this example, since there are equal perturbations in each direction ( $x$  and  $y$ ) for each degree of freedom (prime focus rotations and center of curvature rotations), there should be modes that occur in pairs that are the same, except in the direction of rotation. Each mode in a pair that only differs in the direction should be equally easy to control. Thus, we expect the singular values to come in pairs. This is true and is shown in Figure 7.5.

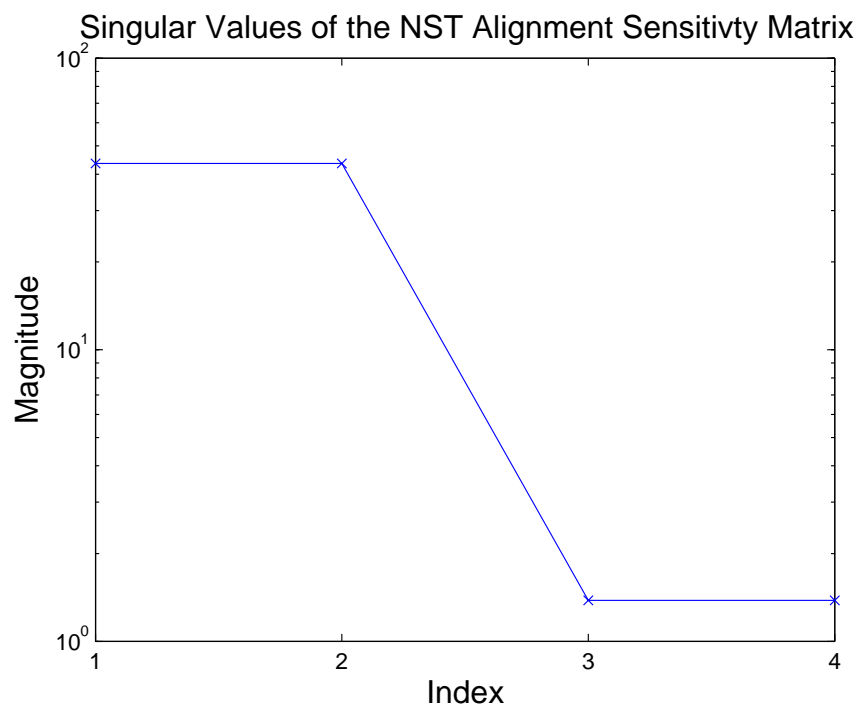


Figure 7.5: Singular values of the axisymmetric Gregorian telescope. The vertical scales depends on the normalization of the  $\mathbf{U}$  and  $\mathbf{V}$  matrices.



The first two singular values are 43.528 and 43.525. The second two singular values are 1.3817 and 1.3815. Thus, the singular values still numerically decrease down the matrix. The ratio of the first two singular values to the second two are  $43.53/1.282 = 33.95$ . This ratio is similar to the ratio of the sensitivities found for constant coma to linear astigmatism in Chapter 4 and shows approximately how much more sensitive the system is to constant coma compared to linear astigmatism.

Each column of the  $\mathbf{U}$  matrix from the SVD contains the eight Zernike polynomial coefficients to describe the wavefront at the grid of 121 field points. The aberrations that we expect to be able to control in the system are constant coma and linear astigmatism. These indeed show up as the control modes of the system in the  $\mathbf{U}$  matrix when the wavefront map is created from the Zernike coefficients. (Instead of plotting the wavefront maps for such a fine sampling of field points, the orthogonal functions of Chapter 6 are fit to the resulting data and then these functions are used to calculate and plot the wavefront maps at a grid of field points seven across.) These modes are shown in Figure 7.6.

The  $\mathbf{V}$  vectors are shown in Figure 7.7. The first two modes consist of center of curvature rotations in the  $x$  and  $y$  directions. This is expected because this degree of freedom was shown in Chapter 4 to produce a large amount of constant coma. The second two modes are composed of prime focus rotations. These are the degrees of freedom shown in Chapter 4 to produce linear astigmatism only with no contribution to constant coma.

As mentioned previously, the orthogonal double Zernike function coefficients can be fit to the aberration data describing the modes in the  $\mathbf{U}$  matrix. When this is

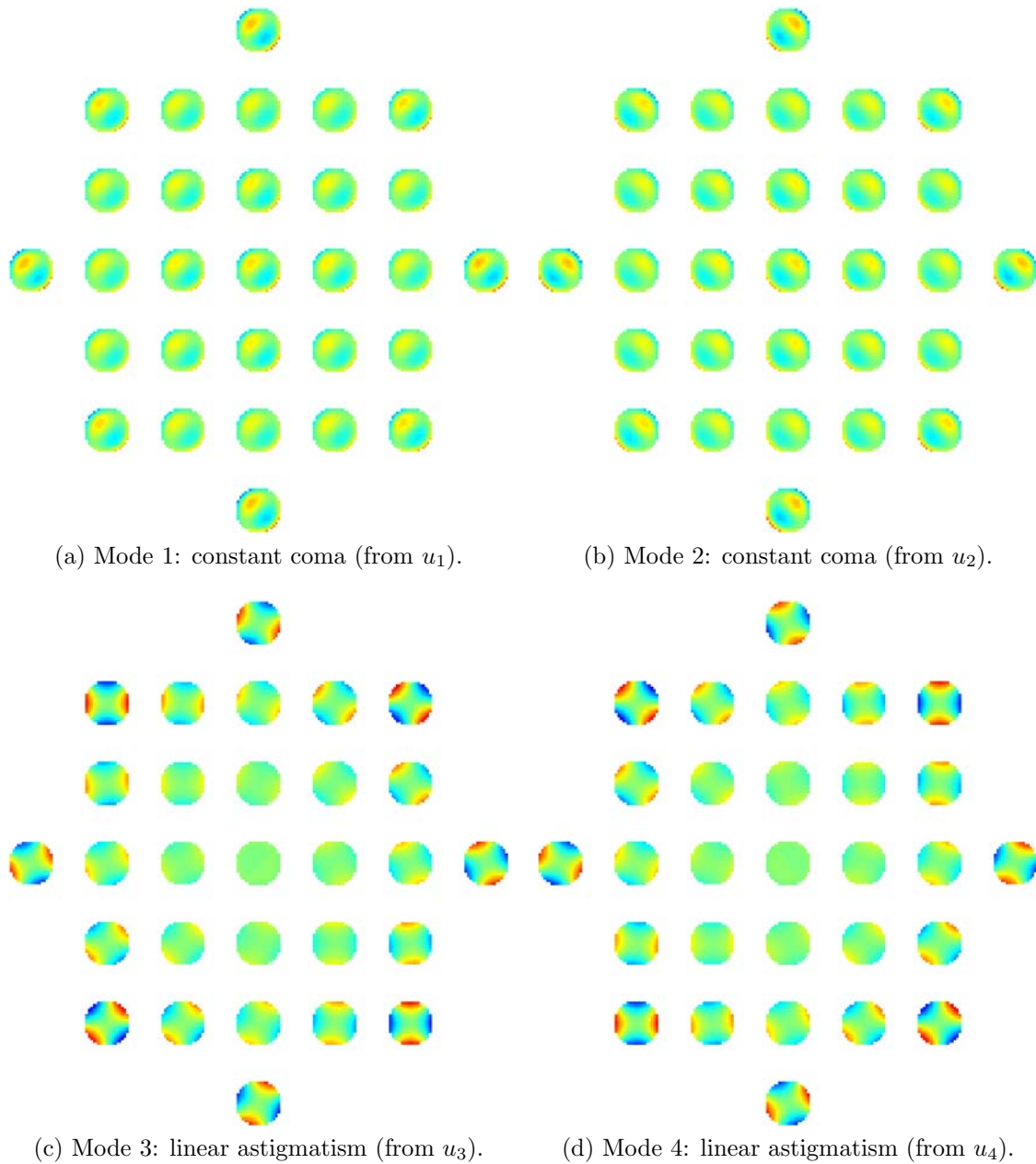


Figure 7.6: Aberration modes from the  $\mathbf{U}$  matrix.

done, most of the resulting functions are very small or zero. The constant coma and linear astigmatism functions have the largest coefficients, as expected. When the modes are expressed in this manner, the vectors from the columns in the  $\mathbf{U}$  and  $\mathbf{V}$

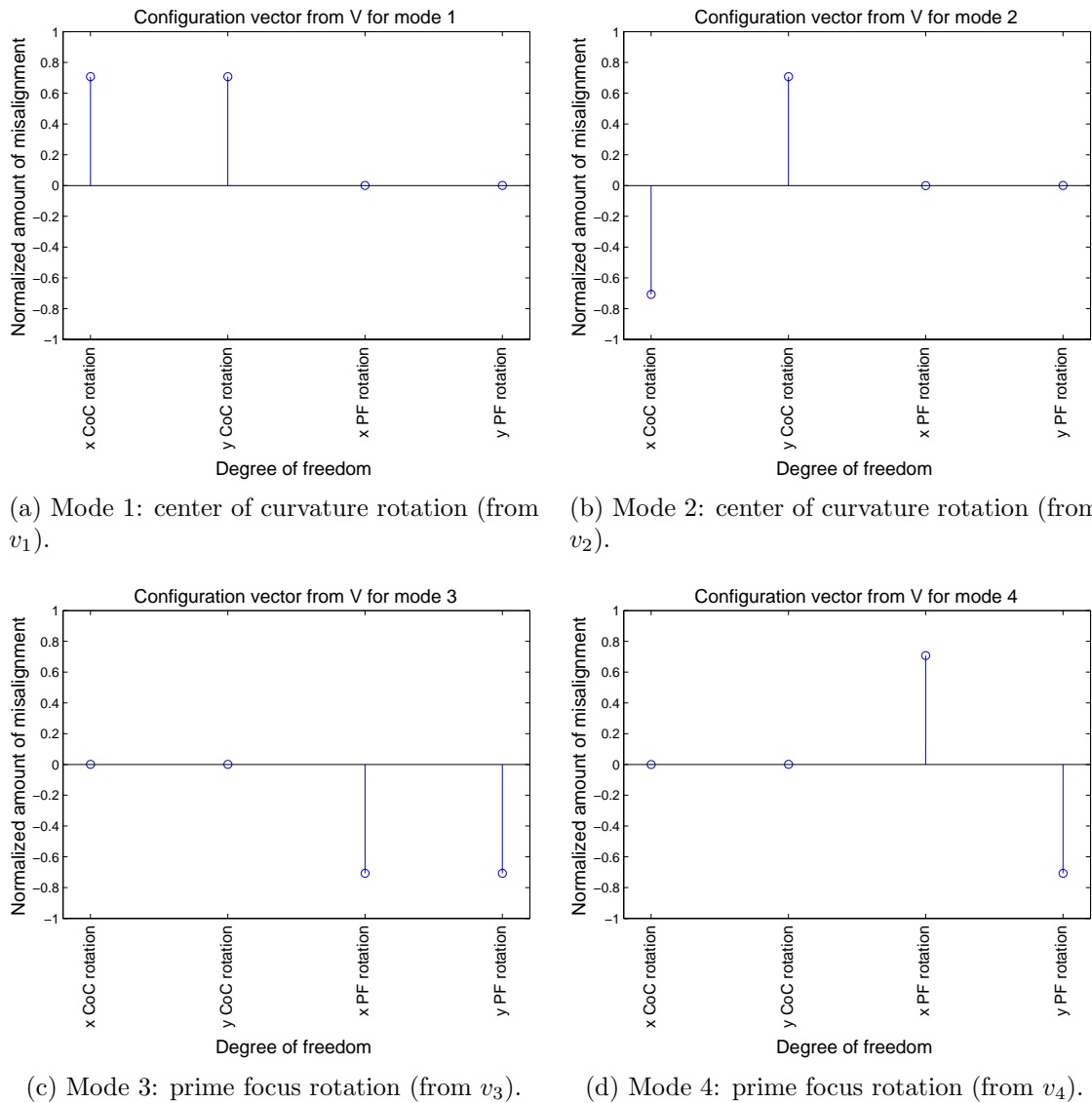


Figure 7.7: Configuration modes from the  $\mathbf{V}$  matrix.

matrices are no longer normalized to one. Alternatively, the orthogonal functions can be fit to the measured data throughout the field and the sensitivity matrix using those coefficients can be formed. All field dependent Zernike polynomials up to  $Z_{13}(h, \theta)$  and pupil dependent Zernike polynomials up to  $Z_{11}(\rho, \phi)$  were included (even the ones not expected, just for simplification in code), and the resulting sen-

sitivity matrix is of size  $104 \times 4$ . An SVD of the matrix results in exactly the same modes in the same order of constant coma and linear astigmatism. However, now the coefficients in the  $\mathbf{U}$  and  $\mathbf{V}$  matrices are normalized. Constant coma, linear astigmatism and the functions with the the next largest coefficients are listed in Table 7.4 for comparison.

Table 7.4: Coefficients for the orthogonal functions from the SVD analysis of the NST sensitivity matrix.

Function	Mode				RSS
	1	2	3	4	
$Z_1(h, \theta)Z_4(\rho, \phi)$	-0.02	0.00	0.01	0.00	0.02
$Z_2(h, \theta)Z_4(\rho, \phi)$	-0.02	0.02	-0.02	0.02	0.04
$Z_3(h, \theta)Z_4(\rho, \phi)$	0.02	0.02	0.02	0.02	0.04
$Z_2(h, \theta)Z_5(\rho, \phi) - Z_3(h, \theta)Z_6(\rho, \phi)$	0.02	0.02	<b>0.70</b>	<b>0.70</b>	<b>0.98</b>
$Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi)$	-0.02	0.02	<b>-0.70</b>	<b>0.70</b>	<b>0.98</b>
$Z_7(h, \theta)Z_5(\rho, \phi) + Z_8(h, \theta)Z_6(\rho, \phi)$	0.00	0.00	-0.09	0.09	0.13
$Z_8(h, \theta)Z_5(\rho, \phi) - Z_7(h, \theta)Z_6(\rho, \phi)$	0.00	0.00	0.09	0.09	0.13
$Z_9(h, \theta)Z_5(\rho, \phi) - Z_{10}(h, \theta)Z_6(\rho, \phi)$	0.00	0.00	-0.08	0.08	0.12
$Z_{10}(h, \theta)Z_5(\rho, \phi) + Z_9(h, \theta)Z_6(\rho, \phi)$	0.00	0.00	-0.08	-0.08	0.12
$Z_1(h, \theta)Z_7(\rho, \phi)$	<b>-0.70</b>	<b>-0.70</b>	0.02	0.02	<b>0.99</b>
$Z_1(h, \theta)Z_8(\rho, \phi)$	<b>0.70</b>	<b>-0.70</b>	-0.02	0.02	<b>0.99</b>
$Z_4(h, \theta)Z_7(\rho, \phi)$	-0.05	-0.05	0.00	0.00	0.07
$Z_4(h, \theta)Z_8(\rho, \phi)$	0.05	-0.05	0.00	0.00	0.07
$Z_{11}(h, \theta)Z_7(\rho, \phi)$	-0.09	-0.09	0.00	0.00	0.13
$Z_{11}(h, \theta)Z_8(\rho, \phi)$	0.09	-0.09	0.00	0.00	0.13

Other aberrations result from the configuration vectors in the  $\mathbf{V}$  matrices in addition to constant coma and linear astigmatism, but they are not controllable. This is indicated by the value of the root sum square (RSS) across all the modes. Constant coma and linear astigmatism which are controllable, as shown by the value of RSS being approximately one. Even with no prior understanding of the aberrations caused by misalignments, this can be determined by the singular value

decomposition.

The result of this study is that since the only aberrations that are controllable are constant coma and linear astigmatism, only these four Zernike coefficients actually need to be measured when performing alignment of this system. Also, since the field dependence of these aberrations is low order (constant and linear for coma and astigmatism, respectively), then only a few points in the field need to be measured. As few as three fields are needed to determine a constant field dependence and a linear field dependence in each of the  $x$  and  $y$  directions.

Therefore, this system can be described by a sensitivity matrix  $\mathbf{M}$  with only 12 rows, not the 968 or 104 used for the example.

## CHAPTER 8

SVD OF THE INFLUENCE MATRIX FOR THE HOBBY-EBERLY  
TELESCOPE WIDE FIELD CORRECTOR

This chapter investigates the orthogonal control modes resulting from a singular value decomposition (SVD) of the influence matrix of the system. The numerical analysis of the HET WFC optical design, given in Chapter 2, is performed in this chapter using ZEMAX and MATLAB. Matrices resulting from the SVD of the influence matrix describe the orthogonal set of aberrations that may be caused by misalignment of the WFC, the combination of mirror misalignments that cause these aberrations and the sensitivities that relate the two effects. These control modes will be examined in detail in this chapter.

## 8.1 OVERVIEW OF METHOD

The influence matrix is a record of the Zernike polynomial coefficients describing the wavefront aberrations (in the pupil) throughout the field for all of the degrees of freedom. Each column of this matrix is an influence function or vector that describes the aberrations throughout the field for one misalignment degree of freedom. Thus, the number of columns in the matrix is equal to the number of degrees of freedom in the system. The number of rows of the influence matrix depends on how many Zernike polynomials are used to describe the wavefront and at how many field points. For the purposes of an investigative study, it is worthwhile to measure “too many”

aberrations on a finely sampled grid of field points. This ensures that the most sensitive field dependent aberrations are discovered in the control modes found by SVD of the influence matrix.

Calculating the control modes can be done in one of two ways. In the first method, an SVD of the influence matrix is calculated after it is expressed in the basis of orthogonal double Zernike polynomial coefficients. In the second method, the SVD of the influence matrix is calculated and then the resulting control modes are put in the basis of orthogonal field-dependent functions. Once the control modes are found in terms of the basis of the double Zernike polynomial coefficients, then the most sensitive double Zernike functions can be chosen to be included in the sensitivity matrix  $\mathbf{M}$ . The process of finding the sensitivity matrix is shown in Figure 8.1. The values that describe the sizes of the matrices in Figure 8.1 are listed in Table 8.1.

Table 8.1: Values describing the sizes of the influence, control and sensitivity matrices.

Variable	Description
$n$	number of field points
$z$	index of largest Zernike coefficient measured
$d$	number of alignment degrees of freedom
$f_I$	number of orthogonal functions calculated from data
$f_M$	number of orthogonal functions used in the sensitivity matrix $\mathbf{M}$
	$d \leq f_M \leq f_I$

It is important that the alignment sensitivity matrix is expressed in the basis of orthogonal double Zernike functions because the correct functions can be chosen that have a linear dependence on the misalignment. It is not helpful to include

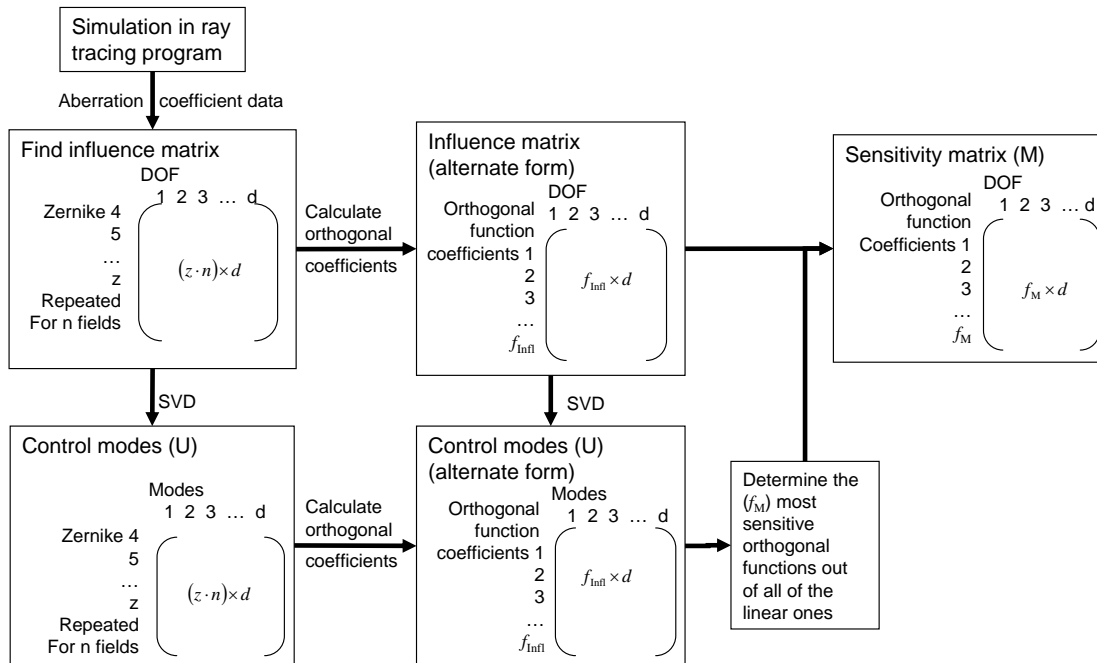


Figure 8.1: Process of calculating the alignment sensitivity matrix.

effects, like constant astigmatism, that depend on the square of the misalignment in the sensitivity matrix because the mathematical calculation of the misalignments from the measured aberrations assumes linearity. Note that this alignment process will still work for a non-linear system, but more iterations are required with the sensitivities of each non-linear effect re-calculated at each iteration if the non-linear effects are important. This method for choosing the aberrations for the alignment degrees of freedom was not used in Chapter 4, since there were so few degrees of freedom and the resulting aberrations that depended linearly on the misalignment were already understood. The expected aberrations were constant coma and linear astigmatism and these were just put directly into a sensitivity matrix. Thus, when there are many degrees of freedom and many field-dependent aberrations, it is important to separate the aberrations that depend linearly on the misalignment and



those that do not.

Once the important field-dependent aberration modes are known, only these aberrations need to be measured in the as-built system to generate the maximum likelihood estimate of the misalignment state of the system. (Therefore the number of Zernike coefficients  $z$  and the number of field points measured  $d$  only need to be large enough to calculate the coefficients of orthogonal double Zernike field dependent functions that are used in  $\mathbf{M}$  and are not necessarily as large as when the original analysis was done to choose what should go in  $\mathbf{M}$ .) The maximum likelihood estimate comes from a least-squares fit of the alignment degrees of freedom to the aberration vector  $Z$ . If the matrix  $\mathbf{M}$  has small singular values, the inverse of  $\mathbf{M}$  may require conditioning, as described in Chapter 7. The conditioning may be done using SVD and in this case, an appropriate threshold needs to be found, depending of the signal to noise of the system that measures that wavefront data. This process is shown in Figure 8.2.

In general, a least-squares fit of the alignment degrees of freedom to the aberration vector  $Z$  results in a vector of values for adjusting the alignment of every degree of freedom. Since some of the degrees of freedom are not as helpful in actually removing the measured aberrations, they may only require very small adjustments that do not significantly improve the system. There is a process, described in detail by Chapman & Sweeney (1998), that may be used to choose the optimum degrees of freedom to use as compensators.

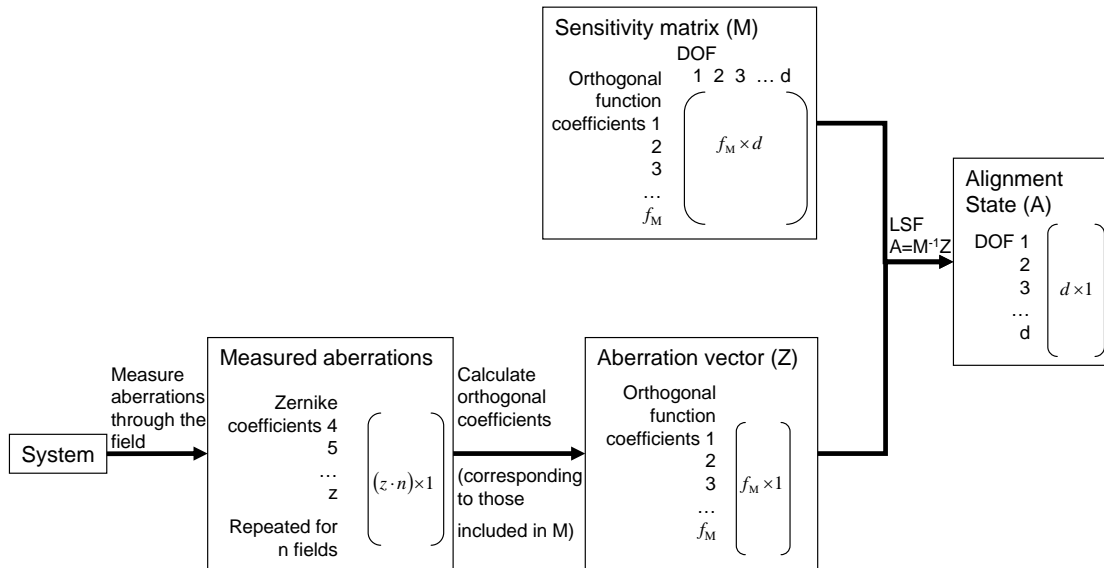


Figure 8.2: Process of calculating the alignment state.

## 8.2 INFLUENCE MATRIX

To do the numerical analysis, the Zernike coefficients across the field resulting from perturbing each of the individual degrees of freedom of the four mirrors were recorded as text files using a ZEMAX macro. A MATLAB script was written to load these files and generate the influence matrix. MATLAB was also used to find the singular value decomposition of the matrix and then examine the results. (This is the same process as shown in Table 7.4.)

### 8.2.1 PERTURBATIONS

Each of the 25 columns in the sensitivity matrix corresponds to a different misalignment. (The primary mirror position is not perturbed since it only affects pointing.) The 25 degrees of freedom, which are listed in Table 8.2, include perturbations of

Table 8.2: Degrees of freedom for the HET WFC alignment sensitivity matrix.

Number	Degree of freedom	Units
20	Five degrees of freedom for each of the four mirrors in the corrector: $x$ and $y$ tilt $x$ and $y$ decenter axial position error	0.5 arcseconds 1 $\mu\text{m}$ 1 $\mu\text{m}$
2	Rigid body motion of the entire WFC, including the focal plane (controlled by the hexapod in the actual system) $x$ and $y$ tilt	0.01°
3	Focal plane rigid body motion $x$ and $y$ tilt axial position (focus) adjustment	0.01° 0.1 mm
25	Total	

all four mirrors in the corrector, and rigid body motions of the entire corrector and of the focal plane.

This analysis assumes misalignment perturbations only (e.g. no mirror bending or shape errors). All of the mirrors are assumed to be fabricated correctly with the appropriate radius of curvature and conic constant or asphericity. In general, additional columns could be included in the sensitivity matrix to find and correct these shape errors, as well, if active control of the mirror surface is available.

### Magnitudes of perturbations for the degrees of freedom

As discussed in Chapter 7, the units of the singular value matrix  $\mathbf{S}$  are the same as those used for the sensitivity matrix. In general, the matrix has units of aberration per misalignment. Since the misalignment degrees of freedom are a combination of distances and angles, the sensitivity matrix and singular values matrix for HET

is neither waves/ $\mu\text{m}$  or waves/arcsecond, but can be considered as waves/unit perturbation, where the unit perturbation depends on the amount of the misalignment simulated. Singular value decompositions of matrices with mixed units are discussed by Chapman & Sweeney (1998) and Descour *et al.* (2000). One way to deal with this is to assume the units are waves (or microns) per unit perturbation. If the perturbation is an angle, the 1 unit perturbation is an angular unit (degrees, arcminutes, arcseconds or radians, as desired). If the perturbation is a decenter, then the 1 unit perturbation is a distance unit (meters or microns). For this analysis, the units of angular and distance perturbations are found in Table 8.2.

The particular results of the SVD depend on the magnitudes of the perturbations in ZEMAX. For this analysis, the perturbations of the tilt and decenters of the mirrors were chosen to be very small: the tilts were 0.5 arcseconds about the vertex and the decenters were 1  $\mu\text{m}$ . This was done so that the resulting coefficients would be exceedingly small for the double Zernike functions that depend on the square or cube of the misalignment. Therefore, when the SVD of the influence matrix is calculated, it would not appear like these modes are controllable. The axial positions of each of the four mirrors had the same perturbations as the vertex decenters in  $x$  or  $y$ ; that is, 1  $\mu\text{m}$ .

If the tilt of the entire WFC and the tilt of the focal plane used exactly the same small perturbations as for the mirrors, there would be a very small effect on the aberration coefficients measured throughout the field. If we take the limit, and assume that the coefficients are precisely zero, then this degree of freedom has no effect in the system in the SVD and the singular value is zero. However, the

system, in actuality, is sensitive to these degrees of freedom, which is why they can be used as compensators. In addition, these degrees of freedom result in, almost purely, the same field-dependent aberrations that are largest for the other degrees of freedom, so they are very good compensators. The tilt of the entire WFC results in constant coma, while focal plane tilt, not surprisingly, results in focal plane tilt. In order for these aberrations to show up in the lower order (most-sensitive) control modes, the aberration coefficients in the influence matrix must be recorded for larger perturbations. The entire WFC and individual focal plane tilt degrees of freedom were each perturbed by  $0.1^\circ$ .

For the same reason, the axial position of the focal plane was perturbed with a larger value than for the decenters. Perturbing the position by 1 mm resulted in constant defocus showing up as mode 3 in the system. For other cases investigated and not included here, with larger axial perturbations of the focal plane (and/or smaller tilts of the WFC), constant defocus was the first mode from the SVD.

### 8.2.2 ABERRATIONS IN TERMS OF ZERNIKE COEFFICIENTS

The rows of the sensitivity matrix contain the Zernike coefficients for all of the field points under consideration. For this analysis, the  $z = 12$  Zernike polynomial coefficients considered (to describe the pupil dependence) are those from Chapter 6. These include focus ( $Z_4$ ), astigmatism ( $Z_5$  and  $Z_6$ ), coma ( $Z_7$  and  $Z_8$ ), trefoil ( $Z_9$  and  $Z_{10}$ ), spherical aberration ( $Z_{11}$ ), secondary astigmatism ( $Z_{12}$  and  $Z_{13}$ ), and secondary coma ( $Z_{16}$  and  $Z_{17}$ ). Each of these aberrations in the exit pupil are described by Zernike polynomials that are up to quartic in field dependence. Thus,

a  $7 \times 7$  grid of field points is sufficient to capture these field dependences. Therefore, the number of rows in the influence matrix is equal to  $n \times z = 49 \text{ fields} \times 12 \text{ coefficients} = 588$  rows.

Measuring all 12 of these aberrations in the pupil using a  $7 \times 7$  grid ensures that all the coefficients for the 24 aberrations in Table 6.9 that depend linearly on the misalignment can be found. In addition, 11 aberrations in rotationally symmetric systems (from terms with  $W$  in Tables 6.1–6.7) can be found. These 35 aberrations are more than sufficient to calculate the major effects from 25 degrees of freedom.

### 8.3 SINGULAR VALUE DECOMPOSITION OF THE ALIGNMENT SENSITIVITY MATRIX

This section investigates the contents of the three matrices ( $\mathbf{U}$ ,  $\mathbf{S}$ , and  $\mathbf{V}$ ) resulting from the SVD of the influence function in the basis of the orthogonal double Zernike coefficients. The vectors from the matrices for each mode are plotted in a way to show what the numbers mean. All of the plots for the control modes are provided in Appendix O.

#### 8.3.1 U MATRIX

Each of the columns in the  $\mathbf{U}$  matrix ( $588 \times 25$  when using the economy SVD) is a singular vector that lists the Zernike coefficients across each of the 49 field points. Each of these resulting particular combinations of field-dependent aberrations is an aberration mode. Since  $\mathbf{U}$  is a column orthogonal matrix, the aberration modes are orthogonal. They also form a complete basis set. That is, combinations of

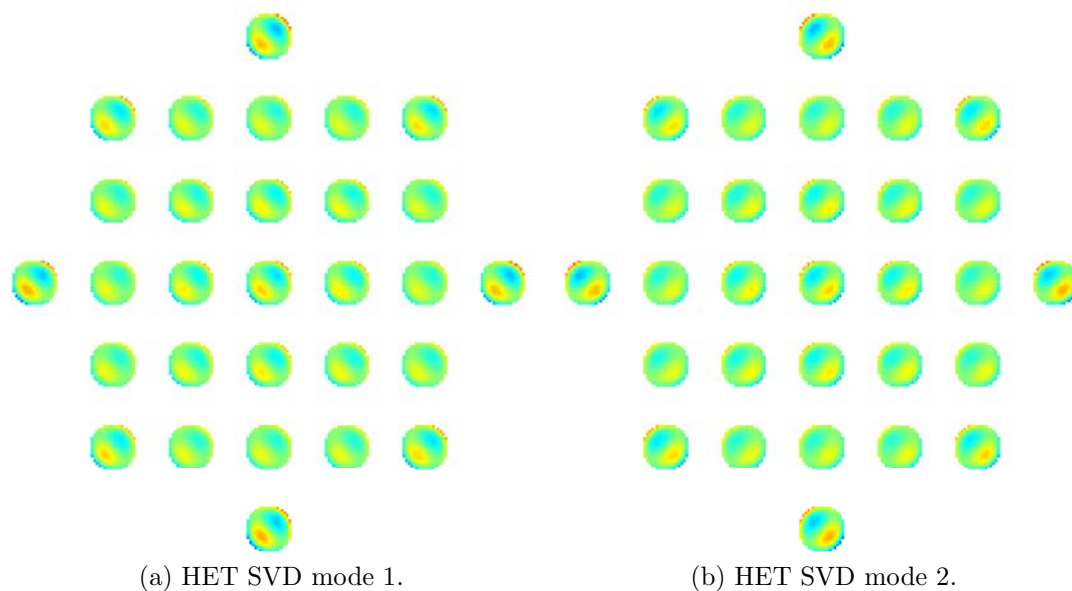
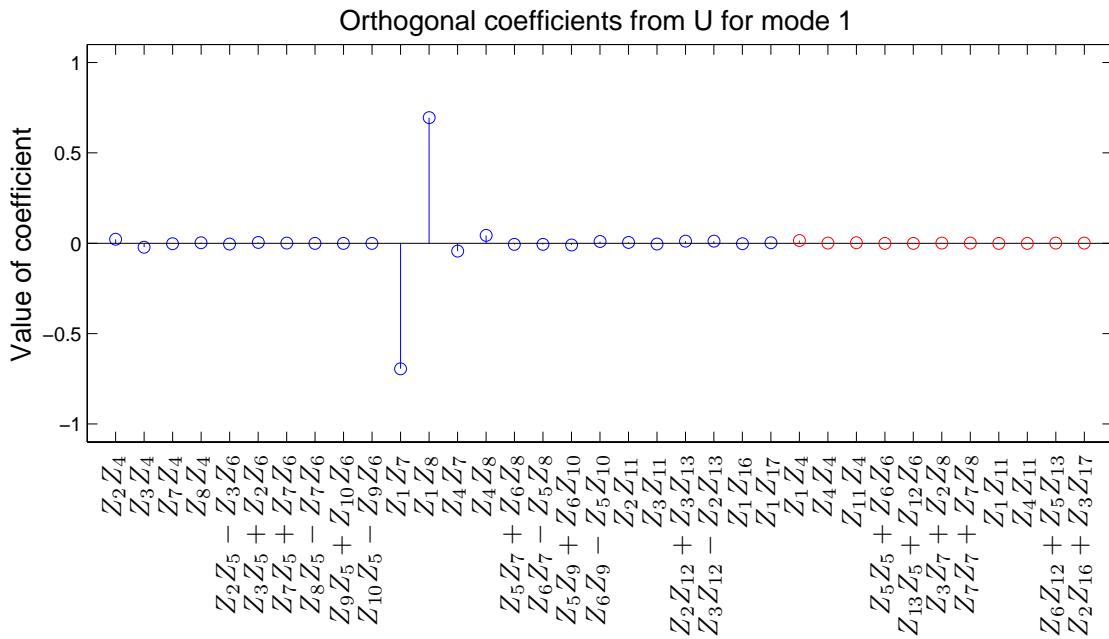


Figure 8.3: Wavefront maps across the field showing constant coma from the first and second control modes.

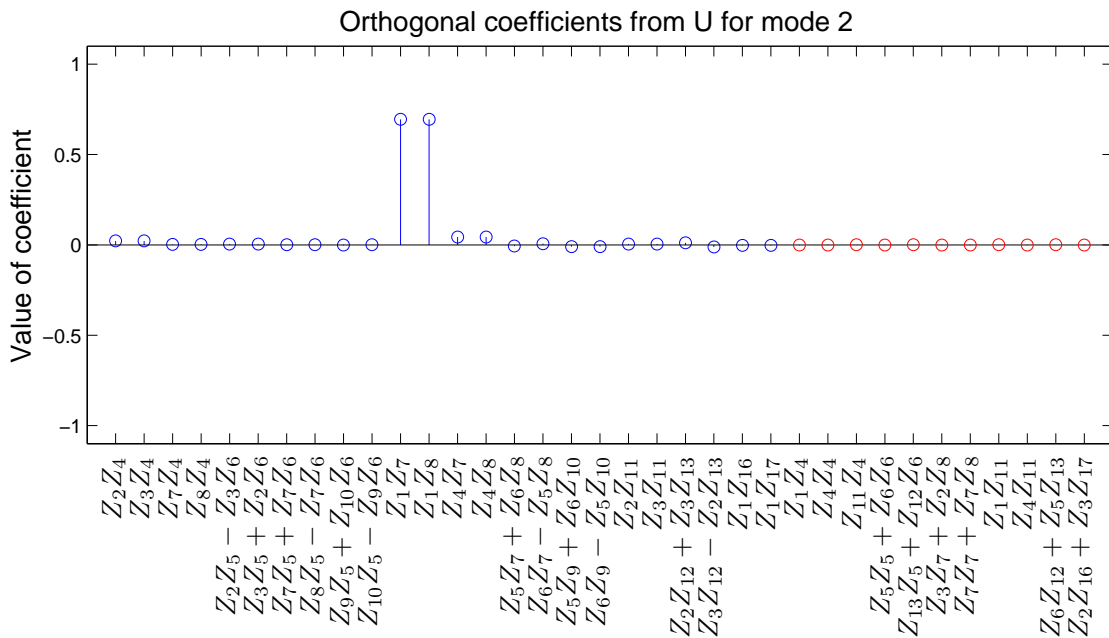
the aberration modes describe all the possible aberrations that can result from misalignments of the WFC.

The first two modes are the easiest ones to control and they each correspond to constant coma. These aberrations are shown in Figure 8.3. The coefficients for the orthogonal double Zernike functions from control modes 1 and 2 are shown in Figure 8.4, and the result shows clearly that the modes are composed of constant coma:  $Z_1(h, \theta)Z_7(\rho, \phi)$  and  $Z_1(h, \theta)Z_8(\rho, \phi)$ . The color blue in the plots of the double Zernike coefficients in Figure 8.4 indicates that a mode occurs in a non-rotationally symmetric system, while the color red indicates a function that occurs in a rotationally symmetric system.

The third mode, shown in Figure 8.5, corresponds to defocus and is the next most easy to control. The coefficients for the orthogonal functions for control mode 3 are



(a) HET SVD mode 1.



(b) HET SVD mode 2.

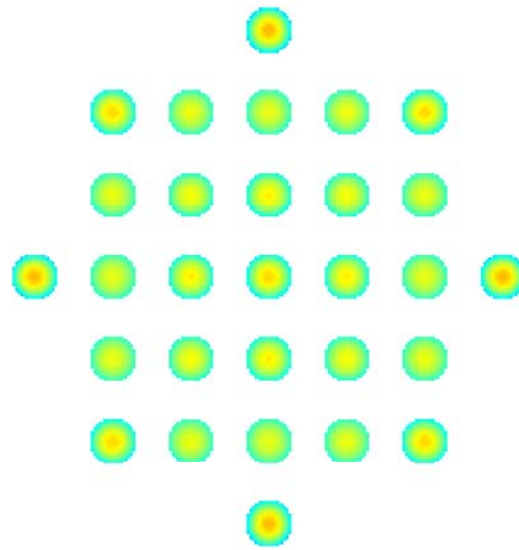
Figure 8.4: Coefficients showing constant coma from the first and second control modes.



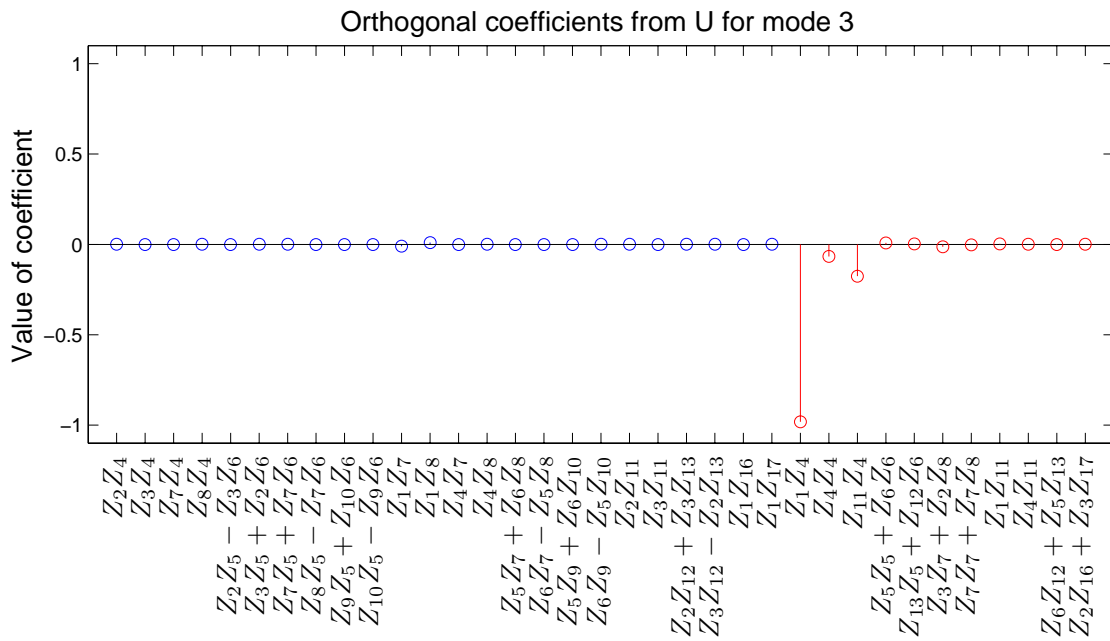
shown in Figure 8.5, where indeed primarily constant defocus is seen, with a little bit of quadratic and quartic defocus.

The fourth and fifth modes can be identified as focal plane tilt, as shown in Figure 8.6. The coefficients for this mode, shown in Figure 8.7, reveals primarily focal plane tilt in these modes.

All 25 of the orthogonal aberration modes across all 49 of the field points and their fits to the double Zernike functions are shown in Appendix O.



(a) HET SVD mode 3.



(b) HET SVD mode 3.

Figure 8.5: Defocus from third control mode.

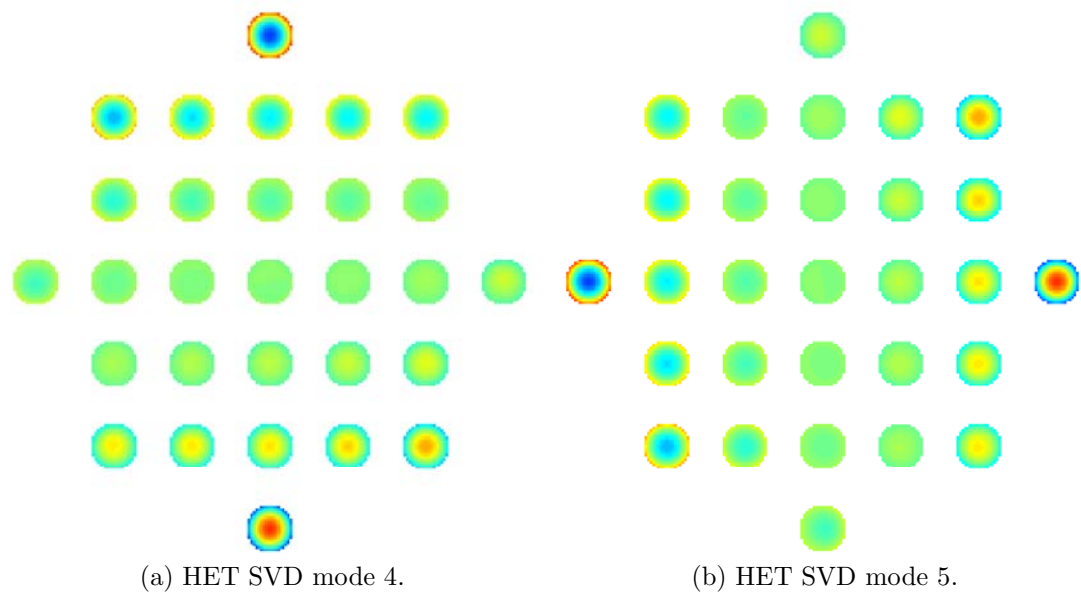
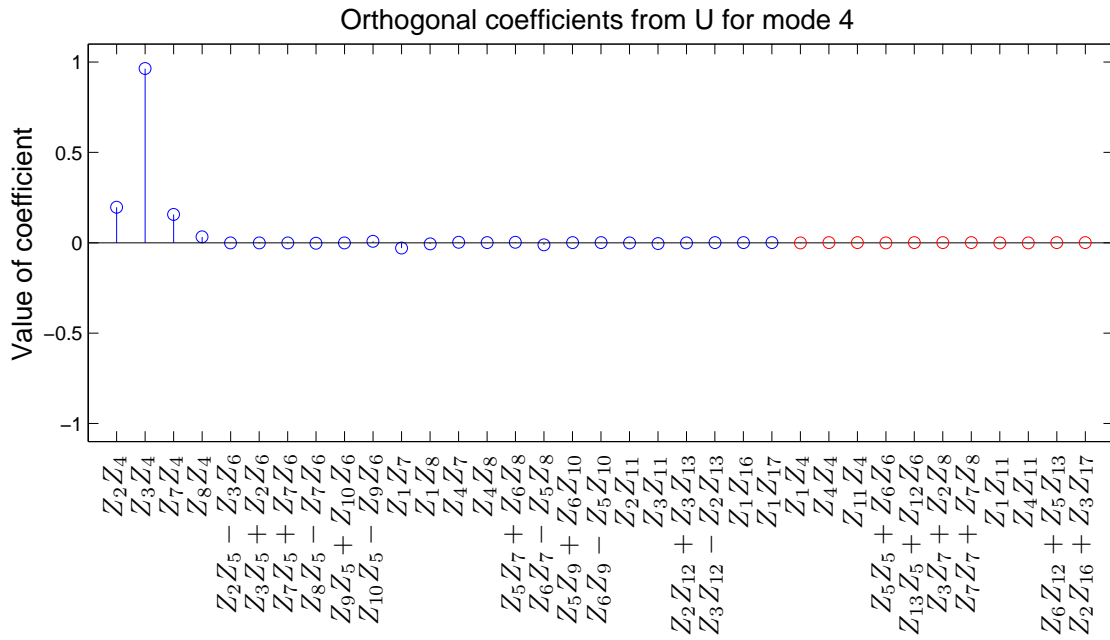
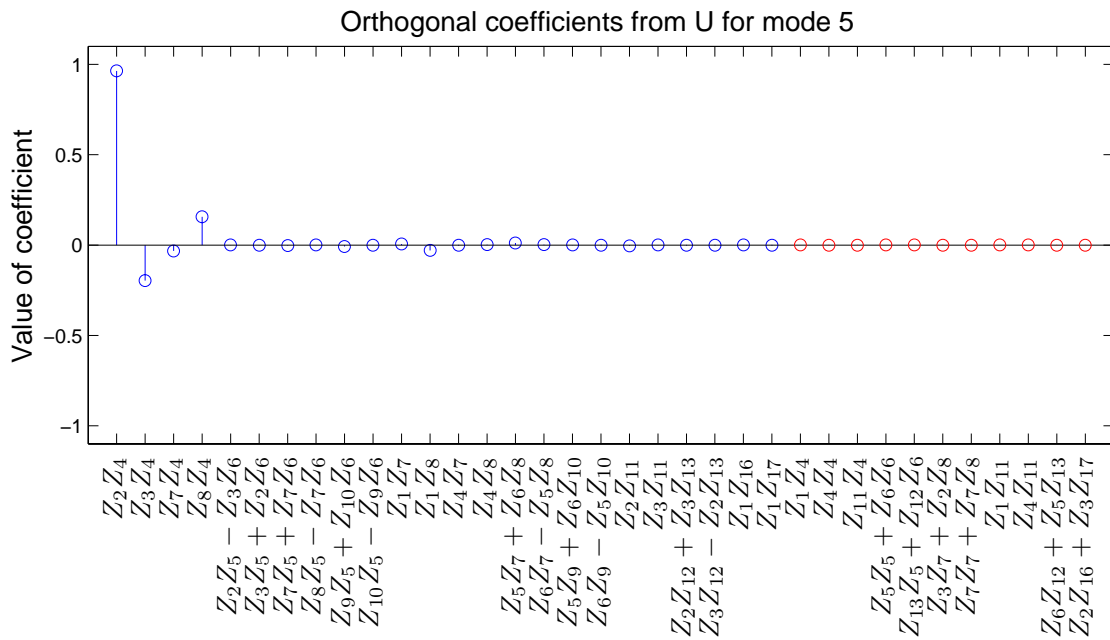


Figure 8.6: Wavefront maps across the field showing focal plane tilt from fourth and fifth control modes.



(a) HET SVD mode 4.



(b) HET SVD mode 5.

Figure 8.7: Coefficients showing focal plane tilt from fourth and fifth control modes.

### 8.3.2 S MATRIX

The  $\mathbf{S}$  matrix ( $25 \times 25$ ) is the diagonal matrix that contains the singular values. There is only one singular value in each column, so there are only 25 singular values. The singular values quantify how sensitive the system is to a particular combination of misalignments. The most sensitive aberration mode is listed first and so forth, which means the singular values are always listed in decreasing order. The modes corresponding to the lowest index singular values are those most easily controlled. The ratios between the singular values describe the dynamic range required by the measurement system to be able to detect all the modes of interest. The lowest order modes will have large amounts of those aberrations because the system is very sensitive to those combinations of degrees of freedom while the highest order modes will have only small amounts of aberrations because the system is less sensitive to those degrees of freedom. The measurement system must be able to measure aberrations in quantities that vary by orders of magnitude. The exact dynamic range required will depend on the number of modes required to be corrected in order to achieve the specification in the optical system. The singular values are plotted in Figure 8.8.

In Figure 8.8, the modes that correspond to aberrations caused by the degrees of freedom that cause the optical system to become non-symmetric come in pairs because each aberration can be caused by perturbations in either the  $x$  or  $y$  direction. These pairs have almost exactly the same singular value because they are equally controllable and they are observed as the flat steps in Figure 8.8. However, when the aberration is caused by a rotationally symmetric degree of freedom (such as

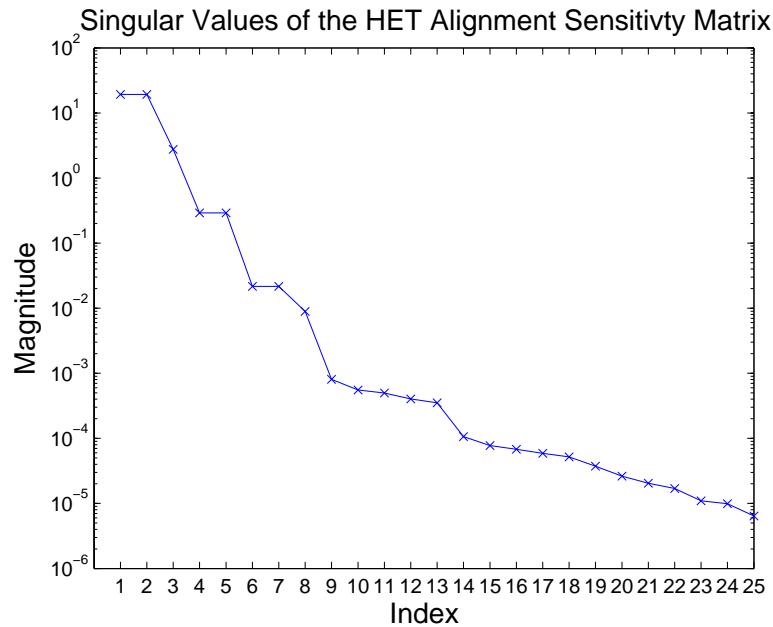


Figure 8.8: HET SVD singular values.

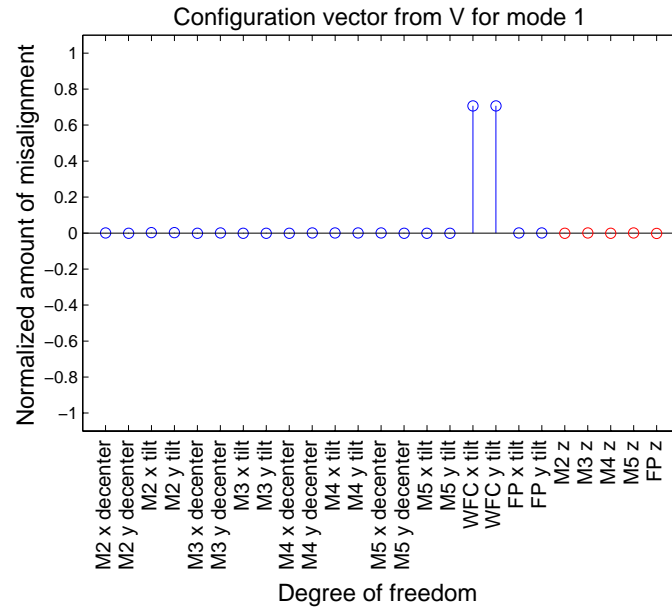
defocus of the focal plane) the mode does not have a partner (for example modes 3 and 8). The highest order control modes correspond to a mixture of all the degrees of freedom and do not come in pairs.

### 8.3.3 V MATRIX

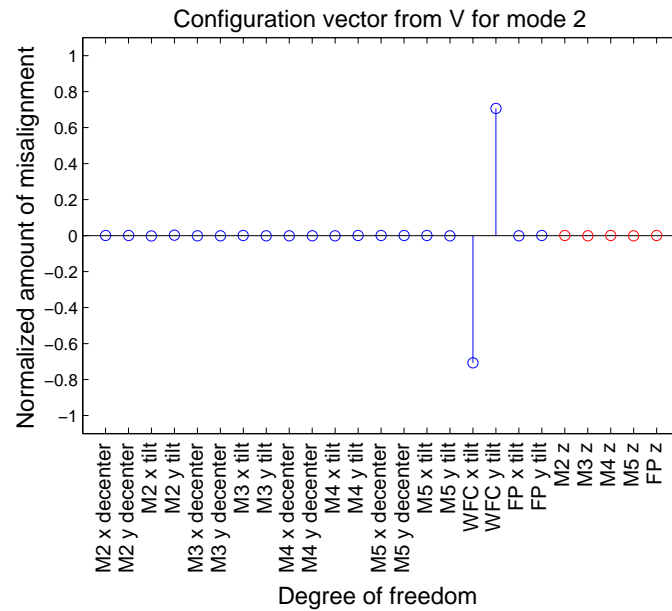
The  $\mathbf{V}$  matrix ( $25 \times 25$ ) describes the misalignments required to generate the orthogonal aberration modes in the  $\mathbf{U}$  matrix. Each column of  $\mathbf{V}$  is a singular vector for a different mode and each row in  $\mathbf{V}$  represents one of the misalignment degrees of freedom.

The misalignment degrees of freedom for the first five control modes are shown in Figures 8.9–8.9, while the remaining control modes are included in Appendix O. Modes 1 and 2 (Figure 8.9), which are constant coma, are caused by tilt of the

entire wide field corrector. This degree of freedom is expected to cause constant coma because the entire WFC is similar to the secondary mirror in a two mirror telescope. Mode 3, which is defocus, is caused by axial defocus of the detector (Figure 8.9). Modes 4 and 5, which are focal plane tilt, are caused by tilting the focal plane (Figure 8.9).



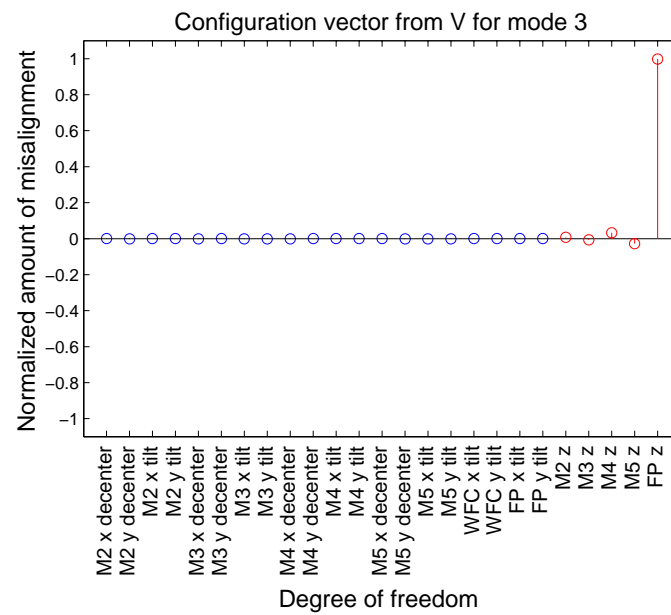
(a) HET SVD mode 1.



(b) HET SVD mode 2.

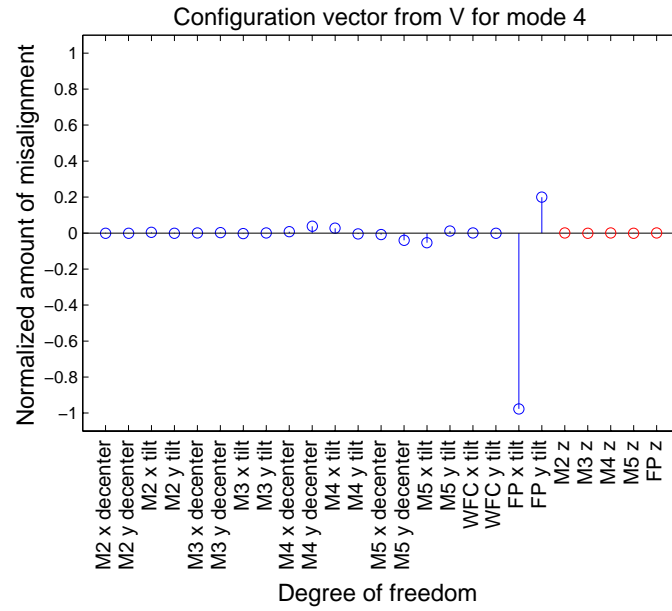
Figure 8.9: Required degrees of freedom to produce HET control modes.



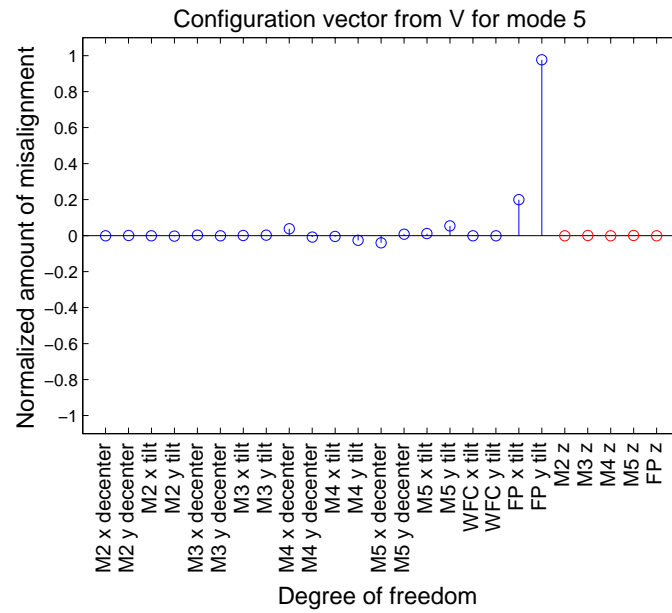


(c) HET SVD mode 3.

Figure 8.9: Continued.



(d) HET SVD mode 4.



(e) HET SVD mode 5.

Figure 8.9: Continued.

## 8.4 SUMMARY OF RESULTING MODES FROM SVD

For reference, Table 8.3 summarizes each of the modes discussed previously and those that are only included in the appendix.

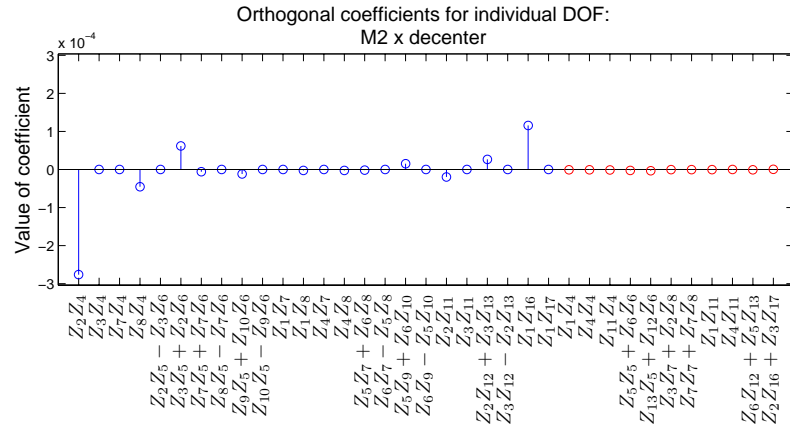
Table 8.3: Summary of HET control modes.

Mode	Main aberration	Misalignment
1, 2	Constant coma	WFC Tilt
3	Defocus	Focal plane axial position
4, 5	Focal plane tilt	Focal plane tilt
6, 7	Lin. astig.	M4 and M5: tilt/decenter
8	Lin. coma, quad. astig.	M4 and M5: axial
9, 10	Const. sec. coma, Lin. sec. astig	M2 and M3: tilt
11	Const. spherical, + other axial sym.	M2 and M3: axial
12, 13	Mixture non-sym	Mixture non-sym
14	Const. spherical, + other axial sym.	M3–M5 axial
15, 16	Mixture non-sym	Mixture non-sym
17–25	Mixture	Mixture

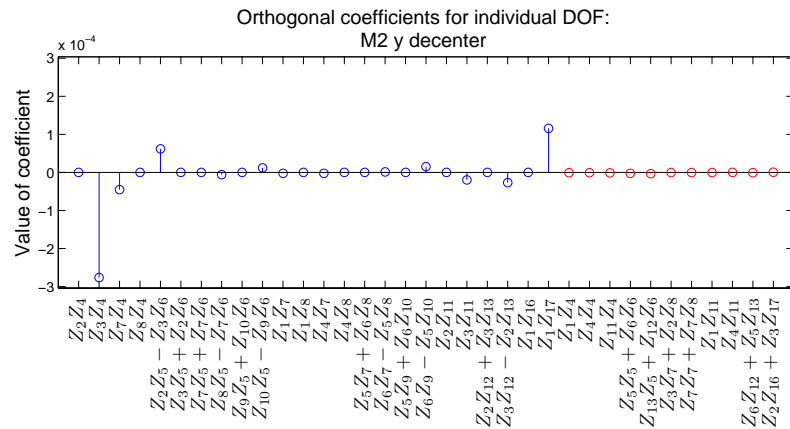
## 8.5 VARIATIONS ON THE SENSITIVITY MATRIX

In another investigation, after each of the degrees of freedom was perturbed by the ZEMAX macro, an optimization of the tilt of the entire WFC was done to correct the constant coma. In this case, the merit function was the Zernike coma (coefficients  $Z_7$  and  $Z_8$ ) for the on-axis field angle. Then the aberration coefficients were recorded, as normal, into the influence matrix and the SVD was calculated. The control modes that resulted from this were practically the same modes as found previously. One of the only differences is that when the aberration coefficients are examined for the individual tilt and decenter degrees of freedom for the mirrors, the effect without the constant coma is more easily seen.

In general, tilt and decenter misalignments of M2 and M3 result in constant coma among other aberrations. When the constant coma is corrected by motion of the entire WFC on the hexapod, then smaller aberrations remain. In general, tilt and decenter misalignments of M4 and M5 do not result in very much constant coma. Therefore, these degrees of freedom can not be compensated by the WFC rigid body motion and were found to have high sensitivities in a previous study (Hvisc & Burge, 2008). The remaining aberrations in all four mirrors M2–M5, after coma is corrected, are focal plane tilt and linear astigmatism. Sometimes the sign of the focal plane tilt and linear astigmatism are the same and sometimes they are not. This is shown in Figure 8.10.

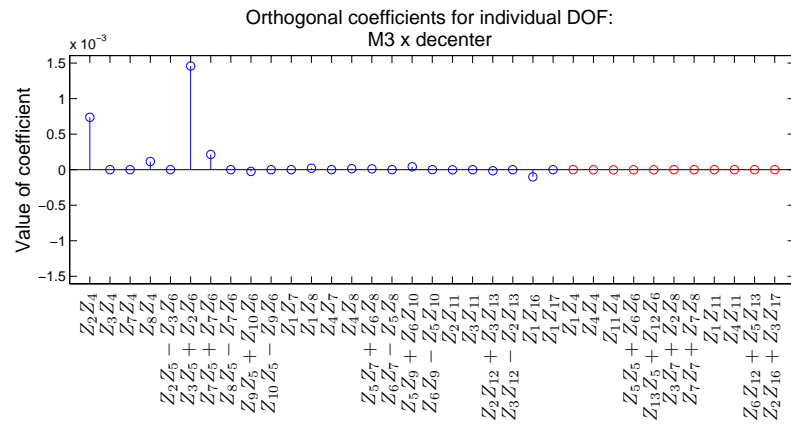


(a) M2 *x* decenter.

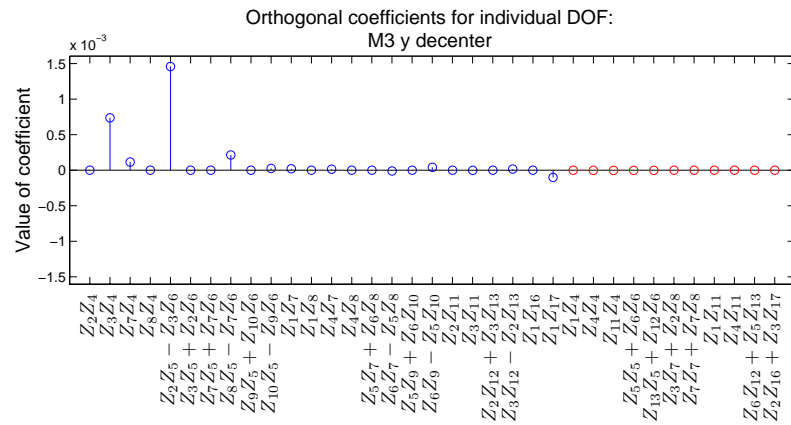


(b) M2 *y* decenter.

Figure 8.10: Aberrations when degrees of freedom are perturbed and the system is corrected for constant coma with the WFC rigid body hexapod tilt.

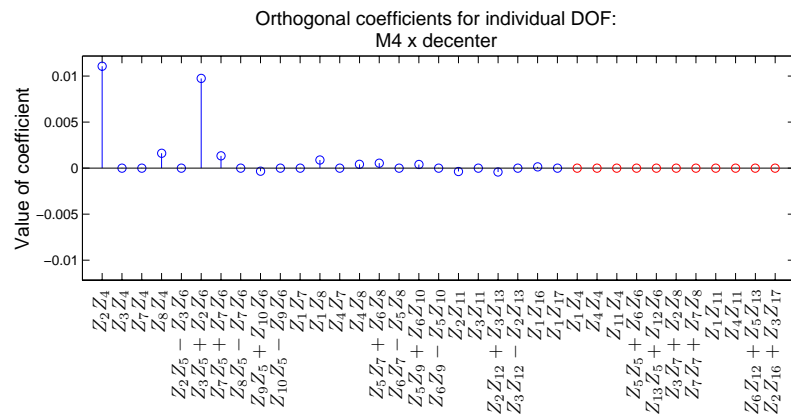


(c) M3 *x* decenter.

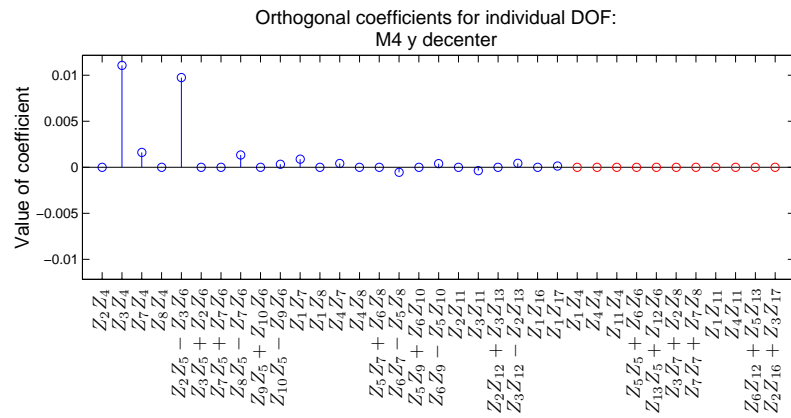


(d) M3 *y* decenter.

Figure 8.10: Continued.

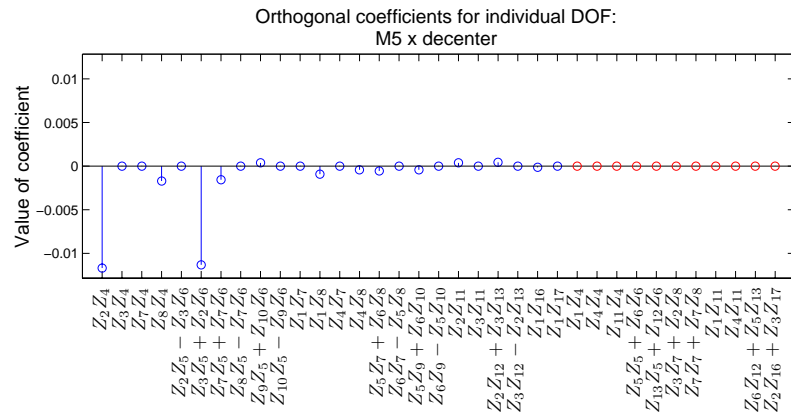


(e) M4 *x* decenter.

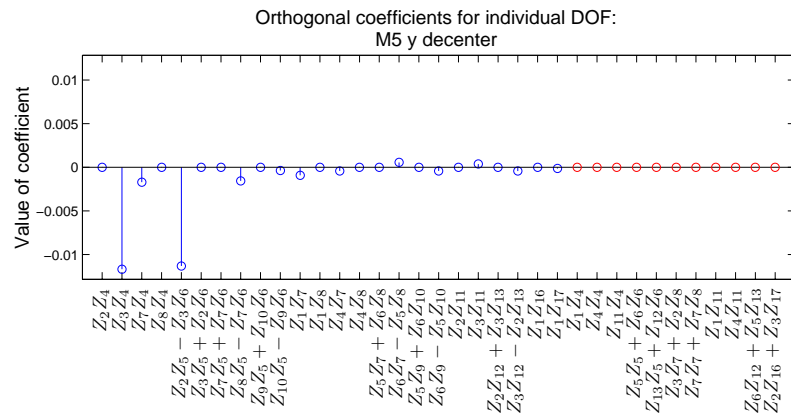


(f) M4 *y* decenter.

Figure 8.10: Continued.



(g) M5 *x* decenter.



(h) M5 *y* decenter.

Figure 8.10: Continued.



## 8.6 WHAT SHOULD GO IN THE SENSITIVITY MATRIX?

In Section 8.5, it was discussed that the order of the modes is somewhat arbitrary, depending on the magnitudes of the perturbations used in the lens design program. (The relative magnitudes of the perturbations for the different degrees of freedom used in the analysis should represent the relative “pain” of the different adjustments. For example, perhaps the perturbation used for the sensitivity matrix SVD can be some small percentage of the total range of motion of that degree of freedom, when the actual system components are known.) Based on that, and the results of Table 8.3, it is still not obvious what orthogonal double Zernike polynomial functions should be chosen to go in the sensitivity matrix used for the system alignment.

Table 8.4 lists the RSS of the values of the coefficients for the double Zernike functions from the first 14 modes. This number of modes was chosen because it represents an approximate number of modes that might be corrected. If an orthogonal double Zernike function was completely controllable by the degrees of freedom in the system (using those 14 modes), the RSS value would be one. The table is listed in order of decreasing RSS value. The functions for the sensitivity matrix should be chosen from closer to the top of the list. The number of functions chosen should be at least equal to the number of degrees of freedom, so there are at least an equal number of measurements as variables in the solution for the alignment vector. It is not harmful to include more orthogonal double Zernike functions. After some point, however, it is not very helpful either. In addition, this consideration should take place for the rotationally-symmetric and non-rotationally symmetric degrees of freedom individually. (It is not possible to make corrections of the tilts and decenters

using wavefront data for axially symmetric system coefficients, for example.)

Another consideration when choosing the functions to include in the sensitivity matrix is the particular field-dependences of the functions. Most of the functions near the top of the list in Table 8.4 are either constant or linear in field. These functions can be measured with only three field points in the system. When there is quadratic field-dependence for the function, four or five field angle measurements are needed. Further down the table, there are functions with cubic or quartic dependence on the field. Since these functions require increasingly many field measurements to determine the coefficients and this increases the difficulty in testing the system, it may make more sense to choose an aberration with less sensitivity, but a lower overall field dependence. For example, linear spherical aberration  $Z_2Z_{11}$  and  $Z_3Z_{11}$  would require only three field measurements, which is easier than measuring the quadratic coma in  $Z_5Z_7 + Z_6Z_8$  and  $Z_6Z_7 - Z_5Z_8$ .

Table 8.4: RSS of coefficients across modes 1–14.

Description	Orthogonal function	RSS for modes 1–14	Axial?
Linear astig.	$Z_3Z_5 + Z_2Z_6$	0.986	
Linear astig.	$Z_2Z_5 - Z_3Z_6$	0.986	
Linear focus	$Z_2Z_4$	0.985	
Linear focus	$Z_3Z_4$	0.985	
Defocus	$Z_1Z_4$	0.983	axial
Constant coma	$Z_1Z_8$	0.983	
Constant coma	$Z_1Z_7$	0.983	
Constant spher. aber.	$Z_1Z_{11}$	0.976	axial
Linear coma	$Z_3Z_7 + Z_2Z_8$	0.975	axial
Quadratic astig.	$Z_5Z_5 + Z_6Z_6$	0.813	axial
Linear sec. astig.	$Z_2Z_{12} + Z_3Z_{13}$	0.762	
Linear sec. astig.	$Z_3Z_{12} - Z_2Z_{13}$	0.761	
Constant sec. coma	$Z_1Z_{17}$	0.682	
Constant sec. coma	$Z_1Z_{16}$	0.674	
Quadratic coma	$Z_5Z_7 + Z_6Z_8$	0.466	
Quadratic coma	$Z_6Z_7 - Z_5Z_8$	0.465	
Quadratic focus	$Z_4Z_4$	0.407	axial
Quadratic coma	$Z_4Z_7$	0.323	
Quadratic coma	$Z_4Z_8$	0.323	
Linear spher. ab.	$Z_2Z_{11}$	0.304	
Linear spher. ab.	$Z_3Z_{11}$	0.304	
Quartic astig.	$Z_{13}Z_5 + Z_{12}Z_6$	0.254	axial
Cubic focus	$Z_7Z_4$	0.252	
Cubic focus	$Z_8Z_4$	0.252	
Quartic astig.	$Z_{13}Z_5 - Z_{12}Z_6$	0.232	
Cubic astig.	$Z_8Z_5 - Z_7Z_6$	0.211	
Cubic astig.	$Z_7Z_5 + Z_7Z_6$	0.210	
Quartic focus.	$Z_{11}Z_4$	0.209	axial
Cubic astig.	$Z_9Z_5 + Z_{10}Z_6$	0.198	
Cubic astig.	$Z_{10}Z_5 - Z_9Z_6$	0.198	
Quartic coma	$Z_{11}Z_8$	0.197	
Quartic coma	$Z_{11}Z_7$	0.197	
Quadratic trefoil	$Z_5Z_9 + Z_6Z_{10}$	0.187	
Quadratic trefoil	$Z_6Z_9 - Z_5Z_{10}$	0.185	
Linear sec. coma	$Z_2Z_{16} + Z_3Z_{17}$	0.180	axial

## 8.7 CONCLUSIONS

This chapter provided an introduction to using singular value decomposition to examine the aberration control modes of a complex optical system. The SVD showed which aberrations result from different combinations of degrees of freedom. The aberrations that were most significant were investigated and the most likely aberrations that would be used in the sensitivity matrix used in the estimate of the alignment of the system were listed. The actual field-dependent aberration functions required depend on how many modes will be controlled and this in turn depends on the specific characteristics of the dynamic range of the system that measures the wavefront aberrations in real life (i.e. wavefront sensor). This was not investigated here and remains a topic for future work.

## CHAPTER 9

## CONCLUSION

This dissertation provided a thorough investigation of the aberrations that occur in misaligned reflective optical systems. Tilting or decentering elements in an optical system results in degradation to the performance. This performance degradation must be predicted in order to choose tolerances before the system is built that will enable the system to meet the performance specifications. Predicting the performance is a type of forward problem and was performed using a Monte Carlo analysis for the Hobby-Eberly Telescope (HET) Wide Field Corrector (WFC). After a system is built, wavefront measurements throughout the field give information about the alignment state of the system. Determining the alignment from measured data is a type of inverse problem and is the main concentration of this dissertation.

The analytical form of the aberrations in optical systems was studied, and a new basis was proposed to describe the functional form of the aberrations that occur in the pupil and the field. The aberration functions in the two-term double Zernike basis separate cleanly into aberrations that occur in an rotationally-symmetric system and those that occur in non-rotationally symmetric systems. From the theory, some aberration functions were predicted to depend directly on the misalignment, while others were predicted to depend on the square or cube of the misalignment. These dependencies were confirmed by simulations of an optical system in a lens design program. While the main concentration of this dissertation was telescope

systems that are misaligned, this basis of functions may also be useful for designing optical systems that are intentionally off-axis. There is more work that can be done to investigate the usefulness of the double Zernike functions while designing off-axis systems and plenty of room for studying the usefulness for refractive systems. Since the basis is complete, it will necessarily be able to describe the field-dependent aberration functions that result in refractive systems. However, the terms that depend linearly on the misalignment and those that depend nonlinearly on the misalignment could be coupled differently.

The double Zernike basis set presented here was found to be useful in describing the aberrations that occur to mirror shape errors in the system as well. The aberrations possible from shape errors sometimes can be compensated by misalignments (for example astigmatism or coma), and sometimes they can not (for trefoil). This means that the aberrations due to misalignment are a subset of those possible from shape errors.

When actually performing a least-squares fit of the system misalignment to the data, the functions that depend linearly on the misalignment should be the ones that are used in the sensitivity matrix. Sometimes the inverse of the sensitivity matrix needed to calculate the misalignments may have small singular values. In this case, measurement noise will be amplified by the least-squares fit. In order to create a better inverse (pseudoinverse) to use for calculations, singular value decomposition (SVD) techniques can be used. Some simple examples were given to show what the SVD means before moving on to a system with many degrees of freedom. An SVD analysis of the HET WFC was performed to study the control modes of the

system. The aberrations that were most easily controlled in this system were found for a given number of modes, but, there is not a general solution that describes the important modes that occur in different types of system. This may depend on the number of elements, the field of view of the system or even the obscuration ratio.

In addition, there may be some flexibility in choosing the modes that can be included in the sensitivity matrix. In order to calculate the field-dependent part of the function, multiple field measurements must be made. In a real-life system, that means having multiple wavefront sensors or interferometers available or moving them around in the field to measure Zernike coefficients. This adds complexity to the system test. Including functions that are higher order in pupil, but lower order in field, can simplify the system test. More work can be done to investigate the practicality or sensitivity of using these lower-order functions in field. In reality, it will depend on the signal to noise ratio in the particular wavefront-measuring system. However, it may be possible to come up with general rules for different ranges of systems.

## APPENDIX A

## HET WFC MONTE CARLO PROCEDURE

Monte Carlo (MC) simulations are used to predict the statistical effect of tolerances on the degrees of freedom in a system by simulating a large number of random systems that meet all tolerances. For this system, the MC trials are performed in ZEMAX using the tolerancing functions with a custom tolerancing script. In one trial, each degree of freedom is randomly perturbed to some number, within a uniform distribution of values allowed by the tolerance. The compensators (which are the five degrees of freedom of the entire WFC with focal plane) are optimized to minimize the rms spot size at a collection of points in the field (on axis and at four at  $(\pm 5 \text{ arcmin}, \pm 5 \text{ arcmin})$ ). The resulting (80%, 85% and 90%) encircled energy spot sizes of this system are then recorded for some points in the field.

**Summary of Monte Carlo Procedure**

1. Fill out tolerance data editor in Zemax for the all of the toleranced degrees of freedom.
  - (a) Make sure the min and max tolerances entered are the same values with opposite signs.
2. Run the Zemax tolerancing routine.



- (a) Can do a sensitivity analysis if desired (takes longer, but advantage is that the values of the tolerances are recorded in the file).
- (b) Run Monte Carlo analysis using a tolerance script.
  - i. # of trials should be at least  $n^2$ , where  $n$  is the number of degrees of freedom.
  - ii. The criterion should be a user script with the following steps:
    - A. Define the compensators.
    - B. Load the merit function for optimization (rms spot size at five field points).
    - C. Optimize the system.
    - D. Perturb the compensators (if desired).
    - E. Load the merit function used for evaluation (encircled energy across the field).
    - F. Print the encircled energy across the field (values from certain rows of the merit function) to the results window.
- (c) Save the data from the window that opens with the results at the end of the Monte Carlo analysis into a text file.

## APPENDIX B

## ZEMAX MACRO FOR HET WFC

```
! This tolerance script is used to run the Monte Carlo simulation for
! the HET WFC file "HET OE 002 MC.zmx"
! Stacie Manuel
! April 28, 2009
! =====

! clear any existing compensators for a clean start
CLEARCOMP

! Load the merit function for optimization
LOADMERIT 5field_rmsspot.mf

! define the compensators
COMP 4 0
CPAR 4 1
CPAR 4 2
CPAR 4 3
CPAR 4 4

! optimize Automatic
OPTIMIZE

! Then perturb the compensators
! 15um cases:
PERTURB 0 4 0 0 1 0 -0.015 0.015
PERTURB 1 4 1 0 1 0 -0.015 0.015
PERTURB 1 4 2 0 1 0 -0.015 0.015
PERTURB 1 4 3 0 1 0 -0.001229 0.001229
PERTURB 1 4 4 0 1 0 -0.001229 0.001229

! Load the encircled energy merit function
LOADMERIT EE_2009.MF
```

```
! Report the Encircled energies
REPORT "80% field 1 = " 2
REPORT "80% field 2 = " 3
REPORT "80% field 3 = " 4
REPORT "80% field 4 = " 5
REPORT "80% field 5 = " 6
REPORT "85% field 1 = " 7
REPORT "85% field 2 = " 8
REPORT "85% field 3 = " 9
REPORT "85% field 4 = " 10
REPORT "85% field 5 = " 11
REPORT "90% field 1 = " 12
REPORT "90% field 2 = " 13
REPORT "90% field 3 = " 14
REPORT "90% field 4 = " 15
REPORT "90% field 5 = " 16
REPORT "EFFL = " 17
```

```
LOADMERIT 5field_rmsspot.mf
```

## APPENDIX C

## MATLAB CODE FOR THE HET WFC FORWARD PROBLEM

## C.1 GETHET\_EE.M MATLAB CODE

This function analyzes one Zemax file with Monte Carlo tolerance results to find the resulting encircled energy diameter across the field for different confidence levels.

```
function [conf]=getHET_EE(d)

Output = readtext2(d);
% Figure out how big the loaded file is: The total number of rows
% changes depending on whether or not there is a sensitivity analysis
% and the number of degrees of freedom tolerated.
[r,c]=size(Output);

%% Find where certain things are located in the loaded data
% The following lines put a text space in the empty cells of the
% Output. This makes some commands of searching for rows that things
% happen easier
for i=1:r
    for j=1:c
        if isempty(Output{i,j})
            Output{i,j}=' ';
        end
    end
end

% Look for certain rows in the text. (Result is a cell array empty
% everywhere except where the strings are found.)
e = strfind(Output(:,1), 'Evaluating');
l = strfind(Output(:,6), 'lens');
s = strfind(Output(:,1), 'Sensitivity');
f = strfind(Output(:,1), 'Fields');

% The next for loop extracts the indices. It is long is because
% I don't see a command like isempty that works for an entire cell
% array. (i.e. It can only return '1' or '0', not a cell array of '1's
% and '0's.)
```

```

count = 0; count2 = 0;
for i = 1:r
    % where the evaluation results are (both sensitivities and MC)
    if isempty(e{i})==0
        count=count+1;
        liste(count)=i;
    end
    % where the only Monte Carlo trials are
    if isempty(l{i})==0
        count2=count2+1;
        trials(count2)=i;
    end
    % if a sensitivity analysis is there
    if isempty(s{i})==0
        sens=i;
    end
    % where the fields are listed
    if isempty(f{i})==0
        fieldloc=i;
    end
end

% nom is where the nominal system is evaluated
nom=liste(1); % It's the first one evaluated

% liste is where the sensitivities are evaluated
liste = setdiff(liste, trials); % removes MC trials from the list
liste=liste(2:end); % removes the nominal system evaluation

% trials is where the MC systems are evaluated
% (no modification from above calculation needed)

%% Get the field points
xfield = zeros(1,5);
yfield = zeros(1,5);

for i = 1:5
    xfield(i) = str2double(Output{fieldloc+1+i,2});
    yfield(i) = str2double(Output{fieldloc+1+i,3});
end

% Multiply field by 60 to put into arcminutes
field = 60*sqrt(xfield.^2 + yfield.^2);
field=[field field field];
%% Get the nominal performance numbers
NominalMF = str2double(Output{nom+17,4}); %The nominal merit function

% Find the nominal EE for all the field points (More interesting for
% Monte Carlo analysis)
NominalEE = zeros(1,5);

```

```

for i = 1:15
    NominalEE(i) = str2double(Output{nom+i,5});
end

% Convert the um radius of the EE to arcseconds diameter
NominalEE = NominalEE/88.4784968;

%% Monte Carlo Analysis
% Don't perform MC analysis if not done in Zemax
if exist('trials','var')

    ntrials = length(trials); % Find the number of Monte Carlo trials
    % If you want to investigate the effect of using a smaller number
    % of trials than in the data, change ntrials here.
    %ntrials = 173;

    % Gather all the data from the Monte Carlo trials into matrices
    % For the: Encircled energy, Effective focal length, Compensators

    % Preallocate variables
    EE=zeros(ntrials,15); % 15 columns (for 5 fields, 3 different EEs)
    EFFL=zeros(ntrials,1);
    thicknesscomp=zeros(ntrials,1);
    xdeccomp=zeros(ntrials,1);
    ydeccomp=zeros(ntrials,1);
    xtiltcomp=zeros(ntrials,1);
    ytiltcomp=zeros(ntrials,1);

    for i = 1:ntrials
        for j = 1:15
            EE(i,j)=str2double(Output{trials(i)+j,5});
        end
        EFFL(i)=str2double(Output{trials(i)+j+1,3});
        thicknesscomp(i)=str2double(Output{trials(i)+j+3,3});
        xdeccomp(i)=str2double(Output{trials(i)+j+4,6});
        ydeccomp(i)=str2double(Output{trials(i)+j+5,6});
        xtiltcomp(i)=str2double(Output{trials(i)+j+6,6});
        ytiltcomp(i)=str2double(Output{trials(i)+j+7,6});
    end

    labels = {'EFFL'; 'thickness'; 'xdec'; 'ydec'; 'xtilt'; 'ytilt'}';
    E=[EFFL, thicknesscomp, xdeccomp, ydeccomp, xtiltcomp, ytiltcomp];
    stdminmaxmean = [std(E)', min(E)', max(E)', mean(E)']';

    % Convert the um radius of the EE to arcseconds diameter
    EE = EE/88.4784968;

    % Define the requirement: (because it will be plotted later)
    fieldreq=[0 5 11];    EEreq = [0.45 0.45 0.8];

```

```

if strcmp(d, '35_0_align_manu_windows.TXT')
% Histogram Analysis
figure('Position',[50 50 1000 200])
for i = 1:5
    subplot(1,5,i),hist(EE(:,10+i)),
        title([num2str(field(10+i)), ' arcmin field'])
        h=max(hist(EE(:,10+i)));
        axis([0 1 0 h])
        hold on, plot([NominalEE(10+i),NominalEE(10+i)], [0,h], 'r'),
        hold off
        ylabel('# trials'),xlabel('90% EE (arcsec)')
end

% Range for histogram plots
r1=max([-min(xdeccomp) -min(ydeccomp)...
        max(xdeccomp) max(ydeccomp)]);
r2=max([-min(xtiltcomp) -min(ytiltcomp)
        max(xtiltcomp) max(ytiltcomp)]);
figure
subplot(2,3,1), hist(thicknesscomp(:)),
title('Thickness compensator')
subplot(2,3,2), hist(xdeccomp(:)),
title('decenter x compensator'),axis([-r1 r1 0 inf])
subplot(2,3,5), hist(ydeccomp(:)),
title('decenter y compensator'),axis([-r1 r1 0 inf])
subplot(2,3,3), hist(xtiltcomp(:)),
title('tilt x compensator'),axis([-0.05 0.05 0 inf])
subplot(2,3,6), hist(ytiltcomp(:)),
title('tilt y compensator'),axis([-0.05 0.05 0 inf])
subplot(2,3,4), hist(EFFL(:)),
title('Effective Focal Length'),
end

% Plot the cumulative distributions
allEE=EE;
allnominalEE=NominalEE;
    for ii=1:3
        index=1+(ii-1)*5;
        EE=allEE(:,index:index+4);

        NominalEE=allnominalEE(:,index:index+4);

        x = linspace(0,max(EE(:)),30);

        for i = 1:5
            n_elements(i,:) = histc(EE(:,i),x);
            c_elements(i,:) = cumsum(n_elements(i,:))/ntrials;
        end
    end

```

```

if (ii==3)&&(strcmp(d,'35_0_align_manu_windows.TXT'))
    figure, plot(x,c_elements,[0,1],[0.9, 0.9],'k:')
    xlabel('90% EE diameter (in arcseconds)')
    ylabel('Confidence Level')
    legend([num2str(field(1)), ' arcmin'],...
           [num2str(field(2)), ' arcmin'],...
           [num2str(field(3)), ' arcmin'],...
           [num2str(field(4)), ' arcmin'],...
           [num2str(field(5)), ' arcmin'],'Location','Southeast')
    title(['Cumulative probabilities for the spot size'...
          ' histograms'])
end

% Find the 25%, 50%, 75%, 90% and 98% Confidence Levels
for i = 1:5
    conf25(i) = interp1q(c_elements(i,:),x',0.25);
    conf50(i) = interp1q(c_elements(i,:),x',0.5);
    conf75(i) = interp1q(c_elements(i,:),x',0.75);
    conf90(i) = interp1q(c_elements(i,:),x',0.90);
    conf95(i) = interp1q(c_elements(i,:),x',0.95);
    conf98(i) = interp1q(c_elements(i,:),x',0.98);
end
conf=[conf25;conf50;conf75;conf90;conf95;conf98];
end
end
end

```

## C.2 READTEXT2.M MATLAB CODE

The function *readtext2.m* is required by *getHET\_EE.m*.

```

function Output = readtext2(root)

% Location of Data File should be given in root

% Open File and count the number of rows in the file
fid=fopen(root);
nRows=0;
while 1
    iString=fgetl(fid);
    if ~ischar(iString)
        break
    end
    nRows=nRows+1;
end
end

```



```

% Return to beginning of file
fseek(fid,0,'bof');

% Preallocate for speed (May need >20 at some point)
Output=cell(nRows,20);

% For each row, assign each space delimited object to a cell in the
% "Output" matrix
for iRow=1:nRows
    iCol=1;
    % Temporary storage of the first object
    % Note: the space delimiter used here can be replaced by any
    % delimiter
    [TempOutput,Rem]=strtok(fgetl(fid),' ');
    % If there is now data on this row, then assign the first object
    % to be an underscore
    if (length(TempOutput) == 0)
        TempOutput='_';
    end
    % Build the "Output" matrix this will be the first column of the
    % iRow-th row
    Output(iRow,iCol)=cellstr(TempOutput);
    % Repeat this only using Rem as the total string and incrementing
    % the iCol counter
    while length(Rem) > 0
        iCol=iCol+1;
        [TempOutput,Rem]=strtok(Rem,' ');
        Output(iRow,iCol)=cellstr(TempOutput);
    end
end
end

```

### C.3 HET\_MC\_ANALYSIS\_OCT09.M MATLAB CODE

This script compiles all of the different Monte Carlo results to find and plot the expected system performance.

```

%% HET_MC_Analysis_Oct09.m
% Written by Stacie Manuel

%% Close all plots and clear all old data
close all
clear all

%% Set flag for saving figures
p = 0; % To save the figures as .eps files, set p = 1.

```

```

% Location for the figures to be saved:
mypath=['C:\Documents and Settings\shvisc\My Documents\',...
        'My dissertation\The best staciethesis\MATLAB\HET_Forward\'];
%% Start timer
tic
%%
field = [0,2.7,5.4,8.1,10.998;];
%% Nominal Performance
nominal=[0.2533,0.2289,0.193,0.3184,0.584;];
nominal5=[nominal;nominal;nominal;nominal;nominal;nominal];
%% Mirror fabrication
MC_Fabr=getHET_EE('35_0_manu.txt');
Fab = real(sqrt(MC_Fabr.^2-nominal5.^2));
%% Alignment
MC_Align=getHET_EE('35_0_align.TXT');
Align = real(sqrt(MC_Align.^2-nominal5.^2));
%% Window
MC_Window=getHET_EE('35_0_windows.TXT');
MC_Window2=getHET_EE('35_0_windows_pt5arcminwedge.TXT');
Window = real(sqrt(MC_Window.^2 - nominal5.^2)); %1 arcmin wedge
Window2 = real(sqrt(MC_Window2.^2 - nominal5.^2)); % 1/2 arcmin wedge
%% Nominal + Alignment + Fabrication + windows
AlignManuold = sqrt(nominal5.^2 + Fab.^2 + Align.^2 + Window.^2);
MC_AlignManu=getHET_EE('35_0_align_manu_windows.TXT');
MC_AlignManu_8comp=[0.2143,0.1899,0.1874,0.3041,0.5749;...
    0.2285,0.2041,0.198,0.32,0.5982;...
    0.2461,0.2239,0.2086,0.3453,0.6252;...
    0.2672,0.2493,0.2217,0.3666,0.6512;...
    0.2839,0.2716,0.2351,0.374,0.6679;...
    0.2939,0.285,0.2493,0.3901,0.6946;];
%% Assembly and Operation
AssOp = 0.25*Align;
%% Compensators -15um accuracy
MC_Compen=getHET_EE('35_0_align_manu_windows_15um.TXT');
Comp = real(sqrt(MC_Compen.^2 - MC_AlignManu.^2));
%% Polishing and figuring
PolFig=[0.25,0.25,0.25,0.25,0.25];
PolFig5 = [PolFig;PolFig;PolFig;PolFig;PolFig;PolFig];
%% Temperature effects
MC_Pos_Temp=getHET_EE('35_0_align_manu_windows_15um_pos_temp.TXT');
Pos_Temp = real(sqrt(MC_Pos_Temp.^2 - MC_Compen.^2));

MC_Neg_Temp=getHET_EE('35_0_align_manu_windows_15um_neg_temp.TXT');
Neg_Temp = real(sqrt(MC_Neg_Temp.^2 - MC_Compen.^2));

Temp = max(Pos_Temp,Neg_Temp);
%% WFC Angle change effects
MC_deg26pt5=getHET_EE('26pt5_0_align_manu_windows_15um.TXT');
deg26=real(sqrt(MC_deg26pt5.^2 - MC_Compen.^2));

```

```

MC_deg43pt5=getHET_EE('43pt5_0_align.manu.windows_15um.TXT');
deg43=real(sqrt(MC_deg43pt5.^2 - MC_Compen.^2));

MC_deg35_8pt5=getHET_EE('35_8pt5_align.manu.windows_15um.TXT');
deg35_8pt5=real(sqrt(MC_deg35_8pt5.^2 - MC_Compen.^2));

MC_deg41_6=getHET_EE('41_6_align.manu.windows_15um.TXT');
deg41_6=real(sqrt(MC_deg41_6.^2 - MC_Compen.^2));

MC_deg29_6=getHET_EE('29_6_align.manu.windows_15um.b.TXT');
deg29_6=real(sqrt(MC_deg29_6.^2 - MC_Compen.^2));

Angle1 = max(deg26,deg43);
Angle2 = max(deg35_8pt5,deg41_6);
Angle2 = max(deg29_6,Angle2);
Angle = max(Angle1,Angle2);
%% RSS
Total = sqrt(nominal5.^2 + Align.^2 + Fab.^2 + AssOp.^2 + Comp.^2 + ...
    Temp.^2 + Angle.^2 + PolFig5.^2 + Window.^2);
%% Specification
f = [0 5 11];
y = [0.45 0.45 0.8];
%% Graphs
labels.x='Field Radius (in arcminutes)';
labels.y='90% Encircled Energy Diameter (in arcseconds)';
text=['25% confidence';'50% confidence';'75% confidence';...
    '90% confidence';'95% confidence';'98% confidence'];
labels.legend={'specification','nominal','alignment','fabrication',...
    'assembly and operation','compensators','temperature',...
    'correlated misalignment','polishing/figuring','window',...
    'expected performance (rss total)'};
figure
for i = 1:6
    subplot(2,3,i)
    plot(f,y,field,nominal,field,Align(i,:), '—', ...
        field,Fab(i,:), '-.', field,AssOp(i,:), '—', ...
        field,Comp(i,:), '-.', field,Temp(i,:), '—', ...
        field,Angle(i,:), '-.', field,PolFig, '—', ...
        field,Window(i,:), 'c-', field,Total(i,:), 'r')
    title(text(i,:)),
    axis([0 11 0 0.9]), xlabel(labels.x), ylabel(labels.y)
end
legend(labels.legend,0)

i=4; % 90% Confidence level
h=figure;
plot(f,y,field,nominal,field,Align(i,:), '—', ...
    field,Fab(i,:), '-.', field,AssOp(i,:), '—', ...
    field,Comp(i,:), '-.', field,Temp(i,:), '—', ...
    field,Angle(i,:), '-.', field,PolFig, '—', ...

```

```

        field,Window(i,:), 'c-', field, Total(i,:), 'r')
title({'Contributions to expected performance',...
      'at the 90% confidence level'}, 'FontSize', 12),
legend(labels.legend, 'Location', 'NorthWest')
axis([0 11 0 1.5]),
xlabel(labels.x, 'FontSize', 12), ylabel(labels.y, 'FontSize', 12)
if p==1 % Save this figure as an .eps file
    set(gcf, 'PaperPositionMode', 'auto')
    print(h, '-depssc', [mypath, 'HET_Fig1.eps']);
end

labels.title={'Net effect of residual design errors, alignment '...
             'tolerances, fabrication', ['tolerances (including windows)'...
             ' and compensator perturbations']};
Total2 = sqrt(nominal5.^2 + Align.^2 + Fab.^2 + Window.^2 + Comp.^2);
h=figure;
plot(f,y, field, MC_Comp(4,:), '—', field, Total2(4,:), '—')
legend('Specification', 'Monte Carlo simulation', 'RSS of components', 0)
axis([0 11 0 inf]),
xlabel(labels.x, 'FontSize', 12), ylabel(labels.y, 'FontSize', 12)
title(labels.title, 'FontSize', 12);
if p==1 % Save this figure as an .eps file
    set(gcf, 'PaperPositionMode', 'auto')
    print(h, '-depssc', [mypath, 'HET_Fig2.eps']);
end

h=figure;
plot(f,y, field, MC_AlignManu(4,:), ':', ...
      field, MC_AlignManu_8comp(4,:), '—')
axis([0 11 0 inf]),
xlabel(labels.x, 'FontSize', 12), ylabel(labels.y, 'FontSize', 12)
legend('Specification', '5 compensators', '8 compensators', 0)
title({'Net effect of residual design errors, alignment',...
      ' tolerances and fabrication tolerances'}, 'FontSize', 12);
if p==1 % Save this figure as an .eps file
    set(gcf, 'PaperPositionMode', 'auto')
    print(h, '-depssc', [mypath, 'HET_8comp.eps']);
end

h=figure;
plot(f,y, field, nominal, field, Total, '—');
legend('Specification', 'Nominal', '25%', '50%', '75%', '90%', '95%', '98%')
set(legend, 'Position', [0.73 0.12 0.24 0.37]);
title('Comparison of confidence levels for different rss totals',...
      'FontSize', 12)
axis([0 11 0 inf]),
xlabel(labels.x, 'FontSize', 12), ylabel(labels.y, 'FontSize', 12)
if p==1 % Save this figure as an .eps file
    set(gcf, 'PaperPositionMode', 'auto')
    print(h, '-depssc', [mypath, 'HET_Fig4.eps']);
end

```

```

end

h=figure;
plot(f,y,field,deg26(4,:), 'o-', field,deg43(4,:), 's-', ...
     field,deg35_8pt5(4,:), 'd-', field,deg41_6(4,:), '^-', ...
     field,deg29_6(4,:), 'v-', 'MarkerSize',6);hold on,
plot(field,Angle(4,:), 'k—', 'LineWidth',2)
legend('Specification', '26.5\circ tilt', '43.5\circ tilt', ...
      '35\circ, 8.5\circ tilt', '41\circ, 6\circ tilt', ...
      '29\circ, 6\circ tilt', 'correlated misalignment error')
set(legend, 'Position', [0.14 0.58 0.41 0.33]);
axis([0 11 0 inf]),
xlabel(labels.x, 'FontSize',12), ylabel(labels.y, 'FontSize',12)
title(['Correlated misalignment error contributions', ...
      ' (90% confidence)'], 'FontSize',12)
if p==1 % Save this figure as an .eps file
    set(gcf, 'PaperPositionMode', 'auto')
    print(h, '-depsc', [mypath, 'HET_Fig3.eps']);
end
figure
plot(f,y,field,Pos_Temp(4,:), 'o:', field,Neg_Temp(4,:), 'o:', ...
     field,Temp(4,:), 'kx-')
legend('Specification', 'Positive change', 'Negative change', ...
      'Temp error for table',0)
title('Temperature effect (90% confidence)')
axis([0 11 0 inf]), xlabel(labels.x), ylabel(labels.y)

nonsystemtest=sqrt(AssOp.^2 + Comp.^2 + Temp.^2 + Angle.^2);
spec= [0.45, 0.45, 0.45+(0.8-0.45)*0.4/6, 0.45+(0.8-0.45)*3.1/6, 0.8];
testspec = sqrt(spec.^2 - nonsystemtest(6,:).^2);

labels.title2=['Comparison of confidence levels for different rss'...
              'totals for the system test'];
labels.legend2={'Specification', 'Nominal', 'System Test Goal', ...
              '25%', '50%', '75%', '90%', '95%', '98%'};
figure
plot(f,y,field,testspec,field,nonsystemtest, ':');
legend(labels.legend2([1,3:end]),0), title(labels.title2)
axis([0 11 0 inf]), xlabel(labels.x), ylabel(labels.y)

systemtest=sqrt(nominal5.^2+Align.^2 + Fab.^2 + PolFig5.^2+Window.^2);
figure
plot(f,y,field,nominal,field,testspec,field, systemtest, ':');
legend(labels.legend2,0), title(labels.title2)
axis([0 11 0 inf]), xlabel(labels.x), ylabel(labels.y)

% Output matrix of 90% confidence results
matrix = [field; nominal; Fab(4, :); Align(4, :); AssOp(4, :); Comp(4, :); ...
          PolFig; Temp(4, :); Angle(4, :); Window(4, :); Total(4, :)];

```

```
% End Timer  
toc
```

## APPENDIX D

## OPTICAL SPECIFICATIONS FOR HET

This appendix includes the report for the Hobby-Eberly Telescope project describing the optical specifications for the mirrors. This University of Arizona internal report is unpublished, so it is provided here in entirety for reference. This document control number for this report is HET OE 001 - REV B.

## Hobby-Eberly Telescope Wide Field Corrector Optical Specifications for the Mirrors

Stacie Manuel  
November 23, 2009

### **Abstract:**

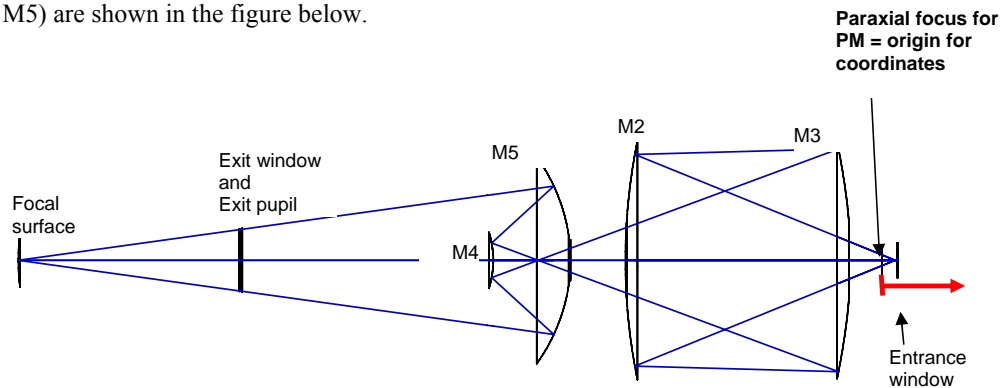
This document reviews the optical design and specifications with tolerances of each of the four mirrors and two windows in the Hobby-Eberly Telescope (HET) Wide Field Corrector (WFC).

The report is divided into the following sections:

- Optical layout
- Mirror sizes and clear apertures
- Radii of curvature and conic constants
- Aspheric terms
- Surface irregularity
- Surface sag
- Aspheric departure
- Optical surface to mechanical part specifications
- Windows

### ***Optical layout***

The WFC is a four mirror corrector designed to correct the aberrations from the large (10 m), spherical primary mirror (M1 or PM) of HET. The positions of each of the four mirrors (M2 – M5) are shown in the figure below.



**Figure 1: Optical layout of the HET WFC**

The locations of each of the elements (as described by the optical design *HET OE 002.zmx*) are listed in Table 1.



**Table 1: Location of optical elements in HET WFC**

Primary Mirror	vertex		13082.590
Paraxial	focus		0.000
Entrance window		front	256.590
Entrance window		rear	254.590
M2	vertex		-932.003
M3	vertex		50.646
M4	vertex		-1509.113
M5	vertex		-1172.003
Exit pupil			-2605.358
Exit window		front	-2615.358
Exit window		rear	-2618.358
Focal surface			-3582.767

***Mirror sizes and clear apertures***

The clear apertures of each mirror were determined by the maximum extent where the rays from all fields fall on the mirror, for the most part. The physical sizes were chosen to be larger than the clear apertures with a small margin (to make polishing easier, etc.). There are two critical places where there is a tradeoff between the mirror being big enough to reflect more light and the mirror being small enough to let light pass by it. These are at the inner diameter of M2 and the outer diameter of M4. In these cases, there needs to be a very small margin, so that as little light is lost as possible. Since it is difficult to polish to the edge of optics, the mirror blank will be fabricated oversized, then polished and finally trimmed to the appropriate final dimension. (Tolerances on the diameters, as listed on the drawings, are  $\pm 0.010$  inch.)

**Table 2: Size and clear aperture specifications - diameters**

	<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>
OD(mm)	1020	1020	248 (final) 298 (blank)	900
ID (mm)	322 (final) 203 (blank)	280	20	220
CA OD(mm)	965	980	244	880
CA ID (mm)	326	325	30	254
Beam footprint diameter (mm)	920	970	150	646

**Table 3: Size and clear aperture specifications - radii**

	<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>
CA outer radius (mm)	482.5	490	122	440
CA inner radius (mm)	163	162.5	15	127
Beam footprint radius (mm)	460	485	75	323
Margin between CA OD and OD (mm) radius	27.5	20	2 (final) 27 (blank)	10
Margin between CA ID and ID (mm) radius	2 (final) 61.5 (blank)	23	5	17

### ***Radii of curvature and conic constants***

Three of the four mirrors are concave and one (M4) is convex. All four mirrors are aspheric, with surfaces described by conic constants. In addition, two of the mirrors (M3 and M5) have higher order aspheric terms, described in the next section. The sag of the two conic mirrors (M2 and M4) is given by the following equation:

$$z(r) = \frac{cr^2}{1 + \sqrt{1 - (1+k)c^2r^2}}$$

where  $c$  is the mirror curvature (inverse of radius of curvature),  $k$  is the conic constant and  $r$  is the radial distance from the center of the mirror.

The tolerances for both the radii and the conic constants have two different components. First, there is an allowable delta, which describes how close the actual surface needs to match the specified value. Second, there is a delta uncertainty, which describes the accuracy of the optical test measuring the surface. The delta uncertainties were determined by the previous sensitivity analysis and Monte Carlo simulations. The PV asphere uncertainty for  $\Delta k$  is the found by the following equation:

$$\text{PV asphere uncertainty for } \Delta k = \frac{\Delta k}{k} \cdot \text{PV aspheric departure,}$$

where the PV aspheric departure is found using the zero index method described in the appendix for the mirror surfaces without the higher order aspheric terms included.

**Table 4: Radii of curvature and conic constants**

	<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>
<u>Radius of curvature <math>R</math> (mm)</u>	2620.8	2032.5	-376.7	742.1
Allowable departure from nominal $\Delta R$ (mm)	0.5	0.5	0.5	0.5
$\Delta R$ measurement uncertainty (mm)	0.075	0.075	0.045	0.04
Sag uncertainty for $\Delta R$ ( $\mu\text{m}$ )	2	4	3	9
<u>Conic Constant <math>k</math></u>	0.663	-7.711	-2.098	-0.2675
PV aspheric departure ( $\mu\text{m}$ ) (Does not include higher order asphere terms)	52	998	241	741
$\Delta k$ measurement uncertainty	0.0011	0.0008	0.0006	0.0001
PV asphere uncertainty for $\Delta k$ (nm)	86	104	69	277

### ***Aspheric terms***

In addition to the conic constants, M3 and M5 require higher order aspheric terms to describe their surfaces. The equation of the sag for these mirrors is similar to the equation above, with the addition of the terms for higher order powers of the radius:

$$z(r) = \frac{cr^2}{1 + \sqrt{1 - (1+k)c^2r^2}} + \alpha_1 r^2 + \alpha_2 r^4 + \alpha_3 r^6 + \alpha_4 r^8 + \alpha_5 r^{10}$$

The terms  $\alpha_1$  and  $\alpha_2$  are not needed because the second order dependence on the radial pupil position is defined by the radius of curvature and the conic already describes the fourth order dependence. In order to optimize M3, three more terms were needed, and two more terms for M5. Since the surfaces are radially symmetric, only even order terms are needed.

**Table 5: Higher order aspheric terms**

		<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>
$\alpha_3$	6 <sup>th</sup> order term (1/mm <sup>5</sup> )	-	-8.263E-17	-	5.762E-19
$\alpha_4$	8 <sup>th</sup> order term (1/mm <sup>7</sup> )	-	8.482E-23	-	2.137E-25
$\alpha_5$	10 <sup>th</sup> order term (1/mm <sup>9</sup> )	-	-3.595E-29	-	-

The aspheric terms are not measured in the optical test. They are just used to help design the optical test. There are no tolerances for these terms, which are essentially rolled into the rms figure error of the manufactured part, as measured during the optical test.

### **Surface irregularity**

The effect of the surface irregularity of the mirrors depends on the frequency of the irregularity.

#### Mid-spatial frequency errors

The mid spatial frequency errors are those errors measured by the interferometer but aren't included with the lower order figure errors. These errors, caused by both mirror polishing and the support of the mirror, affect the spot sizes in the image (encircled energy diameter). The amount of irregularity allowed for the mid spatial frequencies is defined as an rms slope specification for each mirror, including all frequencies up to some cutoff frequency, which varies for each mirror. The allowable surface irregularity (nm/cm rms) for both the polishing and support contributions individually and the net effect for each mirror is listed in Table 6.

The irregularity specifications due to the mirror supports reflect the maximum expected support errors with the current design. The polishing irregularity specifications were chosen such that the net effect of the total irregularity (including support errors) contributes 0.25 arcsecond to the 90% encircled energy diameter.

**Table 6: Mid-spatial frequency irregularity specifications**

		<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>	<b>EnW</b>	<b>ExW</b>	<b>Net</b>
		<u>Surface slopes</u>				<u>Transmitted WF</u>		
Cutoff cycles/diameter		23	22	35	29	41	22	
Surface irregularity (nm/cm rms)	Polishing	9.0	13.0	30.0	10.0	20.0	20.0	
	Support	8.8	6.2	8.5	4.4			
	Total (P&S)	12.6	14.4	31.2	10.9	20.0	20.0	
rms WF ( $\mu$ rad)	Total (P&S)	0.23	0.28	0.09	0.14	0.02	0.05	<b>0.402</b>
90% EE (arcsec)	Total (P&S)	0.144	0.171	0.058	0.087	0.010	0.033	<b>0.249</b>

The cutoff frequencies in Table 7 come from the following calculations:

$$\text{spacing at PM} = \frac{\lambda}{\text{max angle}} = \frac{500 \text{ nm}}{0.225 \text{ arcsec}} = 463 \text{ mm}$$

$$\text{Cutoff frequency} = \frac{\text{system aperture}}{\text{spacing at PM}} = \frac{10000\text{mm}}{463\text{mm}} = 21.6 \text{ cycles across beam footprint}$$

The cutoff frequency is the number of cycles across the footprint of the beam on each of the mirrors. The actual number of cycles across the clear aperture of each mirror is larger by the correct ratio accordingly:

$$\text{Cutoff frequency} = 21.6 \text{ cycles} \frac{\text{clear aperture diameter}}{\text{beam footprint diameter}}.$$

The contribution to wavefront error (angle on the sky) is twice the surface error, however this effect is scaled by the size of the beam footprint on the optic to the diameter of the entrance pupil (primary mirror) due to the Lagrange invariant.

$$WF_{rms} \text{ (in } \mu\text{rad)} = 2 * \left( \frac{\text{Irregularity}_{rms} \text{ (in nm/cm)}}{10} \right) * \left( \frac{\text{beam footprint diameter}}{\text{system aperture diameter}} \right)$$

Assuming Gaussian image statistics, the 90% EE diameter is 3×rms radius:

$$90\% \text{EE diameter (in arcsec)} = \frac{3 \cdot WF_{rms} \text{ (in } \mu\text{rad)}}{4.848 \mu\text{rad/arcsec}}.$$

#### Surface finish

The highest spatial frequencies of the irregularity correspond to surface finish or roughness, and cause wide angle scatter. Wide angle light is scattered out of the image and lost completely. Thus, it appears somewhat similar to lower mirror reflectivity and affects all field angles equally. The total scattered light is  $\sigma^2$  where  $\sigma$  is the rms wavefront in radians. Since the total wavefront error is twice the surface error for a mirror on reflection, the equation for rms surface is:

$$\text{loss} = \sigma^2 = \left( 2 \cdot \text{rms}_{\text{surface in nm}} \frac{2\pi}{\lambda} \right)^2.$$

For a window (or a lens) with an index of approximately 1.5, the transmitted wavefront error is half of the surface errors. However, there are also two surfaces to take into account. The equation for the loss is thus

$$\text{loss} = \sigma^2 = 2 \left( \frac{\text{rms}_{\text{surface in nm}}}{2} \frac{2\pi}{\lambda} \right)^2.$$

The total loss allowed in the system is 10%. The tolerances, shown in Table 7, were set so the total loss is 9.5% (0.5% margin). The rms surface is calculated for each mirror or window, based on the allocated percent loss, assuming a wavelength  $\lambda$  of 500nm.

**Table 7: Surface finish tolerances**

	M2	M3	M4	M5	EnW	ExW
	surface				transmitted WF	
Loss budget (%)	2	3	2	2	0.25	0.25
rms surface finish (nm)	5.6	6.9	5.6	5.6	5.6	5.6

**Surface sag**

A few point pairs of the sag ( $z$ ) vs. radial distance ( $r$ ) on each surface are listed in the following table for an extra check. This is especially important check in order to unambiguously define the signs of the higher order coefficients. The numbers are calculated using the equations above, and verified in Zemax (using the SSAG operand in the merit function). The first and last rows are the clear aperture inner and outer radii (and the two in between points are just in between points).

**Table 8: Sag values for each of the mirrors (all values in mm)**

M2		M3		M4		M5	
$r$ (mm)	$z(r)$ (mm)	$r$ (mm)	$z(r)$ (mm)	$r$ (mm)	$z(r)$ (mm)	$r$ (mm)	$z(r)$ (mm)
163	5.077050	162.5	6.426311	15	-0.298516	127	10.926053
265	13.455065	270	17.402220	50	-3.302396	230	36.292224
375	27.060975	380	33.438868	90	-10.587882	335	78.668223
482.5	45.059276	490	53.322531	122	-19.217538	440	140.136615

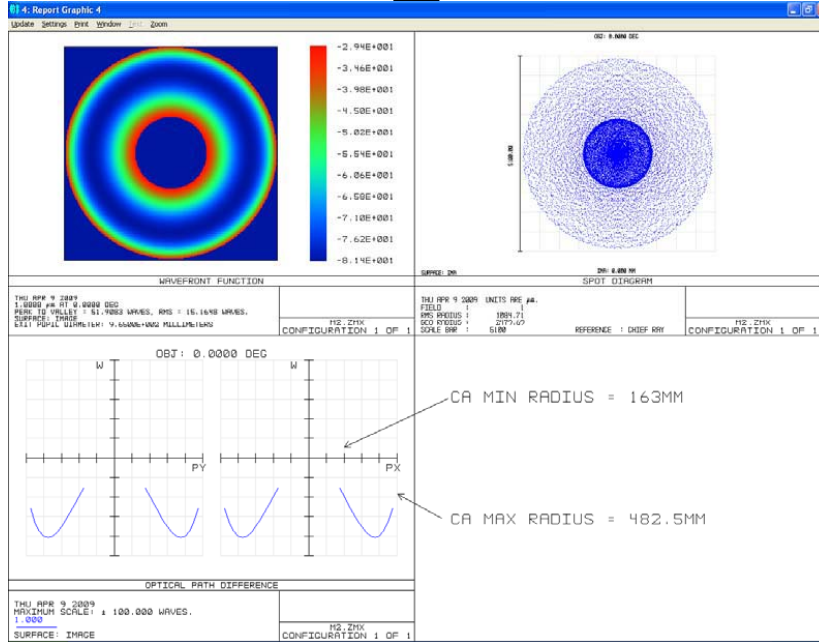
**Aspheric departure**

The follow table lists information about the aspheric departure of each mirror. The method of finding the aspheric departure of each mirror in Zemax is described in the appendix. The wavefront maps showing the peak to valley and root mean square aspheric departure for each surface, the spot diagrams used for calculating the aspheric slopes and the OPD plots for each mirror follows the table.

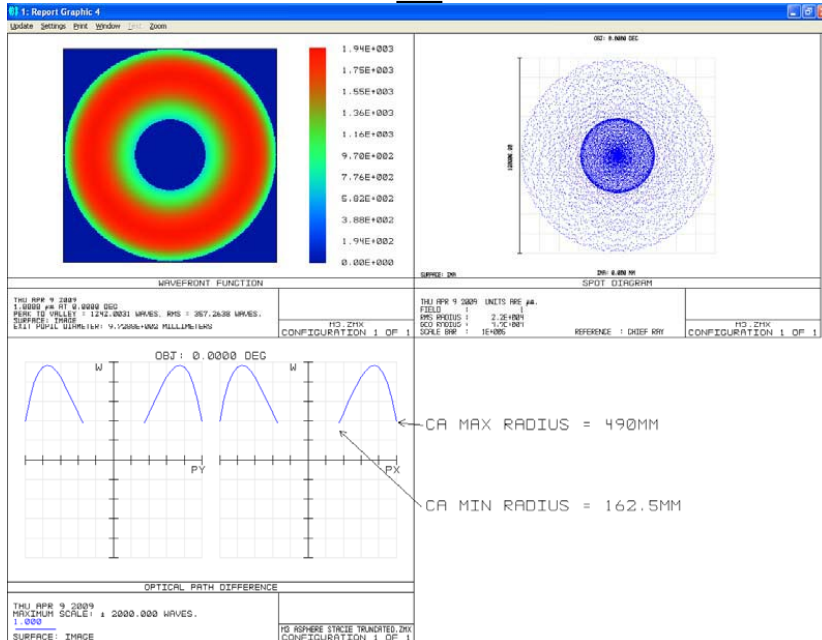
**Table 9: Aspheric departure information**

	M2	M3	M4	M5
Aspheric departure PV ( $\mu\text{m}$ )	52	1242	241	742
Aspheric departure RMS ( $\mu\text{m}$ )	15	357	71	218
Radius of best fit sphere (mm)	2604.04	2308.82	397.00	762.40
Distance from focal plane to plane of minimum OPD (over the clear aperture area) (mm)	-16.76	276.32	20.30	20.30
RMS spot radius in plane of minimum OPD (mm)	1.08	22.07	2.62	4.90
RMS aspheric slope (mrad) (spot radius/best fit sphere)	0.42	9.56	6.61	6.43
Geometric spot radius in plane of minimum OPD (mm)	2.48	49.35	6.14	11.77
Aspheric slope (mrad) (geometric spot radius/best fit sphere)	0.95	21.37	15.46	15.44

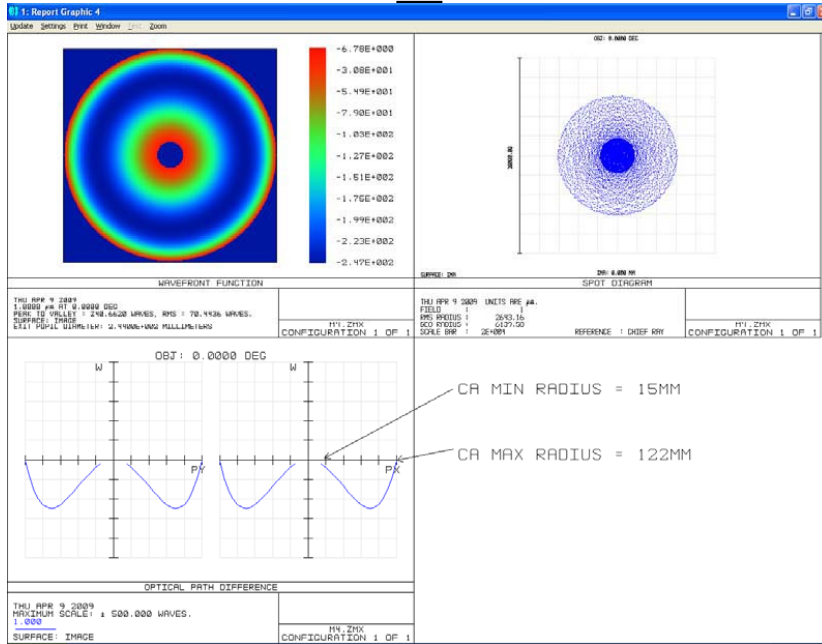
**M2:**



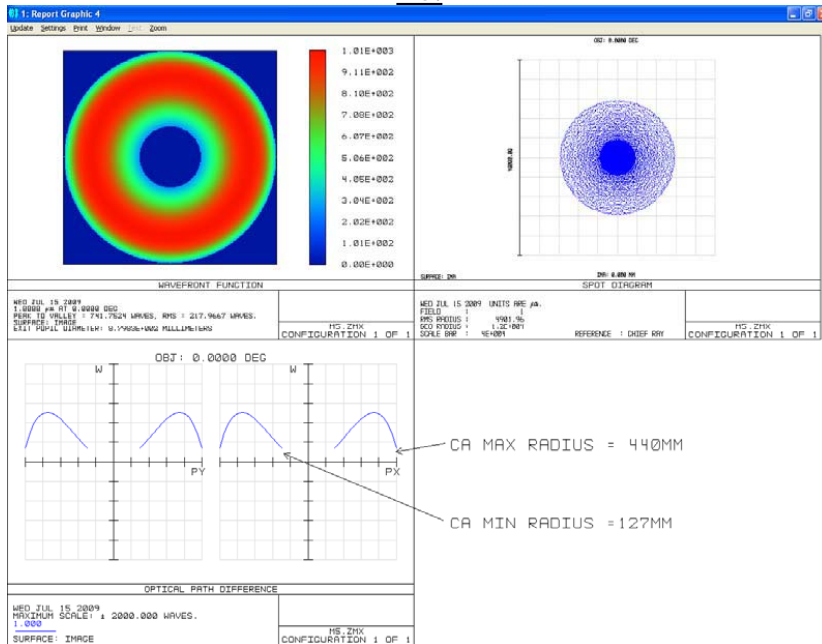
**M3:**



**M4:**



**M5:**



***Optical surface to mechanical part specifications***

This section describes how well the optical surface needs to be matched the mechanical part. Errors in the centration of the optical surface will cause the mirror to mechanically not be centered in the system, even when optically aligned. For areas with tight tolerances, this will cause the mirror substrate to vignette rays that should actually pass by the outside or inside the hole of the mirror. Therefore, the tolerances for the mirrors with the critical apertures (M2 and M4) need to be tighter than for the other two mirrors. A different report will document the method of using the kinematic mounting of mirror plugs on the inside holes of the mirrors during the system alignment process and the repeatability and tolerances involved.

**Table 10: Optical to mechanical specifications**

	<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>
Allowable decenter (centration) in mm	0.5	3	0.5	3

***Windows***

The nominal characteristics of the entrance and exit windows are listed in Table 11.

**Table 11: Entrance and Exit window specifications**

	Entrance window	Exit window
Material	Fused silica	Fused silica
Thickness	2 mm	3 mm
Clear aperture diameter	160 mm	265 mm
Overall part diameter	<b>TBD</b>	<b>TBD</b>

The tolerances for the entrance and exit windows are listed in Table 12. The windows may also have transmitted wavefront errors (from surface irregularities or index variations), but this effect is included in the 0.25 arcsecond figure specification.

**Table 12: Tolerances on the Entrance and Exit windows**

Wedge (amount allowed)	1.0 arcminute
Thickness (measurement uncertainty)	0.1 mm

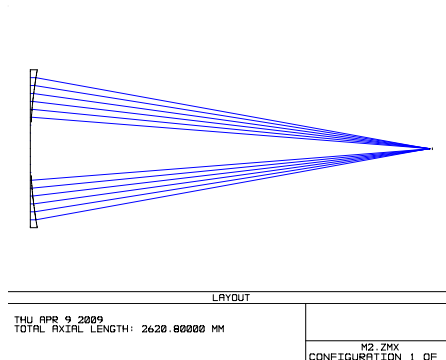


**Appendix:**

How to find the aspheric departure maps:

Create a zero index glass, and the rays leaving it will be perpendicular to the optical surface. The “zero” index glass should actually have a very small index (like 0.000001) so that Zemax does not have other problems, like in finding the exit pupil (the plane for which the OPD is referenced). Then, optimize the distance to the focal plane to minimize the rms wavefront error (OPD). Sufficient sampling for wavefront error should be used in the merit function. (Another choice to minimize the rms spot radius would give a slightly different value for aspheric departure.) This distance is the radius of the best fit sphere and would be equal to the radius of curvature if not for the conic and other higher order aspheric terms. The wavefront map in the focal plane shows the PV and RMS aspheric departure and the spot diagram show numbers that can be used for calculating the aspheric slope. (Also, it is important to set the correct clear aperture on the mirror surface.)

Surf. Type	Comment	Radius	Thickness	Glass	Semi-Diameter	Conic	Par 0 (umrad)	Par 1 (umrad)
OBJ	Standard	Infinity	Infinity		0.000000	0.000000		
STO*	Standard	Infinity	0.000000	0.00,0.0	492.207200	0.000000		
C*	Standard	2620.000000	2620.000000	F	510.000000	0.643000		
S	Standard	Infinity	-16.757585	F	5.661267	0.000000		
IMA	Standard	Infinity	-		2.495424	0.000000		



## APPENDIX E

## HET WFC FINITE ELEMENT ANALYSIS RESULTS

The system is nominally aligned at an elevation angle of  $35^\circ$ . As the pointing changes, the mirrors in the WFC move together in a correlated manner due to the changing gravity vector. These motions have been analyzed by Dr. Rob Stone using FEA models of the mechanical support design.

The following tables list the motions of each of the mirrors for different pointing configurations. The degrees of freedom are vertex decenters ( $U_x$ ,  $U_y$ , and  $U_z$ ) in units of  $\mu\text{m}$  and rotations ( $R_x$ ,  $R_y$ , and  $R_z$ ) about the vertices in units of  $\mu\text{rad}$ .

Table E.1: Mirror displacement at  $35^\circ$  in plane tilt and  $0^\circ$  out of plane tilt (Nominal design).

Mirror	$U_x$	$U_y$	$U_z$	$R_x$	$R_y$	$R_z$
M2	0	0	0	0	0	0
M3	0	0	0	0	0	0
M4	0	0	0	0	0	0
M5	0	0	0	0	0	0

Table E.2: Mirror displacement at  $26.5^\circ$  in plane tilt and  $0^\circ$  out of plane tilt.

Mirror	$U_x$	$U_y$	$U_z$	$R_x$	$R_y$	$R_z$
M2	0.241	16.096	17.518	-11.280	2.743	-1.901
M3	1.270	38.151	12.662	-10.550	-2.577	0.640
M4	-0.605	4.818	15.801	-7.457	0.053	-0.173
M5	-1.634	1.373	17.219	-12.800	3.404	-2.086

Table E.3: Mirror displacement at  $43.5^\circ$  in plane tilt and  $0^\circ$  out of plane tilt.

Mirror	$U_x$	$U_y$	$U_z$	$R_x$	$R_y$	$R_z$
M2	-0.081	-15.253	-20.744	11.480	-2.537	2.332
M3	-1.105	-34.849	-15.334	10.150	2.679	-0.942
M4	0.505	-4.262	-19.182	6.873	-0.076	0.341
M5	1.672	-0.876	-20.277	13.140	-3.144	2.516

Table E.4: Mirror displacement at  $35^\circ$  in plane tilt and  $8.5^\circ$  out of plane tilt.

Mirror	$U_x$	$U_y$	$U_z$	$R_x$	$R_y$	$R_z$
M2	37.846	2.233	-2.183	2.290	26.360	9.913
M3	54.432	-0.965	-1.868	-0.363	15.340	10.030
M4	17.048	0.620	-1.916	-0.008	8.754	9.419
M5	10.188	4.615	-2.158	2.865	31.280	9.841

Table E.5: Mirror displacement at  $29^\circ$  in plane tilt and  $6^\circ$  out of plane tilt.

Mirror	$U_x$	$U_y$	$U_z$	$R_x$	$R_y$	$R_z$
M2	28.702	13.040	11.466	-6.370	21.830	6.047
M3	41.986	25.933	8.131	-7.743	9.720	8.084
M4	12.459	3.813	10.470	-5.231	6.652	6.958
M5	6.518	4.374	11.239	-7.052	26.000	5.861

Table E.6: Mirror displacement at  $41^\circ$  in plane tilt and  $6^\circ$  out of plane tilt.

Mirror	$U_x$	$U_y$	$U_z$	$R_x$	$R_y$	$R_z$
M2	24.567	-9.368	-15.420	9.498	15.360	8.021
M3	34.671	-25.578	-11.516	6.940	11.850	5.924
M4	11.478	-2.675	-14.127	4.906	5.655	6.348
M5	7.803	2.311	-15.105	11.020	18.130	8.103

Table E.7: Positive temperature change of  $\Delta T = 20^\circ\text{C}$ , WFC at nominal orientation (35° in plane tilt, 0° out of plane tilt).

Mirror	$U_x$	$U_y$	$U_z$	$R_x$	$R_y$	$R_z$
M2	0.000	0.000	27.559	0.000	0.000	-3.580
M3	0.000	0.000	-18.476	0.001	0.000	-3.388
M4	0.000	0.000	-5.580	0.003	0.001	4.743
M5	0.000	0.000	-5.573	0.000	0.000	4.132

## APPENDIX F

## ZERNIKE POLYNOMIALS

While Seidel aberrations are simple and useful for some applications, they are not very practical for those engineers that desire to describe their optical systems as a sum of aberrations. Because the Seidel aberrations are not orthogonal, they are not very good for numerically fitting to measured data. That is, in order to describe high frequency aberrations in a system, more terms are needed, but the coefficients for the terms change depend on how many terms are used in the fit. For this reason, few optical engineers use Seidel aberrations to describe real optical systems. Usually, Zernike polynomials are the solution to this problem because they are orthonormal. Orthogonal functions are explained in Chapter 6.

Zernike polynomials were introduced by Frits Zernike for use in phase contrast microscopy (Zernike, 1934, 2002). (Zernike, although a chemist, was awarded the Nobel prize in physics in 1953 for his work in phase contrast microscopy!) Nijboer (1943, 1947) and Nienhuis & Nijboer (1949) used Zernike polynomials in their studies of the diffraction theory of aberrations.

Zernike polynomials are very commonly used because they are two dimensional functions orthogonal over a circle (the most common shape for optical elements) and the lowest order Zernike polynomials are very similar the the Seidel aberrations. Zernike polynomials are functions of two variables which describe the location in the pupil (a unit circle). Typically, polar coordinates are used and the two variables are

$\rho$  and  $\phi$ , although occasionally Cartesian coordinates ( $x$  and  $y$ ) are used.

Zernike polynomials are composed of a product of a function with radial dependence and one with azimuthal dependence. The radial function is  $R_n^m(\rho)$  and the azimuthal function is given by either  $\cos m\phi$  or  $\sin m\phi$ . The positive integers  $n$  and  $m$  describe the maximum degree of  $\rho$  and the azimuthal frequency, respectively. The Zernike polynomials are ordered by an index  $j$  (which is a function of  $n$  and  $m$ ), such that an even value for  $j$  corresponds to a polynomial with angular dependence  $\cos m\phi$  and an odd value for  $j$  corresponds to  $\sin m\phi$ :

$$Z_j(\rho, \phi) = \begin{cases} \sqrt{2(n+1)} R_n^m(\rho) \cos m\phi & \text{even } j, m \neq 0; \\ \sqrt{2(n+1)} R_n^m(\rho) \sin m\phi & \text{odd } j, m \neq 0; \\ \sqrt{n+1} R_n^0(\rho) & m = 0 \end{cases} \quad (\text{F.1})$$

where the radial function is

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+m}{2} - s\right)! \left(\frac{n-m}{2} - s\right)!} \rho^{n-2s}. \quad (\text{F.2})$$

The polynomials are listed in order of increasing index  $n$  and for given  $n$ , in order of increasing  $m$ .

Sometimes the Zernike polynomials are listed as  $Z_n^m(\rho, \phi)$  instead of  $Z_j(\rho, \phi)$  where negative values of  $m$  are used for the terms with  $\sin m\phi$ . This leads to a changed ordering the terms, if the same rule for listing in order of increasing  $m$  is used, which has led to different standards for Zernike polynomials over the years. (To make matters worse, different normalizations for the polynomials have also

been used.) This dissertation uses the notation given above, which is the same as Noll (1976) and the “Standard Zernike Polynomials” in ZEMAX (2009) shown in Table F.1. Zernike polynomials may also be written in Cartesian coordinates, as shown in Table F.2. This form is useful for the derivations in Chapter 6.

Table F.1: Orthonormal Zernike Circle Polynomials in polar form.

$j$	$n$	$m$	$Z_j(\rho, \phi)$	Aberration Name
1	0	0	1	piston
2	1	1	$2\rho \cos \phi$	distortion ( $x$ tilt)
3	1	1	$2\rho \sin \phi$	distortion ( $y$ tilt)
4	2	0	$\sqrt{3}(2\rho^2 - 1)$	field curvature (defocus)
5	2	2	$\sqrt{6}\rho^2 \sin 2\phi$	primary astigmatism at $45^\circ$
6	2	2	$\sqrt{6}\rho^2 \cos 2\phi$	primary astigmatism at $0^\circ$
7	3	1	$\sqrt{8}(3\rho^3 - 2\rho) \sin \phi$	primary $y$ coma
8	3	1	$\sqrt{8}(3\rho^3 - 2\rho) \cos \phi$	primary $x$ coma
9	3	3	$\sqrt{8}\rho^3 \sin 3\phi$	
10	3	3	$\sqrt{8}\rho^3 \cos 3\phi$	
11	4	0	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$	primary spherical
12	4	2	$\sqrt{10}(4\rho^4 - 3\rho^2) \cos 2\phi$	secondary astigmatism at $0^\circ$
13	4	2	$\sqrt{10}(4\rho^4 - 3\rho^2) \sin 2\phi$	secondary astigmatism at $45^\circ$
14	4	4	$\sqrt{10}\rho^4 \cos 4\phi$	
15	4	4	$\sqrt{10}\rho^4 \sin 4\phi$	
16	5	1	$\sqrt{12}(10\rho^5 - 12\rho^3 + 3\rho) \cos \phi$	secondary $x$ coma
17	5	1	$\sqrt{12}(10\rho^5 - 12\rho^3 + 3\rho) \sin \phi$	secondary $y$ coma
18	5	3	$\sqrt{12}(5\rho^5 - 4\rho^3) \cos 3\phi$	
19	5	3	$\sqrt{12}(5\rho^5 - 4\rho^3) \sin 3\phi$	
20	5	5	$\sqrt{12}\rho^5 \cos 5\phi$	
21	5	5	$\sqrt{12}\rho^5 \sin 5\phi$	
22	6	0	$\sqrt{7}(20\rho^6 - 30\rho^4 + 12\rho^2 - 1)$	secondary spherical

As mentioned earlier, Zernike polynomials are useful in discussions of optical aberrations because they are similar to balanced Seidel aberrations for minimum wavefront variance. For example, the Zernike polynomial term  $Z_{11}$  for spherical aberration ( $\rho^4$ ), includes the right amount of defocus ( $\rho^2$ ) to balance the spherical

aberration. Similarly,  $y$  coma ( $\rho^3 \sin \phi$ ) is balanced by the right amount of tilt ( $\rho \sin \phi$ ) for minimum variance in term  $Z_7(\rho, \phi)$ . Minimizing the variance of an optical wavefront is equivalent to maximizing the Strehl ratio, which is a measure of the light in the core part of an image (Mahajan, 1982, 1983, 1993). (To be precise, this is only exactly true for small aberrations.) The equations that convert Seidel aberrations to Zernike polynomials and vice versa are given by Mahajan's *Zernike Polynomials and Wavefront Fitting* chapter in *Optical Shop Testing*, Third Edition (Malacara, 2007).

Table F.2: Orthonormal Zernike Circle Polynomials in Cartesian form. For brevity, all terms that can expand to  $x^2 + y^2 = \rho^2$  are left in polar form.

$j$	$n$	$m$	$Z_j(\rho, \phi)$	<b>Aberration Name</b>
1	0	0	1	piston
2	1	1	$2x$	distortion ( $x$ tilt)
3	1	1	$2y$	distortion ( $y$ tilt)
4	2	0	$\sqrt{3}(2\rho^2 - 1)$	field curvature (defocus)
5	2	2	$2\sqrt{6}xy$	primary astigmatism at $45^\circ$
6	2	2	$\sqrt{6}(x^2 - y^2)$	primary astigmatism at $0^\circ$
7	3	1	$\sqrt{8}y(3\rho^2 - 2)$	primary $y$ coma
8	3	1	$\sqrt{8}x(3\rho^2 - 2)$	primary $x$ coma
9	3	3	$\sqrt{8}y(x^2 - y^2)$	
10	3	3	$\sqrt{8}x(x^2 - 3y^2)$	
11	4	0	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$	primary spherical
12	4	2	$\sqrt{10}(4\rho^2 - 3)(x^2 - y^2)$	secondary astigmatism at $0^\circ$
13	4	2	$2\sqrt{10}(4\rho^2 - 3)xy$	secondary astigmatism at $45^\circ$
14	4	4	$\sqrt{10}(\rho^4 - 8x^2y^2)$	
15	4	4	$4\sqrt{10}xy(x^2 - y^2)$	

The Zernike polynomials are normalized such that the coefficient of a term represents its standard deviation and the mean value of each aberration (except piston) is zero. The sum of the squares of the orthonormal coefficients yields the aberration



variance.

It is helpful to visualize these functions, so the lowest order Zernike polynomials are shown in Figure F.1. Each row in Figure F.1 shows the polynomials with the same value of  $n$ . The figure is triangle-shaped because there are increasing number of  $m$ -values for the Zernike polynomials as  $n$  increases.

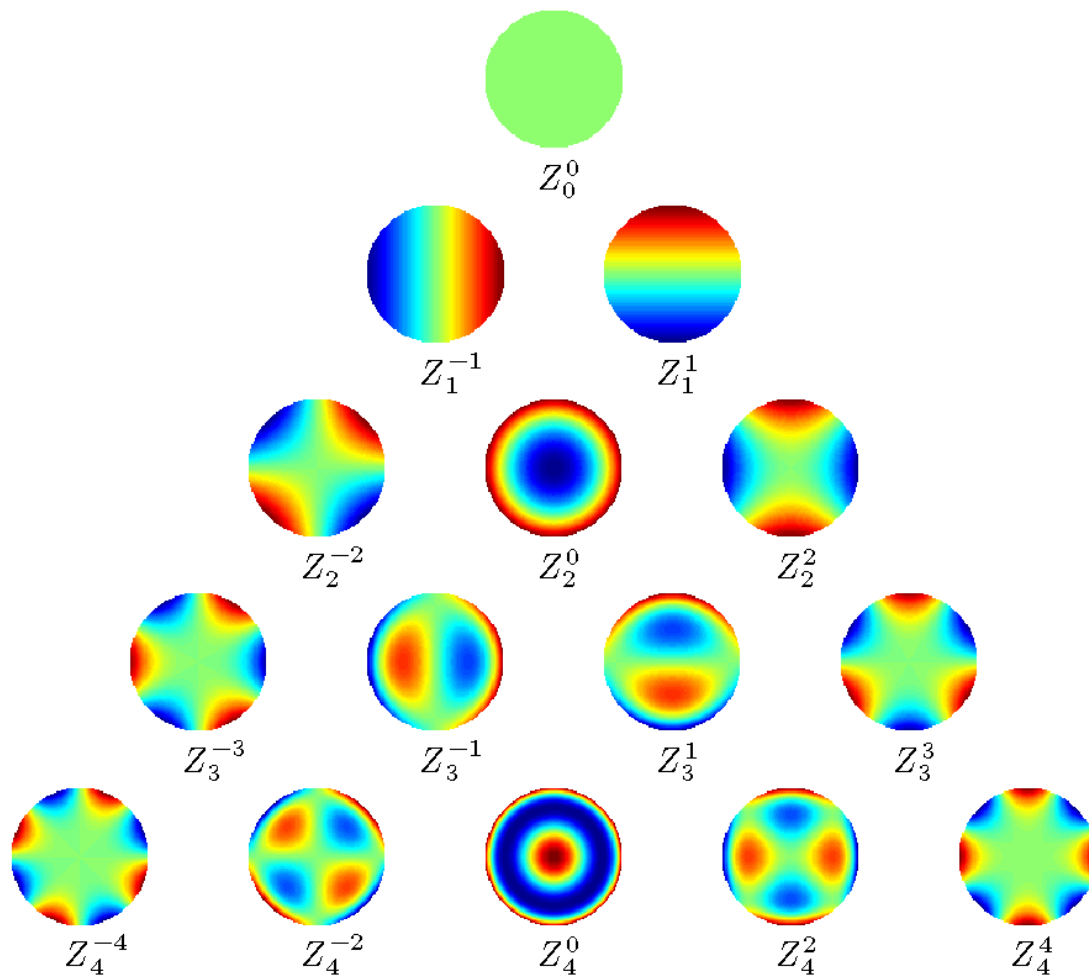


Figure F.1: Zernike polynomial wavefront maps.

Because Zernike polynomials are orthogonal over the unit circle, they are sometimes called Zernike circle polynomials. This is to differentiate them from Zernike

polynomials orthogonal over areas other than unit circle, which have been derived for other applications. Some of these variations, related to telescopes, are elliptical Zernike polynomials, annular Zernike polynomials (useful when the telescope pupil has a central obscuration) (Mahajan, 1981) and hexagonal Zernike polynomials (useful when the telescope mirrors are composed of hexagonal segments) (Mahajan & Dai, 2006).

## APPENDIX G

## VECTOR MULTIPLICATION

Vector multiplication is a mathematical operation, invented by Shack, which is used in the vectorial form of the wavefront expansion (Thompson, 1980, 2005). Vector multiplication is a vector operation which is unique from both dot products and cross products. Multiplying two vectors gives a third vector, coplanar with the first two, where the length of the vector is the product of the lengths of the two original vectors and the angle of the resulting vector is given by the sum of the angles of the two vectors. For example, if the two multiplying vectors are  $\vec{A}$  and  $\vec{B}$ :

$$\vec{A} = a e^{i\alpha} = a_x \hat{i} + a_y \hat{j} \quad (\text{G.1})$$

$$\vec{B} = b e^{i\beta} = b_x \hat{i} + b_y \hat{j} \quad (\text{G.2})$$

where  $\hat{i}$  and  $\hat{j}$  are the unit vectors along  $\vec{x}$  and  $\vec{y}$  respectively, then the multiplication between those two vectors is defined to be

$$\vec{A}\vec{B} = a b e^{i(\alpha+\beta)}. \quad (\text{G.3})$$

Equations G.1–G.3 are always true, no matter what the convention for measuring the angle is. However, depending on whether angles are measured clockwise (CW) from the  $y$ -axis or counterclockwise (CCW) from the  $x$ -axis, the resulting vector

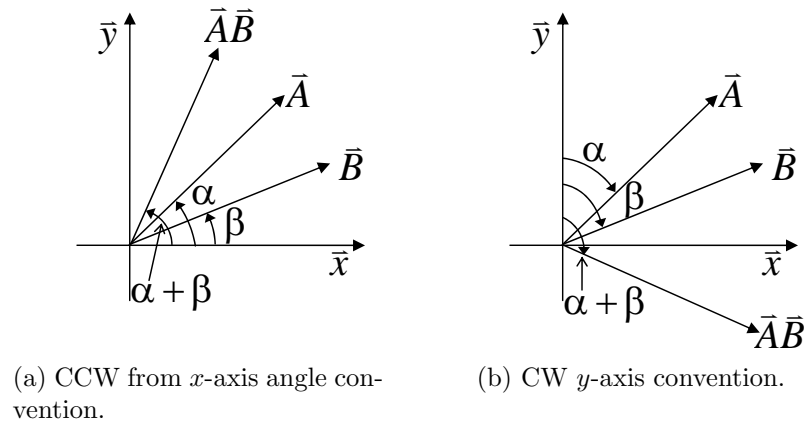


Figure G.1: Vector multiplication displayed in two different coordinate systems. The vectors  $\vec{A}$  and  $\vec{B}$  are each unit length, so the vector product is also unit length.

will be at a different angle, as shown in Figure G.1.

The fact that the resulting vector in a vector multiplication depends on the convention chosen does not invalidate the operation of vector multiplication or any of the equations derived by Thompson. It just means that one must be extra careful when choosing one convention to completely stick with it and use all the correct equations for that convention.

### G.1 COUNTERCLOCKWISE FROM X-AXIS ANGLE CONVENTION

The equations in this section are derived using the convention where the angle is measured counterclockwise from the  $x$ -axis, as shown in Figure G.1a. This convention is used in this dissertation because it is used by optical software programs, including CODE V and ZEMAX.

For a vector  $\vec{A}$ , the  $x$  and  $y$  components can be written for the vector and its

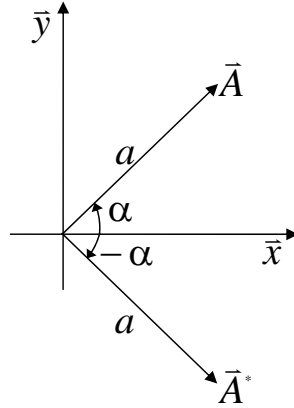


Figure G.2: The angle of the vector is defined counterclockwise from the  $x$ -axis.

complex conjugate:

$$\begin{aligned}\vec{A} &= a e^{i\alpha} = a_x \hat{i} + a_y \hat{j} \\ &= a \cos \alpha \hat{i} + a \sin \alpha \hat{j}\end{aligned}\tag{G.4}$$

$$\begin{aligned}\vec{A}^* &= a e^{-i\alpha} = a \cos(-\alpha) \hat{i} + a \sin(-\alpha) \hat{j} \\ &= a \cos \alpha \hat{i} - a \sin \alpha \hat{j} \\ &= a_x \hat{i} - a_y \hat{j}.\end{aligned}\tag{G.5}$$

In this case, the complex conjugate flips the vector across the  $x$ -axis, as shown in Figure G.2. For the vector product  $\vec{A}\vec{B}$ , Equation G.3 can be written in terms of the  $x$  and  $y$  components of the  $\vec{A}$  and  $\vec{B}$  vectors as follows:

$$\begin{aligned}\vec{A}\vec{B} &= ab e^{i(\alpha+\beta)} \\ &= ab \cos(\alpha + \beta) \hat{i} + ab \sin(\alpha + \beta) \hat{j} \\ &= ab (\cos \alpha \cos \beta - \sin \alpha \sin \beta) \hat{i} + ab (\sin \alpha \cos \beta + \cos \alpha \sin \beta) \hat{j} \\ &= (a_x b_x - a_y b_y) \hat{i} + (a_y b_x + a_x b_y) \hat{j}.\end{aligned}\tag{G.6}$$

Similarly, a vector cubed may be written:

$$\begin{aligned}
 \vec{A}^3 &= a e^{i(3\alpha)} \\
 &= a^3 \cos(3\alpha) \hat{i} + a^3 \sin(3\alpha) \hat{j} \\
 &= a^3 (4 \cos^3 \alpha - 3 \cos \alpha) \hat{i} + a^3 (3 \sin \alpha - 4 \sin^3 \alpha) \hat{j} \quad (\text{G.7}) \\
 &= [4 a_x^3 - 3 a_x (a_x^2 + a_y^2)] \hat{i} + [3 (a_x^2 + a_y^2) a_y - 4 a_y^3] \hat{j} \\
 &= (a_x^3 - 3 a_x a_y^2) \hat{i} + (3 a_x^2 a_y - a_y^3) \hat{j}
 \end{aligned}$$

Therefore, the equations for  $\vec{H}$  are:

$$\vec{H} = H_x \hat{i} + H_y \hat{j} \quad (\text{G.8})$$

$$\vec{H}^2 = (H_x^2 - H_y^2) \hat{i} + 2 H_x H_y \hat{j} \quad (\text{G.9})$$

$$\vec{H}^3 = (H_x^3 - 3 H_x H_y^2) \hat{i} + (3 H_x^2 H_y - H_y^3) \hat{j} \quad (\text{G.10})$$

$$\vec{H}^* = H_x \hat{i} - H_y \hat{j} \quad (\text{G.11})$$

and for  $\vec{\rho}$  are:

$$\vec{\rho} = \rho \cos \phi \hat{i} + \rho \sin \phi \hat{j} \quad (\text{G.12})$$

$$\vec{\rho}^2 = \rho \cos 2\phi \hat{i} + \rho \sin 2\phi \hat{j} \quad (\text{G.13})$$

$$\vec{\rho}^3 = \rho \cos 3\phi \hat{i} + \rho \sin 3\phi \hat{j} \quad (\text{G.14})$$

$$\vec{\rho}^* = \rho \cos \phi \hat{i} - \rho \sin \phi \hat{j}. \quad (\text{G.15})$$

## G.2 CLOCKWISE FROM Y-AXIS ANGLE CONVENTION

Measuring the angle clockwise from the  $y$ -axis is sometimes used in analysis of aberrations by many, including Thompson. This has been done historically because the rays were drawn in the tangential  $y - z$  plane. An angle of zero (measured from the  $y$ -axis) meant that the ray was in the tangential plane. Skew rays, which have nonzero angles, were those that had sagittal components. The equations are needed to show how to convert properly from the CW from  $y$ -axis angle convention into a CCW from  $y$ -axis convention, which will be done in the next section.

For a vector  $\vec{A}$ , the  $x$  and  $y$  components can be written for the vector and its complex conjugate:

$$\begin{aligned}\vec{A} &= a e^{i\alpha} = a_x \hat{i} + a_y \hat{j} \\ &= a \sin \alpha \hat{i} + a \cos \alpha \hat{j}\end{aligned}\tag{G.16}$$

$$\begin{aligned}\vec{A}^* &= a e^{-i\alpha} = a \sin(-\alpha) \hat{i} + a \cos(-\alpha) \hat{j} \\ &= -a \sin \alpha \hat{i} + a \cos \alpha \hat{j} \\ &= -a_x \hat{i} + a_y \hat{j}.\end{aligned}\tag{G.17}$$

In this case, the complex conjugate flips the vector across the  $y$ -axis, as shown in Figure G.3. For the vector product  $\vec{A}\vec{B}$ , Equation G.3 can be written in terms

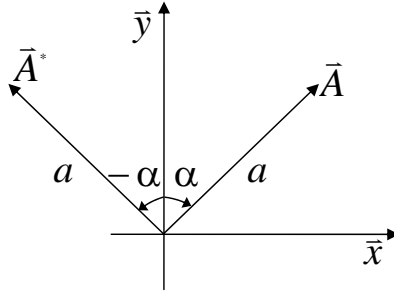


Figure G.3: The angle of the vector is defined clockwise from the  $y$ -axis.

of the  $x$  and  $y$  components of the  $\vec{A}$  and  $\vec{B}$  vectors as follows:

$$\begin{aligned}
 \vec{A}\vec{B} &= ab e^{i(\alpha+\beta)} \\
 &= ab \sin(\alpha + \beta) \hat{i} + ab \cos(\alpha + \beta) \hat{j} \\
 &= ab (\sin \alpha \cos \beta + \cos \alpha \sin \beta) \hat{i} + ab (\cos \alpha \cos \beta - \sin \alpha \sin \beta) \hat{j} \\
 &= (a_x b_y + a_y b_x) \hat{i} + (a_y b_y - a_x b_x) \hat{j}.
 \end{aligned} \tag{G.18}$$

Note that this result is different from that in Equation G.6. Similarly, a vector cubed may be written:

$$\begin{aligned}
 \vec{A}^3 &= a e^{i(3\alpha)} \\
 &= a^3 \sin(3\alpha) \hat{i} + a^3 \cos(3\alpha) \hat{j} \\
 &= a^3 (3 \sin \alpha - 4 \sin^3 \alpha) \hat{i} + a^3 (4 \cos^3 \alpha - 3 \cos \alpha) \hat{j} \\
 &= [3 a_x (a_x^2 + a_y^2) - 4 a_x^3] \hat{i} + [4 a_y^3 - 3 a_y (a_x^2 + a_y^2)] \hat{j} \\
 &= (3 a_x a_y^2 - a_x^3) \hat{i} + (a_y^3 - 3 a_x^2 a_y) \hat{j}.
 \end{aligned} \tag{G.19}$$



Therefore, the equations for  $\vec{H}$  are:

$$\vec{H} = H_x \hat{i} + H_y \hat{j} \quad (\text{G.20})$$

$$\vec{H}^2 = 2 H_x H_y \hat{i} + (H_y^2 - H_x^2) \hat{j} \quad (\text{G.21})$$

$$\vec{H}^3 = (3 H_x H_y^2 - H_x^3) \hat{i} + (H_y^3 - 3 H_x^2 H_y) \hat{j} \quad (\text{G.22})$$

$$\vec{H}^* = -H_x \hat{i} + H_y \hat{j} \quad (\text{G.23})$$

and for  $\vec{\rho}$  are:

$$\vec{\rho} = \rho \sin \phi \hat{i} + \rho \cos \phi \hat{j} \quad (\text{G.24})$$

$$\vec{\rho}^2 = \rho^2 \sin 2\phi \hat{i} + \rho^2 \cos 2\phi \hat{j} \quad (\text{G.25})$$

$$\vec{\rho}^3 = \rho^3 \sin 3\phi \hat{i} + \rho^3 \cos 3\phi \hat{j} \quad (\text{G.26})$$

$$\vec{\rho}^* = -\rho \sin \phi \hat{i} + \rho \cos \phi \hat{j}. \quad (\text{G.27})$$

### G.3 SWITCHING THE ANGLE CONVENTION FROM CW FROM THE Y- AXIS TO CCW FROM THE X-AXIS

In his thesis, Tessieres (2003) states that to consider a right-handed convention (CCW from the  $x$ -axis) “that will lead to replace  $x$  by  $-x$ . However since optical software programs usually “look” at the image plane from behind,...  $x$  is unchanged. Also, here  $\phi$  is now defined from the  $\vec{x}$  axis which will imply to replace  $\cos \phi$  by  $\sin \phi$  and vice versa. Thus to summarize, the only modification will be to exchange the expression of the cosine and sine components for each aberration.” This method does not seem to work.

In order to understand how the equations depend on the change in angle convention, it is helpful to look at some example dot products that are used in the derivations, including  $\vec{H} \cdot \vec{\rho}$  and  $\vec{H}^2 \cdot \vec{\rho}^2$ . The following equations compare the dot product  $\vec{H} \cdot \vec{\rho}$  when it is converted into (unnormalized) Zernike polynomial notation:

$$\begin{aligned} \left(\vec{H} \cdot \vec{\rho}\right)_{CCW} &= H_x \rho \cos \phi + H_y \rho \sin \phi \\ &= H_x Z_2(\rho, \phi) + H_y Z_3(\rho, \phi) \end{aligned} \quad (\text{G.28})$$

$$\begin{aligned} \left(\vec{H} \cdot \vec{\rho}\right)_{CW} &= H_x \rho \sin \phi + H_y \rho \cos \phi \\ &= H_y Z_2(\rho, \phi) + H_x Z_3(\rho, \phi). \end{aligned} \quad (\text{G.29})$$

In each case above, the term  $\rho \cos \phi$  converts to  $Z_2(\rho, \phi)$  and  $\rho \sin \phi$  converts to  $Z_3(\rho, \phi)$  using the numbering of Noll (1976). For this example, it does work to switch the sine and cosine terms to change the angle convention from CW from the  $y$ -axis to CCW from the  $x$ -axis. However, this is not true in general, as the next example will show. The following equations compare the dot product  $\vec{H}^2 \cdot \vec{\rho}^2$  in both conventions when it is converted into (unnormalized) Zernike polynomial notation (using the equations for  $\vec{H}^2$  and  $\vec{\rho}^2$  for each convention given earlier in this appendix):

$$\begin{aligned} \left(\vec{H}^2 \cdot \vec{\rho}^2\right)_{CCW} &= (H_x^2 - H_y^2) \rho^2 \cos 2\phi + 2 H_x H_y \rho^2 \sin 2\phi \\ &= 2 H_x H_y Z_5(\rho, \phi) + (H_x^2 - H_y^2) Z_6(\rho, \phi) \end{aligned} \quad (\text{G.30})$$

$$\begin{aligned} \left(\vec{H}^2 \cdot \vec{\rho}^2\right)_{CW} &= 2 H_x H_y \rho^2 \sin 2\phi + (H_y^2 - H_x^2) \rho^2 \cos 2\phi \\ &= 2 H_x H_y Z_5(\rho, \phi) + (H_y^2 - H_x^2) Z_6(\rho, \phi) \end{aligned} \quad (\text{G.31})$$

where in each case above, the term  $\rho^2 \sin 2\phi$  converts to  $Z_5(\rho, \phi)$  and  $\rho^2 \cos 2\phi$  converts to  $Z_6(\rho, \phi)$  using the numbering of Noll (1976).

This second example shows that it does not suffice to switch the resulting sine and cosine terms to switch the equations for the other convention. The proper way to modify the equations is to switch the  $H_x$  and  $H_y$  terms in each equation, but not do anything to the Zernike polynomial terms. This makes sense because once the  $\rho$  and  $\phi$  dependence is written as a Zernike polynomial, it does not matter what the convention is, as long as the Zernike polynomial coefficient is measured using that convention (and the coefficients do change depending on the convention chosen). This is the same as saying the value of the  $H_x$  and  $H_y$  components depend on the convention. It makes sense that the field vector components written in the equations switch from  $H_x$  to  $H_y$  just considering the that this is the change that happens for a vector by itself to switch from one convention to the other.

This section shows that it is imperative for one to understand which angle convention is being used and stick with it. Ironically, although at first it seems simple to define a convention and use it appropriately, it can be quite difficult to do correctly, even more challenging than a completely new type of math (such as vector multiplication)!

### G.3.1 USEFUL VECTOR DOT PRODUCTS AND VECTOR MULTIPLICATION PRODUCTS

For convenience, this final section lists some vector dot products and vector multiplication products that are used in Appendix J.

### Dot products

$$\vec{H} \cdot \vec{H} = H_x^2 + H_y^2 \quad (\text{G.32})$$

$$\vec{H} \cdot \vec{A} = A_x H_x + A_y H_y \quad (\text{G.33})$$

$$\vec{H} \cdot \vec{\rho} = H_x \rho \cos \phi + H_y \rho \sin \phi \quad (\text{G.34})$$

$$\vec{\rho} \cdot \vec{\rho} = \rho^2 \quad (\text{G.35})$$

$$\vec{A} \cdot \vec{\rho} = A_x \rho \cos \phi + A_y \rho \sin \phi \quad (\text{G.36})$$

### Vector multiplication products

$$\vec{A}\vec{H} = (A_x H_x - A_y H_y)\hat{i} + (A_y H_x + A_x H_y)\hat{j} \quad (\text{G.37})$$

$$\vec{A}\vec{H}^* = (A_x H_x + A_y H_y)\hat{i} + (A_y H_x - A_x H_y)\hat{j} \quad (\text{G.38})$$



```

! Define the 5 output files
outfile$ = path$ + filename$ + ".dat"
outfilexCCrot$ = path$ + filename$ + "_x_cc_rot.dat"
outfileyCCrot$ = path$ + filename$ + "_y_cc_rot.dat"
outfilexpfcrot$ = path$ + filename$ + "_x_pf_rot.dat"
outfileypfcrot$ = path$ + filename$ + "_y_pf_rot.dat"
PRINT outfile$

! Make a vector for the grid of normalized field coordinates
! (1 row only)
DECLARE fields, double, 1, size, 1
FOR i, 1, size, 1
fields(i) = -1+2*(i-1)/(size-1)
NEXT i

! Make two long vectors that list the x, y coordinates respectively
! of the all of the field coordinates in the grid
DECLARE x, double, 1, size*size, 1
DECLARE y, double, 1, size*size, 1
k=1
FOR i, 1, size,1
FOR j, 1, size, 1
x(k)=fields(i)
y(k)=fields(j)
k=k+1
NEXT j
NEXT i

! Loop through all the field points in the grid, saving the
! aberrations after groups of 12 fields are assigned
k=1
nloops =INTE(size*size/12) + 1 # (the function INTE rounds down)
FOR index, 1, nloops, 1
FORMAT 3.0
PRINT,"Set of field points = ",index, " of ",nloops
FOR j, 1, 12, 1
SYSP 102, j, x(k)*max
SYSP 103, j, y(k)*max
k=k+1
NEXT j
UPDATE ALL
GOSUB PERTURB
NEXT index

```

```

PRINT " "
PRINT "All Done!"
END

```

```

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

```

```

SUB PERTURB
! Record aberrations of the nominal sytem.
PRINT "Record aberrations of the nominal sytem."
OUTPUT outfile$ APPEND

! Go to the subroutine Zern to calculate and store the Zernike
! coefficients
GOSUB ZERN

! Usage of SURP, short for SETSURFACEPROPERTY:
! SETSURFACEPROPERTY surface, code, value1, value2
! code 10 changes the parameter values 1-4 (in value2)
! value1 is the amount of perturbation to be applied

```

```

REWIND
PRINT, "CoC Rotation about x of 1 deg for M2"
SURP 8, 10, 1, 3
OUTPUT outfilexCCrot$ APPEND
GOSUB ZERN
SURP 8, 10, 0, 3

```

```

REWIND
PRINT, "CoC Rotation about y of 1 deg for M2"
SURP 8, 10, 1, 4
OUTPUT outfileyCCrot$ APPEND
GOSUB ZERN
SURP 8, 10, 0, 4

```

```

REWIND
PRINT, "Prime Focus Rotation about x of 1 deg for M2"
SURP 4, 10, 1, 3
OUTPUT outfilexfrot$ APPEND
GOSUB ZERN
SURP 4, 10, 0, 3

```

```

REWIND
PRINT, "Prime Focus Rotation about y of 1 deg for M2"

```

```

SURP 4, 10, 1, 4
OUTPUT outfilepfrot$ APPEND
GOSUB ZERN
SURP 4, 10, 0, 4
REWIND

UPDATE EDITORS

RETURN

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

! Subroutine to get the Zernike polynomials and store them into a file
SUB ZERN
UPDATE ALL

! Usage for the GETZERNIKE command:
! GETZERNIKE maxorder,wave,field,sampling,vector,zerntype,epsilon,ref
! sampling = 3 for pupil sampling of 128x128
! zerntype = 1 for "Standard" Zernike terms

f = 1
label 1
GETZERNIKE 11,9,f,3,1,1
FORMAT 16.6
PRINT FLDX(f), FLDY(f),
i = 1
label 2

PRINT vec1(8+i),
i = i + 1
if (i<12) then goto 2
print " "
f = f+1
if (f<13) then goto 1
OUTPUT screen

RETURN

```



## APPENDIX I

## MATLAB CODE

## I.1 FITTINGPOLYNOMIALS.M MATLAB CODE

```

% FittingPolynomials7.m
% Written by Stacie Manuel
%
% Program to fit the data from a ray tracing program or a wavefront
% sensor by using a least squares fit with the equations of the
% Zernike coefficients in misaligned systems.
%
% Output: Coefficients of each field dependence for coma and
% astigmatism and graphical representation of each field dependence.
% More outputs: need to describe
%
% Original for 90" telescope: Regis Tessieres – September 24th, 2003
% Modified for NST telescope: Stacie Manuel – April 17th, 2009
% Modified for General system: Stacie Manuel – October 19th, 2009

clear all
close all

tic % Start timer
scrsz = get(0, 'ScreenSize'); % Will be used to create figures

%% This section is for the user to set some flags

show_fig=0; % Set this flag to 1 in order to show graphs (very slow!)
Hubble = 0; % Set this flag to save the Hubble figures as .eps files
HET = 1; % Set this flag to save the HET figures as .eps files
HST = 0; % Set this flag to plot the linearity of the coefficients
threshold=1e-15; % Alpha and beta coefficients below this threshold
% will be set to zero in the results to make them easier to read.
doSVD=1; % Set this flag to 1 in order to do the SVD
show_SVDfig=1; % Plot the SVD modes across the field
show_SVDfigS=0; show_SVDfigS2=0;
show_SVD_figUV=0;
show_SVDfigU2=1; show_SVDfigU2b=0; show_SVDfigV2=0;
show_DOFFig=0; % To plot the aberrations for the DOF before SVD
o=1; % To plot the orthogonal functions and save them.

```

```

nm=25; % number modes to show
%% Load the files in a structure called allthedata

% 1. Define the filename (will be unique)
%   IMPORTANT: Need to make select the current folder in MATLAB as
%   the location where the files are saved.
% 2. Choose a value for pert to describe which other files to find.
%   These files will be similar for many systems (x and y decenters
%   and tilts, etc.)
% 3. For the Hubble Telescope example where the figures should be
%   saved, select the example number (ex).

%Hubble Telescope files
%filename = 'Hubble_correctfield.dat'; pert = 0; ex = 'Ex1';
%filename = 'Hubble_Z5is3.dat'; pert = 0; ex = 'Ex2';
%filename = 'Hubble_Z5ispt6.dat'; pert = 0; ex = 'Ex3';
%filename = 'Hubble_10mmxdec.dat'; pert = 0; ex = 'Ex4';
%filename = 'Hubble_pt1mmxdec.dat'; pert = 0; ex = 'Ex5';
%filename = 'HST.dat'; pert=7; HST=1;

% NST files
%filename='NST_pt1deg.dat'; pert=2; telescope='NST';
%filename='NST.dat'; pert=4; % Moving the rotation point.

% HET files
% 0.5arcsec tilt, 0.01deg WFC and FP tilt, 0.1mm axial FP
filename = 'HET_wcorrection.dat'; pert=8;telescope='HET';

switch pert
case 0
    extensions = {''};
    perturb = {'nominal'};
case 1
    extensions = {'','xde','yde','xro','yro'};
    perturb = {'nominal','x decenter','y decenter',...
        'x rotation','y rotation'};
case 2
    extensions = {'','_x_cc_rot','_y_cc_rot',...
        '_x_pf_rot','_y_pf_rot'};
    perturb = {'nominal','x CoC rotation','y CoC rotation',...
        'x PF rotation','y PF rotation'};
case 3
    extensions = {'','M2_xde','M2_yde','M2_xro','M2_yro',...
        'M3_xde','M3_yde','M3_xro','M3_yro',...
        'M4_xde','M4_yde','M4_xro','M4_yro',...
        'M5_xde','M5_yde','M5_xro','M5_yro'};
    perturb = {'nominal','M2 x decenter','M2 y decenter',...
        'M2 x rotation','M2 y rotation','M3 x decenter',...
        'M3 y decenter','M3 x rotation','M3 y rotation',...
        'M4 x decenter','M4 y decenter','M4 x rotation',...

```

```

        'M4 y rotation', 'M5 x decenter', 'M5 y decenter', ...
        'M5 x rotation', 'M5 y rotation'};
case 4
    extensions = {'_0mm', '_100mm', '_200mm', '_300mm', ...
        '_400mm', '_573mm', '_3217mm', '_6490mm'};
    perturb = extensions;
case 5
    extensions = {'', 'system', };
    perturb = {'nominal', 'decenter', };
case 6
    extensions = {'', 'M2_xde', 'M2_yde', 'M2_xro', 'M2_yro', ...
        'M3_xde', 'M3_yde', 'M3_xro', 'M3_yro', ...
        'M4_xde', 'M4_yde', 'M4_xro', 'M4_yro', ...
        'M5_xde', 'M5_yde', 'M5_xro', 'M5_yro', 'WFC_xro', 'WFC_yro'};
    perturb = {'nominal', 'M2 x decenter', 'M2 y decenter', ...
        'M2 x rotation', 'M2 y rotation', 'M3 x decenter', ...
        'M3 y decenter', 'M3 x rotation', 'M3 y rotation', ...
        'M4 x decenter', 'M4 y decenter', 'M4 x rotation', ...
        'M4 y rotation', 'M5 x decenter', 'M5 y decenter', ...
        'M5 x rotation', 'M5 y rotation', 'WFC x rotation', ...
        'WFC y rotation'};
case 7
    extensions = {'_0mm', '_0pt001mm', '_0pt01mm', '_0pt1mm', ...
        '_1mm', '_3mm', '_5mm', '_7mm', '_10mm', '_15mm', '_20mm'};
    perturb = extensions;
case 8
    extensions = {'', 'M2_xde', 'M2_yde', 'M2_xro', 'M2_yro', ...
        'M3_xde', 'M3_yde', 'M3_xro', 'M3_yro', ...
        'M4_xde', 'M4_yde', 'M4_xro', 'M4_yro', ...
        'M5_xde', 'M5_yde', 'M5_xro', 'M5_yro', ...
        'WFC_xro', 'WFC_yro', 'FP_xro', 'FP_yro', ...
        'M2_z', 'M3_z', 'M4_z', 'M5_z', 'FP_z'};
    perturb = {'nominal', 'M2 x decenter', 'M2 y decenter', ...
        'M2 x tilt', 'M2 y tilt', 'M3 x decenter', ...
        'M3 y decenter', 'M3 x tilt', 'M3 y tilt', ...
        'M4 x decenter', 'M4 y decenter', 'M4 x tilt', ...
        'M4 y tilt', 'M5 x decenter', 'M5 y decenter', ...
        'M5 x tilt', 'M5 y tilt', 'WFC x tilt', 'WFC y tilt', ...
        'FP x tilt', 'FP y tilt', 'M2 z', 'M3 z', 'M4 z', 'M5 z', ...
        'FP z'};
end

[~,name,ext,~] = fileparts(filename);

for idx = 1:length(extensions)
    file = char(strcat(name,extensions(idx),ext));
    % Import the data from the file and hold it in a structure
    rawData = importdata(file);
    % The last few rows of the data column may contain extra entries
    % for zero field angle, which need to be deleted. The actual

```

```

% number of fields should be a square number.
numfields=(floor(sqrt(length(rawData))))^2;%%%
rawData=rawData(1:numfields,:);%%%
n_z=size(rawData,2)-2; % n_z is the # of Zernikes in the file

if n_z==11
    rawData1=unique(dataset({rawData,'x','y','Z1','Z2','Z3',...
        'Z4','Z5','Z6','Z7','Z8','Z9','Z10','Z11'}));
elseif n_z==17
    rawData1=unique(dataset({rawData,'x','y','Z1','Z2','Z3',...
        'Z4','Z5','Z6','Z7','Z8','Z9','Z10','Z11','Z12','Z13',...
        'Z14','Z15','Z16','Z17'}));
else
    disp('Unrecognized number of Zernikes in data file')
    return
end

allthedata.(genvarname([name char(extensions(idx))]))=rawData1;
numfields = length(rawData1);

% Store the data in a matrix as well (used for the SVD later)
% The 1st two columns are the fields, the next 3 are piston & tilt
% Start recording Z4 at column 6. Skip columns 16-17 (Z14 & Z15).
if n_z==11
    rawData2=double(rawData1(:,6:end));
elseif n_z==17
    rawData2=double(rawData1(:,[6:15,18:19]));
else
    disp('Unrecognized number of Zernikes in data file')
    return
end

%if (idx==16)||(idx==17)
%   rawData2(26,:)=[];
%end

M(:,idx)=rawData2(:);
end
clear file filename idx ext %rawData rawData1 rawData2

%% Loop through all the different perturbations

% Prepare the output dataset called coefficients
NameObs = strcat({'alpha'},num2str((0:4),'-%d'));
NameObs(6:8) = strcat({'beta'},num2str((0:2),'-%d'));
coefficients = dataset({'(1:8)','i'},'ObsNames',NameObs);

for index=1:length(perturb)
    %% Store the grid data in a structure

```

```

a=allthedata.(genvarname([name char(extensions(index))]));
Grid.angles=unique(a.x);

[Grid.X,Grid.Y]=meshgrid(Grid.angles,Grid.angles);
[Grid.THETA,Grid.RHO] = cart2pol(Grid.X,Grid.Y);
for idx=1:n_z % n_z is the # of Zernike polynomials in the file
    Z_num=['Z' num2str(idx)];
    temp=griddata(a.x,a.y,a.(genvarname(Z_num)),Grid.X,Grid.Y);
    temp(isnan(temp))=0;
    Grid.(genvarname(Z_num))=temp;
end
% Normalize the fields to the max field angle
a.x=a.x/max(Grid.angles);
a.y=a.y/max(Grid.angles);
Grid.X=Grid.X/max(Grid.angles);
Grid.Y=Grid.Y/max(Grid.angles);
Grid.RHO=Grid.RHO/max(Grid.angles);
Grid.mask=ones(size(Grid.RHO)); Grid.mask(Grid.RHO>1)=0;
if index==1
    nominal=a;
end
clear idx Z_num

%% Plot the measured coefficients for all Zernike terms

if show_fig==1
    figure('Name','Measured coefficients through the field',...
        'Position',[10 scrsz(4)/3 1100 600])
    for idx=1:n_z-3
        if n_z==11; subplot(2,5,idx); end
        if n_z==17; subplot(3,5,idx); end
        imagesc(Grid.angles,Grid.angles,...
            Grid.(genvarname(['Z' num2str(idx+3)])))
        axis square,title(['Z' num2str(idx+3)]),colorbar,
        set(gca,'YDir','normal')
    end
end

%% Least-squares fit (LSF) to calculate the Zernike coefficients

% Matrix for the LSF for Astig3 (third order astigmatism)
sz=size(a.x);
M5=[(2*a.x.*a.y)    a.x  a.y ones(sz)  zeros(sz)];
M6=[(a.x.^2-a.y.^2) -a.y  a.x zeros(sz) ones(sz)];
MAstig3 = [M5; M6];
% Least-squares fit for Z5 and Z6
RAstig3=MAstig3\[a.Z5; a.Z6]; %x=A\b
%b=MAstig3*RAstig3; diff=b-[a.Z5; a.Z6]; norm(diff) % b=Ax

% Matrix for the LSF for Coma3 (third order coma)
M7=[a.y zeros(sz) ones(sz)];

```

```

M8=[a.x ones(sz) zeros(sz)];
MComa3=[M7; M8];
% Least-squares fit for Z7 and Z8
RComa3=MComa3\[a.Z7; a.Z8]; %x=A\b
%b2=MComa3*RComa3; diff2=b2-[a.Z7; a.Z8]; norm(diff2)% b=Ax

% Delete coefficients below the threshold value
RAstig3(abs(RAstig3)<threshold)=0;
RComa3(abs(RComa3)<threshold)=0;

% Save the coefficients to the dataset called coefficients
thistrial=[name char(extensions(index))];
coefficients.(genvarname(thistrial))=[RAstig3; RComa3];

%% Calculate astigmatism contributions from coefficients

% Measured astigmatism
astig.meas.Z5=Grid.Z5;
astig.meas.Z6=Grid.Z6;
B=ones(size(Grid.X));

% Constant part
astig.cst.Z5=RAstig3(4)*B;
astig.cst.Z6=RAstig3(5)*B;
%Linear part
astig.lin.Z5 = RAstig3(2).*Grid.X+RAstig3(3).*Grid.Y;
astig.lin.Z6 = -RAstig3(2).*Grid.Y+RAstig3(3).*Grid.X;
%Quadratic part
astig.quad.Z5=2*RAstig3(1).*Grid.X.*Grid.Y;
astig.quad.Z6=RAstig3(1).*(Grid.X.^2-Grid.Y.^2);

%Residual part
astig.res.Z5=Grid.Z5-astig.cst.Z5-astig.lin.Z5-astig.quad.Z5;
astig.res.Z6=Grid.Z6-astig.cst.Z6-astig.lin.Z6-astig.quad.Z6;

% limits for the figure with the 5 graphs
xmin = 1.1*min(Grid.X(:)); xmax = 1.1*max(Grid.X(:));
ymin = 1.1*min(Grid.Y(:)); ymax = 1.1*max(Grid.Y(:));

%% Graphic representation of astigmatism contributions
if show_fig==1

    all_fields={'meas', 'cst', 'lin', 'quad', 'res'};
    graphlabels={'Total ', 'Constant ', 'Linear ', ...
                'Quadratic ', 'Residual '};

    figure('Name', [thistrial, ...
                  ' image orientation'], ...
          'Position', [10 scrsz(4)/3 scrsz(3)-10 scrsz(4)/2])
    h1=gcf;

```

```

figure('Name',[thistrial,...
' astigmatism coefficient magnitudes'],...
'Position',[10 scrsz(4)/5 scrsz(3)-10 scrsz(4)/1.5])
h2=gcf;

% Find the scale factor for the graphs so that the measured
% astigmatism has line images that look good.
temp=Astigmatism(astig.meas.Z5,astig.meas.Z6);
scale = 0.3/max(max(temp.x));

% 5 it. in loop for the meas, quad, linear, const, residual:
for idx = 1:length(all_fields)
    C5=astig.(char((genvarname(all_fields(idx))))).Z5;
    C6=astig.(char((genvarname(all_fields(idx))))).Z6;
    temp=Astigmatism(C5,C6);
    figure(h1),subplot(2,5,idx)
    switch 1 % Change this to 0 or 1
        case 0 % This option scales the lines so they always
            % show without overlapping (no matter how small or
            % large)
            quiv(Grid.X,Grid.Y,temp.x,temp.y);
        case 1 % This option lets the user scale the data
            temp.x=temp.x*scale; temp.y=temp.y*scale;
            quiv(Grid.X,Grid.Y,temp.x,temp.y,0);
    end

    axis square, axis([xmin xmax ymin ymax]);...
    title([graphlabels(idx) 'astigmatism'])
    xlabel('X position'); ylabel('Y position');

    figure(h2)
    subplot(3,5,idx)
    imagesc(Grid.angles,Grid.angles,temp.amp)
    axis square; colorbar; set(gca,'YDir','normal')
    title([graphlabels(idx) 'sqrt(C5^2+C6^2)'])
    subplot(3,5,5+idx)
    imagesc(Grid.angles,Grid.angles,C5)
    axis square; colorbar; set(gca,'YDir','normal')
    title([graphlabels(idx) 'C5'])
    subplot(3,5,10+idx)
    imagesc(Grid.angles,Grid.angles,C6)
    axis square; colorbar; set(gca,'YDir','normal')
    title([graphlabels(idx) 'C6'])
end
clear idx temp C5 C6
end
%% Calculate coma contributions from coefficients

% Measured part
coma.meas.Z7=Grid.Z7;

```

```

coma.meas.Z8=Grid.Z8;
% Constant part
coma.cst.Z7=RComa3(3).*B;
coma.cst.Z8=RComa3(2).*B;
% Linear part
coma.lin.Z7=RComa3(1).*Grid.Y;
coma.lin.Z8=RComa3(1).*Grid.X;
% Residual part
coma.res.Z7=coma.meas.Z7-coma.cst.Z7-coma.lin.Z7;
coma.res.Z8=coma.meas.Z8-coma.cst.Z8-coma.lin.Z8;

%% Graphic representation of coma contributionss

if show_fig==1
    figure('Name',[thistrial,...
        ' coma coefficient magnitudes'],...
        'Position',[10 scrsz(4)/5 scrsz(3)-10 scrsz(4)/1.5])
    h3=gcf;
    for idx=1:5
        if idx==4
            figure(h1),subplot(2,5,5+idx)
            text(xmin,0,'Does not exist.');
```

axis off

```

            figure(h3),subplot(3,5,idx)
            text(xmin,0,'Does not exist.');
```

axis off

```

            axis([xmin xmax ymin ymax]);
        else
            C7=coma.(char((genvarname(all_fields(idx))))).Z7;
            C8=coma.(char((genvarname(all_fields(idx))))).Z8;
            temp=Coma(C7,C8);

            figure(h1),subplot(2,5,5+idx)
            switch 1 % Change this to 0 or 1
                case 0 % This option scales the lines so they
                    % always show without overlapping (no matter
                    % how small or large)
                    quiv(Grid.X,Grid.Y,temp.x,temp.y);
                case 1 % This option lets the user scale the data
                    temp.x=temp.x*scale; temp.y=temp.y*scale;
                    quiv(Grid.X,Grid.Y,temp.x,temp.y,0);
            end
            xlabel('X position'); ylabel('Y position');
```

```

            figure(h3)
            subplot(3,5,idx),
            imagesc(Grid.angles,Grid.angles,temp.amp)
            colorbar; set(gca,'YDir','normal'), axis square
            title([graphlabels(idx) 'sqrt(C7^2+C8^2)'])

            subplot(3,5,5+idx),
```



```

        imagesc(Grid.angles,Grid.angles,C7)
        colorbar; set(gca,'YDir','normal'), axis square
        title([graphlabels(idx) 'C7'])

        subplot(3,5,10+idx),
        imagesc(Grid.angles,Grid.angles,C8)
        colorbar; set(gca,'YDir','normal'), axis square
        title([graphlabels(idx) 'C8'])
    end
    figure(h1); axis square, axis([xmin xmax ymin ymax]);
    title([graphlabels(idx) 'coma'])
end
clear idx graphlabels C7 C8
end

%% Creation of low order Zernikes

% This will create the Zernike polynomials on a grid that is the
% same size as the measured data.
if index==1
    Zernike=ZernikePolysonGrid(Grid);
    if show_fig==1
        figure('Name','Zernike polynomials 1-8'),
        for idx = 1:11
            subplot(3,4,idx),
            imagesc(Grid.angles,Grid.angles,...
                Zernike.(genvarname(['Z' num2str(idx)]))),
            axis square;set(gca,'YDir','normal');colorbar
        end
        clear idx m
    end
end

%% Coefficients for orthogonal functions

aberrations={'focus';'astig';'coma';'trefoil';'spherical'};
aberrations_num1={'Z4';'Z5';'Z7';'Z9';'Z11'};
aberrations_num2={' '; 'Z6';'Z8';'Z10';' '};

if n_z==17
    aberrations(6:7)={'oblsph';'seccoma'};
    aberrations_num1(6:7)={'Z12';'Z16'};
    aberrations_num2(6:7)={'Z13';'Z17'};
end
coeff=[];
for idx=1:length(aberrations)
    % The variable (of type = structure) Pupil_Z holds the
    % Zernike coefficients recorded in the lens design program
    % which are measured in the pupil (for a grid of fields).
    Pupil_Z.f1=Grid.(genvarname(char(aberrations_num1(idx))));
    if idx==1||idx==5 % focus or spherical

```

```

        % No other dependence to add.
    else % for aberrations coming in pairs
        % Add other term.
        Pupil_Z.f2=Grid.(genvarname(char(aberrations_num2(idx))));
    end
    c=DoubleZernikeFit(Pupil_Z,Zernike,threshold);
    coeff=[coeff; c];
    clear Pupil_Z
end

% Create a list of the coefficients:
ortho_coefficients.(genvarname(char(perturb(index))))=coeff;

% Create a list of the induced coefficients
% = Perturbed system coefficients – nominal system coefficients
% Don't need to find difference coefficients for the first trial,
% which is the aligned system. (The coefficients would be all
% zeros)
if index≠1
    ortho_diff_coefficients.(genvarname(char(perturb(index))))=...
        coeff-ortho_coefficients.(genvarname(char(perturb(1))));
end
clear f1 f2 c coeff aberrations_num1 aberrations_num2 idx
end

%% Create datasets for the orthogonal functions

NameAb(1:13)=aberrations(1);          % focus
NameAb(14:39)=aberrations(2);        % astig
NameAb(40:65)=aberrations(3);        % coma
NameAb(66:91)=aberrations(4);        % trefoil
NameAb(92:104)=aberrations(5);       % spherical
if n_z==17
    NameAb(105:130)=aberrations(6);   % oblsph
    NameAb(131:156)=aberrations(7);   % seccoma
end
NameAb=NameAb';

List1={'Constant1','Constant2',...
    'Linear1','Linear2','Linear3','Linear4',...
    'Focus1','Focus2',...
    'Astig1','Astig2','Astig3','Astig4',...
    'Coma1','Coma2','Coma3','Coma4',...
    'Trefoil1','Trefoil2','Trefoil3','Trefoil4',...
    'Quartic1','Quartic2','OblSph1','OblSph2','OblSph3','OblSph4'};
List2={'Constant','Linear1','Linear2','Focus','Astig1','Astig2',...
    'Coma1','Coma2','Trefoil1','Trefoil2','Quartic',...
    'OblSph1','OblSph2'};
Listall=[List2;List1;List1;List1;List2];
if n_z==17; Listall=[Listall;List1;List1]; end

```

```

Names=strcat(Listall,NameAb);
clear NameAb List1 List2 Listall

ortho_coeff_DS=dataset(ortho_coefficients,'ObsNames',Names');

if length(perturb)>1
    orth_diff_coeff_DS=dataset(ortho_diff_coefficients,...
        'ObsNames',Names');
end

exp=[0; 1;1; 0; 2;2; 1;1; 4;4; 0; 4;4;...%focus
2;2; 3;1;1;3; 2;2; 0;4;4;4; 1;4;4;1; 1;4;4;1; 4;4; 4;4;0;4;...%astig
1;1; 2;4;0;2; 1;1; 1;4;4;1; 0;4;4;4; 4;4;4;4; 4;4; 4;4;4;4;...%coma
3;3; 4;2;2;4; 4;4; 1;4;4;1; 4;4;4;4; 4;4;4;4; 4;4; 4;4;4;4;...%trefoil
0; 1;1; 0; 4;4; 4;4; 4;4; 4; 4;4;...%spherical
2;2; 1;4;4;1; 4;4; 4;4;0;4; 4;4;4;4; 4;4;4;4; 4;4; 4;4;4;4;...%oblsph
1;1; 0;4;4;4; 4;4; 4;4;4;4; 4;4;4;4; 4;4;4;4; 4;4; 4;4;4;4];...%2coma

% Sort the data to examine the results for the largest contributions

for idx=1:length(perturb)
    if idx==1
        data=ortho_coeff_DS.(genvarname(char(perturb(1))));
    else
        data=orth_diff_coeff_DS.(genvarname(char(perturb(idx))));
    end
    [I,IX] = sort(abs(data),1,'descend');
    result.(genvarname(char(perturb(idx)))) = ...
        dataset(data(IX),'ObsNames',Names(IX));
end
clear idx IX data
% Print figures to .eps files for Latex

% For Hubble astigmatism example:
if Hubble==1
    Latexnames = {'HST_lines.eps',...
        'HST_Z5Z6_Matlab.eps',...
        'HST_Z7Z8_Matlab.eps'};

    Latexnames = strcat(ex,'_',Latexnames);

    for idx=1:length(Latexnames)
        h=figure(idx);
        set(gcf,'PaperPositionMode','auto')
        print(h,'-depsc', Latexnames{idx});
    end
end
end

```

```

%% SVD analysis of sensitivity matrix to find control modes

if doSVD==1

    % In order to plot the resulting wavefronts from the SVD, the
    % Zernike polynomials over a reasonable size grid for plotting
    % nicely needs to be created.
    n=20;
    grid2.angles=linspace(-1,1,n);
    [grid2.X,grid2.Y]=meshgrid(grid2.angles,grid2.angles);
    [grid2.THETA,grid2.RHO] = cart2pol(grid2.X,grid2.Y);
    grid2.mask=ones(size(grid2.RHO)); grid2.mask(grid2.RHO>1)=0;

    Zernike2=ZernikePolysonGrid(grid2);

    % Take difference to find sensitivity:
    Msens=M-repmat(M(:,1), [1,length(extensions)]);
    Msens=Msens(:,2:end); % delete the first column (all zeros)
    [U,S,V] = svd(Msens, 'econ');

    % Look at the Singular Values
    s = nonzeros(S); % The singular values
    if show_SVDfigS==1
        figure, semilogy(s, 'x'), hold on, semilogy(s), hold off
        xlabel('Index', 'FontSize', 14),
        ylabel('Magnitude', 'FontSize', 14)
        title(['Singular Values of the ', telescope, ...
            ' Alignment Sensitivity Matrix'], 'FontSize', 14)
        set(gca, 'XTick', 1:length(s))
    end
    col=1;
    for idx=4:n_z
        % Create a Zernike matrix where the rows are all the field
        % points and the columns are terms 4 (focus) through 11
        % (spherical) or 17 (secondary coma) as the case may be,
        % except skip terms Z14, Z15
        if (idx<14)|| (idx>15)
            Zernmat(:,col) = ...
                Zernike2.(genvarname(char(['Z', num2str(idx)])))(:);
            col=col+1;
        end
    end

    for idx2=1:length(extensions)-1
        temp = reshape(U(:,idx2), numfields, []);
        surf_vecs = Zernmat*temp';

        grid2.Z4=reshape(temp(:,1), sqrt(numfields), sqrt(numfields));
        grid2.Z5=reshape(temp(:,2), sqrt(numfields), sqrt(numfields));
        grid2.Z6=reshape(temp(:,3), sqrt(numfields), sqrt(numfields));
    end
end

```

```

grid2.Z7=reshape(temp(:,4),sqrt(numfields),sqrt(numfields));
grid2.Z8=reshape(temp(:,5),sqrt(numfields),sqrt(numfields));
grid2.Z9=reshape(temp(:,6),sqrt(numfields),sqrt(numfields));
grid2.Z10=reshape(temp(:,7),sqrt(numfields),sqrt(numfields));
grid2.Z11=reshape(temp(:,8),sqrt(numfields),sqrt(numfields));
if n_z==17
grid2.Z12=reshape(temp(:,9),sqrt(numfields),sqrt(numfields));
grid2.Z13=reshape(temp(:,10),sqrt(numfields),sqrt(numfields));
grid2.Z16=reshape(temp(:,11),sqrt(numfields),sqrt(numfields));
grid2.Z17=reshape(temp(:,12),sqrt(numfields),sqrt(numfields));
end
%%%
aberrations={'focus';'astig';'coma';'trefoil';'spherical'};
aberrations_1={'Z4';'Z5';'Z7';'Z9';'Z11'};
aberrations_2={' ';'Z6';'Z8';'Z10';' '};
if n_z==17
    aberrations(6:7)={'oblsph';'seccoma'};
    aberrations_1(6:7)={'Z12';'Z16'};
    aberrations_2(6:7)={'Z13';'Z17'};
end

coeff=[];
for idx=1:length(aberrations)
    % The variable (of type = structure) Pupil_Z holds the
    % Zernike coefficients recorded in the lens design program
    % which are measured in the pupil (for a grid of fields).
    Pupil_Z.f1=grid2.(genvarname(char(aberrations_1(idx))));
    if idx==1||idx==5 % focus or spherical
        % No other dependence to add.
    else % for aberrations coming in pairs
        % Add other term.
        Pupil_Z.f2=grid2.(genvarname(char(aberrations_2(idx))));
    end
    c=DoubleZernikeFit(Pupil_Z,Zernike,threshold);
    coeff=[coeff; c];
    clear Pupil_Z
end
% Create a list of the coefficients:
ortho_coeff_U.(genvarname(char(['mode',num2str(idx2)])))=coeff;
end
ortho_coeff_U_ds=dataset(ortho_coeff_U);
dortho_coeff_U_ds=double(ortho_coeff_U_ds);

if show_SVDfig==1
    if show_SVD_figUV==1
        for index=1:min(nm,size(U,2)) % show up to 10 modes
            figure('Name', ['Mode ',num2str(index)],...
                'Position',[50 50 600 600])
            c=dortho_coeff_U_ds(:,index);
            [map]=DoubleZernikeCreate2(c,Zernike2,n_z,[-0.75 0.75]);
        end
    end
end

```

```

figure('Name', ['Configuration for mode ',...
    num2str(index)], 'Position', [50 50 600 500])
stem(V(:,index))
title(['Configuration vector from V for mode ',...
    num2str(index)], 'FontSize',14)
xlabel('Degree of freedom', 'FontSize',14)
axis([0.8 length((V(:,index)))+0.2 -1 1])
xticklabel_rotate(1:size(V,2),90,perturb(2:end),...
    'FontSize',12)
ylabel('Normalized amount of misalignment',...
    'FontSize',14)
end
end
end

%}
%% Plot the aberrations resulting from each degree of freedom
for idx2=1:length(extensions)-1
    temp = reshape(Msens(:,idx2),numfields,[]);
    surf_vecs = Zernmat*temp';

    grid3.Z4=reshape(temp(:,1),sqrt(numfields),sqrt(numfields));
    grid3.Z5=reshape(temp(:,2),sqrt(numfields),sqrt(numfields));
    grid3.Z6=reshape(temp(:,3),sqrt(numfields),sqrt(numfields));
    grid3.Z7=reshape(temp(:,4),sqrt(numfields),sqrt(numfields));
    grid3.Z8=reshape(temp(:,5),sqrt(numfields),sqrt(numfields));
    grid3.Z9=reshape(temp(:,6),sqrt(numfields),sqrt(numfields));
    grid3.Z10=reshape(temp(:,7),sqrt(numfields),sqrt(numfields));
    grid3.Z11=reshape(temp(:,8),sqrt(numfields),sqrt(numfields));
    if n_z==17
    grid3.Z12=reshape(temp(:,9),sqrt(numfields),sqrt(numfields));
    grid3.Z13=reshape(temp(:,10),sqrt(numfields),sqrt(numfields));
    grid3.Z16=reshape(temp(:,11),sqrt(numfields),sqrt(numfields));
    grid3.Z17=reshape(temp(:,12),sqrt(numfields),sqrt(numfields));
    end

    coeff=[];
    for idx=1:length(aberrations)
        % The variable (of type = structure) Pupil_Z holds the
        % Zernike coefficients recorded in the lens design program
        % which are measured in the pupil (for a grid of fields).
        Pupil_Z.f1=grid3.(genvarname(char(aberrations_1(idx))));
        if idx==1||idx==5 % focus or spherical
            % No other dependence to add.
        else % for aberrations coming in pairs
            % Add other term.
            Pupil_Z.f2=grid3.(genvarname(char(aberrations_2(idx))));
        end
        c=DoubleZernikeFit(Pupil_Z,Zernike,threshold);
    end
end

```

```

        coeff=[coeff; c];
        clear Pupil_Z
    end
    % Create a list of the coefficients:
    ortho_coeff_M.(genvarname(char(['mode', num2str(idx2)])))=coeff;
end
ortho_coeff_M_ds=dataset(ortho_coeff_M);
dortho_coeff_M_ds=double(ortho_coeff_M_ds);

if show_DOFfig==1
    if(gcf<40)
        for index=1:min(10,size(M,2)) %show up to 10 dof
            %a=find(exp<=1);
            figure('Name', ['DOF ', num2str(index)], ...
                'Position', [50 50 900 450])
            Functionnames; %Create string of names
            a0=find(exp==0);
            a1=find(exp==1);
            a01=[a1; a0];
            subplot(6,1,2:5) % to have room for function labels
            h1=stem(1:length(a1), dortho_coeff_M_ds(a1, index), 'b');
            hold on
            h2=stem((length(a1)+1):length(a01), ...
                dortho_coeff_M_ds(a0, index), 'r');
            ylimit = 1.1*max(abs(dortho_coeff_M_ds(a01, index)));
            axis([0 length(a01)+1 -ylimit ylimit])
            xticklabel_rotate(1:length(a01), 90, ...
                FunctionNames(a01), 'FontSize', 14, ...
                'interpreter', 'latex')
            ylabel('Value of coefficient', 'FontSize', 14)
            title(['Orthogonal coefficients for ' ...
                'individual DOF:', perturb(index+1)], 'FontSize', 14)
        end
    end
end

%% SVD analysis of orthogonal coefficients to find control modes

[U2,S2,V2] = svd(double(orth_diff_coeff_DS), 'econ');

if show_SVDfig==1
    if show_SVDfigS2==1
        % Look at the Singular Values
        s2 = nonzeros(S2); % The singular values
        figure, semilogy(s2, 'x'), hold on, semilogy(s2), hold off
        xlabel('Index', 'FontSize', 14),
        ylabel('Magnitude', 'FontSize', 14)
        title(['Singular Values of the ', telescope, ...
            ' Alignment Sensitivity Matrix'], 'FontSize', 14)
        set(gca, 'XTick', 1:length(s2))
    end
end

```

```

end
% Plot the wavefront maps across the field.
for index=1:min(nm,size(U2,2)) %show up to 10 modes of the dof
    if(gcf<40 %Too many figures will crash MATLAB
        if show_SVDfigU2==1
            figure('Name', ['Mode ',num2str(index)],...
                'Position',[50 50 600 600])
            c=U2(:,index);
            [map]=DoubleZernikeCreate2(c,Zernike2,n_z,[-7 7]);
        end
% Find the indices for all coefficients in a rotationally symmetric
% system or those that vary linearly with misalignment:
        Functionnames; %Create string of names
        a0=find(exp==0);
        a1=find(exp==1);
        a01=[a1; a0];
        if show_SVDfigU2b==1
            figure('Name', ['Mode ',num2str(index)],...
                'Position',[50 50 900 450])
            subplot(5,1,1:4) % to have room for function labels
            h1=stem(1:length(a1),U2(a1,index),'b'); hold on
            h2=stem((length(a1)+1):length(a01),U2(a0,index),'r');
            axis([0 length(a01)+1 -1.1 1.1])
            xticklabel_rotate(1:length(a01),90,...
                FunctionNames(a01),'FontSize',14,...
                'interpreter','latex')
            ylabel('Value of coefficient','FontSize',14)
            title(['Orthogonal coefficients from U for mode ',...
                num2str(index)],'FontSize',14), hold off
            %xlabel('Function','FontSize',14)
        end
    end

    if show_SVDfigV2==1
        figure('Name', ['Configuration for mode ',...
            num2str(index)],'Position',[50 50 600 500])
        h1=stem(1:20,V2(1:20,index),'b'); hold on
        h2=stem(21:25,V2(21:25,index),'r');
        title(['Configuration vector from V for mode ',...
            num2str(index)],'FontSize',14)
        xlabel('Degree of freedom','FontSize',14)
        axis([0 size(V2,2)+1 -1.1 1.1])
        xticklabel_rotate(1:size(V2,2),90,perturb(2:end),...
            'FontSize',12)
        ylabel('Normalized amount of misalignment',...
            'FontSize',14)
    end
end
end
end
end
end

```



```

%% Print figures to .eps files for Latex

% For HET mode analysis:
if HET==1
    Latexnames = {'1', '2', '3', '4', '5', '6', '7', '8', '9', '10', ...
                  '11', '12', '13', '14', '15', '16', '17', '18', '19', '20', ...
                  '21', '22', '23', '24', '25'};
    ex='HET_mode'; ext='.eps';
    Latexnames = strcat(ex, Latexnames, ext);

    for idx=1:length(Latexnames)
        h=figure(idx+1);
        set(gcf, 'PaperPositionMode', 'auto')
        print(h, '-depsc', Latexnames{idx});
    end

end

end

%% Plot the orthogonal functions

if o==1
    for index=121:156
        h=figure('Name', ['Orthogonal Mode ', num2str(index)], ...
                'Position', [50 50 600 600]);
        c=zeros(156,1); %12x13=156
        c(index)=1;
        [map]=DoubleZernikeCreate2(c, Zernike2, n_z, [-8 8]);
        Latexnames = num2str(index);
        ex='Orthmode'; ext='.eps';
        Latexnames = strcat(ex, Latexnames, ext);
        set(gcf, 'PaperPositionMode', 'auto')
        print(h, '-depsc', Latexnames);
    end

end

%% Plot the linearity of the HST coefficients

if HST==1
    values=[0; 0.001; 0.01; 0.1; 1; 3; 5; 7; 10; 15; 20];
    aaa=double(orth_coeff_DS);

    r=9; % Include perturbations up to, including 10mm decenters

    for idx=0:3
        a=find(exp==idx);
        figure('Position', [10 scrsz(4)/3 1100 350])
        subplot(1,2,1), plot(values(1:r), aaa(a,1:r), 'x-'),
        xlabel('Secondary mirror decenter (in mm)', 'FontSize', 12)
        ylabel({'Value of orthogonal coefficient', ...

```

```

        '(waves of 587.8 nm)', 'FontSize', 12)
title('Linear plot', 'FontSize', 12)
if idx==0;
    subplot(1,2,2), semilogx(values(1:r), aaa(a,1:r), 'x-')
    axis([1e-3 1e1 -0.1 0.6]),
    title('Semi-log plot', 'FontSize', 12);
    ylabel({'Value of orthogonal coefficient', ...
           '(waves of 587.8 nm)'}, 'FontSize', 12)
elseif idx==1;
    subplot(1,2,2), loglog(values(1:r), abs(aaa(a,1:r)), 'x-')
    axis([1e-3 1e1 1e-7 inf]),
    title('Log-log plot', 'FontSize', 12);
    ylabel({'Absolute value of orthogonal coefficient', ...
           '(waves of 587.8 nm)'}, 'FontSize', 12)
elseif idx==2;
    subplot(1,2,2), loglog(values(1:r), abs(aaa(a,1:r)), 'x-')
    axis([1e-3 1e1 1e-7 inf]),
    title('Quadratic', 'FontSize', 12);
    ylabel({'Absolute value of orthogonal coefficient', ...
           '(waves of 587.8 nm)'}, 'FontSize', 12)
elseif idx==3;
    subplot(1,2,2), loglog(values(1:r), abs(aaa(a,1:r)), 'x-')
    axis([1e-3 1e1 1e-7 inf]),
    title('Log-log plot', 'FontSize', 12);
    ylabel({'Absolute value of orthogonal coefficient', ...
           '(waves of 587.8 nm)'}, 'FontSize', 12)
end
xlabel('Secondary mirror decenter (in mm)', 'FontSize', 12)
end
end

%%
toc

```

## I.2 ASTIGMATISM.M MATLAB CODE

```
function [dataout]=Astigmatism(Z5,Z6)

dataout.amp=sqrt(Z6.^2+Z5.^2);
%dataout.angle=(1/2)*atan2(Z5,Z6)-pi/4;
dataout.angle=(1/2)*atan2(Z5,Z6);
[dataout.x,dataout.y] = pol2cart(dataout.angle,dataout.amp);
```

## I.3 COMA.M MATLAB CODE

```
function [dataout]=Coma(Z7,Z8)

dataout.amp=sqrt(Z7.^2+Z8.^2);
dataout.angle=atan2(Z7,Z8);

[dataout.x,dataout.y] = pol2cart(dataout.angle,dataout.amp);
```

## I.4 ZERNIKEPOLYSONGRID.M MATLAB CODE

```
function [Zernike]=ZernikePolysonGrid(grid)

m=grid.mask;
r=grid.RHO;
a=grid.THETA;

Zernike.Z1 =m;
Zernike.Z2 =m.*(2*r.*cos(a));
Zernike.Z3 =m.*(2*r.*sin(a));
Zernike.Z4 =m.*(sqrt(3)*(2*r.^2-1));
Zernike.Z5 =m.*(sqrt(6)*(r.^2.*sin(2*a)));
Zernike.Z6 =m.*(sqrt(6)*(r.^2.*cos(2*a)));
Zernike.Z7 =m.*(sqrt(8)*(3*r.^3-2*r).*sin(a));
Zernike.Z8 =m.*(sqrt(8)*(3*r.^3-2*r).*cos(a));
Zernike.Z9 =m.*(sqrt(8)*r.^3.*(sin(3*a)));
Zernike.Z10=m.*(sqrt(8)*r.^3.*(cos(3*a)));
Zernike.Z11=m.*(sqrt(5)*(6*r.^4-6*r.^2+1));
Zernike.Z12=m.*(sqrt(10)*(4*r.^4-3*r.^2).*cos(2*a));
Zernike.Z13=m.*(sqrt(10)*(4*r.^4-3*r.^2).*sin(2*a));
Zernike.Z14=m.*(sqrt(10)*(r.^4).*cos(4*a));
```

```
Zernike.Z15=m.*(sqrt(10)*(r.^4).*sin(4*a));
Zernike.Z16=m.*(sqrt(12)*(10*r.^5-12*r.^3+3*r).*cos(a));
Zernike.Z17=m.*(sqrt(12)*(10*r.^5-12*r.^3+3*r).*sin(a));
```

## I.5 DOUBLEZERNIKEFIT.M MATLAB CODE

```
function [c]=DoubleZernikeFit(PupilZern,Zernike,threshold)

f1=PupilZern.f1;

% If there are a pair of Zernike polynomial coefficients throughout
% the field (Zernike ):
if isfield(PupilZern,'f2')
    f2=PupilZern.f2;

    c=zeros(26,1);

    c(1) = sum(sum(Zernike.Z1.*f1));
    c(2) = sum(sum(Zernike.Z1.*f2));

    c(3) = sum(sum(Zernike.Z2.*f1 + Zernike.Z3.*f2));
    c(4) = sum(sum(Zernike.Z2.*f1 - Zernike.Z3.*f2));
    c(5) = sum(sum(Zernike.Z3.*f1 + Zernike.Z2.*f2));
    c(6) = sum(sum(Zernike.Z3.*f1 - Zernike.Z2.*f2));

    c(7) = sum(sum(Zernike.Z4.*f1));
    c(8) = sum(sum(Zernike.Z4.*f2));

    c(9) = sum(sum(Zernike.Z5.*f1 + Zernike.Z6.*f2));
    c(10) = sum(sum(Zernike.Z5.*f1 - Zernike.Z6.*f2));
    c(11) = sum(sum(Zernike.Z6.*f1 + Zernike.Z5.*f2));
    c(12) = sum(sum(Zernike.Z6.*f1 - Zernike.Z5.*f2));

    c(13) = sum(sum(Zernike.Z7.*f1 + Zernike.Z8.*f2));
    c(14) = sum(sum(Zernike.Z7.*f1 - Zernike.Z8.*f2));
    c(15) = sum(sum(Zernike.Z8.*f1 + Zernike.Z7.*f2));
    c(16) = sum(sum(Zernike.Z8.*f1 - Zernike.Z7.*f2));

    c(17) = sum(sum(Zernike.Z9.*f1 + Zernike.Z10.*f2));
    c(18) = sum(sum(Zernike.Z9.*f1 - Zernike.Z10.*f2));
    c(19) = sum(sum(Zernike.Z10.*f1 + Zernike.Z9.*f2));
    c(20) = sum(sum(Zernike.Z10.*f1 - Zernike.Z9.*f2));

    c(21) = sum(sum(Zernike.Z11.*f1));
    c(22) = sum(sum(Zernike.Z11.*f2));
```

```
c(23) = sum(sum(Zernike.Z12.*f1 + Zernike.Z13.*f2));
c(24) = sum(sum(Zernike.Z12.*f1 - Zernike.Z13.*f2));
c(25) = sum(sum(Zernike.Z13.*f1 + Zernike.Z12.*f2));
c(26) = sum(sum(Zernike.Z13.*f1 - Zernike.Z12.*f2));

else % Only one Zernike pupil dependence
    c=zeros(13,1);
    for idx2=1:13
        Z=Zernike.(genvarname(char(['Z',num2str(idx2)])));
        c(idx2) = sum(sum(Z.*f1));
    end
end

c=c/length(find(Zernike.Z1==1)); % Normalize by the field sampling
c(abs(c)<threshold)=0; % Delete coefficients below the threshold value
```

## APPENDIX J

## DERIVATIONS OF ORTHOGONAL FUNCTIONS

In Section 6.4, the orthogonal functions for third order astigmatism in a misaligned system were derived. In this appendix, the other aberrations are derived. Before this is done, there are two sections that show how the  $H_x$  and  $H_y$  field dependencies turn into Zernike polynomials in field space and the  $\rho$  and  $\phi$  dependence turn into Zernike polynomials in pupil space. Some of the vector dot products and vector multiplication products used in this appendix are listed in Appendix G.

Section J.3 includes the third order aberration derivations, which will be listed in Zernike polynomial order ( $Z_4$ : field curvature;  $Z_5, Z_6$ : astigmatism;  $Z_7, Z_8$ : coma; and  $Z_{11}$ : spherical). Section J.4 includes the fifth order aberration derivations, again in Zernike polynomial order ( $Z_4$ : quartic field curvature;  $Z_5, Z_6$ : fifth order astigmatism;  $Z_7, Z_8$ : field cubed coma;  $Z_9, Z_{10}$ : trefoil;  $Z_{11}$ : fifth order field curvature;  $Z_{12}, Z_{13}$ : oblique spherical aberration; and  $Z_{16}, Z_{17}$ : fifth order coma).

All of the equations will start in the form derived by Thompson, then they will be derived into Tessieres's form to show the Greek letter coefficients and finally into orthogonal polynomial form. The orthogonal polynomials will each be underlined. These equations will all use the convention where the angle is defined counterclockwise from the  $x$ -axis.

## J.1 ZERNIKE POLYNOMIALS TO DESCRIBE FIELD DEPENDENCE

In this section, the Zernike polynomials that describe the field appendix are discussed. It may be helpful to refer to the table showing the Zernike polynomials in Cartesian form (Table F.2). Arrows will be used to show the Zernike polynomials are not exactly equal to the previous expression due to the normalizations.

If an aberration has no field dependence (constant throughout the field), it may be described by the piston Zernike polynomial.

$$1 = Z_1(h, \theta) \tag{J.1}$$

Linear field dependence ( $H_x$  or  $H_y$  field terms just by themselves) correspond to the tip and tilt Zernike polynomials.

$$H_x = h \cos \theta \Rightarrow Z_2(h, \theta) \tag{J.2}$$

$$H_y = h \sin \theta \Rightarrow Z_3(h, \theta) \tag{J.3}$$

The term for  $h^2$  corresponds to Zernike polynomial number four:  $Z_4(h, \theta) = \sqrt{3}(2h^2 - 1)$ . This polynomial has a piston term in the field, which can be written as  $Z_1(h, \theta)$  and corresponds to a constant field dependent aberration that can be measured. We do not care about the exact relation between  $Z_4$  and  $Z_1$ , so the constants  $a$  and  $b$  will be used. (All the constants  $a$ ,  $b$ ,  $c$  etc. used in this appendix should be assumed to be different scaling factors valid for terms in individual equations and not as constants that are equal throughout all the different equations that

use them.)

$$H_x^2 + H_y^2 = h^2 \Rightarrow a Z_4(h, \theta) + b Z_1(h, \theta) \quad (\text{J.4})$$

The quadratic field dependent terms that turn into astigmatism terms are straightforward using Table F.2.

$$2 H_x H_y = Z_5(h, \theta) \quad (\text{J.5})$$

$$H_x^2 - H_y^2 = Z_6(h, \theta) \quad (\text{J.6})$$

The next quadratic field dependent terms require a combination of Zernike polynomials.

$$\begin{aligned} H_y^2 &\rightarrow -(H_x^2 - H_y^2) + (H_x^2 + H_y^2 - 1) + 1 \\ &\Rightarrow -a Z_6(h, \theta) + b Z_4(h, \theta) + c Z_1(h, \theta) \end{aligned} \quad (\text{J.7})$$

$$\begin{aligned} H_x^2 &\rightarrow (H_x^2 - H_y^2) + (H_x^2 + H_y^2 - 1) + 1 \\ &\Rightarrow a Z_6(h, \theta) + b Z_4(h, \theta) + c Z_1(h, \theta) \end{aligned} \quad (\text{J.8})$$

The following cubic terms correspond to coma in field, balanced with linear field dependent terms.

$$H_x^2 H_y + H_y^3 = H_y (H_x^2 + H_y^2) \Rightarrow a Z_7(h, \theta) + b Z_3(h, \theta) \quad (\text{J.9})$$

$$H_x^3 + H_x H_y^2 = H_x (H_x^2 + H_y^2) \Rightarrow a Z_8(h, \theta) + b Z_2(h, \theta) \quad (\text{J.10})$$

The following cubic terms correspond to trefoil in field. For these equations, it



is helpful to look at Table F.2.

$$3 H_x^2 H_y - H_y^3 = H_y(3 H_x^2 - H_y^2) = Z_9(h, \theta) \quad (\text{J.11})$$

$$H_x^3 - 3 H_x H_y^2 = H_x(H_x^2 - 3 H_y^2) = Z_{10}(h, \theta) \quad (\text{J.12})$$

The following cubic terms also correspond to trefoil in field, but they are also balanced by coma and linear field dependent terms.

$$3 H_x^2 H_y + H_y^3 \Rightarrow a Z_9(h, \theta) + a Z_7(h, \theta) + b Z_3(h, \theta) \quad (\text{J.13})$$

$$H_x^3 + 3 H_x H_y^2 \Rightarrow -a Z_{10}(h, \theta) + a Z_8(h, \theta) + b Z_2(h, \theta) \quad (\text{J.14})$$

$$H_x^3 \Rightarrow a Z_{10}(h, \theta) + a Z_8(h, \theta) + b Z_2(h, \theta) \quad (\text{J.15})$$

$$H_y^3 \Rightarrow -a Z_9(h, \theta) + a Z_7(h, \theta) + b Z_3(h, \theta) \quad (\text{J.16})$$

This next term for quartic field dependence is similar to the derivation of  $Z_4(h, \theta)$ .

$$(H_x^2 + H_y^2)^2 = h^4 \Rightarrow a Z_{11}(h, \theta) + b Z_4(h, \theta) + c Z_1(h, \theta) \quad (\text{J.17})$$

The following quartic terms in field correspond to field dependence with a secondary astigmatism-like functional form.

$$H_x^4 - H_y^4 = (H_x^2 - H_y^2)(H_x^2 + H_y^2) \Rightarrow a Z_{12}(h, \theta) + b Z_6(h, \theta) \quad (\text{J.18})$$

$$H_x^3 H_y + H_x H_y^3 = H_x H_y (H_x^2 + H_y^2) \Rightarrow a Z_{13}(h, \theta) + b Z_5(h, \theta) \quad (\text{J.19})$$

## J.2 ZERNIKE POLYNOMIALS TO DESCRIBE PUPIL DEPENDENCE

This section lists some of the conversions of  $\rho$  and  $\phi$  into Zernike polynomials. Table F.1 is correct, however since piston in the wavefront does not affect the image and the distortion (tilt) terms can not be easily measured with a wavefront sensor, these terms will be dropped. For example, the  $\rho^2$  term turns into the Zernike polynomial  $Z_4(\rho, \phi) = \sqrt{3}(2\rho^2 - 1)$ . Although this Zernike polynomial includes a piston term (in the pupil), it can be dropped because piston in the wavefront does not affect the image.

$$\rho^2 \Rightarrow Z_4(\rho, \phi) \tag{J.20}$$

$$\rho^3 \sin \phi \Rightarrow Z_7(\rho, \phi) \tag{J.21}$$

$$\rho^3 \cos \phi \Rightarrow Z_8(\rho, \phi) \tag{J.22}$$

$$\rho^5 \cos \phi \Rightarrow a Z_{16}(\rho, \phi) + b Z_8(\rho, \phi) \tag{J.23}$$

$$\rho^5 \sin \phi \Rightarrow a Z_{17}(\rho, \phi) + b Z_7(\rho, \phi) \tag{J.24}$$

## J.3 THIRD ORDER ABERRATION DERIVATIONS

This section includes the derivations of the orthogonal functions corresponding to the terms in Tables 3.1 and 3.4. The derivations use the Zernike polynomial functions listed in Section J.1 in field space and in Section J.2 in pupil space. The lines in the equations marked with  $W_T$  correspond to the form of the equations of Tessieres and the underlined terms are the orthogonal functions.

### J.3.1 THIRD ORDER FIELD CURVATURE

Field curvature is a rotationally symmetric aberration, described by the “focus” Zernike aberration term,  $Z_4(\rho, \phi)$ .

#### Quadratic field curvature

$$W = W_{220_M} \left( \vec{H} \cdot \vec{H} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.25})$$

$$= W_{220_M} (H_x^2 + H_y^2) \rho^2 \quad (\text{J.26})$$

$$W_T = \gamma_0 (H_x^2 + H_y^2) Z_4(\rho, \phi) \quad (\text{J.27})$$

$$\Rightarrow \gamma_0 \left[ a \underline{Z_4(h, \theta)Z_4(\rho, \phi)} + b \underline{Z_1(h, \theta)Z_4(\rho, \phi)} \right] \quad (\text{J.28})$$

#### Linear field curvature

$$W = -2 \left( \vec{H} \cdot \vec{A}_{220_M} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.29})$$

$$= -2 (H_x A_{220_Mx} + H_y A_{220_My}) \rho^2 \quad (\text{J.30})$$

$$W_T = \gamma_1 H_x Z_4(\rho, \phi) + \gamma_2 H_y Z_4(\rho, \phi) \quad (\text{J.31})$$

$$\Rightarrow \gamma_1 \underline{Z_2(h, \theta)Z_4(\rho, \phi)} + \gamma_2 \underline{Z_3(h, \theta)Z_4(\rho, \phi)} \quad (\text{J.32})$$

#### Constant field curvature

$$W = B_{220_M} (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.33})$$

$$W_T = \gamma_3 Z_4(\rho, \phi) \quad (\text{J.34})$$

$$\Rightarrow \gamma_3 \underline{Z_1(h, \theta)Z_4(\rho, \phi)} \quad (\text{J.35})$$

$$(\text{J.36})$$

### J.3.2 ASTIGMATISM

#### Quadratic astigmatism

$$W = \frac{1}{2} W_{222} \vec{H}^2 \cdot \vec{\rho}^2 \quad (\text{J.37})$$

$$= \frac{1}{2} W_{222} [2 H_x H_y \rho^2 \sin 2\phi + (H_x^2 - H_y^2) \rho^2 \cos 2\phi] \quad (\text{J.38})$$

$$W_T = \alpha_0 [2 H_x H_y Z_5(\rho, \phi) + (H_x^2 - H_y^2) Z_6(\rho, \phi)] \quad (\text{J.39})$$

$$\Rightarrow \alpha_0 \left[ \underline{Z_5(h, \theta) Z_5(\rho, \phi) + Z_6(h, \theta) Z_6(\rho, \phi)} \right] \quad (\text{J.40})$$

#### Linear astigmatism

$$W = - \left( \vec{H} \vec{A}_{222} \right) \cdot \vec{\rho}^2 \quad (\text{J.41})$$

$$= - (A_{222x} H_x - A_{222y} H_y) \rho^2 \cos 2\phi - (A_{222y} H_x + A_{222x} H_y) \rho^2 \sin 2\phi \quad (\text{J.42})$$

$$W_T = (\alpha_1 H_x - \alpha_2 H_y) Z_6(\rho, \phi) + (\alpha_2 H_x + \alpha_1 H_y) Z_5(\rho, \phi) \quad (\text{J.43})$$

$$\begin{aligned} \Rightarrow \alpha_1 \left[ \underline{Z_3(h, \theta) Z_5(\rho, \phi) + Z_2(h, \theta) Z_6(\rho, \phi)} \right] \\ + \alpha_2 \left[ \underline{Z_2(h, \theta) Z_5(\rho, \phi) - Z_3(h, \theta) Z_6(\rho, \phi)} \right] \end{aligned} \quad (\text{J.44})$$

#### Constant astigmatism

$$W = \frac{1}{2} \vec{B}_{222}^2 \cdot \vec{\rho}^2 \quad (\text{J.45})$$

$$= \frac{1}{2} [(B_{222x}^2 - B_{222y}^2) \rho^2 \cos 2\phi + 2 B_{222x} B_{222y} \rho^2 \sin 2\phi] \quad (\text{J.46})$$

$$W_T = \alpha_3 Z_5(\rho, \phi) + \alpha_4 Z_6(\rho, \phi) \quad (\text{J.47})$$

$$\Rightarrow \alpha_3 \underline{Z_1(h, \theta) Z_5(\rho, \phi)} + \alpha_4 \underline{Z_1(h, \theta) Z_6(\rho, \phi)} \quad (\text{J.48})$$

$$(\text{J.49})$$

## J.3.3 COMA

**Linear coma**

$$W = W_{131} \left( \vec{H} \cdot \vec{\rho} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.50})$$

$$= W_{131} (H_x \rho \cos \phi + H_y \rho \sin \phi) \rho^2 \quad (\text{J.51})$$

$$= W_{131} (H_x \rho^3 \cos \phi + H_y \rho^3 \sin \phi) \quad (\text{J.52})$$

$$W_T = \beta_0 [H_y Z_7(\rho, \phi) + H_x Z_8(\rho, \phi)] \quad (\text{J.53})$$

$$\Rightarrow \beta_0 \left[ \underline{Z_3(h, \theta) Z_7(\rho, \phi) + Z_2(h, \theta) Z_8(\rho, \phi)} \right] \quad (\text{J.54})$$

**Constant coma**

$$W = - \left( \vec{A}_{131} \cdot \vec{\rho} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.55})$$

$$= -A_{131x} \rho^3 \cos \phi - A_{131y} \rho^3 \sin \phi \quad (\text{J.56})$$

$$W_T = \beta_1 Z_7(\rho, \phi) + \beta_2 Z_8(\rho, \phi) \quad (\text{J.57})$$

$$\Rightarrow \beta_1 \underline{Z_1(h, \theta) Z_7(\rho, \phi)} + \beta_2 \underline{Z_1(h, \theta) Z_8(\rho, \phi)} \quad (\text{J.58})$$

### J.3.4 SPHERICAL ABERRATION

#### Constant spherical aberration

$$W = W_{040} (\vec{\rho} \cdot \vec{\rho})^2 \quad (\text{J.59})$$

$$= W_{040} \rho^4 \quad (\text{J.60})$$

$$W_T = \nu_0 Z_{11}(\rho, \phi) \quad (\text{J.61})$$

$$= \nu_0 [a Z_{11}(\rho, \phi) + b Z_4(\rho, \phi)] \quad (\text{J.62})$$

$$\Rightarrow \nu_0 \left[ a \underline{Z_1(h, \theta) Z_{11}(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_4(\rho, \phi)} \right] \quad (\text{J.63})$$

$$(\text{J.64})$$

## J.4 FIFTH ORDER ABERRATION DERIVATIONS

### J.4.1 QUARTIC FIELD CURVATURE

#### Quartic field curvature

$$W = W_{420_M} \left( \vec{H} \cdot \vec{H} \right)^2 (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.65})$$

$$= W_{420_M} (H_x^2 + H_y^2)^2 \rho^2 \quad (\text{J.66})$$

$$W_T = \psi_0 (H_x^4 + H_y^4 + 2 H_x^2 H_y^2) Z_4(\rho, \phi) \quad (\text{J.67})$$

$$= \psi_0 h^4 Z_4(\rho, \phi) \quad (\text{J.68})$$

$$\Rightarrow \psi_0 \left[ a \underline{Z_{11}(h, \theta) Z_4(\rho, \phi)} + b \underline{Z_4(h, \theta) Z_4(\rho, \phi)} + c \underline{Z_1(h, \theta) Z_4(\rho, \phi)} \right] \quad (\text{J.69})$$

### Cubic field curvature

$$W = -4 \left( \vec{H} \cdot \vec{H} \right) \left( \vec{H} \cdot \vec{A}_{420M} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.70})$$

$$= -4 (H_x^2 + H_y^2) (A_{420Mx} H_x + A_{420My} H_y) \rho^2 \quad (\text{J.71})$$

$$W_T = [\psi_1 (H_x^3 + H_x H_y^2) + \psi_2 (H_x^2 H_y + H_y^3)] Z_4(\rho, \phi) \quad (\text{J.72})$$

$$\begin{aligned} \Rightarrow \psi_1 & \left[ a \underline{Z_8(h, \theta) Z_4(\rho, \phi)} + b \underline{Z_2(h, \theta) Z_4(\rho, \phi)} \right] \\ & + \psi_2 \left[ a \underline{Z_7(h, \theta) Z_4(\rho, \phi)} + b \underline{Z_3(h, \theta) Z_4(\rho, \phi)} \right] \end{aligned} \quad (\text{J.73})$$

### Quadratic field curvature 1

$$W = B_{420M} \left( \vec{H} \cdot \vec{H} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.74})$$

$$W_T = \psi_3 (H_x^2 + H_y^2) Z_4(\rho, \phi) \quad (\text{J.75})$$

$$\Rightarrow \psi_3 \left[ a \underline{Z_4(h, \theta) Z_4(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_4(\rho, \phi)} \right] \quad (\text{J.76})$$

### Quadratic field curvature 2

$$W = 2 \left( \vec{H}^2 \cdot \vec{B}_{420M}^2 \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.77})$$

$$= 2 \left[ (H_x^2 - H_y^2) (B_{420Mx}^2 - B_{420My}^2) + (2 H_x H_y) (2 B_{420Mx} B_{420My}) \right] \rho^2 \quad (\text{J.78})$$

$$W_T = [\psi_5 (H_x^2 - H_y^2) + \psi_4 (H_x H_y)] Z_4(\rho, \phi) \quad (\text{J.79})$$

$$\Rightarrow \psi_4 \underline{Z_5(h, \theta) Z_4(\rho, \phi)} + \psi_5 \underline{Z_6(h, \theta) Z_4(\rho, \phi)} \quad (\text{J.80})$$

### Linear field curvature

$$W = -4 \left( \vec{H} \cdot \vec{C}_{420M} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.81})$$

$$= -4 (H_x C_{420Mx} + H_y C_{420My}) Z_4(\rho, \phi) \quad (\text{J.82})$$

$$W_T = (\psi_6 H_x + \psi_7 H_y) Z_4(\rho, \phi) \quad (\text{J.83})$$

$$\Rightarrow \psi_6 \underline{Z_2(h, \theta) Z_4(\rho, \phi)} + \psi_7 \underline{Z_3(h, \theta) Z_4(\rho, \phi)} \quad (\text{J.84})$$

$$(\text{J.85})$$

### Constant field curvature

$$W = D_{420M} (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.86})$$

$$W_T = \psi_8 Z_4(\rho, \phi) \quad (\text{J.87})$$

$$\Rightarrow \psi_8 \underline{Z_1(h, \theta) Z_4(\rho, \phi)} \quad (\text{J.88})$$

#### J.4.2 FIFTH ORDER ASTIGMATISM

In Thompson's equation for  $W_{422}$ , there are two terms that are cubic in field that include the vector  $\vec{A}_{422}$ . Tessieres had two Greek letter coefficients from each of those equations (four total), when only two total are needed ( $\chi_1 = -\vec{A}_{422x}$  and  $\chi_2 = -\vec{A}_{422y}$ ). Thus, the Greek letters subscripts will be slightly different for this aberration than how he derived them. There are no lines with  $W_T$  for cubic astigmatism 1 or 2. The  $W_T$  line will be found with the total cubic astigmatism.



### Quartic astigmatism

$$W = \frac{1}{2} W_{422} \left( \vec{H} \cdot \vec{H} \right) \left( \vec{H}^2 \cdot \vec{\rho}^2 \right) \quad (\text{J.89})$$

$$= \frac{1}{2} W_{422} (H_x^2 + H_y^2) [2 H_x H_y Z_5(\rho, \phi) + (H_x^2 - H_y^2) Z_6(\rho, \phi)] \quad (\text{J.90})$$

$$W_T = \chi_0 [(2 H_x^3 H_y + 2 H_x H_y^3) Z_5(\rho, \phi) + (H_x^4 - H_y^4) Z_6(\rho, \phi)] \quad (\text{J.91})$$

$$= \chi_0 \left\{ [a Z_{13}(h, \theta) + b Z_5(h, \theta)] Z_5(\rho, \phi) \right. \\ \left. + [a Z_{12}(h, \theta) + b Z_6(h, \theta)] Z_6(\rho, \phi) \right\} \quad (\text{J.92})$$

$$\Rightarrow \chi_0 \left\{ a \left[ \underline{Z_{13}(h, \theta) Z_5(\rho, \phi) + Z_{12}(h, \theta) Z_6(\rho, \phi)} \right] \right. \\ \left. + b \left[ \underline{Z_5(h, \theta) Z_5(\rho, \phi) + Z_6(h, \theta) Z_6(\rho, \phi)} \right] \right\} \quad (\text{J.93})$$

### Cubic astigmatism 1

$$W = - \left( \vec{H} \cdot \vec{H} \right) \left[ \left( \vec{H} \vec{A}_{422} \right) \cdot \vec{\rho}^2 \right] \quad (\text{J.94})$$

$$= -(H_x^2 + H_y^2) [(A_{422x} H_x - A_{422y} H_y) \rho^2 \cos 2\phi \\ + (A_{422y} H_x + A_{422x} H_y) \rho^2 \sin 2\phi] \quad (\text{J.95})$$

$$= (H_x^2 + H_y^2) \{ [(-A_{422x}) H_x - (-A_{422y}) H_y] \rho^2 \cos 2\phi \\ + [(-A_{422y}) H_x + (-A_{422x}) H_y] \rho^2 \sin 2\phi \} \quad (\text{J.96})$$

$$= (H_x^2 + H_y^2) [(\chi_1 H_x - \chi_2 H_y) Z_6(\rho, \phi) + (\chi_2 H_x + \chi_1 H_y) Z_5(\rho, \phi)] \quad (\text{J.97})$$

$$= [\chi_1 (H_x^2 H_y + H_y^3) + \chi_2 (H_x^3 + H_x H_y^2)] Z_5(\rho, \phi) \\ + [\chi_1 (H_x^3 + H_x H_y^2) - \chi_2 (H_x^2 H_y + H_y^3)] Z_6(\rho, \phi) \quad (\text{J.98})$$

$$= \chi_1 [(H_x^2 H_y + H_y^3) Z_5(\rho, \phi) + (H_x^3 + H_x H_y^2) Z_6(\rho, \phi)] \\ + \chi_2 [(H_x^3 + H_x H_y^2) Z_5(\rho, \phi) - (H_x^2 H_y + H_y^3) Z_6(\rho, \phi)] \quad (\text{J.99})$$

### Cubic astigmatism 2

$$W = - \left( \vec{H} \cdot \vec{A}_{422} \right) \left( \vec{H}^2 \cdot \vec{\rho}^2 \right) \quad (\text{J.100})$$

$$= - (A_{422x} H_x + A_{422y} H_y) \left[ 2 H_x H_y Z_5(\rho, \phi) + (H_x^2 - H_y^2) Z_6(\rho, \phi) \right] \quad (\text{J.101})$$

$$= \left[ 2 H_x^2 H_y (-A_{422x}) + 2 H_x H_y^2 (-A_{422y}) \right] Z_5(\rho, \phi) \quad (\text{J.102})$$

$$+ \left[ (-A_{422x})(H_x^3 - H_x H_y^2) + (-A_{422y})(H_x^2 H_y - H_y^3) \right] Z_6(\rho, \phi)$$

$$= \left[ \chi_1 (2 H_x^2 H_y) + \chi_2 (2 H_x H_y^2) \right] Z_5(\rho, \phi) \quad (\text{J.103})$$

$$+ \left[ \chi_1 (H_x^3 - H_x H_y^2) + \chi_2 (H_x^2 H_y - H_y^3) \right] Z_6(\rho, \phi)$$

$$= \chi_1 \left[ 2 H_x^2 H_y Z_5(\rho, \phi) + (H_x^3 - H_x H_y^2) Z_6(\rho, \phi) \right] \quad (\text{J.104})$$

$$+ \chi_2 \left[ 2 H_x H_y^2 Z_5(\rho, \phi) + (H_x^2 H_y - H_y^3) Z_6(\rho, \phi) \right]$$

**Total cubic astigmatism (1 and 2)**

$$\begin{aligned}
 W = & \left[ \chi_1 (H_x^2 H_y + H_y^3 + 2 H_x^2 H_y) \right. \\
 & \left. + \chi_2 (H_x^3 + H_x H_y^2 + 2 H_x H_y^2) \right] Z_5(\rho, \phi)
 \end{aligned} \tag{J.105}$$

$$\begin{aligned}
 & + \left[ \chi_1 (H_x^3 + H_x H_y^2 + H_x^3 - H_x H_y^2) \right. \\
 & \left. + \chi_2 (-H_x^2 H_y - H_y^3 + H_x^2 H_y - H_y^3) \right] Z_6(\rho, \phi)
 \end{aligned}$$

$$W_T = \left[ \chi_1 (3 H_x^2 H_y + H_y^3) + \chi_2 (H_x^3 + 3 H_x H_y^2) \right] Z_5(\rho, \phi) \tag{J.106}$$

$$+ \left[ \chi_1 (2 H_x^3) + \chi_2 (-2 H_y^3) \right] Z_6(\rho, \phi)$$

$$= \chi_1 \left[ (3 H_x^2 H_y + H_y^3) Z_5(\rho, \phi) + (2 H_x^3) Z_6(\rho, \phi) \right] \tag{J.107}$$

$$+ \chi_2 \left[ (H_x^3 + 3 H_x H_y^2) Z_5(\rho, \phi) + (-2 H_y^3) Z_6(\rho, \phi) \right]$$

$$= \chi_1 \left\{ [a Z_9(h, \theta) + a Z_7(h, \theta) + b Z_3(h, \theta)] Z_5(\rho, \phi) \right.$$

$$\left. + [a Z_{10}(h, \theta) + a Z_8(h, \theta) + b Z_2(h, \theta)] Z_6(\rho, \phi) \right.$$

(J.108)

$$+ \chi_2 \left\{ [-a Z_{10}(h, \theta) + a Z_8(h, \theta) + b Z_2(h, \theta)] Z_5(\rho, \phi) \right.$$

$$\left. + [a Z_9(h, \theta) - a Z_7(h, \theta) - b Z_3(h, \theta)] Z_6(\rho, \phi) \right\}$$

$$\Rightarrow \chi_1 \left\{ a \left[ \underline{Z_9(h, \theta) Z_5(\rho, \phi) + Z_{10}(h, \theta) Z_6(\rho, \phi)} \right] \right.$$

$$\left. + a \left[ \underline{Z_7(h, \theta) Z_5(\rho, \phi) + Z_8(h, \theta) Z_6(\rho, \phi)} \right] \right.$$

$$\left. + b \left[ \underline{Z_3(h, \theta) Z_5(\rho, \phi) + Z_2(h, \theta) Z_6(\rho, \phi)} \right] \right\}$$

(J.109)

$$+ \chi_2 \left\{ -a \left[ \underline{Z_{10}(h, \theta) Z_5(\rho, \phi) - Z_9(h, \theta) Z_6(\rho, \phi)} \right] \right.$$

$$\left. + a \left[ \underline{Z_8(h, \theta) Z_5(\rho, \phi) - Z_7(h, \theta) Z_6(\rho, \phi)} \right] \right.$$

$$\left. + b \left[ \underline{Z_2(h, \theta) Z_5(\rho, \phi) - Z_3(h, \theta) Z_6(\rho, \phi)} \right] \right\}$$

### Quadratic astigmatism 1

$$W = \frac{3}{2} (\vec{H} \cdot \vec{H}) (\vec{B}_{422}^2 \cdot \vec{\rho}^2) \quad (\text{J.110})$$

$$= \frac{3}{2} (H_x^2 + H_y^2) [(B_{222x}^2 - B_{222y}^2) \rho^2 \cos 2\phi + 2 B_{222x} B_{222y} \rho^2 \sin 2\phi] \quad (\text{J.111})$$

$$W_T = \chi_3 (H_x^2 + H_y^2) Z_5(\rho, \phi) + \chi_4 (H_x^2 + H_y^2) Z_6(\rho, \phi) \quad (\text{J.112})$$

$$= \chi_3 [a Z_4(h, \theta) + b Z_1(h, \theta)] Z_5(\rho, \phi) \quad (\text{J.113})$$

$$+ \chi_4 [a Z_4(h, \theta) + b Z_1(h, \theta)] Z_6(\rho, \phi)$$

$$\Rightarrow \chi_3 \left[ a \underline{Z_4(h, \theta) Z_5(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_5(\rho, \phi)} \right] \quad (\text{J.114})$$

$$+ \chi_4 \left[ a \underline{Z_4(h, \theta) Z_6(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_6(\rho, \phi)} \right]$$

### Quadratic astigmatism 2

$$W = \frac{3}{2} B_{422} (\vec{H}^2 \cdot \vec{\rho}^2) \quad (\text{J.115})$$

$$= \frac{3}{2} B_{422} (2 H_x H_y Z_5(\rho, \phi) + (H_x^2 - H_y^2) Z_6(\rho, \phi)) \quad (\text{J.116})$$

$$W_T = \chi_5 [2 H_x H_y Z_5(\rho, \phi) + (H_x^2 - H_y^2) Z_6(\rho, \phi)] \quad (\text{J.117})$$

$$\Rightarrow \chi_5 \left[ \underline{Z_5(h, \theta) Z_5(\rho, \phi)} + \underline{Z_6(h, \theta) Z_6(\rho, \phi)} \right] \quad (\text{J.118})$$

### Linear astigmatism 1

$$W = -\frac{1}{2} \left[ \left( \vec{C}_{422}^3 \vec{H}^* \right) \cdot \vec{\rho}^2 \right] \quad (\text{J.119})$$

$$\begin{aligned} &= -\frac{1}{2} \left\{ \left[ (C_{422x}^3 - 3 C_{422x} C_{422y}^2) H_x \right. \right. \\ &\quad \left. \left. + (3 C_{422x}^2 C_{422y} - C_{422y}^3) H_y \right] \rho^2 \cos 2\phi \right. \\ &\quad \left. + \left[ (3 C_{422x}^2 C_{422y} - C_{422y}^3) H_x \right. \right. \\ &\quad \left. \left. - (C_{422x}^3 - 3 C_{422x} C_{422y}^2) H_y \right] \rho^2 \sin 2\phi \right\} \end{aligned} \quad (\text{J.120})$$

$$W_T = (\chi_6 H_x + \chi_7 H_y) Z_6(\rho, \phi) + (\chi_7 H_x - \chi_6 H_y) Z_5(\rho, \phi) \quad (\text{J.121})$$

$$\begin{aligned} &\Rightarrow \chi_6 \left[ -Z_3(h, \theta) Z_5(\rho, \phi) + Z_2(h, \theta) Z_6(\rho, \phi) \right] \\ &\quad + \chi_7 \left[ Z_2(h, \theta) Z_5(\rho, \phi) + Z_3(h, \theta) Z_6(\rho, \phi) \right] \end{aligned} \quad (\text{J.122})$$

### Linear astigmatism 2

$$W = \frac{3}{2} \left[ \left( \vec{H} \vec{C}_{422} \right) \cdot \vec{\rho}^2 \right] \quad (\text{J.123})$$

$$\begin{aligned} &= \frac{3}{2} \left[ (C_{422x} H_x - C_{422y} H_y) \rho^2 \cos 2\phi \right. \\ &\quad \left. + (C_{422y} H_x + C_{422x} H_y) \rho^2 \sin 2\phi \right] \end{aligned} \quad (\text{J.124})$$

$$W_T = (\chi_9 H_x + \chi_8 H_y) Z_5(\rho, \phi) + (\chi_8 H_x - \chi_9 H_y) Z_6(\rho, \phi) \quad (\text{J.125})$$

$$\begin{aligned} &\Rightarrow \chi_8 \left[ Z_3(h, \theta) Z_5(\rho, \phi) + Z_2(h, \theta) Z_6(\rho, \phi) \right] \\ &\quad + \chi_9 \left[ Z_2(h, \theta) Z_5(\rho, \phi) - Z_3(h, \theta) Z_6(\rho, \phi) \right] \end{aligned} \quad (\text{J.126})$$

### Constant astigmatism

$$W = \frac{1}{2} \left( \vec{D}_{422}^2 \cdot \vec{\rho}^2 \right) \quad (\text{J.127})$$

$$= \frac{1}{2} (2 D_{422x} D_{422y} Z_5(\rho, \phi) + (D_{422x}^2 - D_{422y}^2) Z_6(\rho, \phi)) \quad (\text{J.128})$$

$$W_T = \psi_{10} Z_5(\rho, \phi) + \psi_{11} Z_6(\rho, \phi) \quad (\text{J.129})$$

$$\Rightarrow \psi_{10} \underline{Z_1(h, \theta) Z_5(\rho, \phi)} + \psi_{11} \underline{Z_1(h, \theta) Z_6(\rho, \phi)} \quad (\text{J.130})$$

$$(\text{J.131})$$

### J.4.3 FIELD CUBED COMA

#### Cubic coma

$$W = W_{331M} \left( \vec{H} \cdot \vec{H} \right) \left( \vec{H} \cdot \vec{\rho} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.132})$$

$$= W_{331M} [(H_x^2 + H_y^2)(H_x \rho \cos \phi + H_y \rho \sin \phi)] \rho^2 \quad (\text{J.133})$$

$$= W_{331M} [H_x (H_x^2 + H_y^2) \rho^3 \cos \phi + H_y (H_x^2 + H_y^2) \rho^3 \sin \phi] \quad (\text{J.134})$$

$$W_T = \xi_0 [(H_x^3 + H_x H_y^2) Z_8(\rho, \phi) + (H_x^2 H_y + H_y^3) Z_7(\rho, \phi)] \quad (\text{J.135})$$

$$= \xi_0 [a Z_8(h, \theta) Z_8(\rho, \phi) + b Z_2(h, \theta) Z_8(\rho, \phi)] \quad (\text{J.136})$$

$$+ a Z_7(h, \theta) Z_7(\rho, \phi) + b Z_3(h, \theta) Z_7(\rho, \phi)]$$

$$\Rightarrow \xi_0 \left\{ a \left[ \underline{Z_7(h, \theta) Z_7(\rho, \phi) + Z_8(h, \theta) Z_8(\rho, \phi)} \right] \right. \quad (\text{J.137})$$

$$\left. + b \left[ \underline{Z_3(h, \theta) Z_7(\rho, \phi) + Z_2(h, \theta) Z_8(\rho, \phi)} \right] \right\}$$

### Quadratic coma 1

$$W = -2 \left( \vec{H} \cdot \vec{A}_{331_M} \right) \left( \vec{H} \cdot \vec{\rho} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.138})$$

$$= -2 (A_{331_Mx} H_x + A_{331_My} H_y) (H_x \rho \cos \phi + H_y \rho \sin \phi) \rho^2 \quad (\text{J.139})$$

$$= (\xi_2 H_x^2 + \xi_1 H_x H_y) \rho^3 \cos \phi + (\xi_2 H_x H_y + \xi_1 H_y^2) \rho^3 \sin \phi \quad (\text{J.140})$$

$$W_T = (\xi_2 H_x^2 + \xi_1 H_x H_y) Z_8(\rho, \phi) + (\xi_2 H_x H_y + \xi_1 H_y^2) Z_7(\rho, \phi) \quad (\text{J.141})$$

$$= \xi_1 [H_x H_y Z_8(\rho, \phi) + H_y^2 Z_7(\rho, \phi)] \quad (\text{J.142})$$

$$+ \xi_2 [H_x^2 Z_8(\rho, \phi) + H_x H_y Z_7(\rho, \phi)]$$

$$= \xi_1 \left\{ a Z_5(h, \theta) Z_8(\rho, \phi) \right. \\ \left. + [-a Z_6(h, \theta) + b Z_4(h, \theta) + c Z_1(h, \theta)] Z_7(\rho, \phi) \right\} \quad (\text{J.143})$$

$$+ \xi_2 \left\{ [a Z_6(h, \theta) + b Z_4(h, \theta) + c Z_1(h, \theta)] Z_8(h, \theta) \right.$$

$$\left. + a Z_5(h, \theta) Z_7(\rho, \phi) \right\}$$

$$\Rightarrow \xi_1 \left\{ a \left[ \underline{Z_5(h, \theta) Z_8(\rho, \phi) - Z_6(h, \theta) Z_7(\rho, \phi)} \right] \right. \\ \left. + b \underline{Z_4(h, \theta) Z_7(\rho, \phi)} + c \underline{Z_1(h, \theta) Z_7(\rho, \phi)} \right\} \quad (\text{J.144})$$

$$+ \xi_2 \left\{ a \left[ \underline{Z_6(h, \theta) Z_8(\rho, \phi) + Z_5(h, \theta) Z_7(\rho, \phi)} \right] \right.$$

$$\left. + b \underline{Z_4(h, \theta) Z_8(\rho, \phi)} + c \underline{Z_1(h, \theta) Z_8(\rho, \phi)} \right\}$$

### Linear coma 1

$$W = 2 B_{331_M} \left( \vec{H} \cdot \vec{\rho} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.145})$$

$$= 2 B_{331_M} (H_x \rho \cos \phi + H_y \rho \sin \phi) \rho^2 \quad (\text{J.146})$$

$$W_T = \xi_3 H_y Z_7(\rho, \phi) + \xi_3 H_x Z_8(\rho, \phi) \quad (\text{J.147})$$

$$\Rightarrow \xi_3 \left[ \underline{Z_3(h, \theta) Z_7(\rho, \phi) + Z_2(h, \theta) Z_8(\rho, \phi)} \right] \quad (\text{J.148})$$

### Quadratic coma 2

$$W = - \left( \vec{H} \cdot \vec{H} \right) \left( \vec{A}_{331_M} \cdot \vec{\rho} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.149})$$

$$= - (H_x^2 + H_y^2) (A_{331_Mx} \rho \cos \phi + A_{331_My} \rho \sin \phi) \rho^2 \quad (\text{J.150})$$

$$W_T = \xi_4 (H_x^2 + H_y^2) Z_7(\rho, \phi) + \xi_5 (H_x^2 + H_y^2) Z_8(\rho, \phi) \quad (\text{J.151})$$

$$\Rightarrow \xi_4 \left[ a \underline{Z_4(h, \theta) Z_7(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_7(\rho, \phi)} \right] \quad (\text{J.152})$$

$$+ \xi_5 \left[ a \underline{Z_4(h, \theta) Z_8(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_8(\rho, \phi)} \right]$$



## Linear coma 2

$$W = \left[ \left( \vec{B}_{331M}^2 \vec{H}^* \right) \cdot \vec{\rho} \right] (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.153})$$

$$= \left\{ \left[ \left( B_{331Mx}^2 - B_{331My}^2 \right) \hat{i} + \left( 2 B_{331Mx} B_{331My} \right) \hat{j} \right] \left( H_x \hat{i} - H_y \hat{j} \right) \right\} \cdot \vec{\rho} (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.154})$$

$$= \left[ \left( B_{331Mx}^2 - B_{331My}^2 \right) H_x - \left( 2 B_{331Mx} B_{331My} \right) \left( -H_y \right) \right] \rho^3 \cos \phi \quad (\text{J.155})$$

$$+ \left[ \left( 2 B_{331Mx} B_{331My} \right) H_x + \left( B_{331Mx}^2 - B_{331My}^2 \right) \left( -H_y \right) \right] \rho^3 \sin \phi$$

$$W_T = \left( -\xi_7 H_x + \xi_6 H_y \right) Z_8(\rho, \phi) + \left( \xi_6 H_x + \xi_7 H_y \right) Z_7(\rho, \phi) \quad (\text{J.156})$$

$$= \xi_7 \left[ -H_x Z_8(\rho, \phi) + H_y Z_7(\rho, \phi) \right] + \xi_6 \left[ H_y Z_8(\rho, \phi) + H_x Z_7(\rho, \phi) \right] \quad (\text{J.157})$$

$$\Rightarrow \xi_7 \left[ \underline{Z_3(h, \theta) Z_7(\rho, \phi) - Z_2(h, \theta) Z_8(\rho, \phi)} \right] \quad (\text{J.158})$$

$$+ \xi_6 \left[ \underline{Z_2(h, \theta) Z_7(\rho, \phi) + Z_3(h, \theta) Z_8(\rho, \phi)} \right]$$

## Constant coma

$$W = - \left( \vec{C}_{331M} \cdot \vec{\rho} \right) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.159})$$

$$= - \left( C_{331Mx} \rho \cos \phi + C_{331My} \rho \sin \phi \right) \rho^2 \quad (\text{J.160})$$

$$= \xi_{10} \rho^3 \cos \phi + \xi_9 \rho^3 \sin \phi \quad (\text{J.161})$$

$$W_T = \xi_9 Z_7(\rho, \phi) + \xi_{10} Z_8(\rho, \phi) \quad (\text{J.162})$$

$$\Rightarrow \xi_9 \underline{Z_1(h, \theta) Z_7(\rho, \phi)} + \xi_{10} \underline{Z_1(h, \theta) Z_8(\rho, \phi)} \quad (\text{J.163})$$

$$(\text{J.164})$$

## J.4.4 TREFOIL

**Cubic trefoil**

$$W = \frac{1}{4} W_{333} \left( \vec{H}^3 \cdot \vec{\rho}^3 \right) \quad (\text{J.165})$$

$$= \frac{1}{4} W_{333} \left[ (H_x^3 - 3 H_x H_y^2) \rho^3 \cos 3\phi + (3 H_x^2 H_y - H_y^3) \rho^3 \sin 3\phi \right] \quad (\text{J.166})$$

$$W_T = \mu_0 \left[ (3 H_x^2 H_y - H_y^3) Z_9(\rho, \phi) + (H_x^3 - 3 H_x H_y^2) Z_{10}(\rho, \phi) \right] \quad (\text{J.167})$$

$$\Rightarrow \mu_0 \left[ \underline{Z_9(h, \theta) Z_9(\rho, \phi) + Z_{10}(h, \theta) Z_{10}(\rho, \phi)} \right] \quad (\text{J.168})$$

**Quadratic trefoil**

$$W = -\frac{3}{4} \left( \vec{H}^2 \vec{A}_{333} \right) \cdot \vec{\rho}^3 \quad (\text{J.169})$$

$$= -\frac{3}{4} \left\{ [(H_x^2 - H_y^2) \hat{i} + 2 H_x H_y \hat{j}] (A_{333x} \hat{i} + A_{333y} \hat{j}) \right\} \cdot \vec{\rho}^3 \quad (\text{J.170})$$

$$= -\frac{3}{4} \left\{ [(H_x^2 - H_y^2) A_{333x} - 2 H_x H_y A_{333y}] \rho^3 \cos 3\phi \right. \quad (\text{J.171})$$

$$\left. + [2 H_x H_y A_{333x} + (H_x^2 - H_y^2) A_{333y}] \rho^3 \sin 3\phi \right\}$$

$$W_T = [\mu_2 (H_x^2 - H_y^2) - \mu_1 (2 H_x H_y)] Z_{10}(\rho, \phi) \quad (\text{J.172})$$

$$+ [\mu_2 (2 H_x H_y) + \mu_1 (H_x^2 - H_y^2)] Z_9(\rho, \phi)$$

$$\Rightarrow \mu_1 \left[ \underline{Z_6(h, \theta) Z_9(\rho, \phi) - Z_5(h, \theta) Z_{10}(\rho, \phi)} \right] \quad (\text{J.173})$$

$$+ \mu_2 \left[ \underline{Z_5(h, \theta) Z_9(\rho, \phi) + Z_6(h, \theta) Z_{10}(\rho, \phi)} \right]$$

### Linear trefoil

$$W = \frac{3}{4} \left( \vec{H} \vec{B}_{333}^2 \right) \cdot \vec{\rho}^3 \quad (\text{J.174})$$

$$\begin{aligned} &= \frac{3}{4} \left\{ (B_{333x}^2 - B_{333y}^2) H_x - (2 B_{333x} B_{333y}) H_y \right\} \rho^3 \cos 3\phi \\ &\quad + \left\{ (2 B_{333x} B_{333y}) H_x + (B_{333x}^2 - B_{333y}^2) H_y \right\} \rho^3 \sin 3\phi \end{aligned} \quad (\text{J.175})$$

$$W_T = (\mu_3 H_y + \mu_4 H_x) Z_9(\rho, \phi) + (\mu_3 H_x - \mu_4 H_y) Z_{10}(\rho, \phi) \quad (\text{J.176})$$

$$\begin{aligned} &\Rightarrow \mu_3 \left[ \underline{Z_3(h, \theta) Z_9(\rho, \phi) + Z_2(h, \theta) Z_{10}(\rho, \phi)} \right] \\ &\quad + \mu_4 \left[ \underline{Z_2(h, \theta) Z_9(\rho, \phi) - Z_3(h, \theta) Z_{10}(\rho, \phi)} \right] \end{aligned} \quad (\text{J.177})$$

### Constant trefoil

$$W = -\frac{1}{4} \vec{C}_{333}^3 \cdot \vec{\rho}^3 \quad (\text{J.178})$$

$$\begin{aligned} &= -\frac{1}{4} \left[ (C_{333x}^3 - 3 C_{333x} C_{333y}^2) \rho^3 \cos 3\phi \right. \\ &\quad \left. + (3 C_{333x}^2 C_{333y} - C_{333y}^3) \rho^3 \sin 3\phi \right] \end{aligned} \quad (\text{J.179})$$

$$W_T = \mu_5 Z_9(\rho, \phi) + \mu_6 Z_{10}(\rho, \phi) \quad (\text{J.180})$$

$$\Rightarrow \mu_5 \underline{Z_1(h, \theta) Z_9(\rho, \phi)} + \mu_6 \underline{Z_1(h, \theta) Z_{10}(\rho, \phi)} \quad (\text{J.181})$$

#### J.4.5 FIELD CURVATURE FOR OBLIQUE SPHERICAL

Tessieres changed  $\rho^4$  into only  $Z_{11}(\rho, \phi)$ . Since  $Z_{11}(\rho, \phi)$  includes a  $\rho^2$  term,  $W_{240}$  and the other perturbation vectors ( $\vec{A}_{240_m}$  and  $B_{240_M}$ ) also contribute to the focus aberration  $Z_4(\rho, \phi)$ . Thus in the line after  $W_T$ ,  $Z_{11}(\rho, \phi)$  changes into  $a Z_{11}(\rho, \phi) + b Z_4(\rho, \phi)$ . This does not change the functional form of the  $Z_4(\rho, \phi)$  aberration derived by Tessieres. This only helps describe in more detail the source of each of

the aberrations when written as Zernike polynomials.

### Quadratic field curvature for oblique spherical

$$W = W_{240_M} \left( \vec{H} \cdot \vec{H} \right) (\vec{\rho} \cdot \vec{\rho})^2 \quad (\text{J.182})$$

$$= W_{240_M} (H_x^2 + H_y^2) \rho^4 \quad (\text{J.183})$$

$$W_T = \delta_0 (H_x^2 + H_y^2) Z_{11}(\rho, \phi) \quad (\text{J.184})$$

$$= \delta_0 [a Z_4(h, \theta) + b Z_1(h, \theta)] [c Z_{11}(\rho, \phi) + d Z_4(\rho, \phi)] \quad (\text{J.185})$$

$$\begin{aligned} \Rightarrow \delta_0 & \left[ a c \underline{Z_4(h, \theta) Z_{11}(\rho, \phi)} + a d \underline{Z_4(h, \theta) Z_4(\rho, \phi)} \right. \\ & \left. + b c \underline{Z_1(h, \theta) Z_{11}(\rho, \phi)} + b d \underline{Z_1(h, \theta) Z_4(\rho, \phi)} \right] \end{aligned} \quad (\text{J.186})$$

### Linear field curvature for oblique spherical

$$W = -2 \vec{H} \cdot \vec{A}_{240_M} (\vec{\rho} \cdot \vec{\rho})^2 \quad (\text{J.187})$$

$$= -2 (A_{240_Mx} H_x + A_{240_My} H_y) \rho^4 \quad (\text{J.188})$$

$$W_T = (\delta_1 H_y + \delta_2 H_x) Z_{11}(\rho, \phi) \quad (\text{J.189})$$

$$= [\delta_1 Z_3(h, \theta) + \delta_2 Z_2(h, \theta)] [a Z_{11}(\rho, \phi) + b Z_4(\rho, \phi)] \quad (\text{J.190})$$

$$\begin{aligned} \Rightarrow \delta_1 & \left[ a \underline{Z_3(h, \theta) Z_{11}(\rho, \phi)} + b \underline{Z_3(h, \theta) Z_4(\rho, \phi)} \right] \\ & + \delta_2 \left[ a \underline{Z_2(h, \theta) Z_{11}(\rho, \phi)} + b \underline{Z_2(h, \theta) Z_4(\rho, \phi)} \right] \end{aligned} \quad (\text{J.191})$$

### Constant field curvature for oblique spherical

$$W = B_{240_M} (\vec{\rho} \cdot \vec{\rho})^2 \quad (\text{J.192})$$

$$W_T = \delta_3 Z_{11}(\rho, \phi) \quad (\text{J.193})$$

$$= \delta_3 [a Z_{11}(\rho, \phi) + b Z_4(\rho, \phi)] \quad (\text{J.194})$$

$$\Rightarrow \delta_3 \left[ a \underline{Z_1(h, \theta) Z_{11}(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_4(\rho, \phi)} \right] \quad (\text{J.195})$$

#### J.4.6 OBLIQUE SPHERICAL ABERRATION

As for the case of field curvature for oblique spherical aberration, Tessieres changed  $\rho^3 \cos 2\phi$  into only  $Z_{12}(\rho, \phi)$ , but this Zernike polynomial term is also balanced with primary astigmatism  $Z_6(\rho, \phi)$ . Thus in the line after  $W_T$ ,  $Z_{12}(\rho, \phi)$  will also be converted into  $Z_6(\rho, \phi)$  and similarly,  $Z_{13}(\rho, \phi)$  will also be converted in  $Z_5(\rho, \phi)$ .

#### Quadratic oblique spherical aberration

$$W = \frac{1}{2} W_{242} (\vec{H}^2 \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.196})$$

$$= \frac{1}{2} W_{242} [(H_x^2 - H_y^2) \rho^2 \cos 2\phi + 2 H_x H_y \rho^2 \sin 2\phi] \rho^2 \quad (\text{J.197})$$

$$= \eta_0 [(H_x^2 - H_y^2) \rho^4 \cos 2\phi + 2 H_x H_y \rho^4 \sin 2\phi] \quad (\text{J.198})$$

$$W_T = \eta_0 [(H_x^2 - H_y^2) Z_{12}(\rho, \phi) + 2 H_x H_y Z_{13}(\rho, \phi)] \quad (\text{J.199})$$

$$= \eta_0 \{ (H_x^2 - H_y^2) [a Z_{12}(\rho, \phi) + b Z_6(\rho, \phi)] \quad (\text{J.200})$$

$$+ 2 H_x H_y [a Z_{13}(\rho, \phi) + b Z_5(\rho, \phi)] \}$$

$$\Rightarrow \eta_0 \left\{ a \left[ \underline{Z_6(h, \theta) Z_{12}(\rho, \phi)} + \underline{Z_5(h, \theta) Z_{13}(\rho, \phi)} \right] \quad (\text{J.201}) \right.$$

$$\left. + b \left[ \underline{Z_5(h, \theta) Z_5(\rho, \phi)} + \underline{Z_6(h, \theta) Z_6(\rho, \phi)} \right] \right\}$$

### Linear oblique spherical aberration

$$W = -\vec{H} \vec{A}_{242} \cdot \vec{\rho} (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.202})$$

$$= -(A_{242x} H_x - A_{242y} H_y) \rho^3 \cos \phi - (A_{242y} H_x + A_{242x} H_y) \rho^3 \sin \phi \quad (\text{J.203})$$

$$W_T = (\eta_2 H_x - \eta_1 H_y) Z_{12}(\rho, \phi) + (\eta_1 H_x + \eta_2 H_y) Z_{13}(\rho, \phi) \quad (\text{J.204})$$

$$= \eta_1 \{ H_x [a Z_{13}(\rho, \phi) + b Z_5(\rho, \phi)] - H_y [a Z_{12}(\rho, \phi) + b Z_6(\rho, \phi)] \} \quad (\text{J.205})$$

$$+ \eta_2 \{ H_x [a Z_{12}(\rho, \phi) + b Z_6(\rho, \phi)] + H_y [a Z_{13}(\rho, \phi) + b Z_5(\rho, \phi)] \} \\ \Rightarrow \eta_1 \left\{ a \left[ \underline{Z_2(h, \theta) Z_{13}(\rho, \phi) - Z_3(h, \theta) Z_{12}(\rho, \phi)} \right] \right. \\ \left. + b \left[ \underline{Z_2(h, \theta) Z_5(\rho, \phi) - Z_3(h, \theta) Z_6(\rho, \phi)} \right] \right\} \quad (\text{J.206})$$

$$+ \eta_2 \left\{ a \left[ \underline{Z_2(h, \theta) Z_{12}(\rho, \phi) + Z_3(h, \theta) Z_{13}(\rho, \phi)} \right] \right. \\ \left. + b \left[ \underline{Z_2(h, \theta) Z_6(\rho, \phi) + Z_3(h, \theta) Z_5(\rho, \phi)} \right] \right\}$$

### Constant oblique spherical aberration

$$W = \frac{1}{2} \vec{B}_{242}^2 \cdot \vec{\rho} (\vec{\rho} \cdot \vec{\rho}) \quad (\text{J.207})$$

$$= \frac{1}{2} [(H_x^2 - H_y^2) \rho^3 \cos \phi + 2 H_x H_y \rho^3 \sin \phi] \quad (\text{J.208})$$

$$W_T = \eta_3 Z_{12}(\rho, \phi) + \eta_4 Z_{13}(\rho, \phi) \quad (\text{J.209})$$

$$= \eta_3 [a Z_{12}(\rho, \phi) + b Z_6(\rho, \phi)] + \eta_4 [a Z_{13}(\rho, \phi) + b Z_5(\rho, \phi)] \quad (\text{J.210})$$

$$\Rightarrow \eta_3 \left[ a \underline{Z_1(h, \theta) Z_{12}(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_6(\rho, \phi)} \right] \quad (\text{J.211}) \\ + \eta_4 \left[ a \underline{Z_1(h, \theta) Z_{13}(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_5(\rho, \phi)} \right]$$

#### J.4.7 FIFTH ORDER (SECONDARY) COMA

Tessieres converted  $\rho^5 \cos \phi$  into  $Z_{16}(\rho, \phi)$ , but this Zernike polynomial term is also balanced by  $Z_8(\rho, \phi)$ . (Similarly, for  $Z_{17}(\rho, \phi)$  and  $Z_7(\rho, \phi)$ .) Thus, the extra terms will be added back in the line after  $W_T$ .

##### Linear fifth order coma

$$W = W_{151} \vec{H} \cdot \vec{\rho} (\vec{\rho} \cdot \vec{\rho})^2 \quad (\text{J.212})$$

$$= W_{151} (H_x \rho \cos \phi + H_y \rho \sin \phi) \rho^4 \quad (\text{J.213})$$

$$= \kappa_0 (H_x \rho^5 \cos \phi + H_y \rho^5 \sin \phi) \quad (\text{J.214})$$

$$W_T = \kappa_0 [H_x Z_{16}(\rho, \phi) + H_y Z_{17}(\rho, \phi)] \quad (\text{J.215})$$

$$= \kappa_0 \{ Z_2(h, \theta) [a Z_{16}(\rho, \phi) + b Z_8(\rho, \phi)] \quad (\text{J.216})$$

$$+ Z_3(h, \theta) [a Z_{17}(\rho, \phi) + b Z_7(\rho, \phi)] \}$$

$$\Rightarrow \kappa_0 \left\{ a \left[ \underline{Z_2(h, \theta) Z_{16}(\rho, \phi) + Z_3(h, \theta) Z_{17}(\rho, \phi)} \right] \quad (\text{J.217}) \right.$$

$$\left. + b \left[ \underline{Z_2(h, \theta) Z_8(\rho, \phi) + Z_3(h, \theta) Z_7(\rho, \phi)} \right] \right\}$$

##### Constant fifth order coma

$$W = -\vec{A}_{151} \cdot \vec{\rho} (\vec{\rho} \cdot \vec{\rho})^2 \quad (\text{J.218})$$

$$= -(A_{151x} \rho^5 \cos \phi + A_{151y} \rho^5 \sin \phi) \quad (\text{J.219})$$

$$W_T = \kappa_1 Z_{16}(\rho, \phi) + \kappa_2 Z_{17}(\rho, \phi) \quad (\text{J.220})$$

$$\Rightarrow \kappa_1 \left[ a \underline{Z_1(h, \theta) Z_{16}(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_8(\rho, \phi)} \right] \quad (\text{J.221})$$

$$+ \kappa_2 \left[ a \underline{Z_1(h, \theta) Z_{17}(\rho, \phi)} + b \underline{Z_1(h, \theta) Z_7(\rho, \phi)} \right]$$

## APPENDIX K

## ORTHOGONAL ABERRATION FUNCTIONS

This chapter shows wavefront maps across the field for a variety of double Zernike functions that are introduced in Chapter 6. Tables 6.1–6.7 list the dependence that these functions have on the misalignment.

One can observe that the functions that exist in a rotationally symmetric system have wavefront maps that are rotationally symmetric. These include defocus ( $Z_1(h, \theta)Z_4(\rho, \phi)$ ), quadratic astigmatism ( $Z_5(h, \theta)Z_5(\rho, \phi) + Z_6(h, \theta)Z_6(\rho, \phi)$ ), and linear coma ( $Z_3(h, \theta)Z_7(\rho, \phi) + Z_2(h, \theta)Z_8(\rho, \phi)$ ), among many others for example.



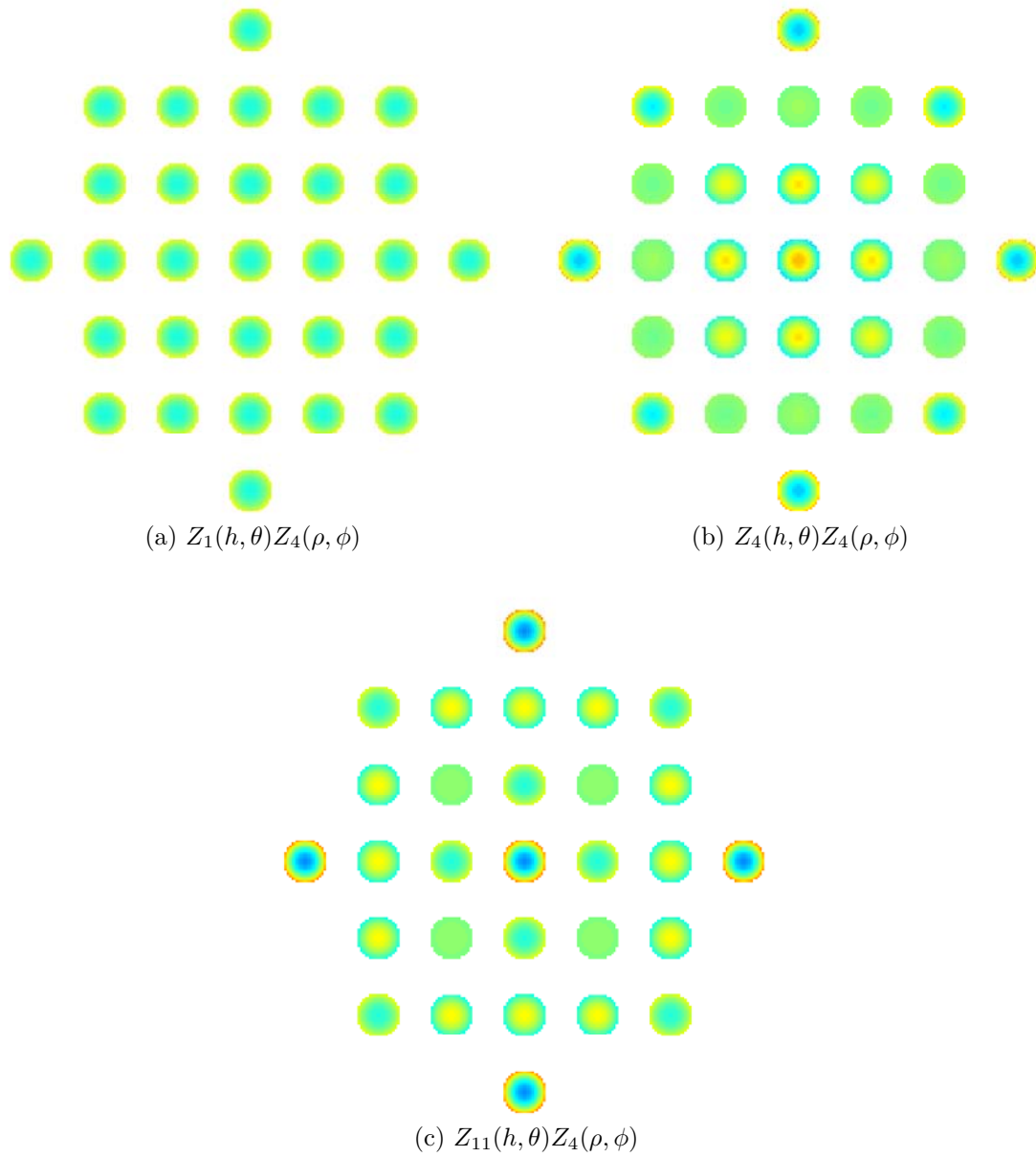


Figure K.1: Orthogonal focus polynomials.

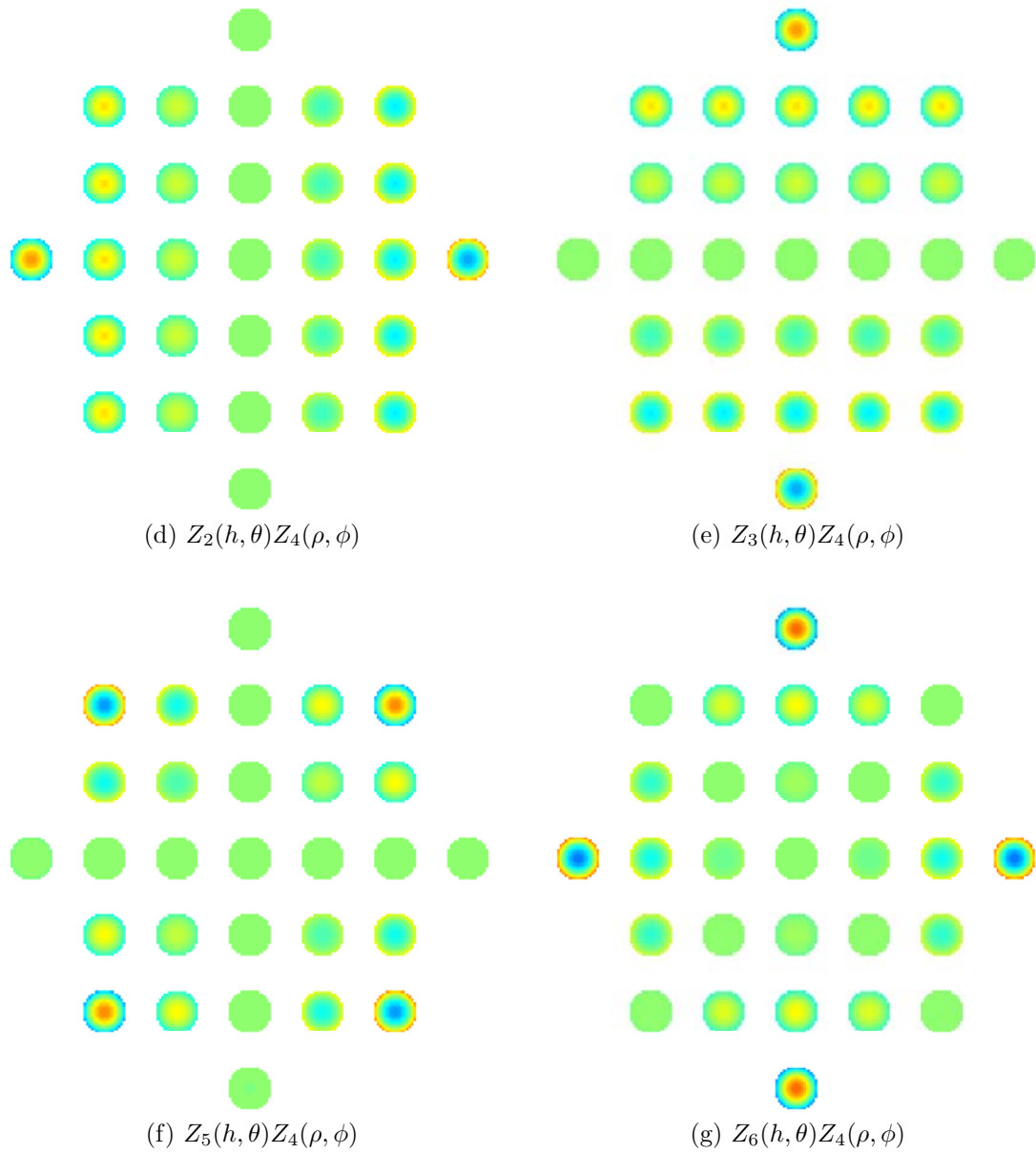


Figure K.1: Orthogonal focus polynomials continued.

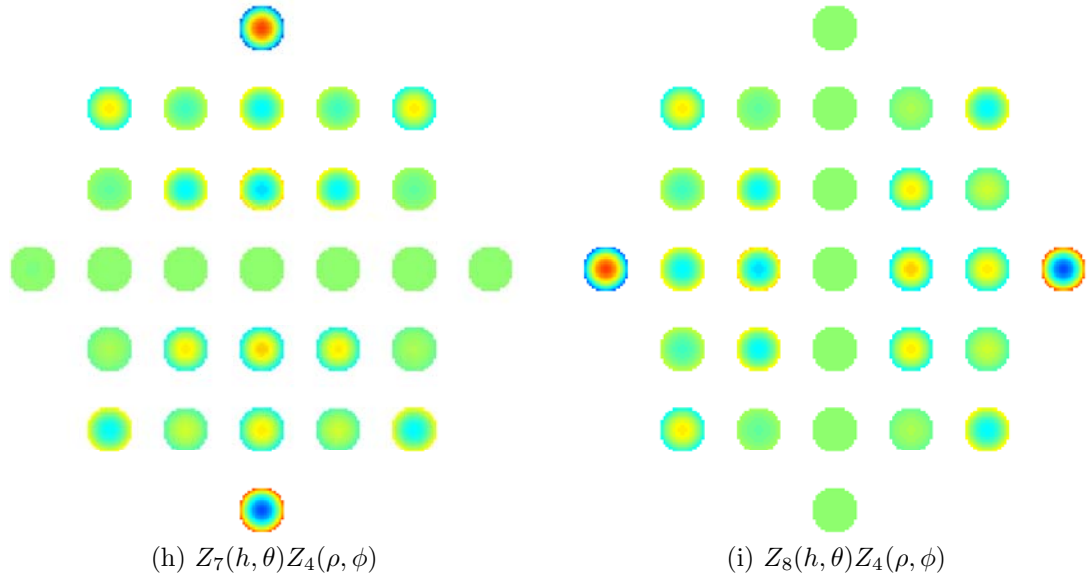


Figure K.1: Orthogonal focus polynomials continued.

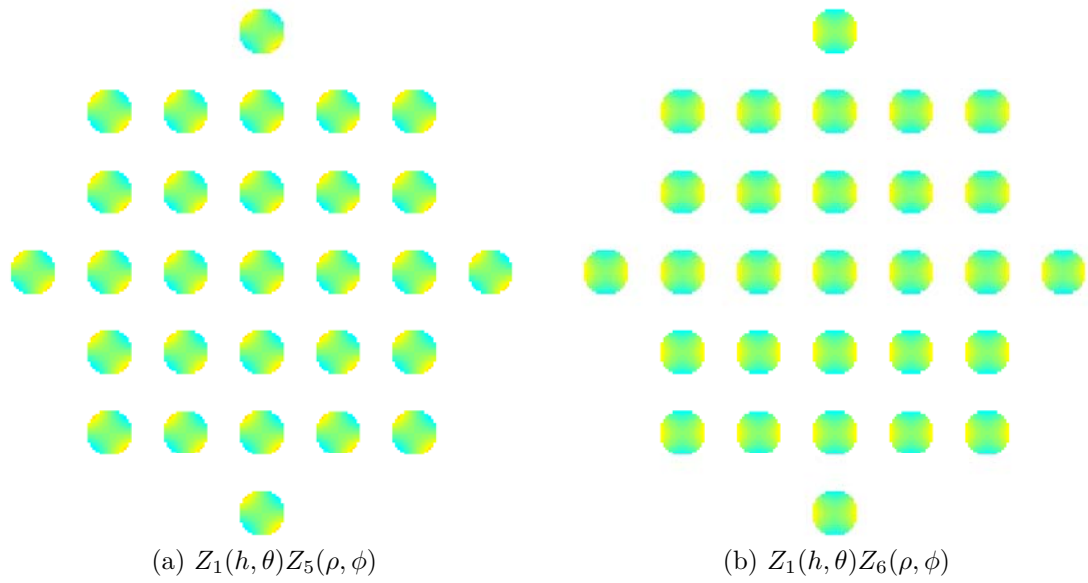


Figure K.2: Orthogonal astigmatism polynomials.

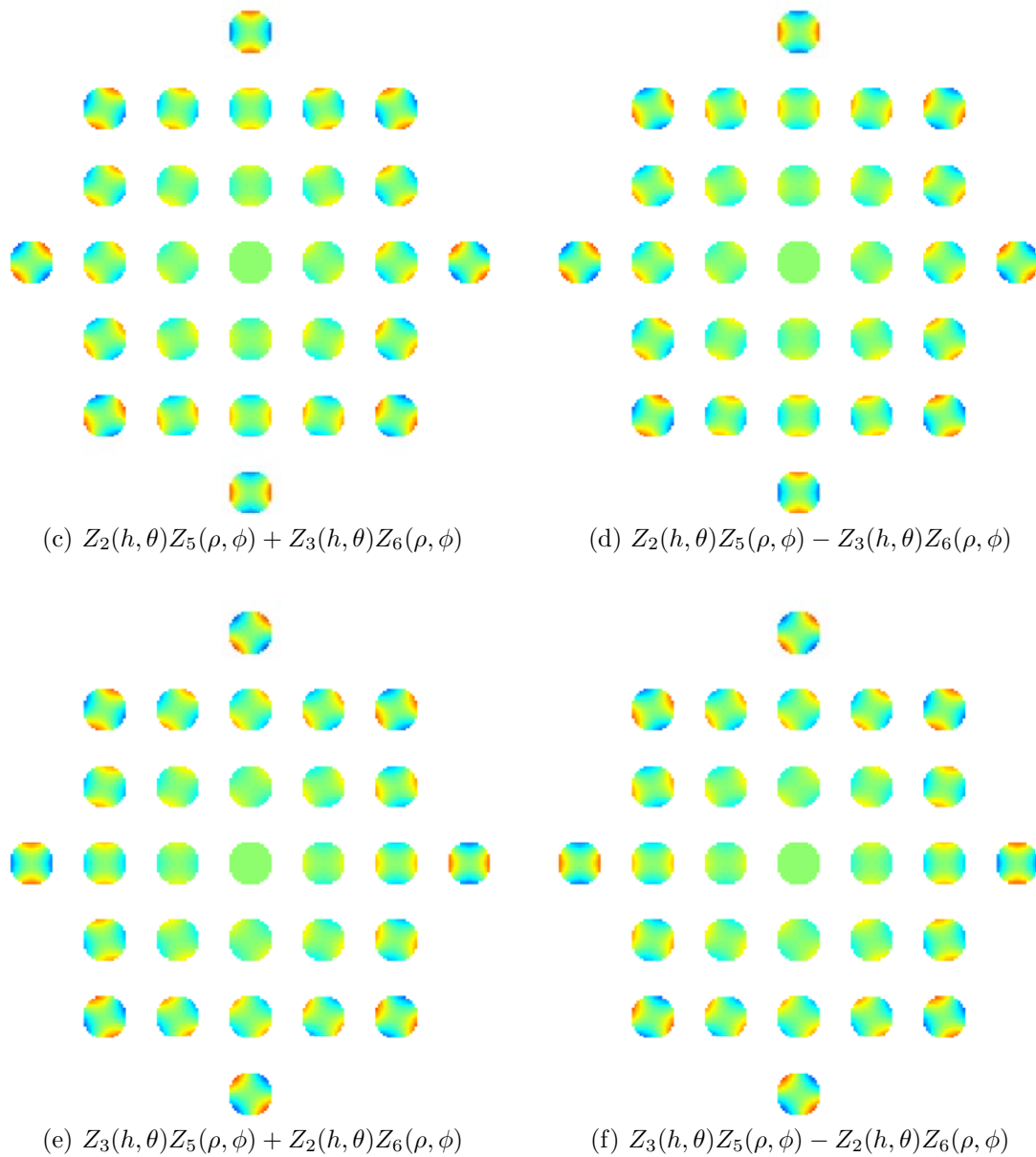


Figure K.2: Orthogonal astigmatism polynomials continued.

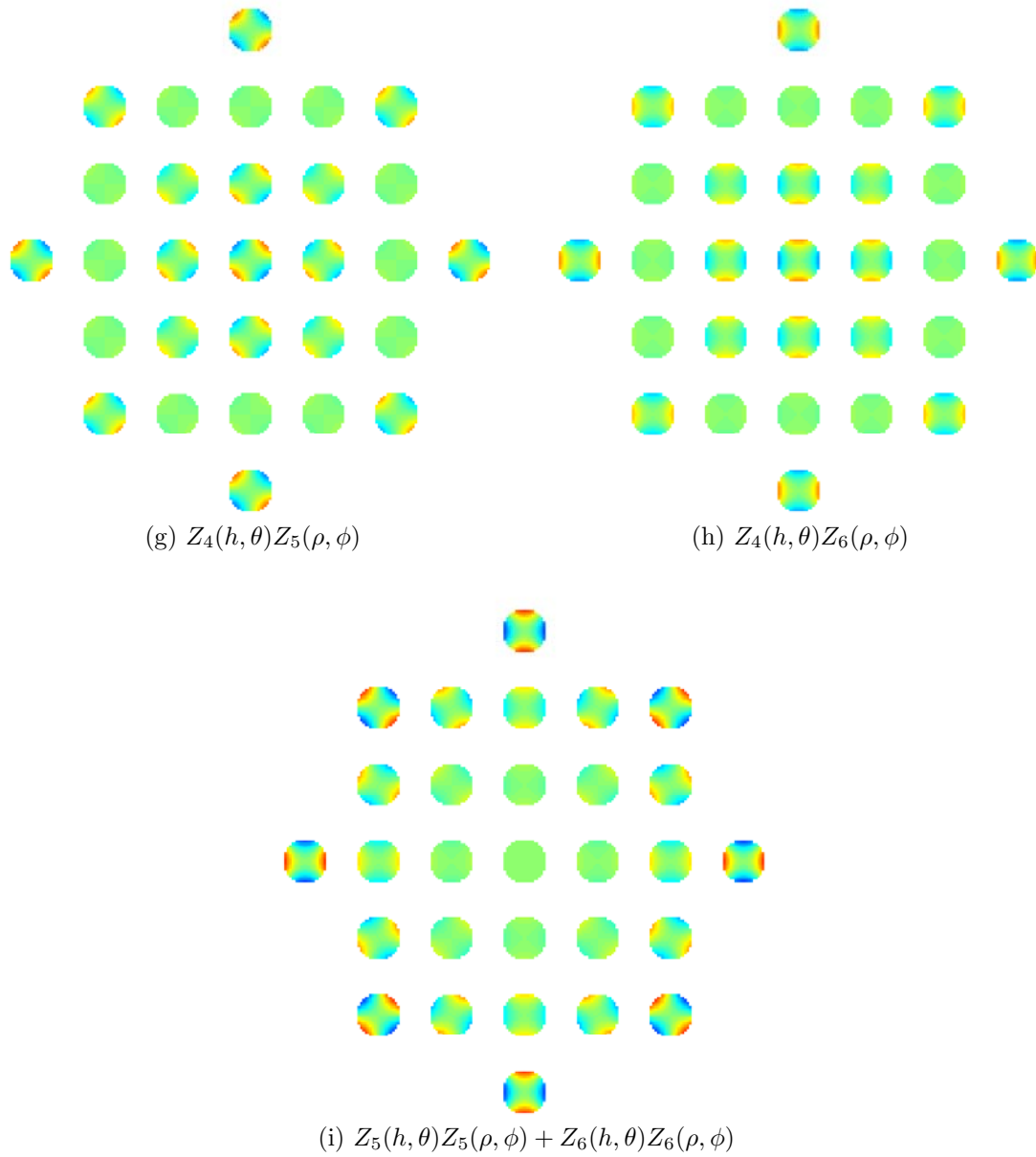


Figure K.2: Orthogonal astigmatism polynomials continued.

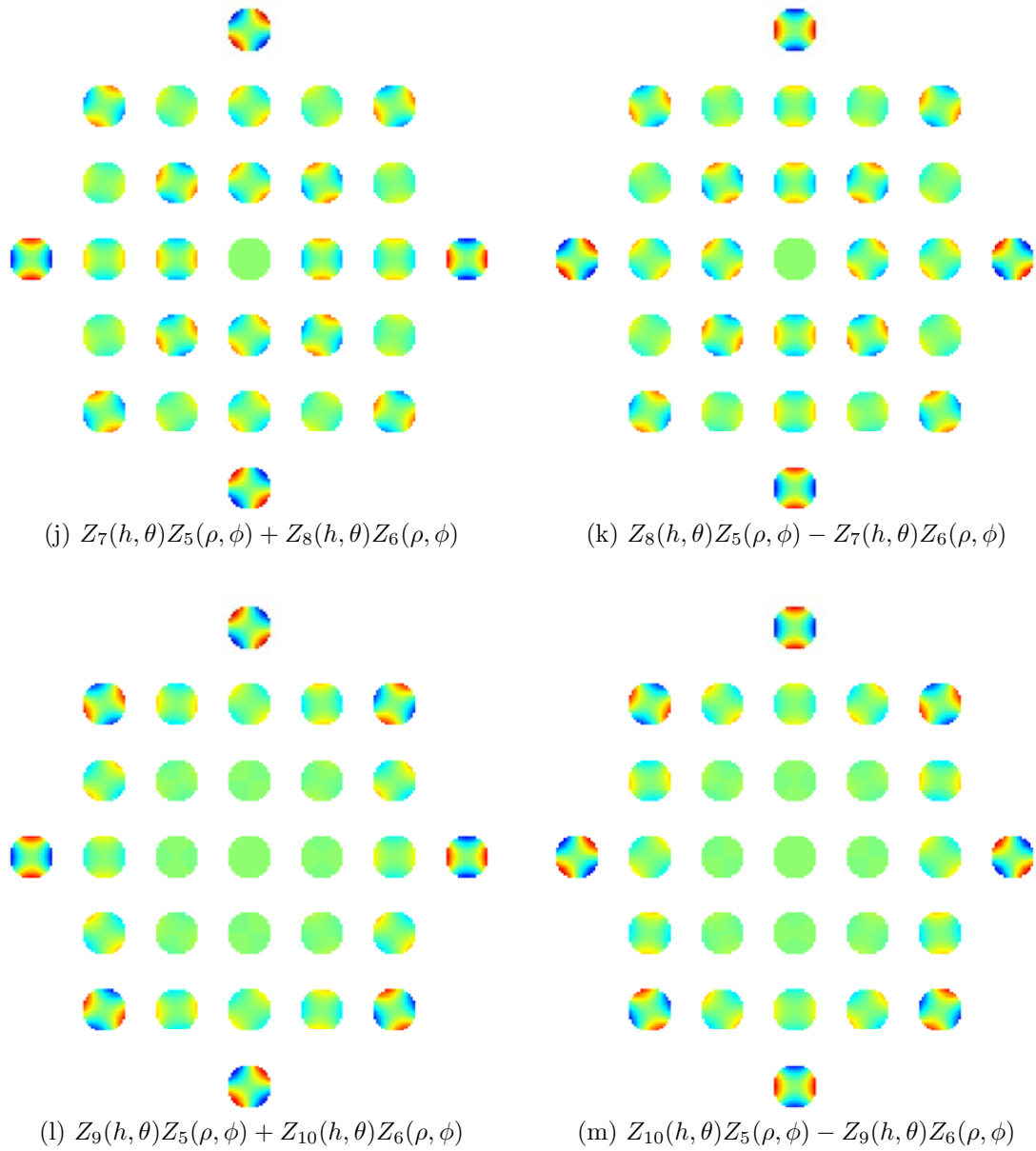
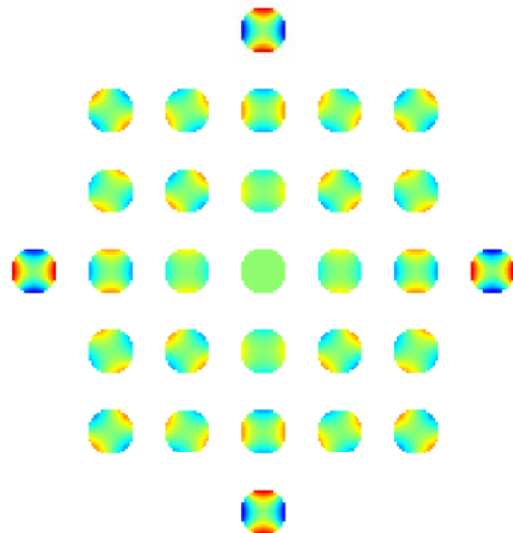
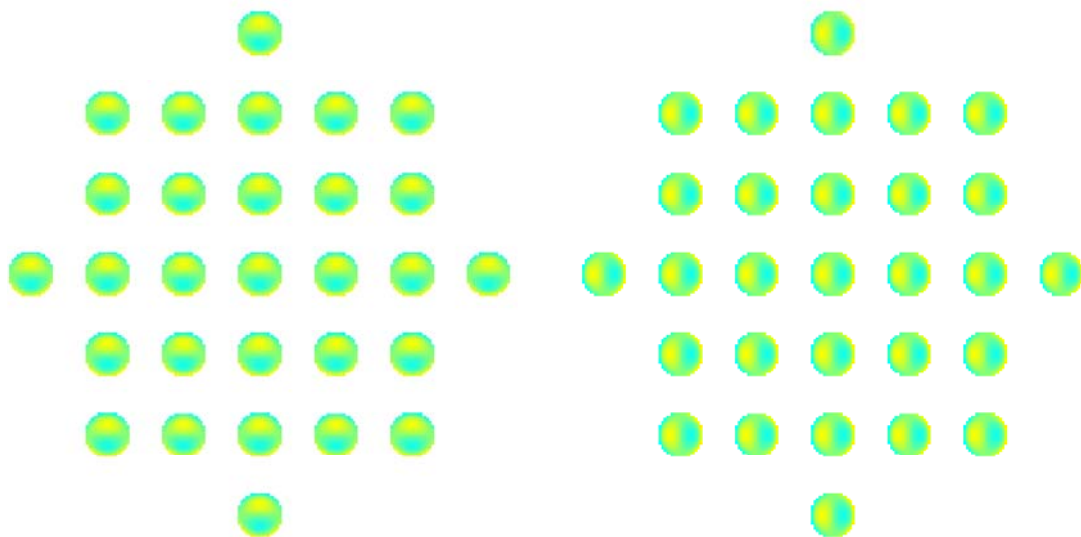


Figure K.2: Orthogonal astigmatism polynomials continued.



$$(n) Z_{13}(h, \theta)Z_5(\rho, \phi) + Z_{12}(h, \theta)Z_6(\rho, \phi)$$

Figure K.2: Orthogonal astigmatism polynomials continued.



$$(a) Z_1(h, \theta)Z_7(\rho, \phi)$$

$$(b) Z_1(h, \theta)Z_8(\rho, \phi)$$

Figure K.3: Orthogonal coma polynomials.

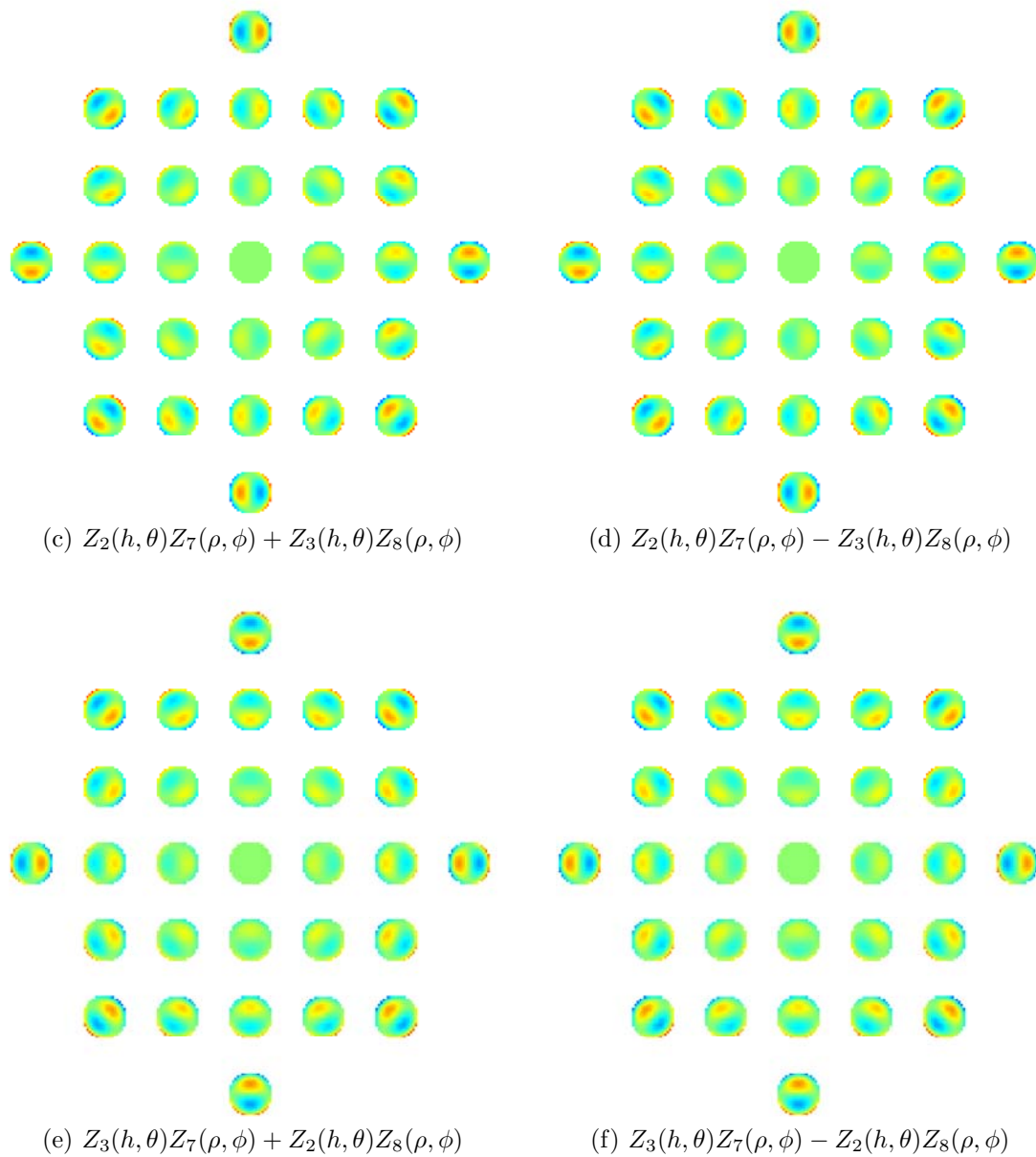


Figure K.3: Orthogonal coma polynomials continued.



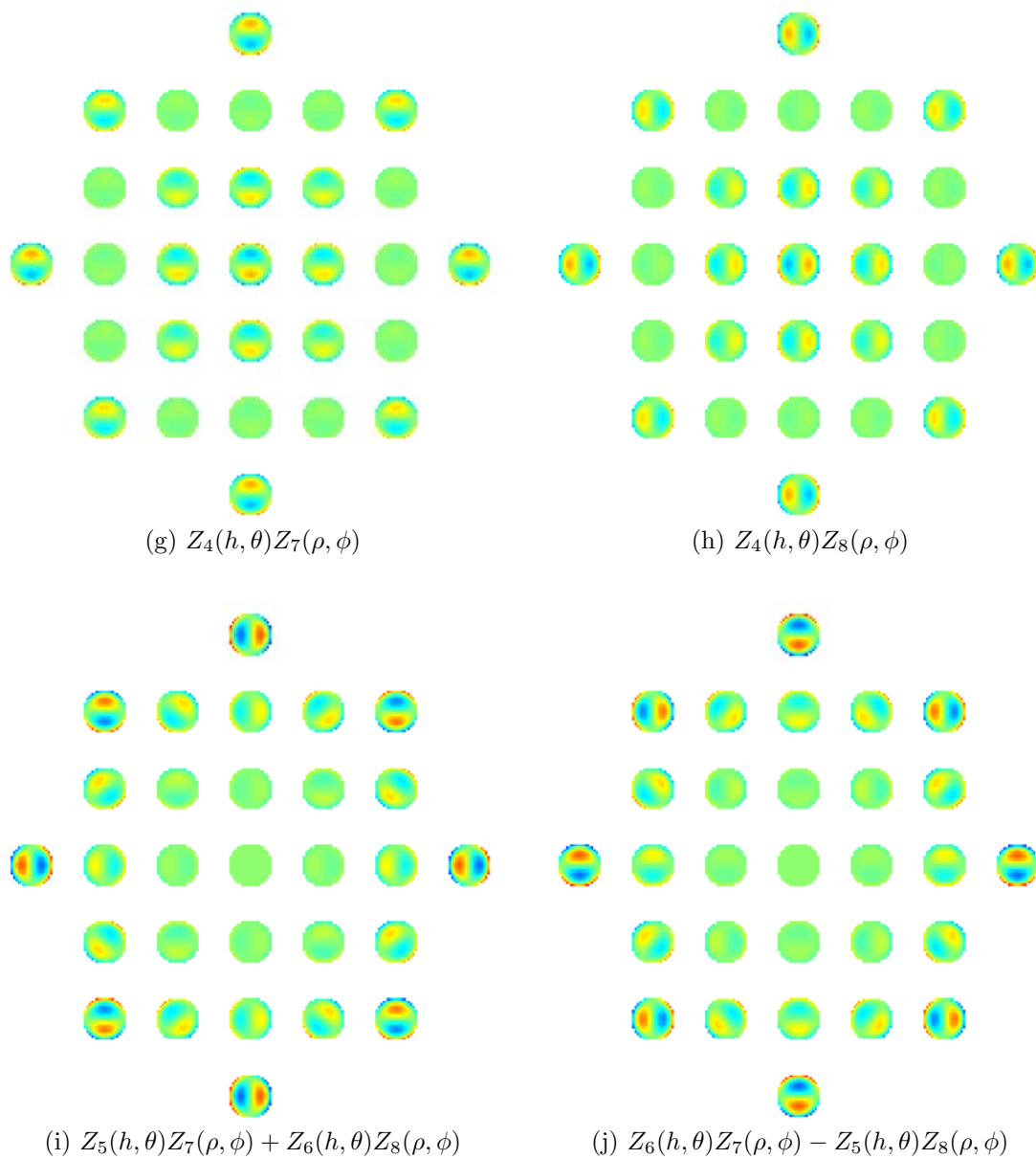


Figure K.3: Orthogonal coma polynomials continued.

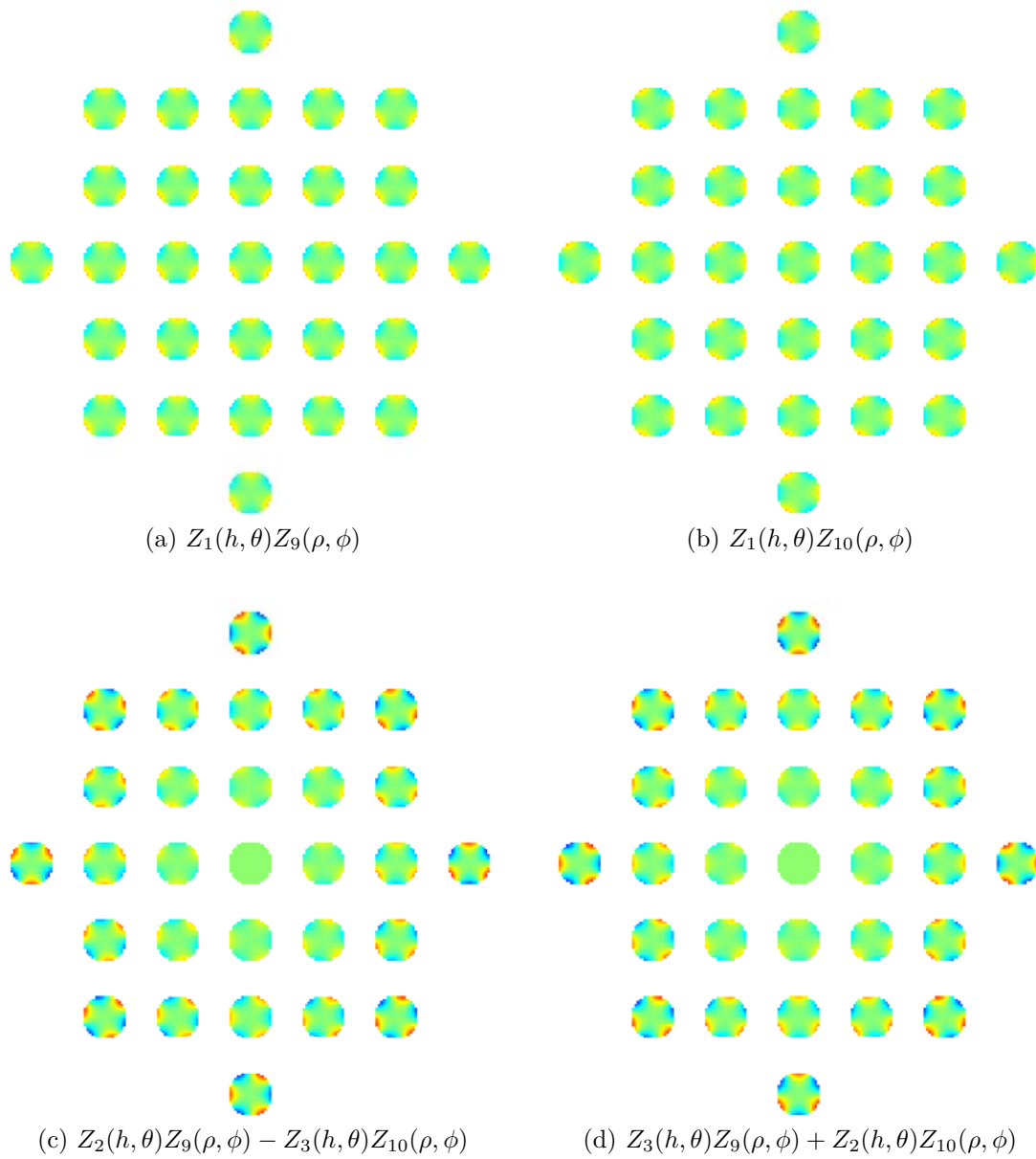


Figure K.4: Orthogonal trefoil polynomials.

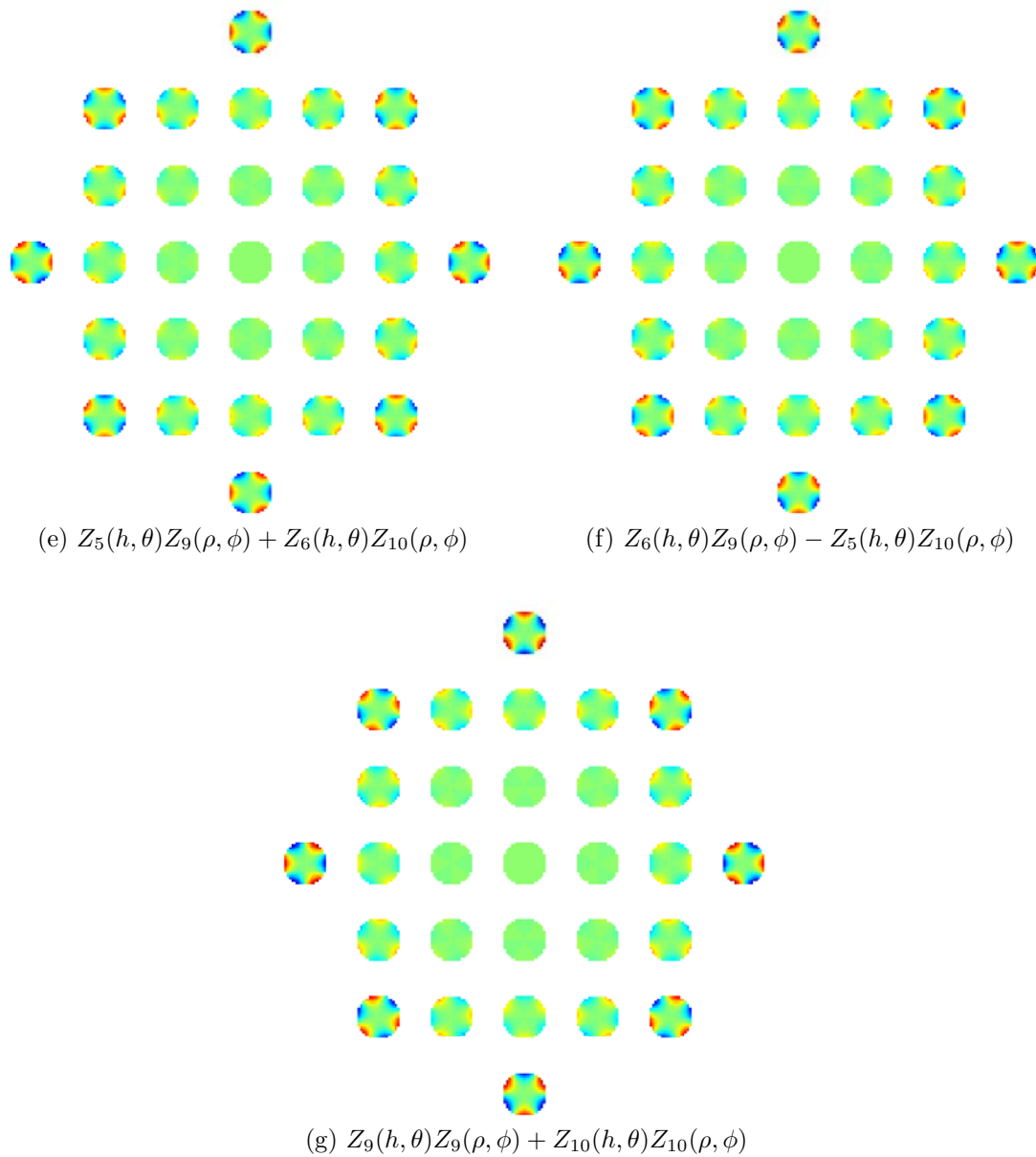


Figure K.4: Orthogonal trefoil polynomials continued.

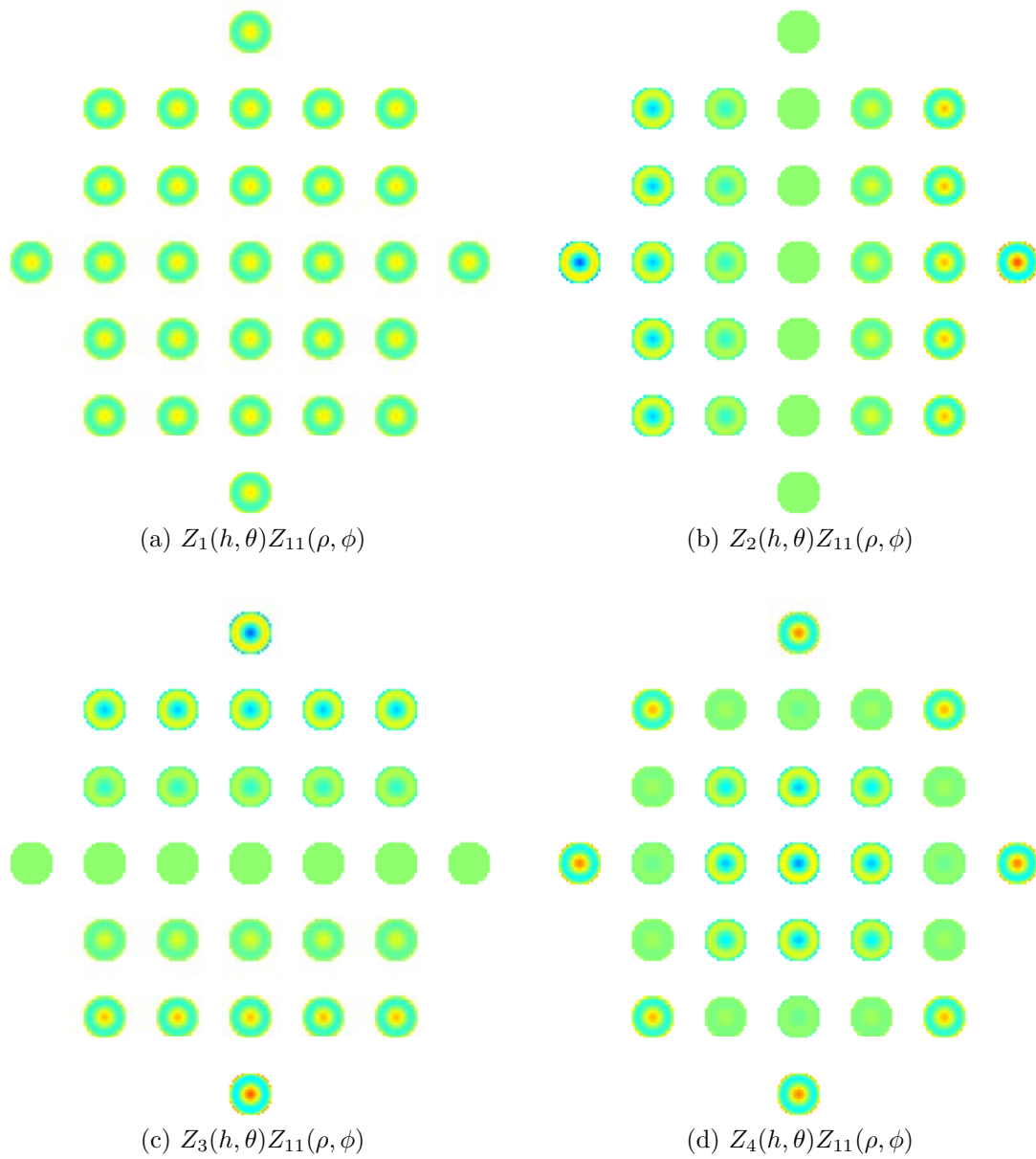


Figure K.5: Orthogonal spherical aberration polynomials.

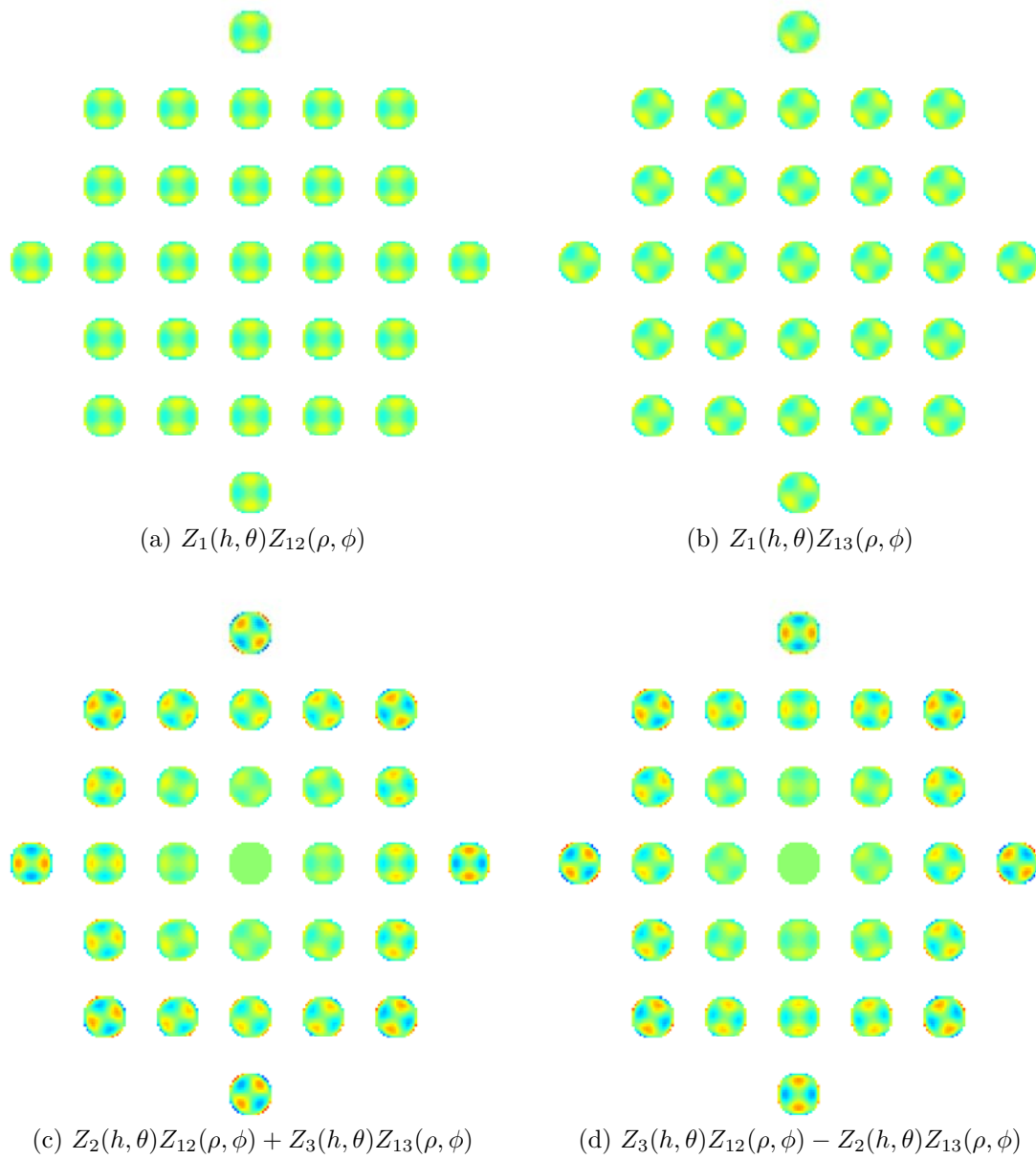
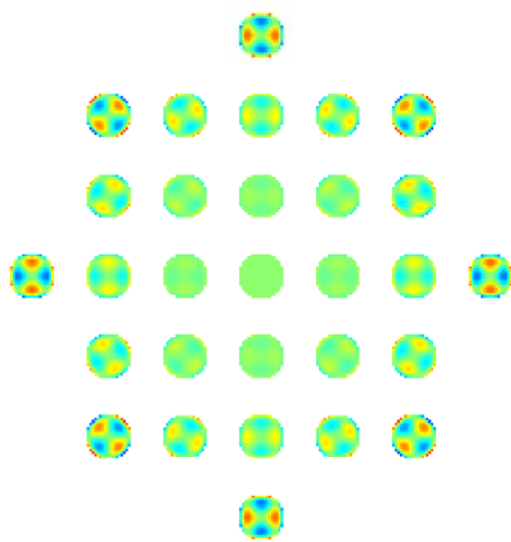


Figure K.6: Orthogonal oblique spherical aberration polynomials.



(e)  $Z_6(h, \theta)Z_{12}(\rho, \phi) + Z_5(h, \theta)Z_{13}(\rho, \phi)$

Figure K.6: Orthogonal oblique spherical aberration polynomials continued.

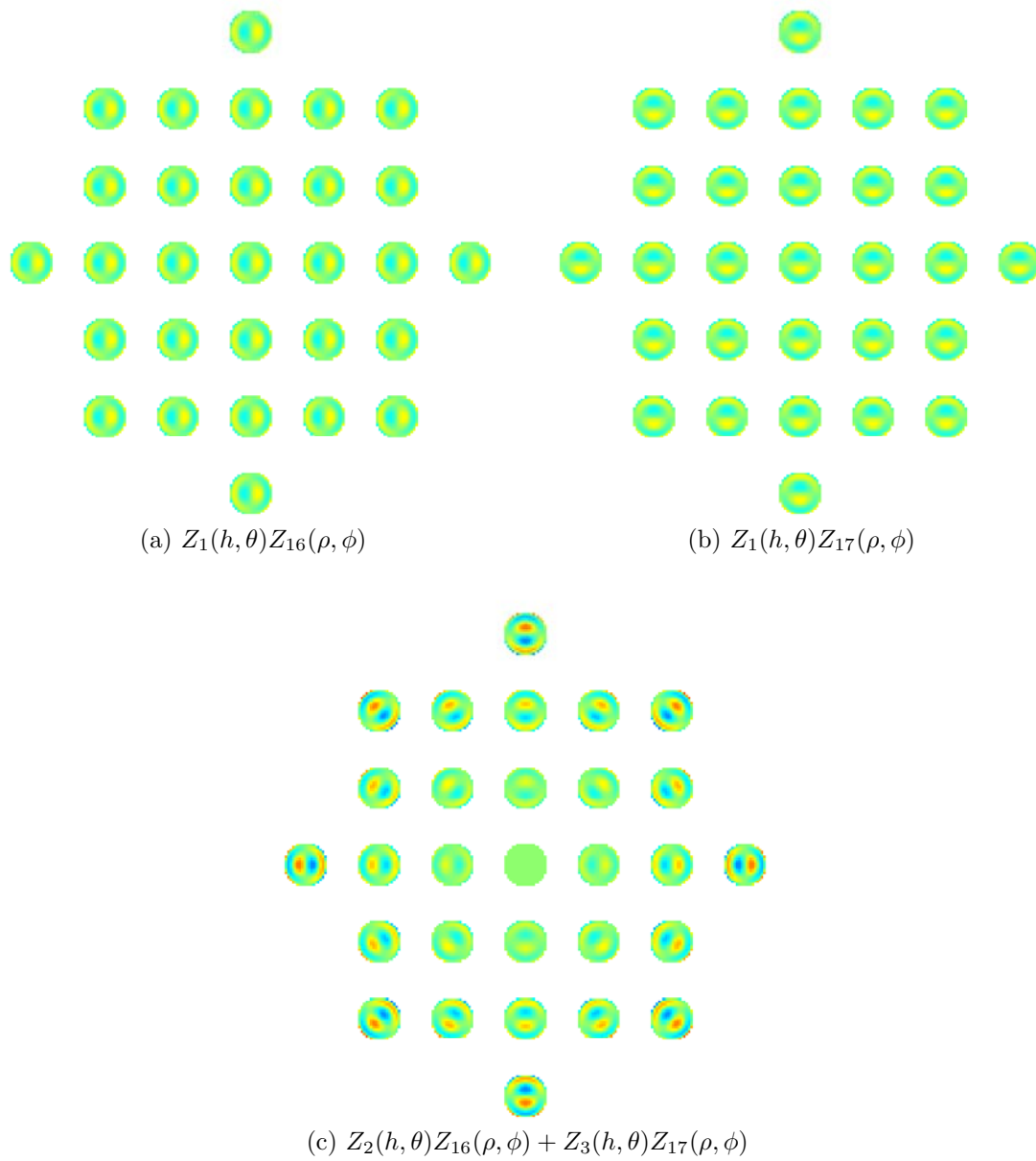


Figure K.7: Orthogonal secondary coma polynomials.

## APPENDIX L

## DOUBLE ZERNIKE EXPANSIONS

Double expansions of Zernike polynomials in both pupil space and field have been previously suggested in the literature (Kwee & Braat, 1993; Agurok, 1998, 2000) to describe the field dependence of aberrations in a misaligned system. The functions listed in these two papers are orthogonal (just as Zernike polynomials are) and are complete. However, neither of these papers suggested functions with two double Zernike terms to describe the aberrations in a misaligned system, as in Chapter 6. Also, the description of the field dependences of aberrations is not as complete as provided in Chapter 6.

Neither author seems to be aware of the other's publications, and they use different notations. Section L.1 discusses the notation and concepts introduced by Kwee & Braat (1993), while Section L.2 discusses the notation and concepts introduced by Agurok (1998, 2000).

## L.1 KWEE AND BRAAT'S DOUBLE EXPANSION

Kwee & Braat (1993) published the first paper proposing a double Zernike expansion of the optical aberration function. They expand the wavefront function into:

$$W(\rho, \phi; h, \theta) = \sum_{n,m} P_n^m R_n^m(\rho) [a_{n,m}(h, \theta) \cos(m\phi) + b_{n,m}(h, \theta) \sin(m\phi)] \quad (\text{L.1})$$



where

$$a_{n,m}(h, \theta) = \sum_{l,k} p_l^k R_l^k(h) [A_{nm} a_{lk} \cos(m\theta) + A_{nm} b_{lk} \sin(m\theta)] \quad (\text{L.2})$$

$$b_{n,m}(h, \theta) = \sum_{l,k} p_l^k R_l^k(h) [B_{nm} a_{lk} \cos(m\theta) + B_{nm} b_{lk} \sin(m\theta)]. \quad (\text{L.3})$$

and  $p_n^m$  is the normalization for the radial polynomial  $R_n^m(\rho)$  in the standard Zernike polynomial:

$$Z_n^m(\rho, \phi) = p_n^m R_n^m(\rho) \begin{cases} \cos(m\phi) \\ \sin(m\phi) \end{cases}. \quad (\text{L.4})$$

The four subscript coefficients like  $A_{nm} b_{lk}$  “should be read as a single quantity and not as a product of two individual coefficients.” For example,  $A_{nm} b_{lk}$  is the coefficient for the double Zernike function  $Z_l^k(h, \theta) Z_n^m(\rho, \phi)$ , although they do not write it out in such a way. They discuss some coefficients that are equal in a rotationally symmetric system.

For example,  $A_{22} a_{22} = B_{22} b_{22}$ . Since the terms with  $A$  or  $a$  correspond to the cosine Zernike terms, the  $A_{22}$  part refers to  $Z_2^2(\rho, \phi) = Z_6(\rho, \phi) = \rho^2 \cos 2\phi$  and the  $a_{22}$  part refers to  $Z_2^2(h, \theta) = Z_6(h, \theta) = h^2 \cos 2\theta$ . Since the terms with  $B$  or  $b$  correspond to the sine Zernike terms, the  $B_{22}$  part refers to  $Z_2^{-2}(\rho, \phi) = Z_5(\rho, \phi) = \rho^2 \sin 2\phi$  and the  $b_{22}$  part refers to  $Z_2^{-2}(h, \theta) = Z_5(h, \theta) = h^2 \sin 2\theta$ . Thus, these two coefficients correspond to the one function in Chapter 6 that describes the quadratic astigmatism that exists in a rotationally symmetric optical system:

$$Z_5(h, \theta) Z_5(\rho, \phi) + Z_6(h, \theta) Z_6(\rho, \phi).$$

As another example,  $A_{31}a_{11} = B_{31}b_{11}$ . Following the same technique in the previous paragraph, the  $A_{31}$  part refers to  $Z_3^1(\rho, \phi) = Z_8(\rho, \phi)$ , the  $a_{11}$  part refers to  $Z_1^1(h, \theta) = Z_2(h, \theta)$ , the  $B_{31}$  part refers to  $Z_3^{-1}(\rho, \phi) = Z_7(\rho, \phi)$  and the  $b_{11}$  part refers to  $Z_1^{-1}(h, \theta) = Z_3(h, \theta)$ . Thus, these two coefficients correspond to the one function in Chapter 6 that describes the linear coma that exists in a rotationally symmetric optical system:  $Z_3(h, \theta)Z_7(\rho, \phi) + Z_2(h, \theta)Z_8(\rho, \phi)$ .

A table is provided listing all the coefficients and some of these coefficients are marked with symbols to show that they are equal. The table includes columns to show which terms are possible in systems with different types of symmetry ( $y = 0$ ,  $x = 0$ ,  $y = 0$  and  $x = 0$ , and rotational symmetry). Unfortunately, the rotational symmetry column is missing from the second page of the table.

Also, there is no reference to any additional terms that are predicted to be equal in this dissertation. For example, linear astigmatism  $Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi)$  would be described by the coefficients  $B_{22}b_{11}$  and  $A_{22}a_{11}$ , but these coefficients are only checked as existing and there is no mention that they are equal. This is probably because those terms ( $Z_3(h, \theta)Z_5(\rho, \phi)$  and  $Z_2(h, \theta)Z_6(\rho, \phi)$ ) are only equal to first order and not precisely equal. Small contributions from the  $\vec{C}_{422}$  perturbation vector using Thompson's notation cause  $Z_3(h, \theta)Z_5(\rho, \phi)$  to not be exactly equal to  $Z_2(h, \theta)Z_6(\rho, \phi)$ , which is why the additional functions, like  $Z_3(h, \theta)Z_5(\rho, \phi) - Z_2(h, \theta)Z_6(\rho, \phi)$  are required to describe the higher order dependence on misalignment in Chapter 6.

For the pupil aberrations of distortion through spherical aberration, the table includes field dependences that are derived from the third-order Seidel aberrations.

For example, there are terms for the Zernike aberration describing the power in the pupil  $Z_4(\rho, \phi)$  up through astigmatism in field, but the coma terms in field  $Z_7(h, \theta)Z_4(\rho, \phi)$  and  $Z_8(h, \theta)Z_4(\rho, \phi)$  that come from the  $\vec{A}_{420}$  perturbation vector are not listed. Neither listed are the terms for cubic field dependence for astigmatism in the pupil (from  $\vec{A}_{420}$ ) or linear spherical aberration  $Z_2(h, \theta)Z_{11}(\rho, \phi)$  from the oblique spherical term  $\vec{A}_{240_M}$ , among others.

In addition, the table only describes contributions to the terms that scale directly with the misalignment. Terms that are predicted to have quadratic dependence on the misalignment, like  $Z_5(h, \theta)Z_4(\rho, \phi)$  (corresponding to the coefficient  $A_{20}b_{22}$  using their notation) which depends on Thompson's vector  $\vec{B}_{420_M}$  are not checked as possible.

The paper then continues on with some vector representations of distortion and astigmatic focal lines for astigmatism, but the images are reproduced poorly. Finally, the paper ends with some image quality considerations, such as calculating the wavefront variance and field-averaged blur from the coefficients.

## L.2 AGUROK'S DOUBLE EXPANSION

Agurok (1998, 2000) published two papers also on double Zernike expansions in the pupil and field of view. This paper has a wavefront expansion of the form:

$$W(h, \theta, \rho, \phi) = \sum_{m_1=0}^N \sum_{m_2=\pm 1} \sum_{n=m_1}^N \sum_{l=m_2}^N D_{nm_1}^{lm_2} P_n^{m_1}(\rho) P_l^{m_2}(h) \cos(m_1\phi - m_2\theta) \quad (\text{L.5})$$

where  $D_{nm_1}^{lm_2}$  is the four subscript coefficient of the double expansion. Some terms

that are possible are listed, but the list is not very complete, and the explanation is not thorough. Equation L.5 is valid for systems with a misalignment in one direction. If misalignments are allowed in any directions, Agurok uses the variables  $H$ ,  $V$ ,  $VC$ ,  $HC$ ,  $VS$ , and  $HS$  to describe the coefficients instead of  $D$ , as shown in Table L.1 (which skips the distortion terms). This table shows the corresponding orthogonal functions from Chapter 6 to these coefficients. All of the terms they listed do have linear dependence on the misalignment, as shown by the  $\vec{A}$  vectors in Tables 6.1–6.7. Higher order terms not included by Agurok are included at the end of the table. Agurok (1998) correctly mentions after the table that constant astigmatism is observed when the perturbation is significant and this dependence has quadratic dependence on the misalignment. No other aberrations with this dependence on the misalignment are mentioned.

Next Agurok shows an example of how to calculate the astigmatism coefficients from measurement of five points in the field of view. This is possible because the astigmatism there are only constant, linear and quadratic field dependences. At first this makes sense. To measure a quadratic 2-D function, three points are needed. Since there are two directions ( $x$  and  $y$ ), it seems like three points on each axis would be needed. Because the point  $(0, 0)$  is on each axis, then only five points are needed instead of six. However, this is not true in general. The five possible field dependences (e.g.  $\alpha_0$ – $\alpha_4$  for one quadratic astigmatism, two linear astigmatism contribution and two constant astigmatism contributions) can be fit to the data from three field points because each field point actually gives two components of the astigmatism when Zernike polynomials are measured ( $Z_5$  and  $Z_6$ ). In the second

Table L.1: Aberrations predicted by Agurok in a perturbed optical system.

Agurok coefficients	Corresponding orthogonal functions	Source of aberration
$V_{20}^{11}, H_{20}^{11}$	$Z_2(h, \theta)Z_4(\rho, \phi), Z_3(h, \theta)Z_4(\rho, \phi)$	$\vec{A}_{220_M}, \vec{A}_{420_M}, \vec{A}_{240_M}$
$V_{40}^{11}, H_{40}^{11}$	$Z_2(h, \theta)Z_{11}(\rho, \phi), Z_3(h, \theta)Z_{11}(\rho, \phi)$	$\vec{A}_{240_M}$
$V_{31}^{00}, H_{31}^{00}$	$Z_1(h, \theta)Z_7(\rho, \phi), Z_1(h, \theta)Z_8(\rho, \phi)$	$\vec{A}_{131}, \vec{A}_{331_M}, \vec{A}_{151}$
$VC_{22}^{11}, HC_{22}^{11},$ $VS_{22}^{11}, HS_{22}^{11}$	$Z_2(h, \theta)Z_5(\rho, \phi) + Z_3(h, \theta)Z_6(\rho, \phi),$ $Z_2(h, \theta)Z_5(\rho, \phi) - Z_3(h, \theta)Z_6(\rho, \phi),$ $Z_3(h, \theta)Z_5(\rho, \phi) + Z_2(h, \theta)Z_6(\rho, \phi),$ $Z_3(h, \theta)Z_5(\rho, \phi) - Z_2(h, \theta)Z_6(\rho, \phi)$	$\vec{A}_{222}, \vec{A}_{422}, \vec{A}_{242}$
$VC_{33}^{22}, HC_{33}^{22},$ $VS_{33}^{22}, HS_{33}^{22}$	$Z_5(h, \theta)Z_9(\rho, \phi) + Z_6(h, \theta)Z_{10}(\rho, \phi),$ $Z_5(h, \theta)Z_9(\rho, \phi) - Z_6(h, \theta)Z_{10}(\rho, \phi),$ $Z_6(h, \theta)Z_9(\rho, \phi) + Z_5(h, \theta)Z_{10}(\rho, \phi),$ $Z_6(h, \theta)Z_9(\rho, \phi) - Z_5(h, \theta)Z_{10}(\rho, \phi)$	$\vec{A}_{333}$
	$Z_7(h, \theta)Z_4(\rho, \phi), Z_8(h, \theta)Z_4(\rho, \phi)$	$\vec{A}_{420_M}$
	$Z_7(h, \theta)Z_5(\rho, \phi) + Z_8(h, \theta)Z_6(\rho, \phi),$ $Z_7(h, \theta)Z_5(\rho, \phi) - Z_8(h, \theta)Z_6(\rho, \phi),$ $Z_7(h, \theta)Z_5(\rho, \phi) + Z_8(h, \theta)Z_6(\rho, \phi),$ $Z_7(h, \theta)Z_5(\rho, \phi) - Z_8(h, \theta)Z_6(\rho, \phi)$	$\vec{A}_{420_M}$
	$Z_4(h, \theta)Z_7(\rho, \phi), Z_4(h, \theta)Z_8(\rho, \phi)$	$\vec{A}_{331_M}$
	$Z_5(h, \theta)Z_7(\rho, \phi) + Z_6(h, \theta)Z_8(\rho, \phi),$ $Z_5(h, \theta)Z_7(\rho, \phi) - Z_6(h, \theta)Z_8(\rho, \phi),$ $Z_6(h, \theta)Z_7(\rho, \phi) + Z_5(h, \theta)Z_8(\rho, \phi),$ $Z_6(h, \theta)Z_7(\rho, \phi) - Z_5(h, \theta)Z_8(\rho, \phi)$	$\vec{A}_{331_M}$
	$Z_2(h, \theta)Z_{12}(\rho, \phi) + Z_3(h, \theta)Z_{13}(\rho, \phi),$ $Z_3(h, \theta)Z_{12}(\rho, \phi) - Z_2(h, \theta)Z_{13}(\rho, \phi)$	$\vec{A}_{242}$
	$Z_1(h, \theta)Z_{16}(\rho, \phi), Z_1(h, \theta)Z_{17}(\rho, \phi)$	$\vec{A}_{151}$

paper (Agurok, 2000), nine field points are used to calculate aberrations that have a cubic dependence on the field.

## APPENDIX M

## ORTHOGONALITY OF DOUBLE ZERNIKE FUNCTIONS

This appendix shows that the two following functions are orthogonal:

$$S_A(h, \theta, \rho, \phi) = Z_k(h, \theta) Z_i(\rho, \phi) + Z_l(h, \theta) Z_j(\rho, \phi) \quad (\text{M.1})$$

$$S_B(h, \theta, \rho, \phi) = Z_k(h, \theta) Z_i(\rho, \phi) - Z_l(h, \theta) Z_j(\rho, \phi). \quad (\text{M.2})$$

For simplicity, the explicit dependence of  $Z_k$  and  $Z_l$  on  $(h, \theta)$  and of  $Z_i$  and  $Z_j$  on  $(\rho, \phi)$  will not be shown.

$$\int_0^{2\pi} \int_0^1 \int_0^{2\pi} \int_0^1 S_A(h, \theta, \rho, \phi) S_B(h, \theta, \rho, \phi) \rho d\rho d\phi h dh d\theta \quad (\text{M.3})$$

$$= \int_0^{2\pi} \int_0^1 \int_0^{2\pi} \int_0^1 (Z_k Z_i + Z_l Z_j) (Z_k Z_i - Z_l Z_j) \rho d\rho d\phi h dh d\theta \quad (\text{M.4})$$

$$= \int_0^{2\pi} \int_0^1 \int_0^{2\pi} \int_0^1 (Z_k^2 Z_i^2 - Z_k Z_l Z_i Z_j + Z_l Z_j Z_k Z_i - Z_l^2 Z_j^2) \rho d\rho d\phi h dh d\theta \quad (\text{M.5})$$

$$= \int_0^{2\pi} \int_0^1 Z_k^2 h dh d\theta \cdot \int_0^{2\pi} \int_0^1 Z_i^2 \rho d\rho d\phi - \int_0^{2\pi} \int_0^1 Z_l^2 h dh d\theta \cdot \int_0^{2\pi} \int_0^1 Z_j^2 \rho d\rho d\phi \quad (\text{M.6})$$

$$= \delta_{kk} \delta_{ii} - \delta_{ll} \delta_{jj} \quad (\text{M.7})$$

$$= 1 \cdot 1 - 1 \cdot 1 \quad (\text{M.8})$$

$$= 0 \quad (\text{M.9})$$

Since the orthogonality integral is equal to zero, the functions are orthogonal.



## APPENDIX N

## SHAPE ERRORS

Table N.1 lists the first 11 Zernike polynomials in vector and scalar form. This is provided for convenience since this the report where Tessieres (2004) lists the first 37 terms is currently unpublished.

Table N.1: Zernike standard polynomials in a scalar and vectorial form. The odd number after the name of each aberration (3 or 5) refers to the order of the aberration.

Zernike term	$Z(r, \theta)$	$Z(\vec{r}, \hat{i}, \hat{j})$
$Z_1$ (Piston)	1	1
$Z_2$ (Tilt $x$ )	$\sqrt{4}r \cos \theta$	$\sqrt{4}\vec{r} \cdot \hat{i}$
$Z_3$ (Tilt $y$ )	$\sqrt{4}r \sin \theta$	$\sqrt{4}\vec{r} \cdot \hat{j}$
$Z_4$ (Defocus)	$\sqrt{3}(2r^2 - 1)$	$\sqrt{3}(2\vec{r} \cdot \vec{r} - 1)$
$Z_5$ (Astigmatism 3 at $45^\circ$ )	$\sqrt{6}(r^2 \sin 2\theta)$	$2\sqrt{6}[(\vec{r} \cdot \hat{i})(\vec{r} \cdot \hat{j})]$
$Z_6$ (Astigmatism 3 at $0^\circ$ )	$\sqrt{6}(r^2 \cos 2\theta)$	$\sqrt{6}[2(\vec{r} \cdot \hat{i})^2 - \vec{r} \cdot \vec{r}]$
$Z_7$ (Coma 3 at $90^\circ$ )	$\sqrt{8}(3r^3 - 2r) \sin \theta$	$\sqrt{8}(\vec{r} \cdot \hat{j})[3(\vec{r} \cdot \vec{r}) - 2]$
$Z_8$ (Coma 3 at $0^\circ$ )	$\sqrt{8}(3r^3 - 2r) \cos \theta$	$\sqrt{8}(\vec{r} \cdot \hat{i})[3(\vec{r} \cdot \vec{r}) - 2]$
$Z_9$ (Trefoil 5 at $30^\circ$ )	$\sqrt{8}r^3 \sin 3\theta$	$\sqrt{8}[3(\vec{r} \cdot \vec{r})(\vec{r} \cdot \hat{j}) - 4(\vec{r} \cdot \hat{j})^3]$
$Z_{10}$ (Trefoil 5 at $0^\circ$ )	$\sqrt{8}r^3 \cos 3\theta$	$\sqrt{8}[4(\vec{r} \cdot \hat{i})^3 - 3(\vec{r} \cdot \vec{r})(\vec{r} \cdot \hat{i})]$
$Z_{11}$ (Spherical aberration)	$\sqrt{5}(6r^4 - 6r^2 + 1)$	$\sqrt{5}[6(\vec{r} \cdot \vec{r})^2 - 6(\vec{r} \cdot \vec{r}) + 1]$

Table N.2: Table of the expressions to describe the field dependence of aberrations induced by the first 10 Zernike polynomial bending modes.

	Shape error on mirror						
	$Z_4$	$Z_5$	$Z_6$	$Z_7$	$Z_8$	$Z_9$	$Z_{10}$
$Z_4(\rho, \phi)$	$a^2$	0	0	$2\sqrt{6}cH_y a^2$	$2\sqrt{6}a^2 cH_x$	0	0
$Z_5(\rho, \phi)$		$a^2$	0	$2\sqrt{3}a^2 cH_x$	$2\sqrt{3}a^2 cH_y$	$2\sqrt{3}a^2 cH_x$	$-2\sqrt{3}a^2 cH_y$
$Z_6(\rho, \phi)$			$a^2$	$-2\sqrt{3}a^2 cH_y$	$2\sqrt{3}a^2 cH_x$	$2\sqrt{3}a^2 cH_y$	$2\sqrt{3}a^2 cH_x$
$Z_7(\rho, \phi)$				$a^3$	0	0	0
$Z_8(\rho, \phi)$					$a^3$	0	0
$Z_9(\rho, \phi)$						$a^3$	0
$Z_{10}(\rho, \phi)$							$a^3$

Table N.3: Expression of the third order spherical aberration introduced as a bending mode.

	$Z_{11}$ Shape error
$Z_4(\rho, \phi)$	$\sqrt{3}\sqrt{5}((4H_y^2 + 4H_x^2)c^2 + a^2 - 1)a^2$
$Z_5(\rho, \phi)$	$4\sqrt{6}\sqrt{5}a^2 c^2 H_x H_y$
$Z_6(\rho, \phi)$	$2\sqrt{6}\sqrt{5}a^2 c^2 (H_x^2 - H_y^2)$
$Z_7(\rho, \phi)$	$2\sqrt{2}\sqrt{5}a^3 cH_y$
$Z_8(\rho, \phi)$	$2\sqrt{2}\sqrt{5}a^3 cH_x$
$Z_9(\rho, \phi)$	0
$Z_{10}(\rho, \phi)$	0
$Z_{11}(\rho, \phi)$	$a^4$

## APPENDIX O

## HET CONTROL MODES

This appendix shows all the modes resulting from the SVD of the HET WFC influence matrix and includes a sample calculation for mode 1 to show how the modes are quantified using the singular values. The singular values are listed in Table O.1. The singular values scale the value of the double Zernike coefficients given in the  $\mathbf{U}$  matrix for the combination of misalignments in the  $\mathbf{V}$  matrix where 1 unit of misalignment is given by the values in Table 8.2.

For example, in mode 1, assume that one unit perturbation is applied. Therefore the WFC  $x$  tilt and  $y$  tilt are both  $0.7 \times 0.01^\circ = 0.07^\circ$ . (The 0.7 is from Figure O.1b, while the  $0.01^\circ$  is from Table 8.2.) The field-dependent aberration expected is  $13.5 \mu\text{m} \times (-Z_1(h, \theta)Z_7(\rho, \phi) + Z_1(h, \theta)Z_8(\rho, \phi))$  (where the 13.5 comes from the value of the coefficient in Figure O.1c times the first singular value:  $13.5 = 0.7 \times 19.3$ ). A few sample points to verify this, modeling the tilts in ZEMAX, are listed in Table O.2. The numbers are not exactly equivalent because mode 1 actually includes a small amount of quadratic coma ( $Z_4(h, \theta)Z_7(\rho, \phi)$  and  $Z_4(h, \theta)Z_8(\rho, \phi)$ ). Also, only the two degrees of freedom with the largest values, which account for most of the mode, are perturbed. This calculation shows the technique that can be used to quantify the aberrations in the other modes.

The wavefront maps across the field (subfigure (a) for each mode) are one representation of the columns from the  $\mathbf{U}$  matrix, while the stem plots of the orthogonal

Table O.1: HET SVD singular values.

Mode	Singular values (waves rms/unit perturbation)
1	19.2736160
2	19.2724396
3	2.7517446
4	0.2912215
5	0.2912215
6	0.0216100
7	0.0216096
8	0.0089162
9	0.0008064
10	0.0005542
11	0.0004976
12	0.0004026
13	0.0003507
14	0.0001061
15	0.0000774
16	0.0000679
17	0.0000588
18	0.0000519
19	0.0000373
20	0.0000261
21	0.0000204
22	0.0000169
23	0.0000109
24	0.0000098
25	0.0000064

Table O.2: HET ZEMAX verification of SVD mode 1. If mode 1 is approximated only as constant coma (even though there are small amounts of higher order coma), then the expected values are  $Z_7 = -13.5$  waves and  $Z_8 = 13.5$  waves for any field angle. The fields are normalized such that  $H_x = H_y = 11$  arcmin and the Zernike coefficients are in waves ( $\lambda = 1 \mu\text{m}$ ).

$H_x$	$H_y$	$Z_7$	$Z_8$
0	0	-13.5	13.5
-1	0	-13.2	13.0
1	0	-13.2	12.8
0	-1	-12.8	13.2
0	1	-13.0	13.2

double Zernike coefficients are another representation of the same data (subfigure (c) for each mode). For space considerations, the field and pupil dependence was not printed on the graph labels. Here, the functions are written in the same order as in the rest of the dissertation, with the field-dependent Zernike first and the pupil-dependent Zernike polynomial second. For example,  $Z_1 Z_4 = Z_1(h, \theta)Z_4(\rho, \phi)$  for constant defocus and  $Z_2 Z_5 - Z_3 Z_6 = Z_2(h, \theta)Z_5(\rho, \phi) - Z_3(h, \theta)Z_6(\rho, \phi)$  for linear astigmatism. The configuration of the system to produce the mode is given by columns in the  $V$  matrix. These are plotted in subfigure (b) for each mode.

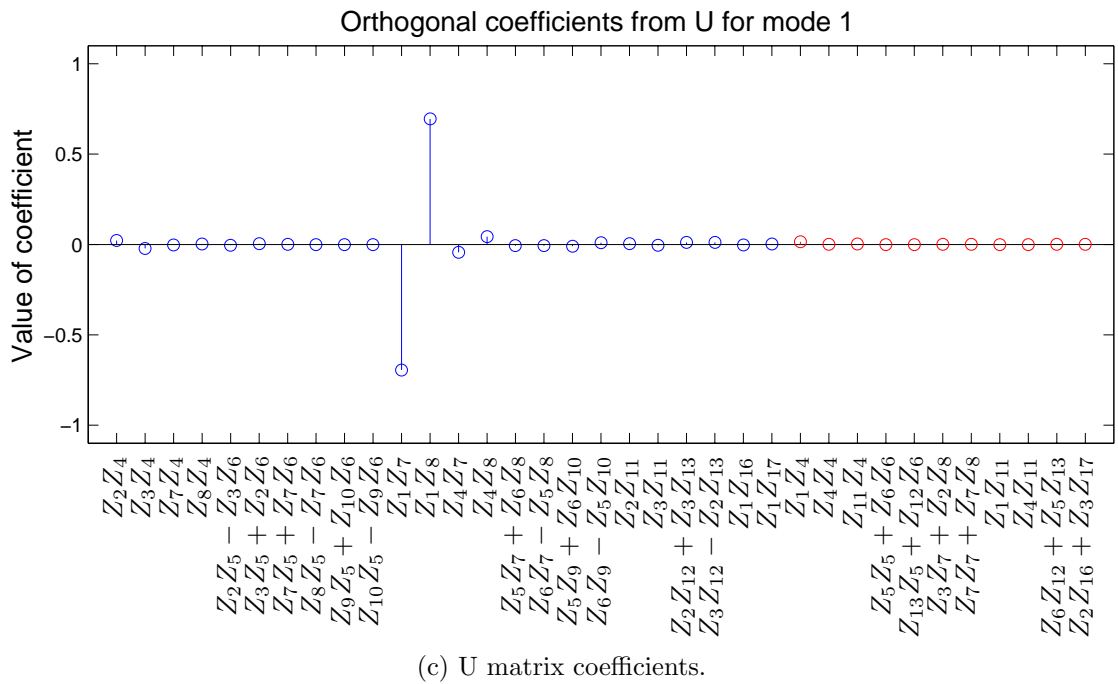
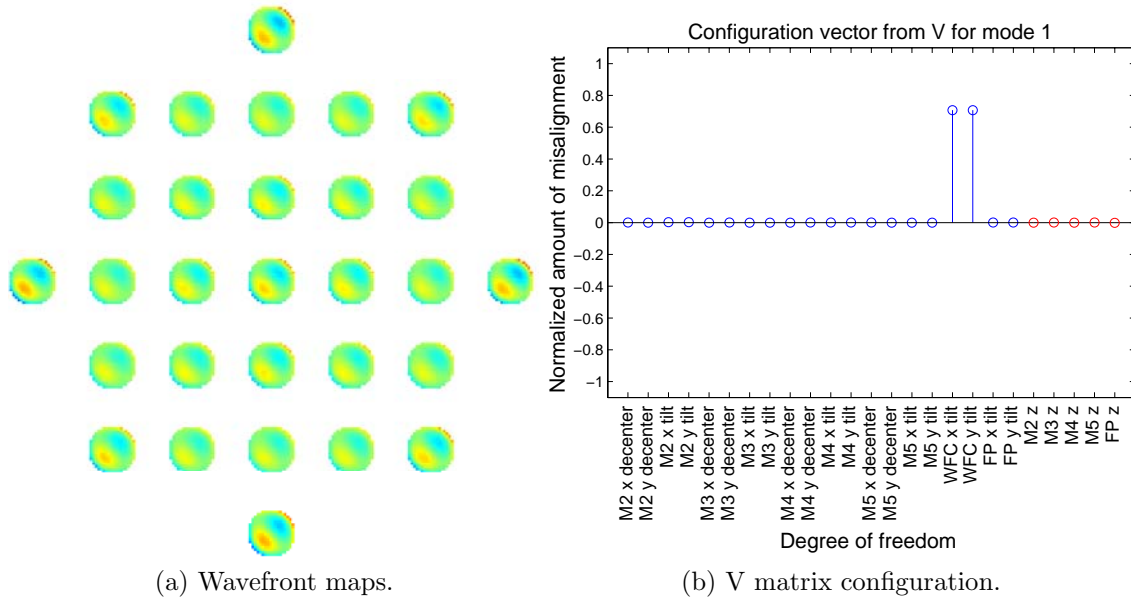


Figure O.1: HET SVD mode 1: constant coma.

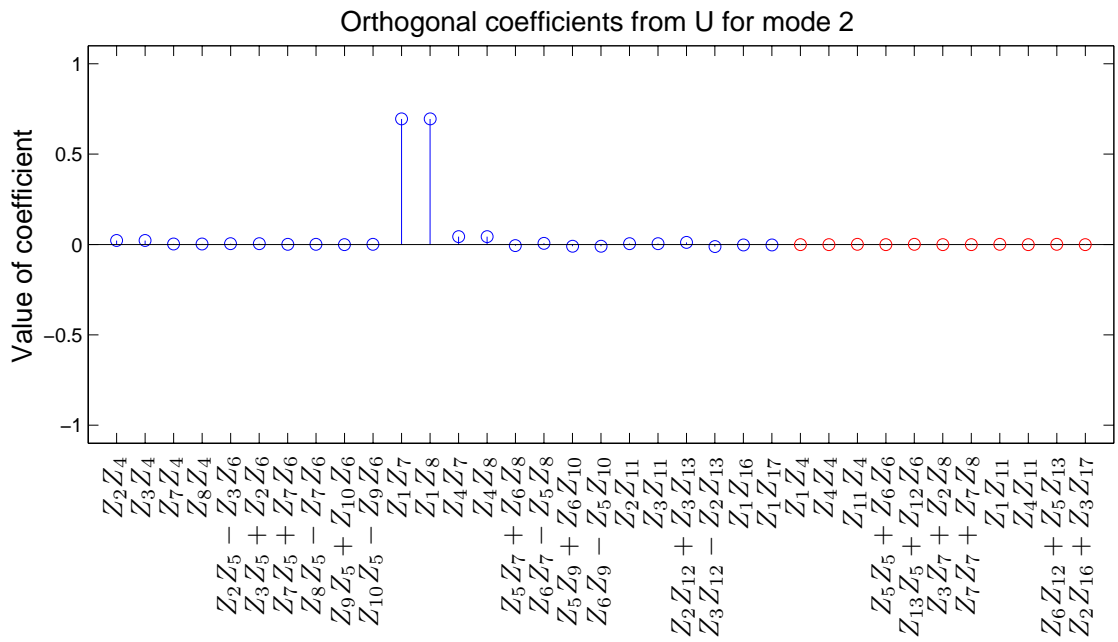
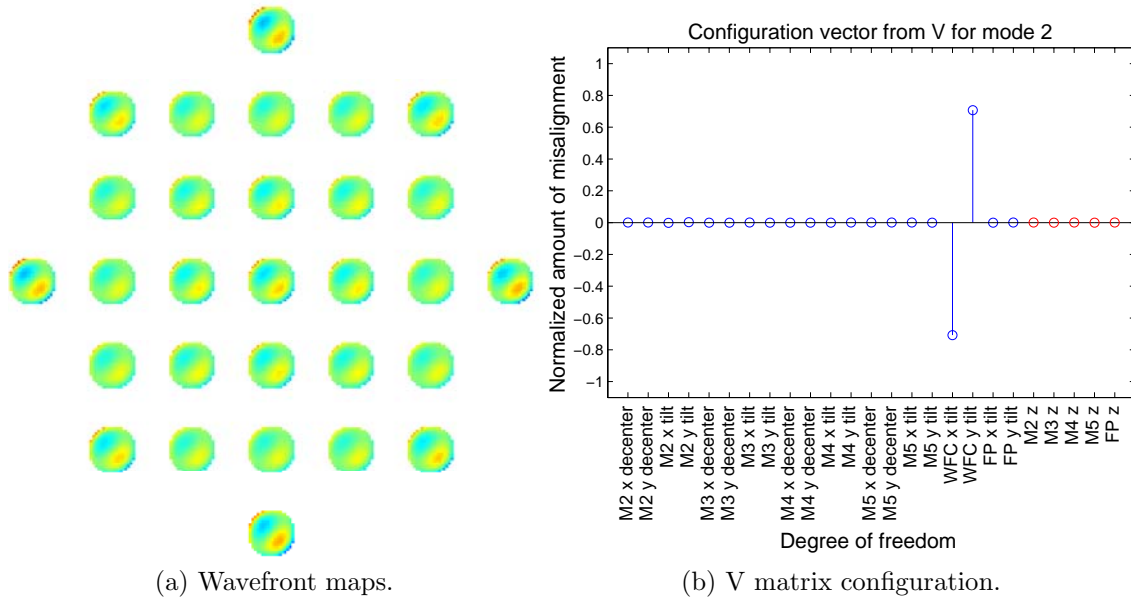
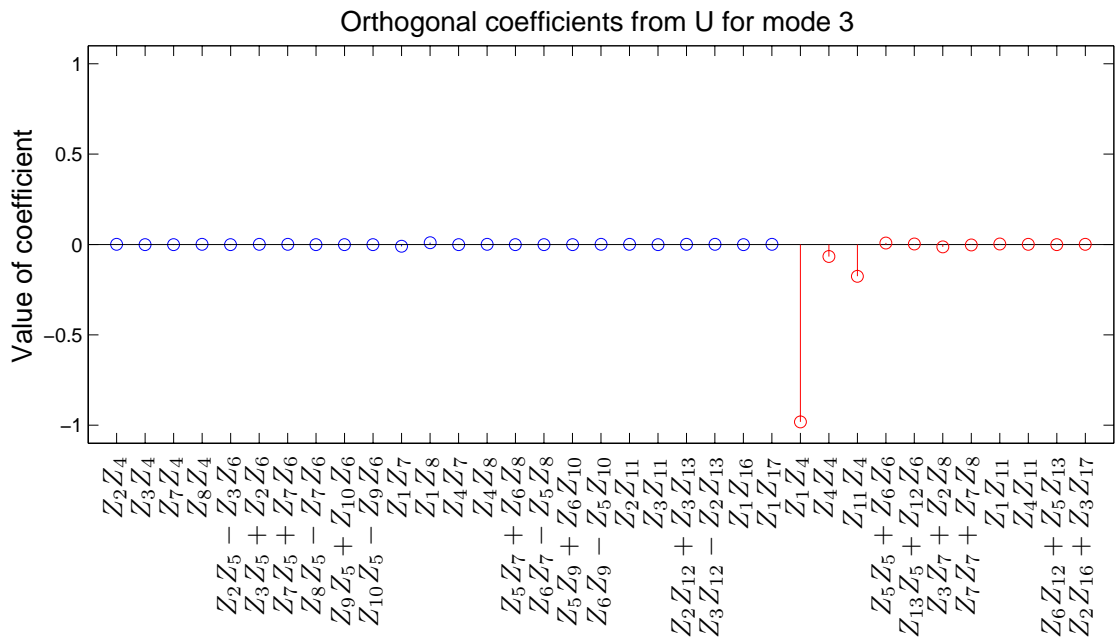
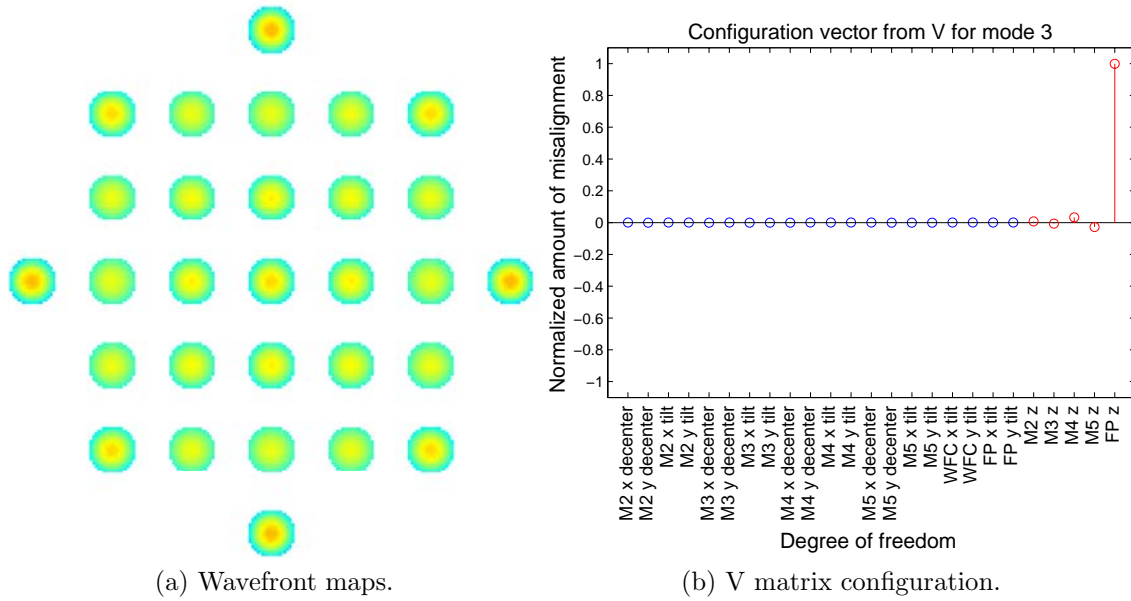


Figure O.2: HET SVD mode 2: constant coma.



(c) U matrix coefficients.

Figure O.3: HET SVD mode 3: defocus.



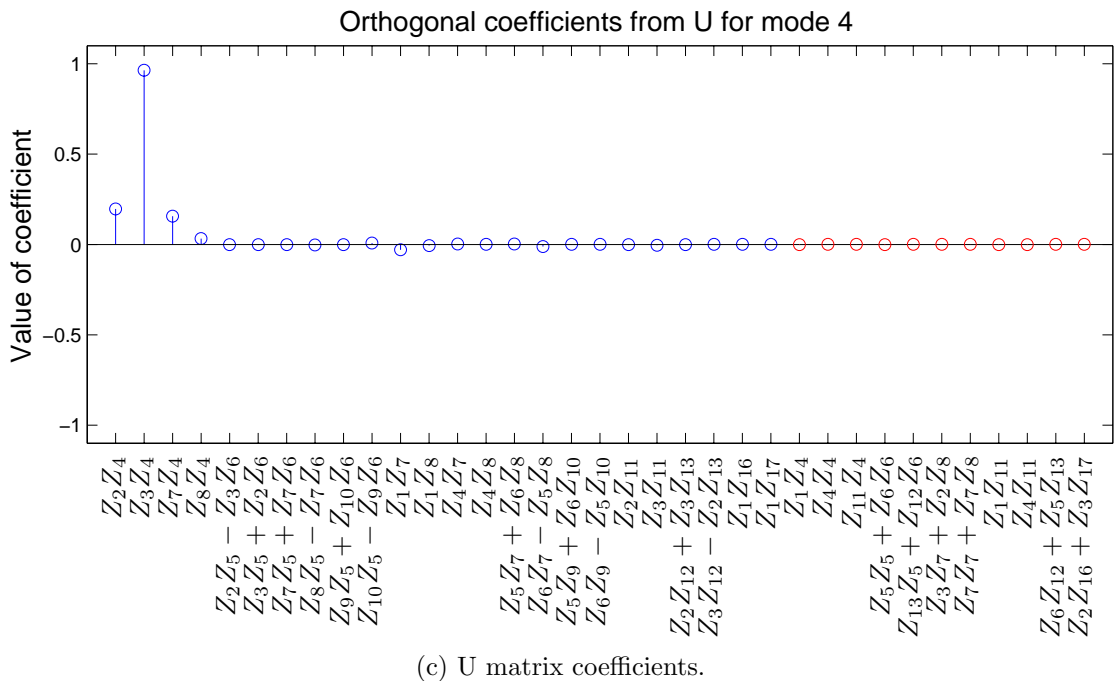
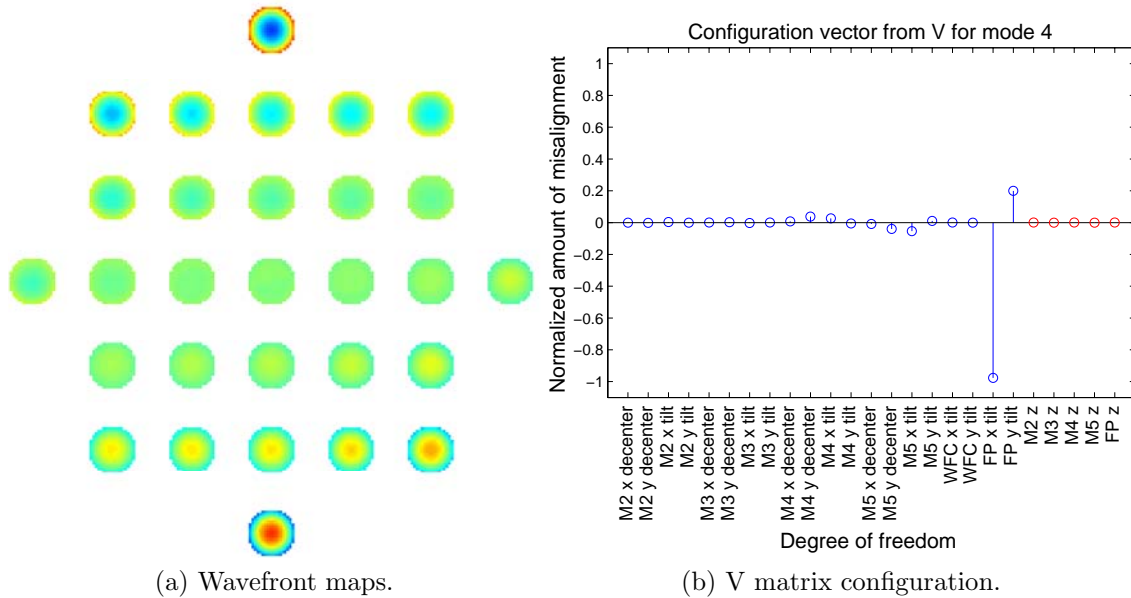
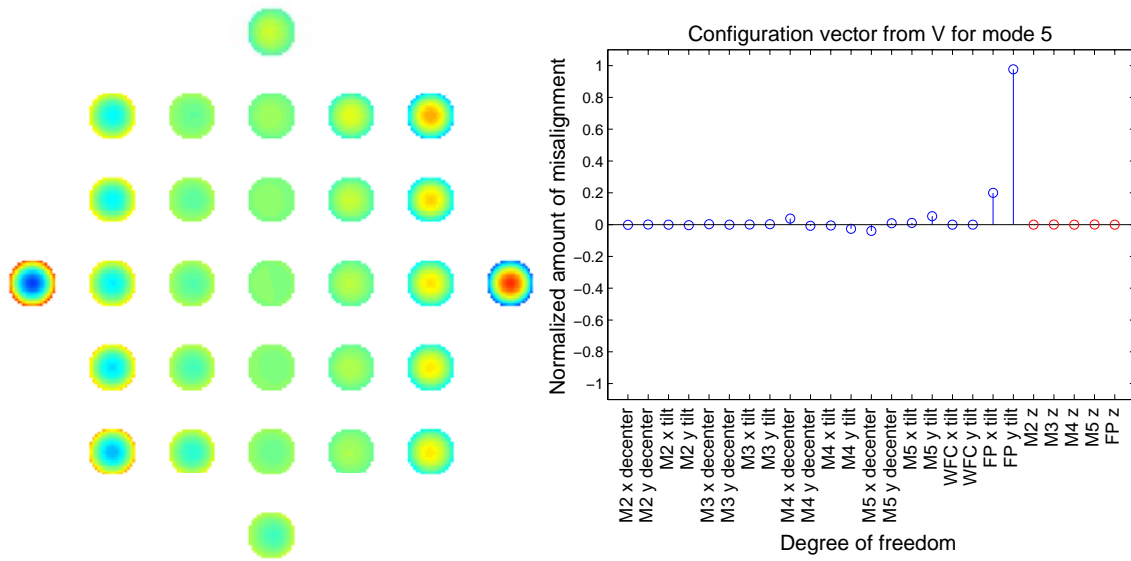
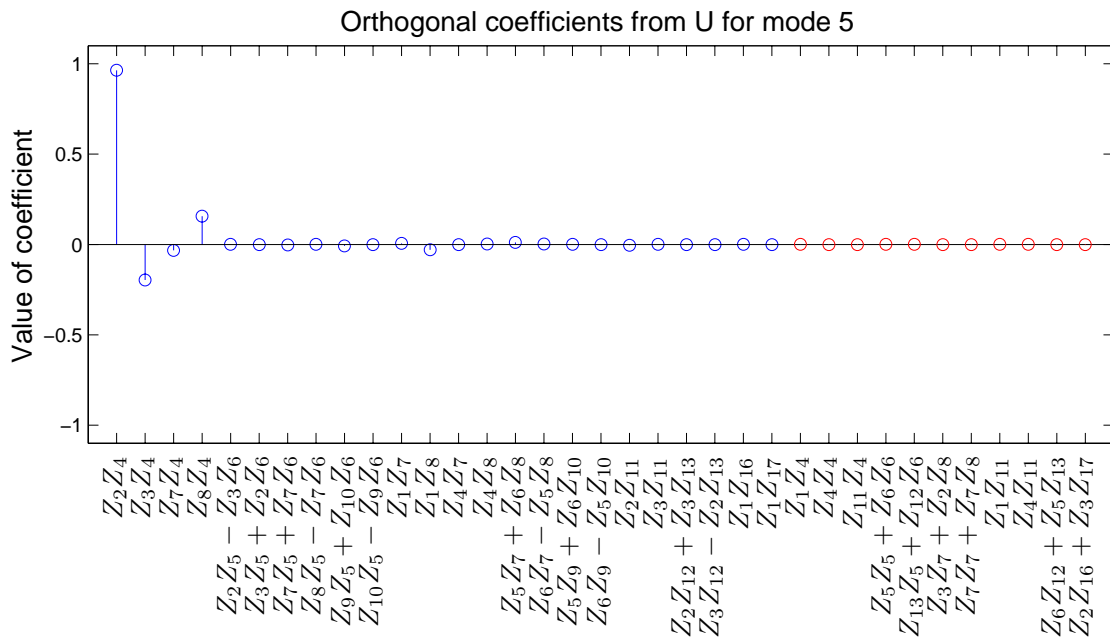


Figure O.4: HET SVD mode 4: focal plane tilt (linear defocus).



(a) Wavefront maps.

(b) V matrix configuration.



(c) U matrix coefficients.

Figure O.5: HET SVD mode 5: focal plane tilt (linear defocus).

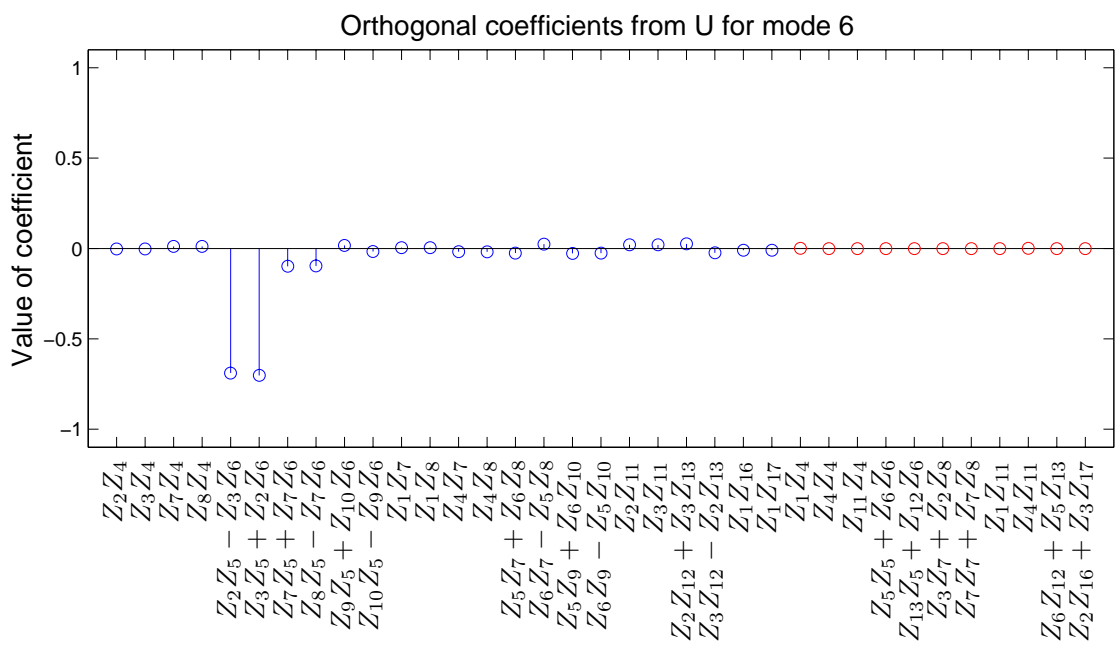
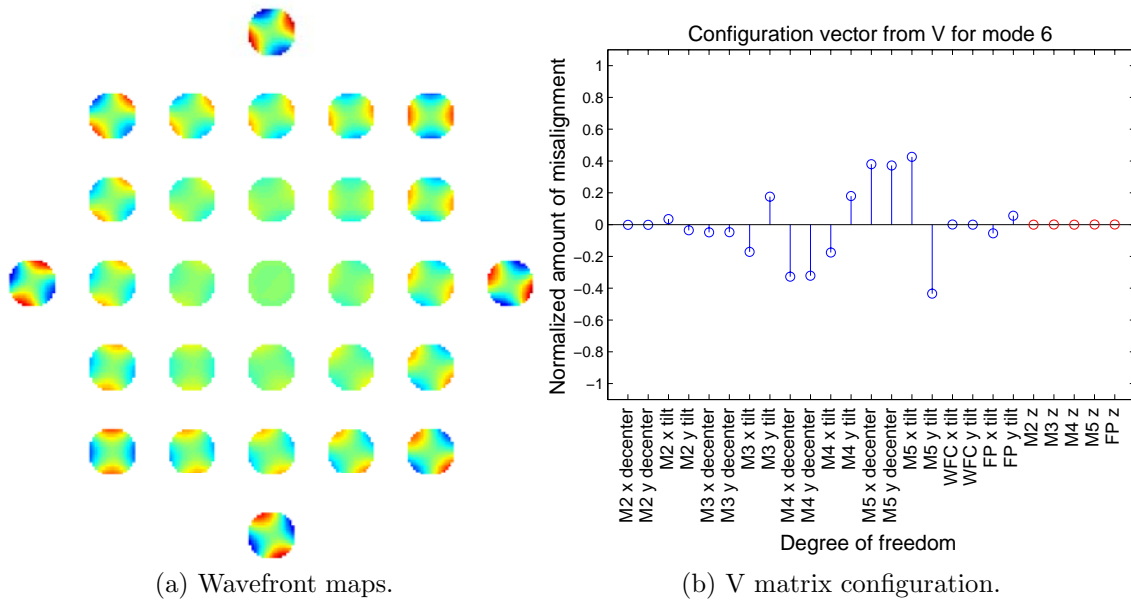


Figure O.6: HET SVD mode 6: linear astigmatism.

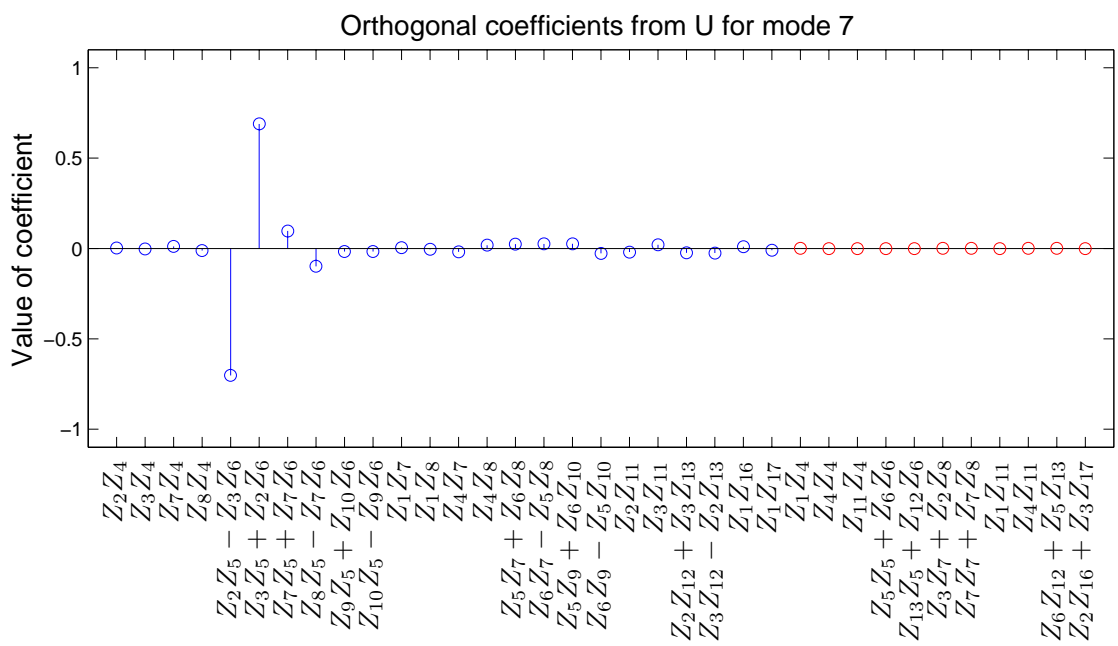
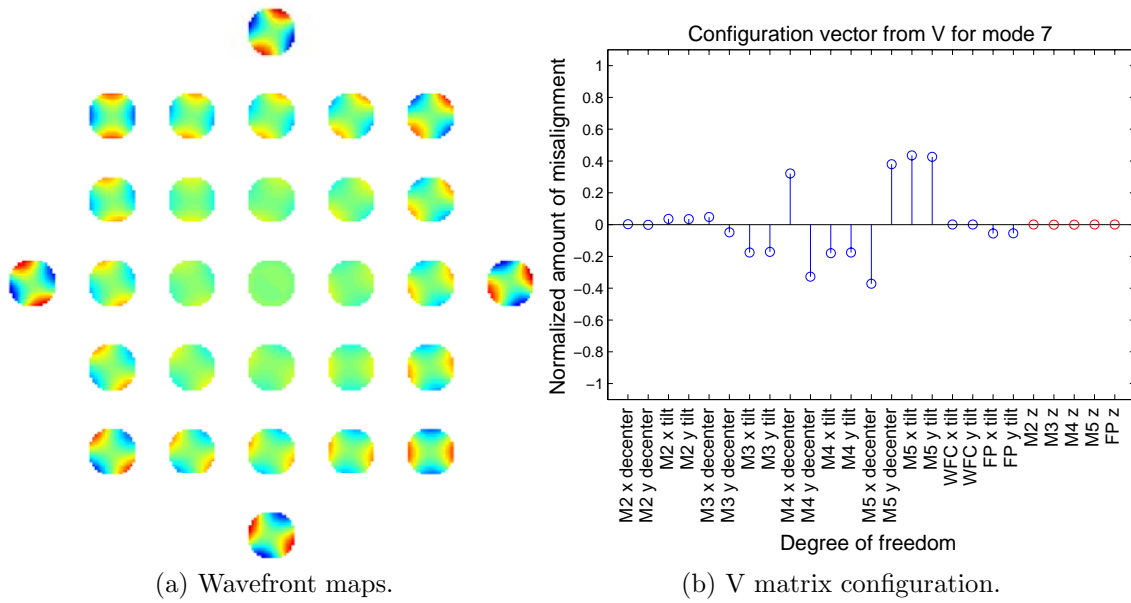


Figure O.7: HET SVD mode 7: linear astigmatism.

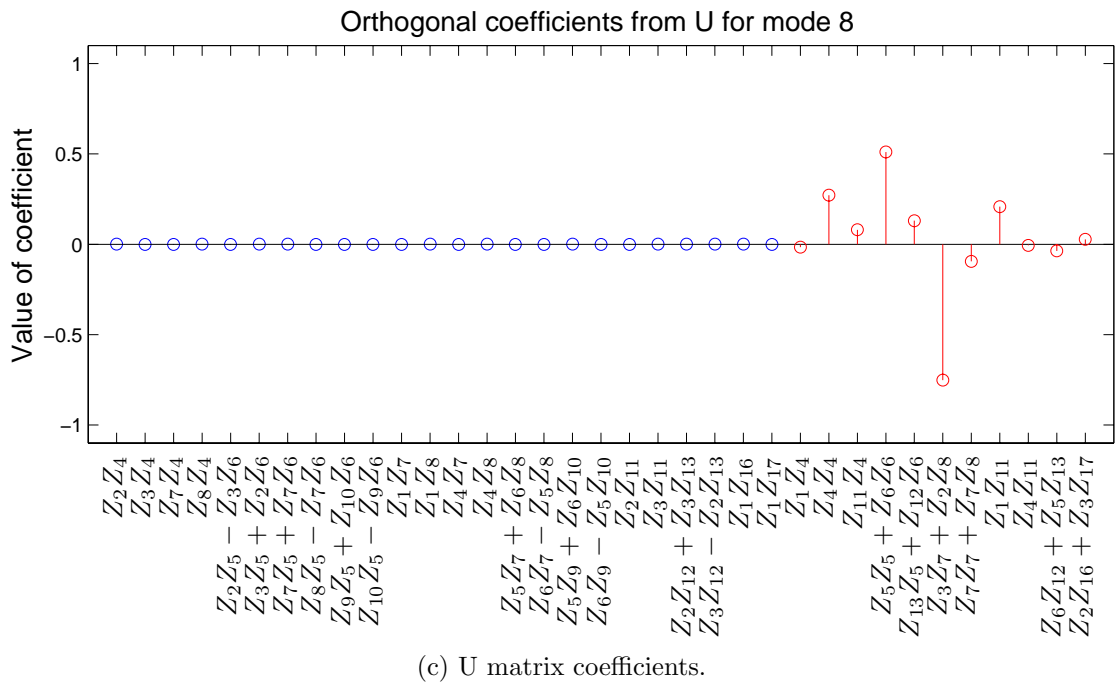
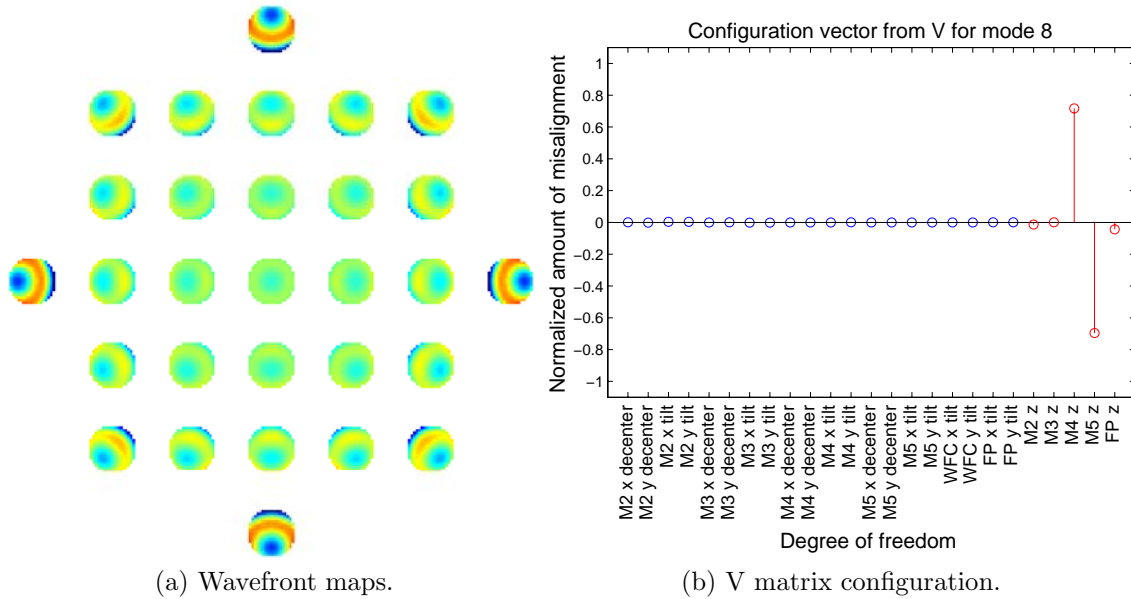
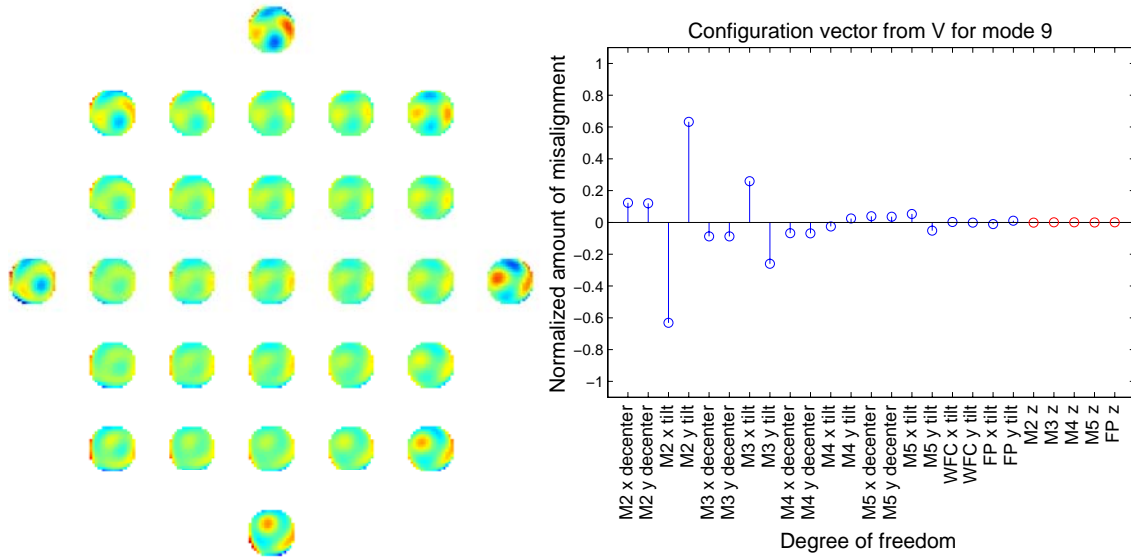
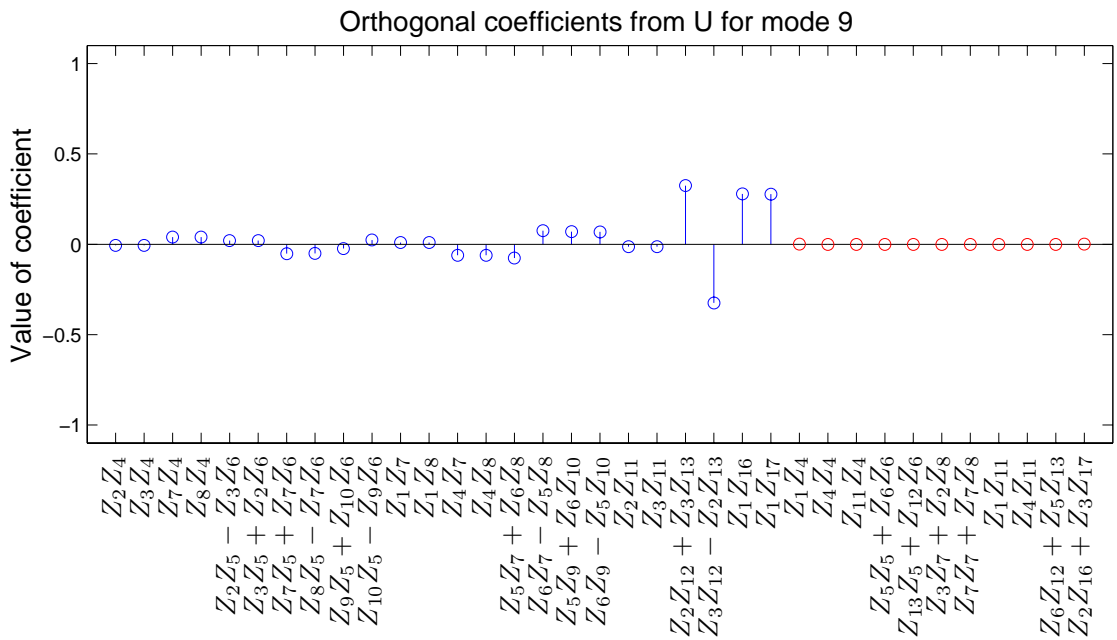


Figure O.8: HET SVD mode 8: linear coma, quadratic astigmatism (axisymmetric errors).



(a) Wavefront maps.

(b) V matrix configuration.



(c) U matrix coefficients.

Figure O.9: HET SVD mode 9: constant secondary coma, linear secondary astigmatism.

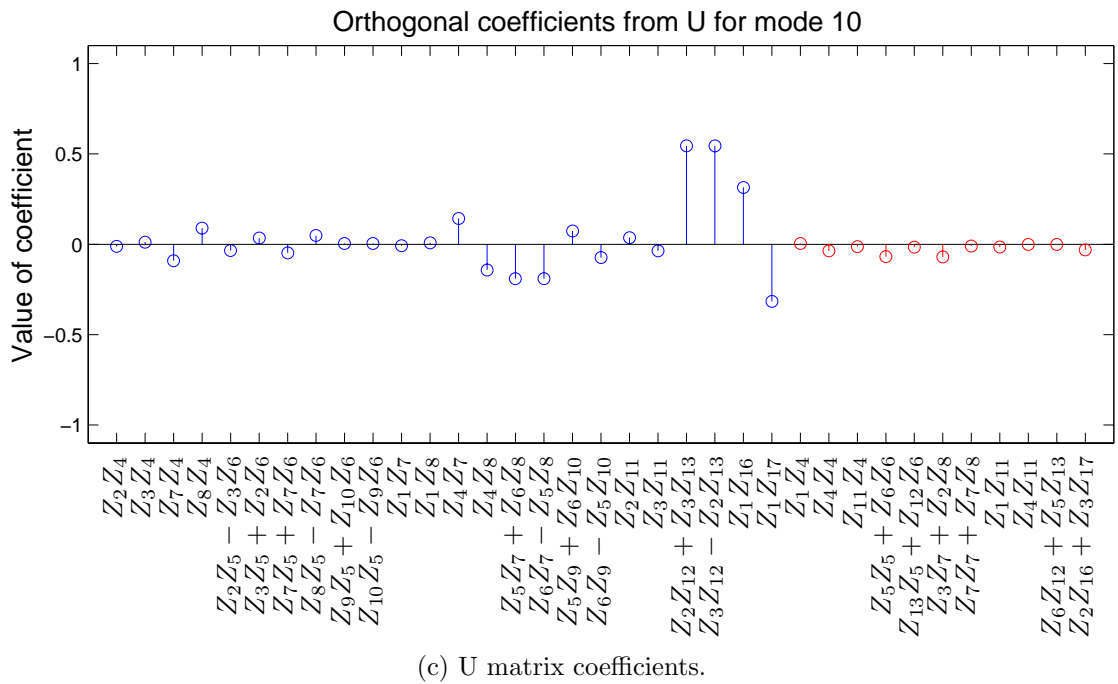
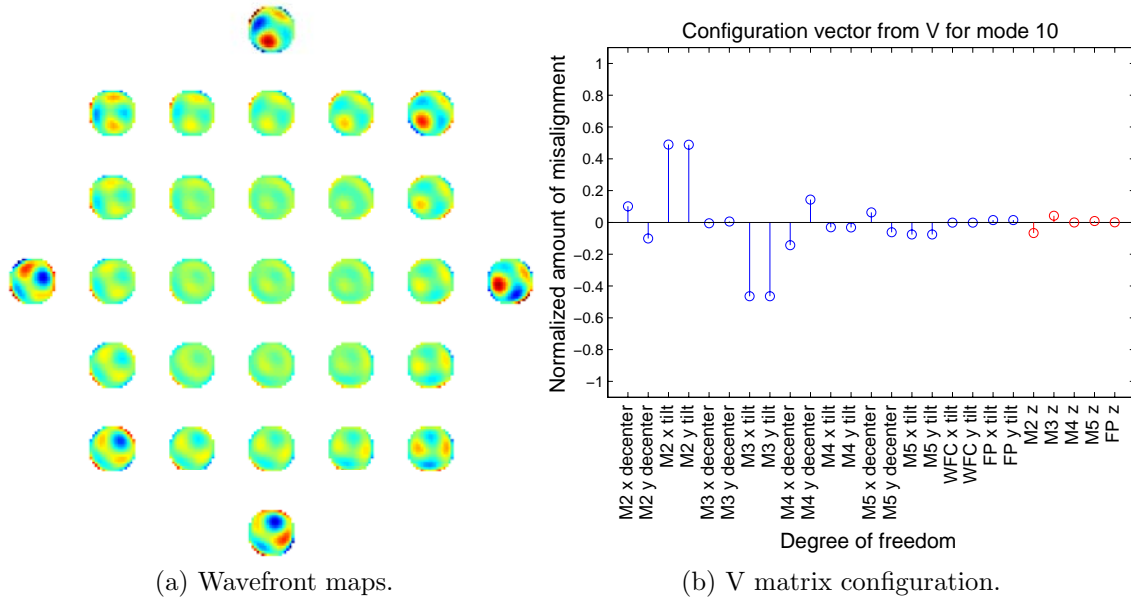
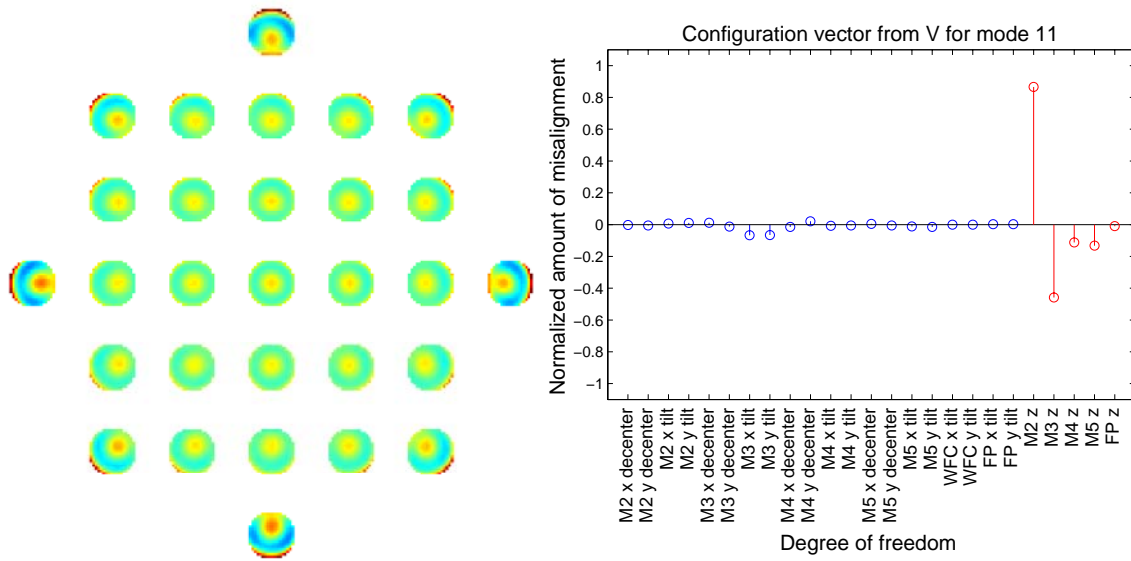
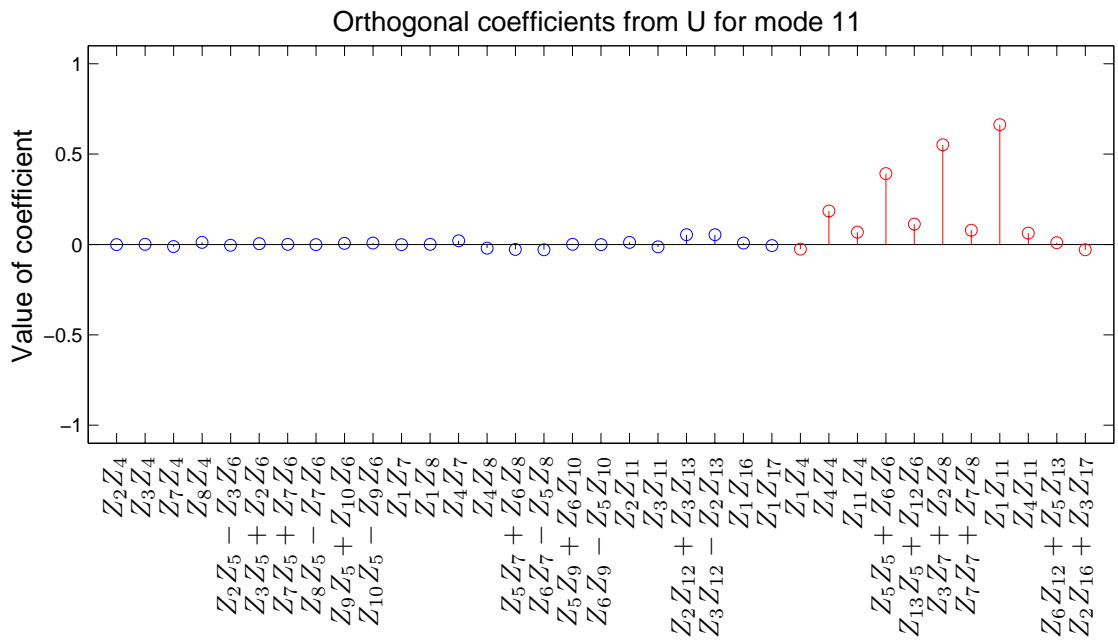


Figure O.10: HET SVD mode 10: constant secondary coma, linear secondary astigmatism.



(a) Wavefront maps.

(b) V matrix configuration.



(c) U matrix coefficients.

Figure O.11: HET SVD mode 11: constant spherical and other axisymmetric error.



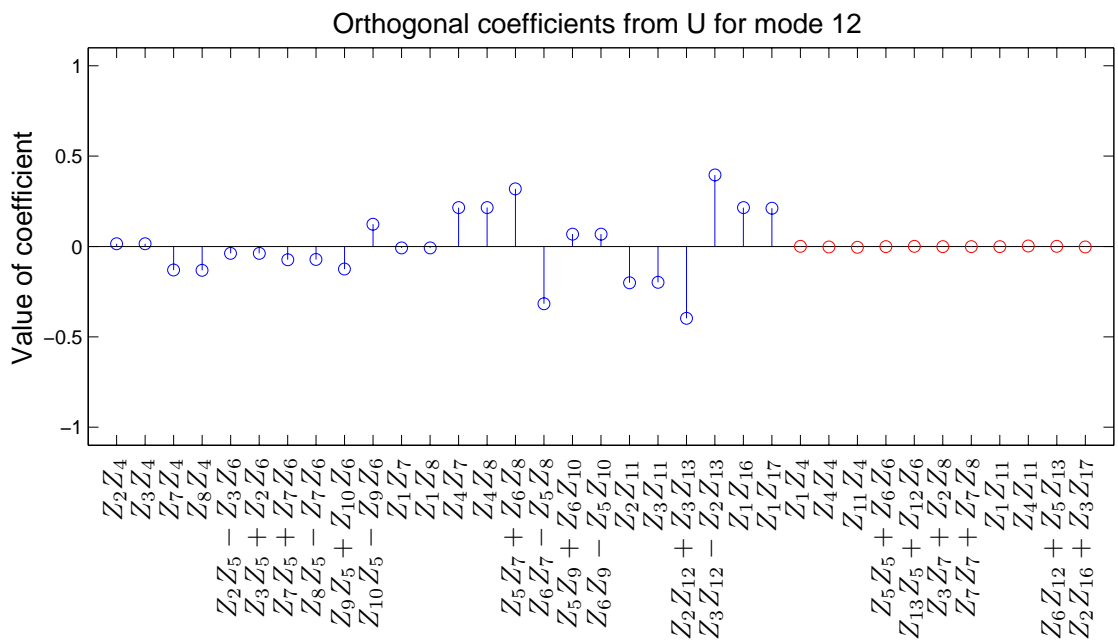
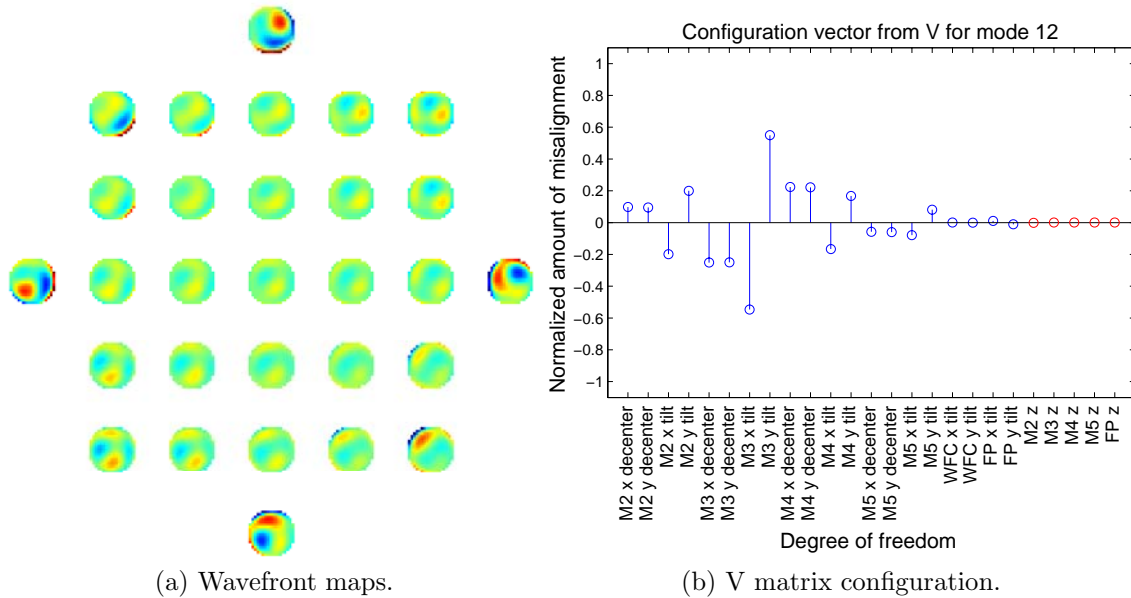


Figure O.12: HET SVD mode 12.

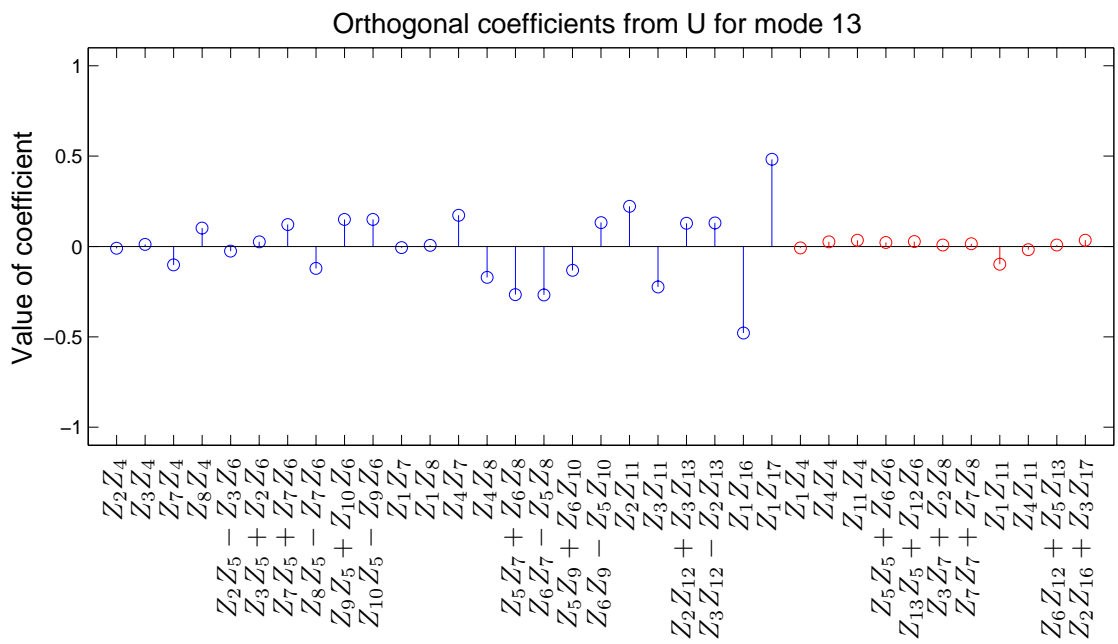
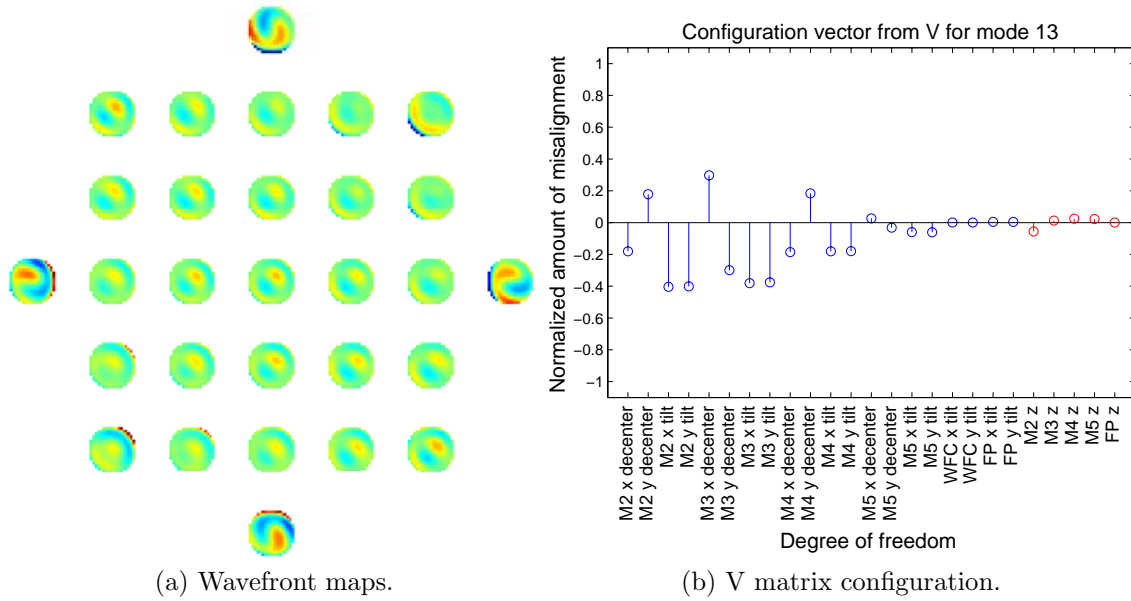


Figure O.13: HET SVD mode 13.

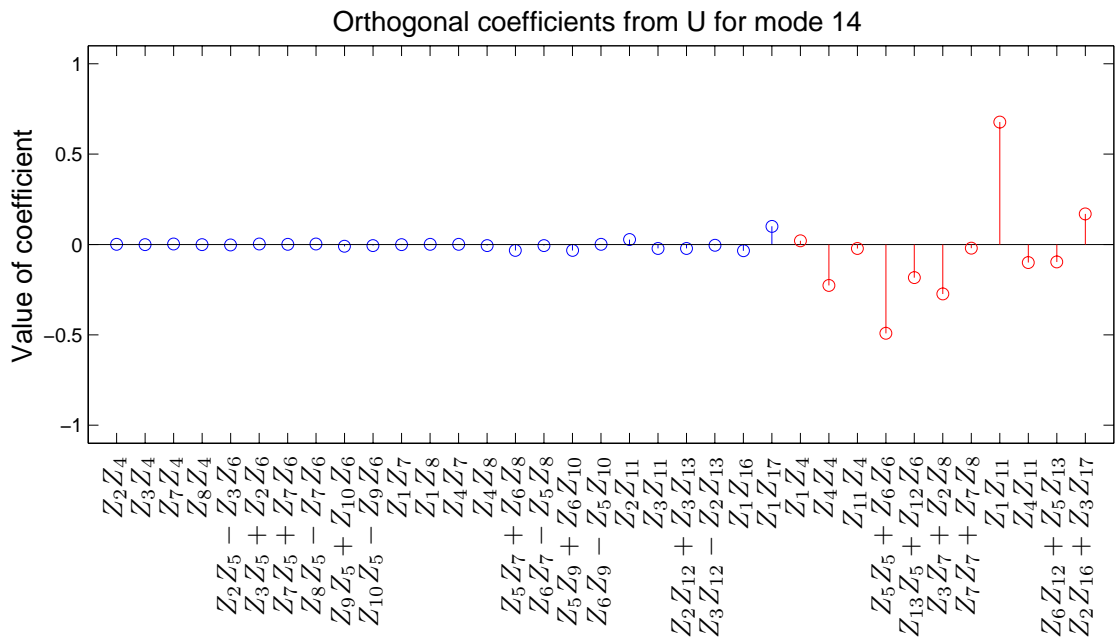
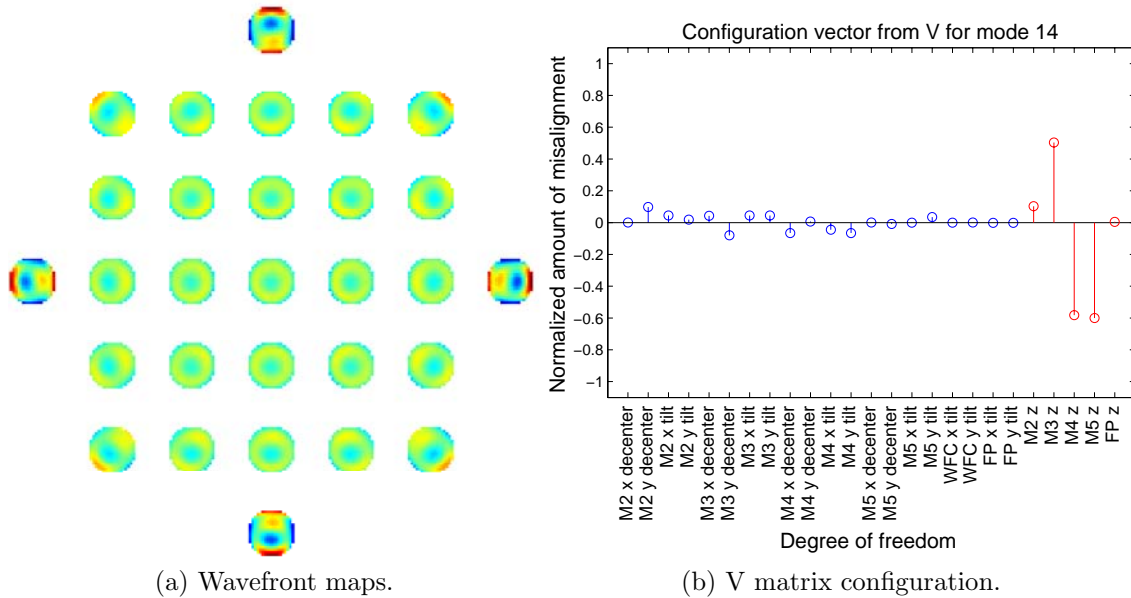


Figure O.14: HET SVD mode 14.

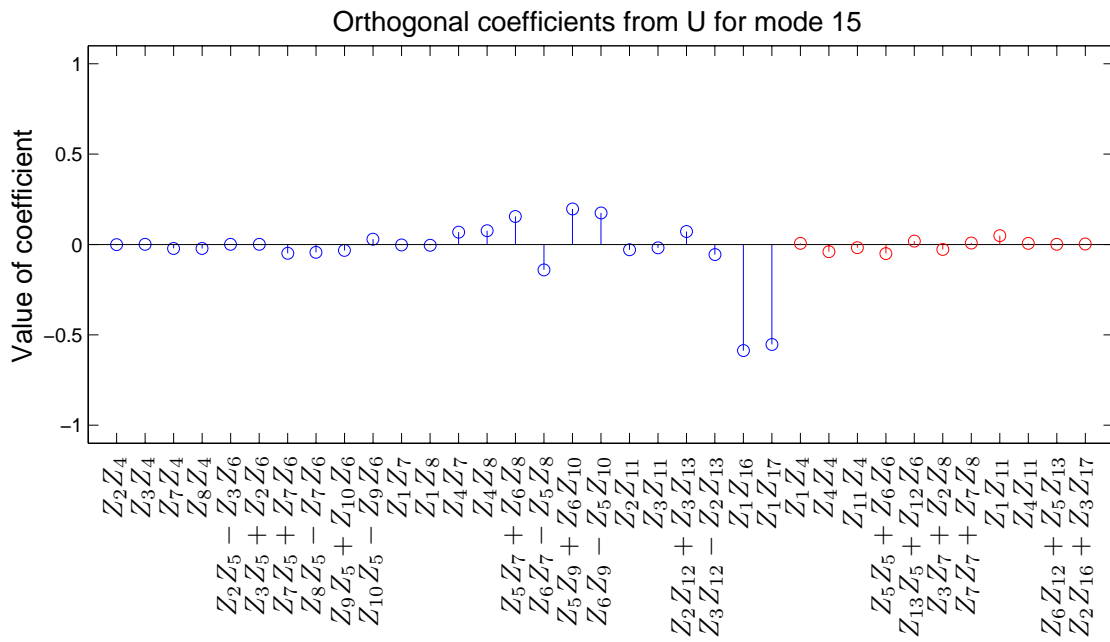
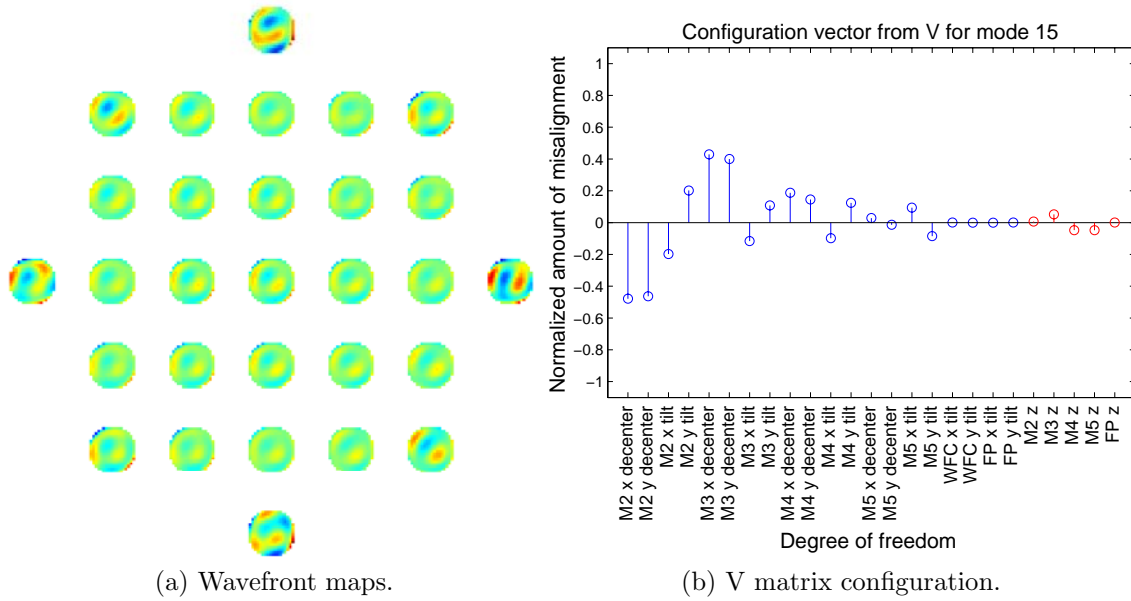


Figure O.15: HET SVD mode 15.

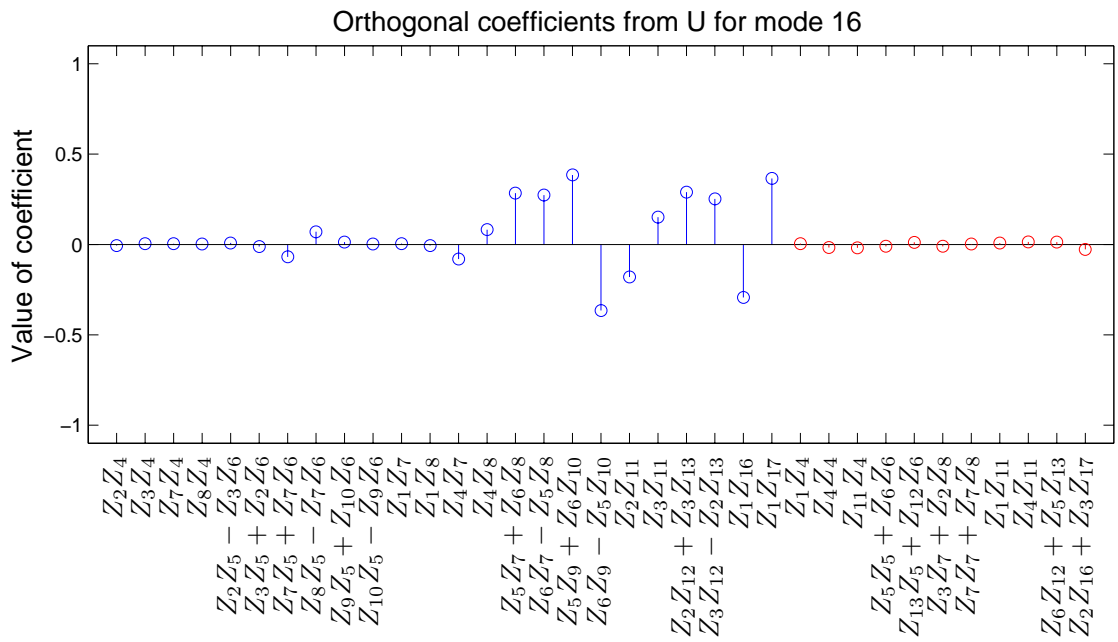
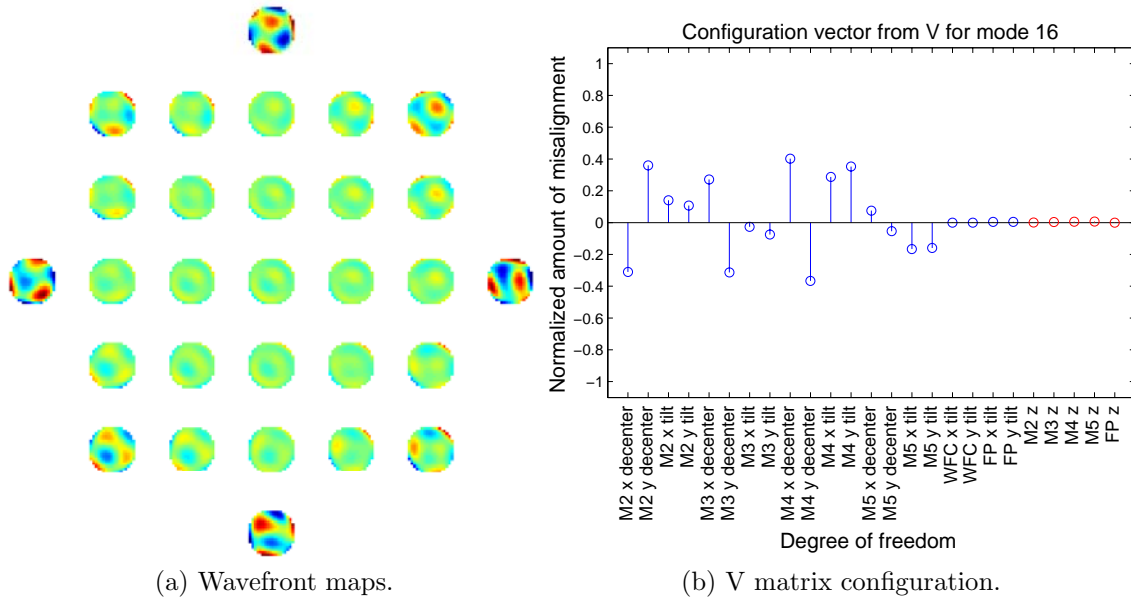


Figure O.16: HET SVD mode 16.

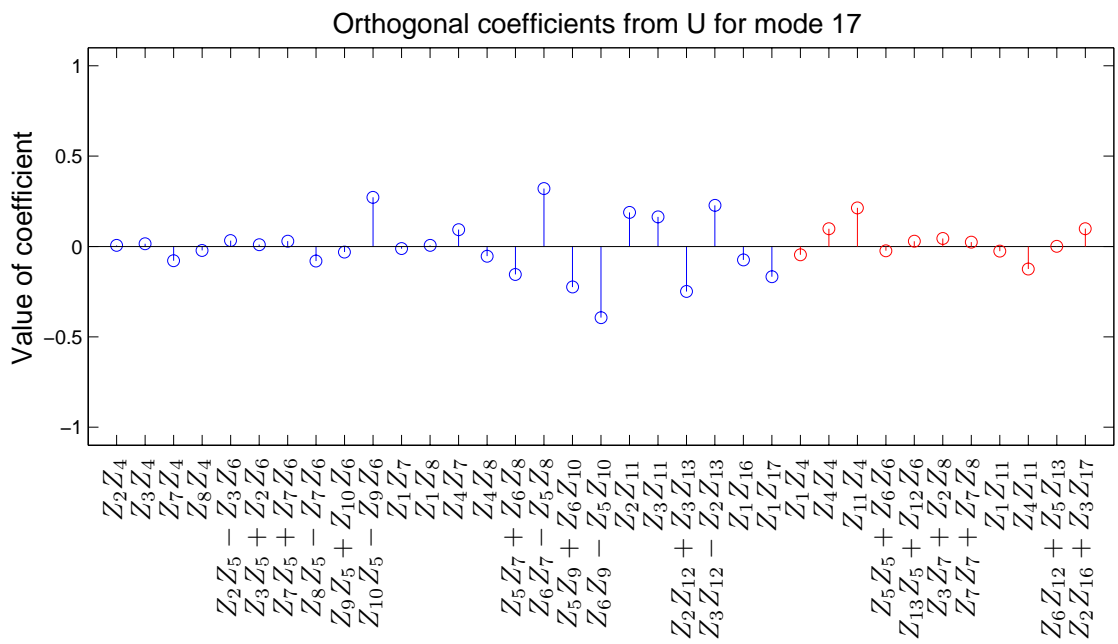
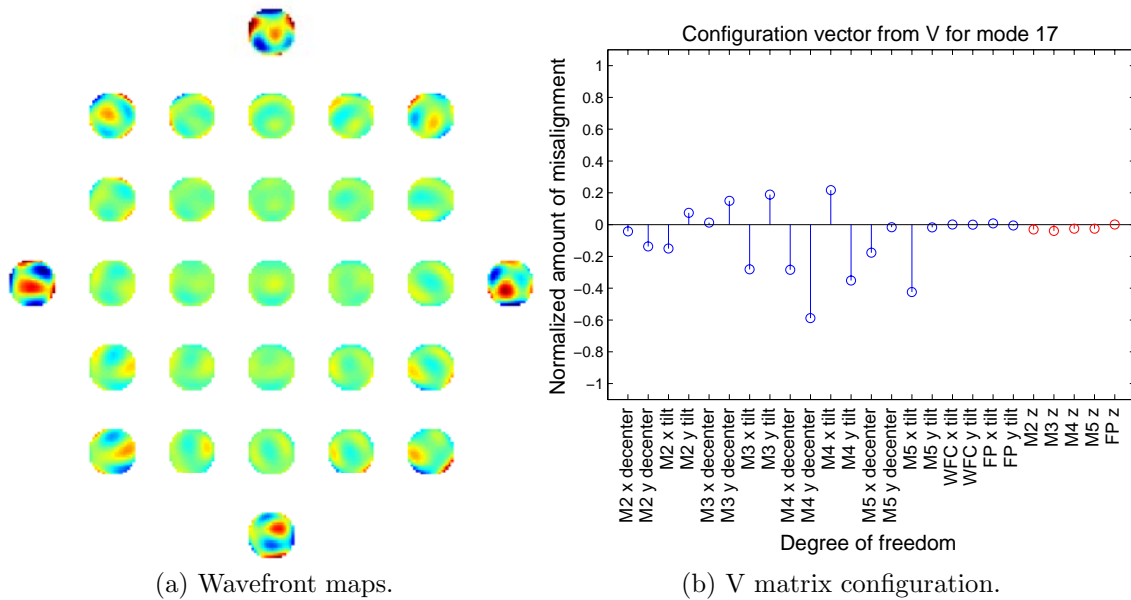


Figure O.17: HET SVD mode 17.

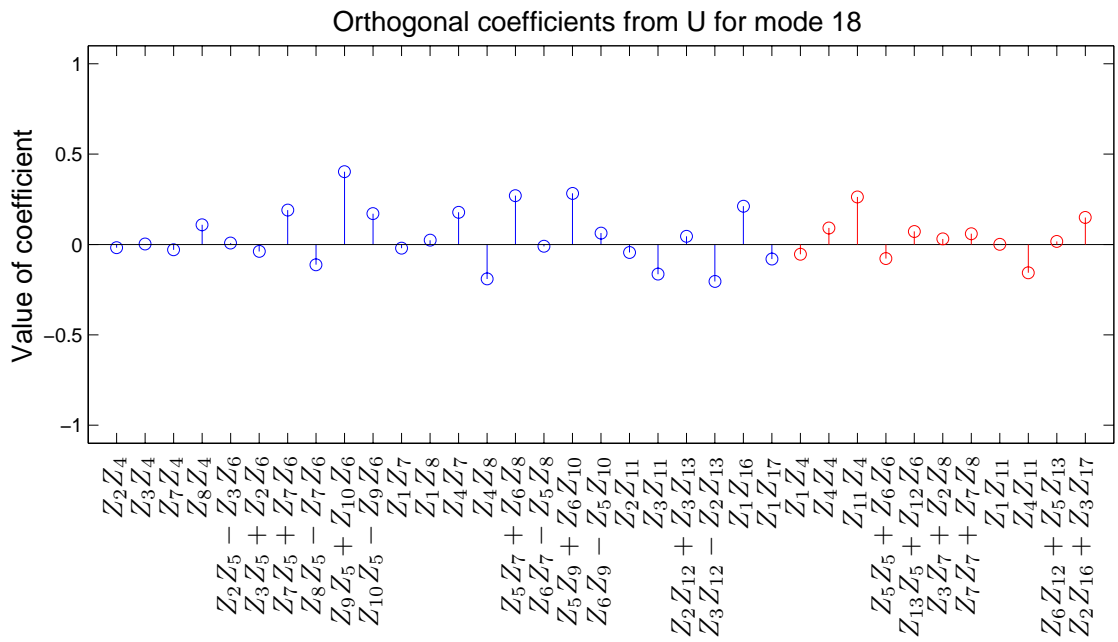
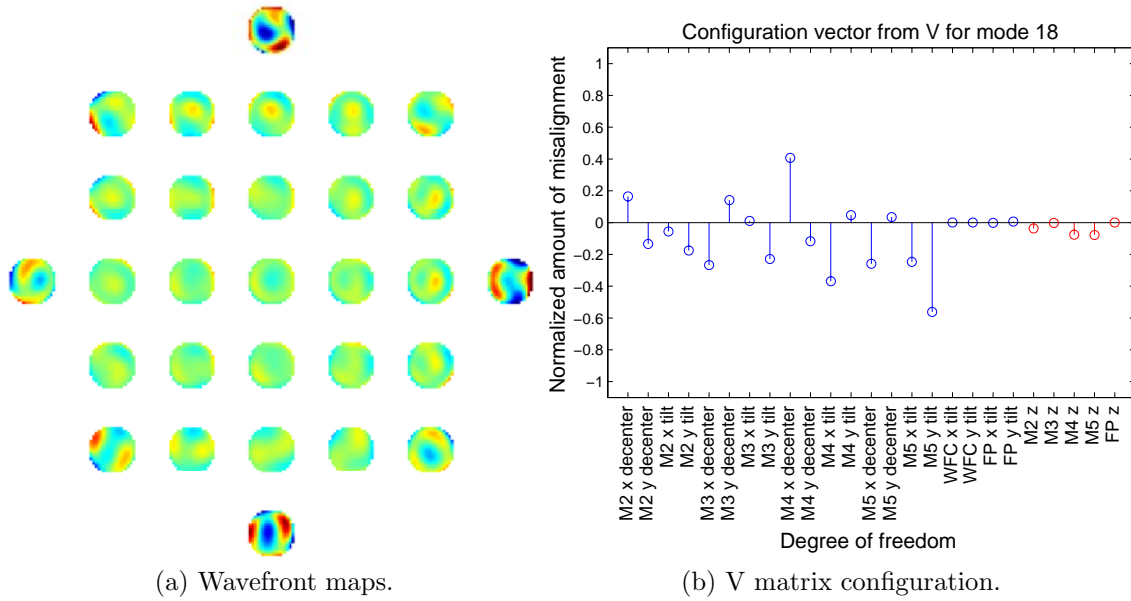


Figure O.18: HET SVD mode 18.

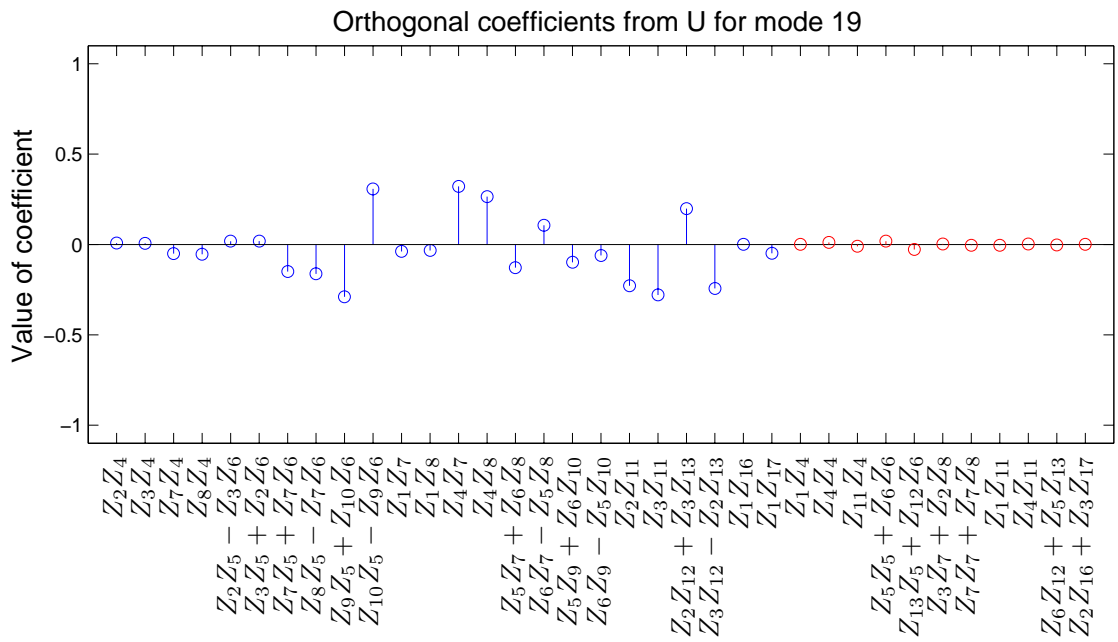
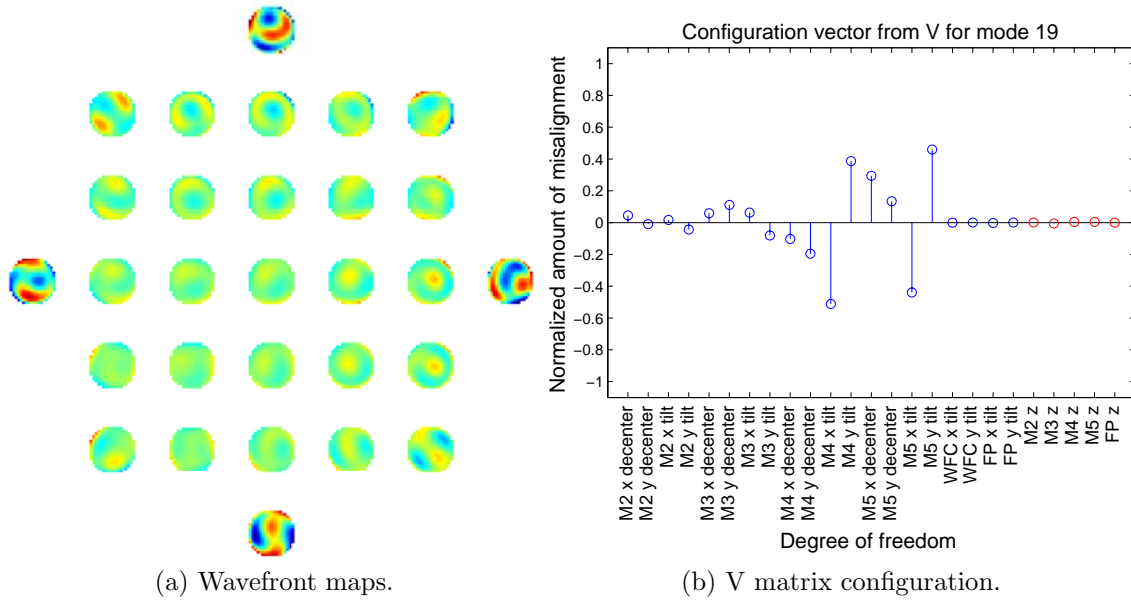


Figure O.19: HET SVD mode 19.



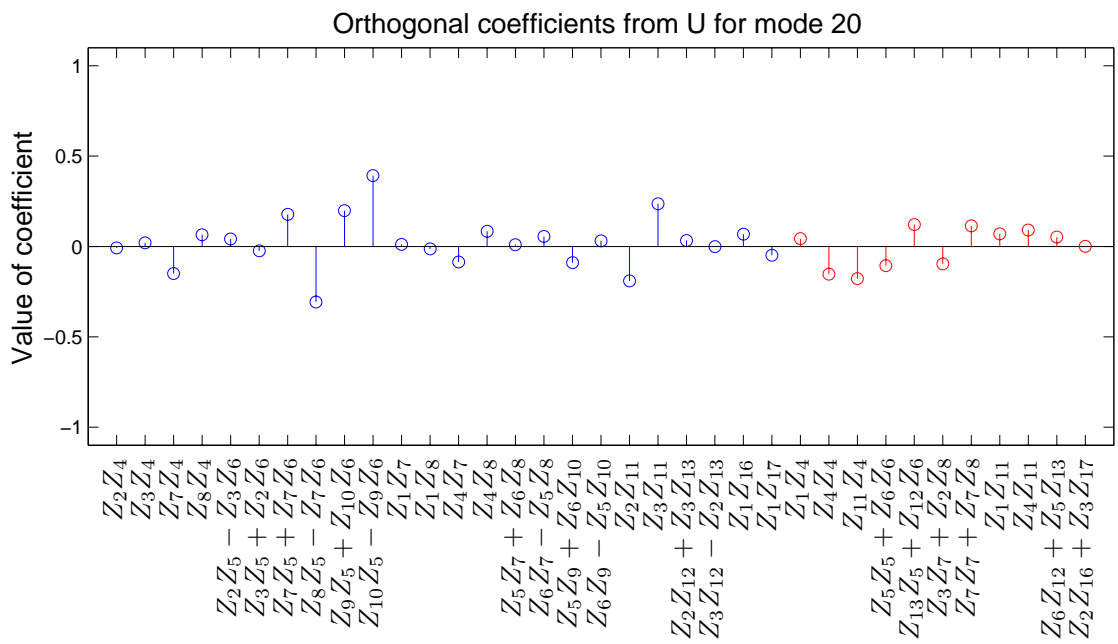
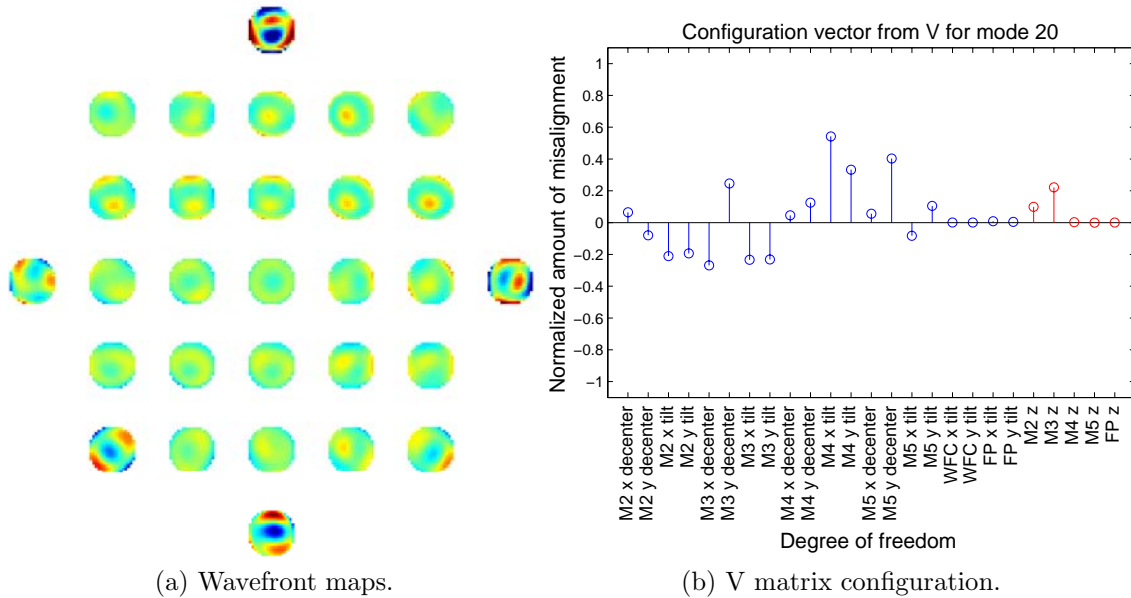


Figure O.20: HET SVD mode 20.

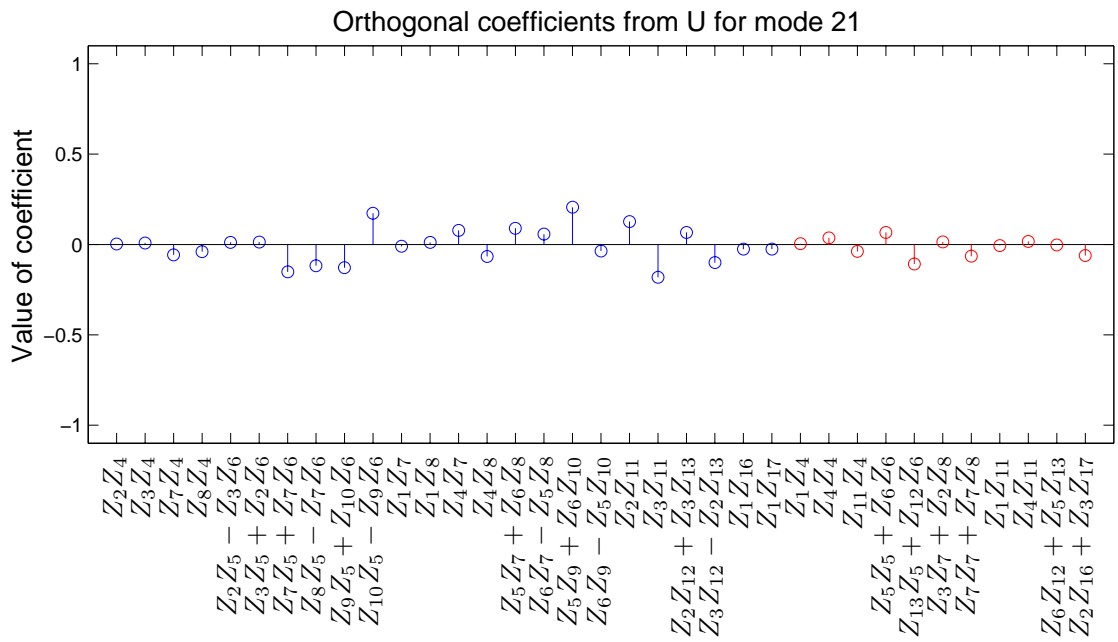
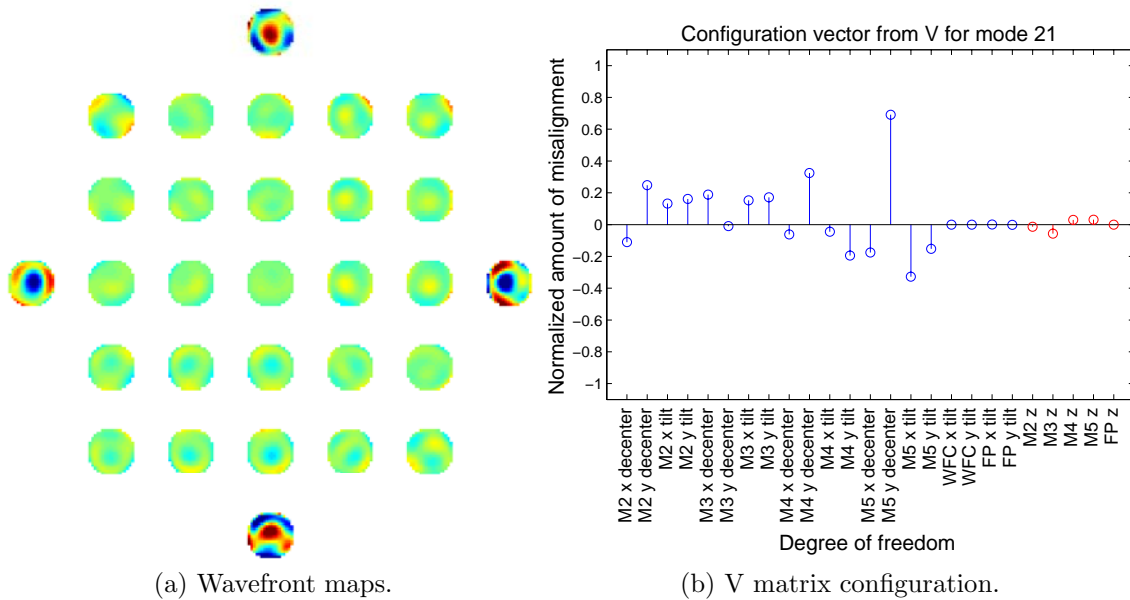


Figure O.21: HET SVD mode 21.

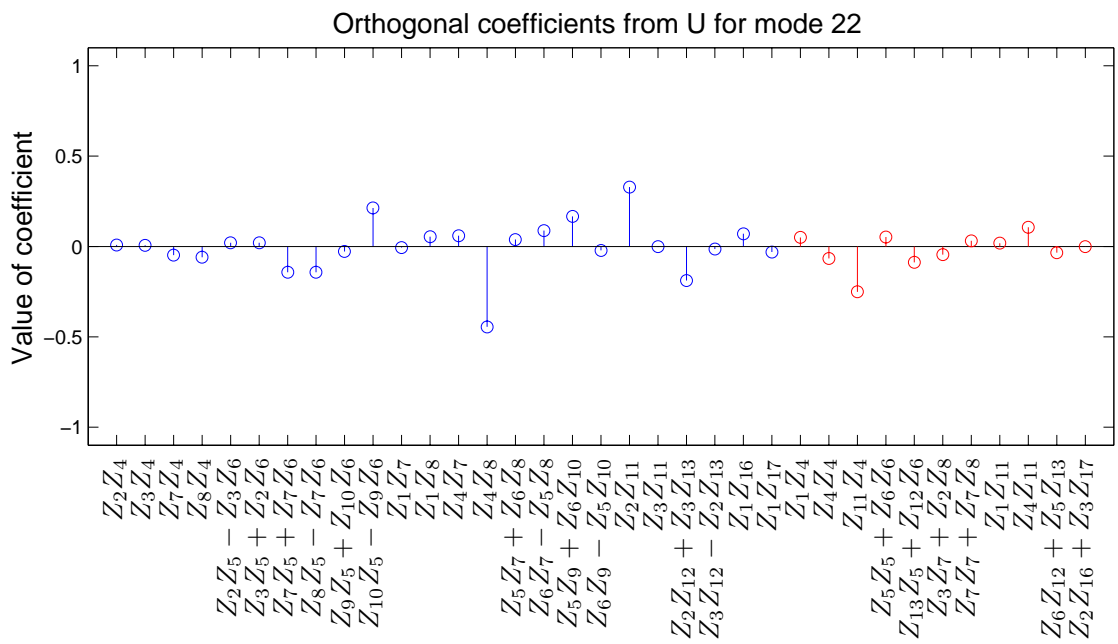
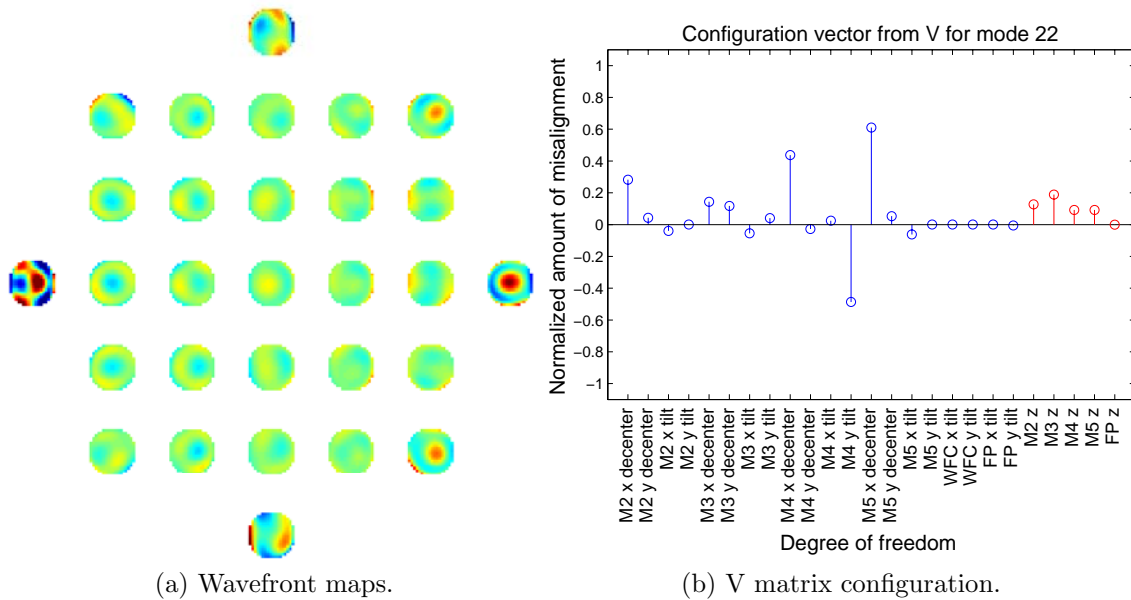


Figure O.22: HET SVD mode 22.

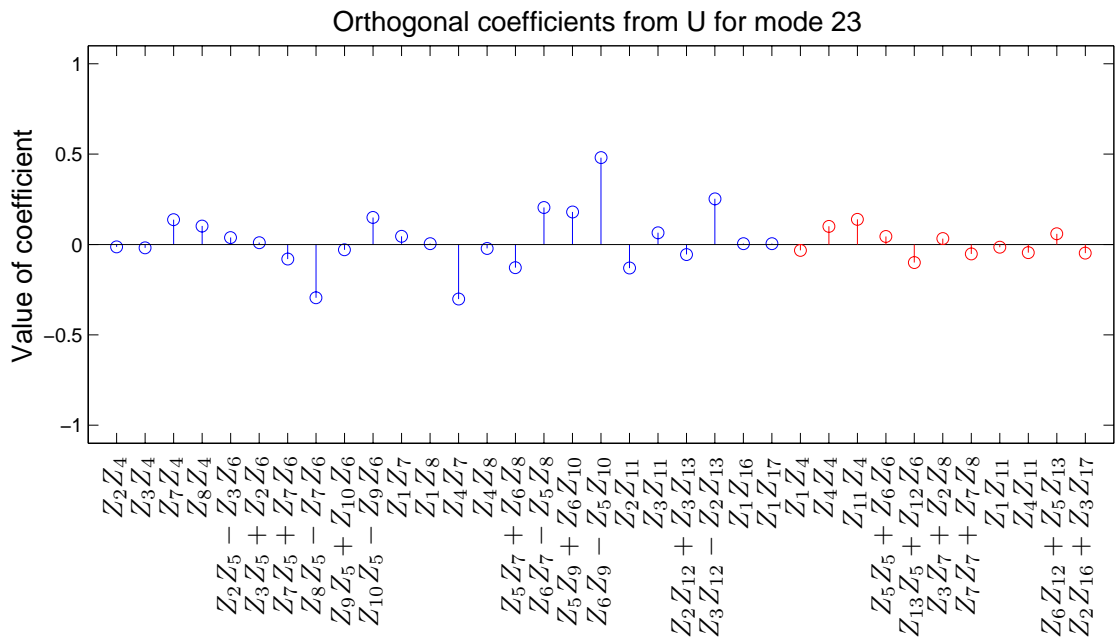
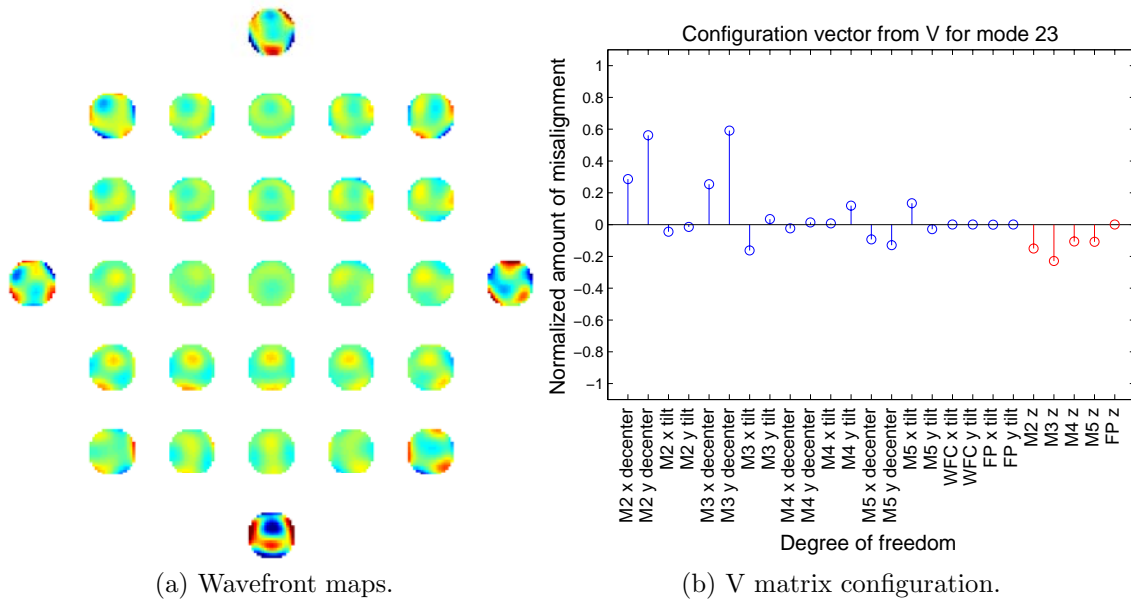


Figure O.23: HET SVD mode 23.

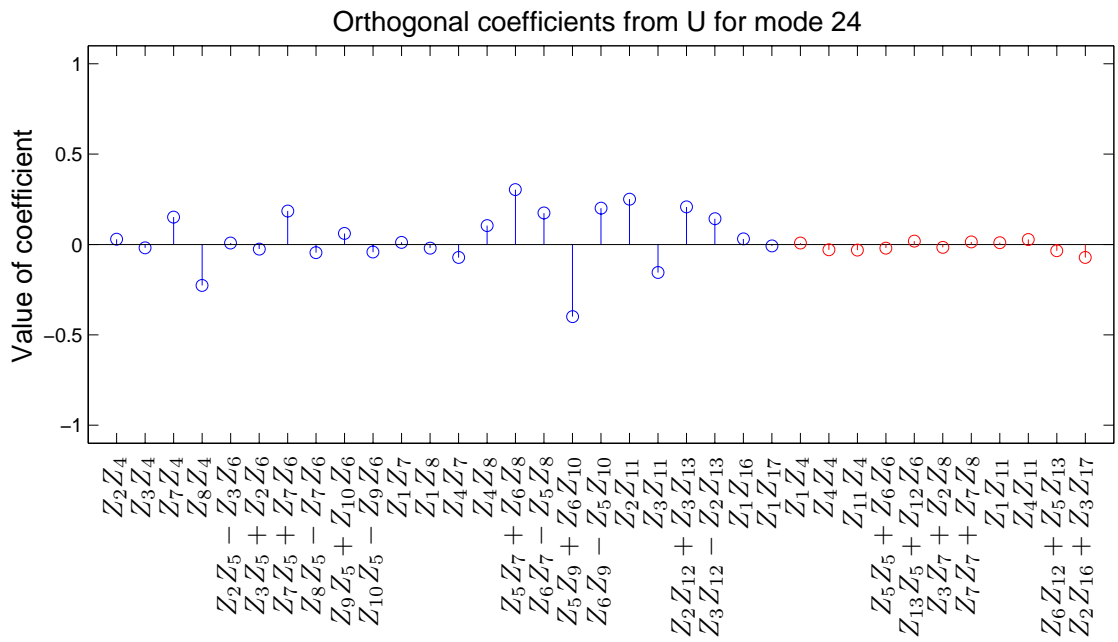
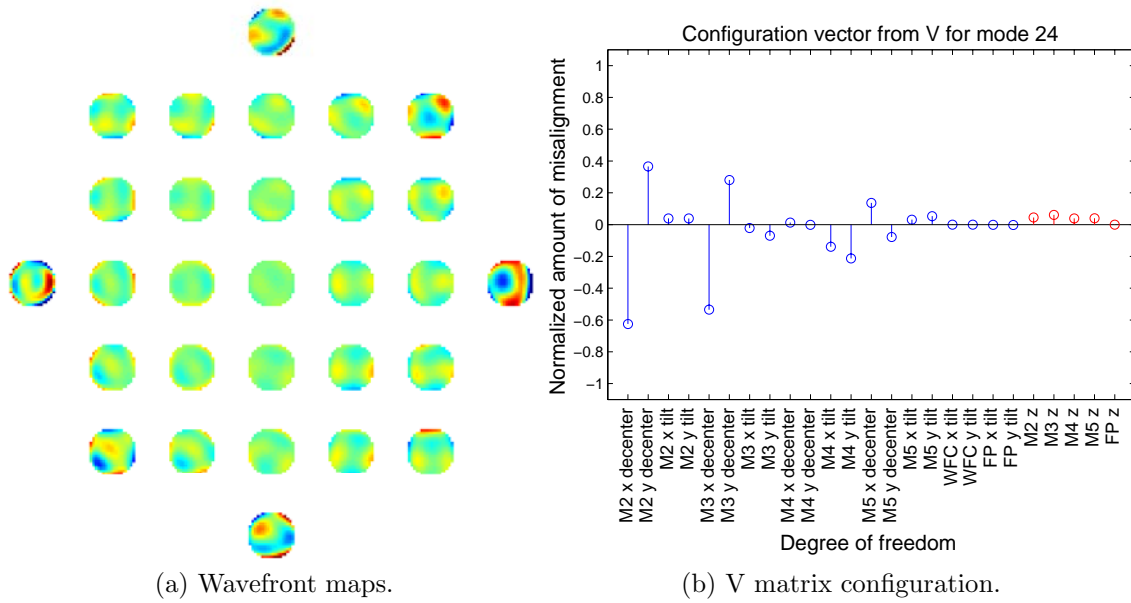


Figure O.24: HET SVD mode 24.

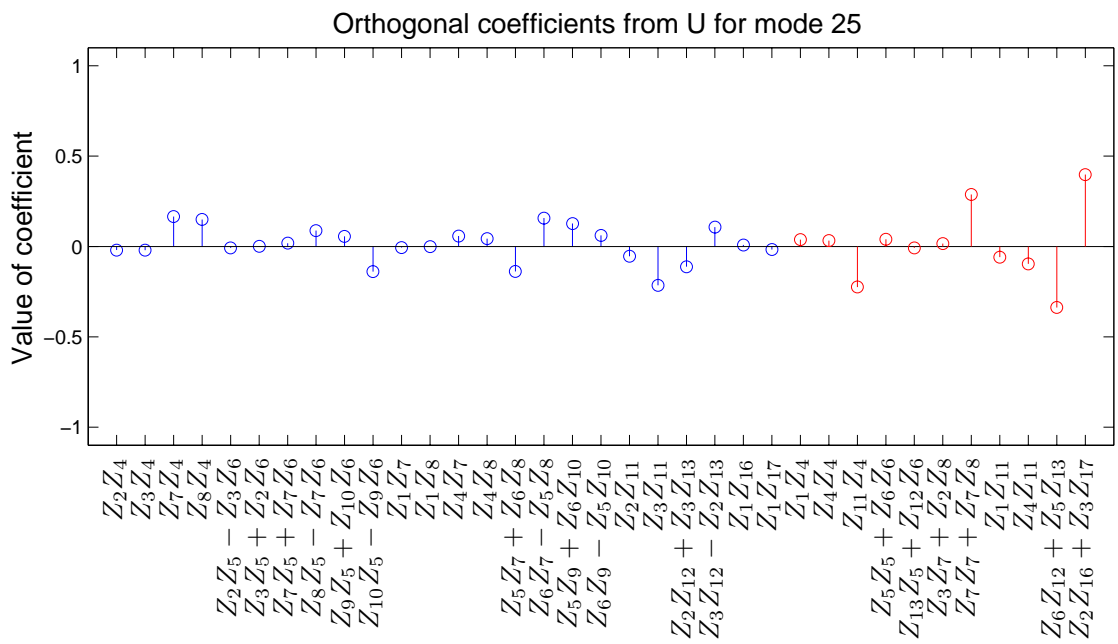
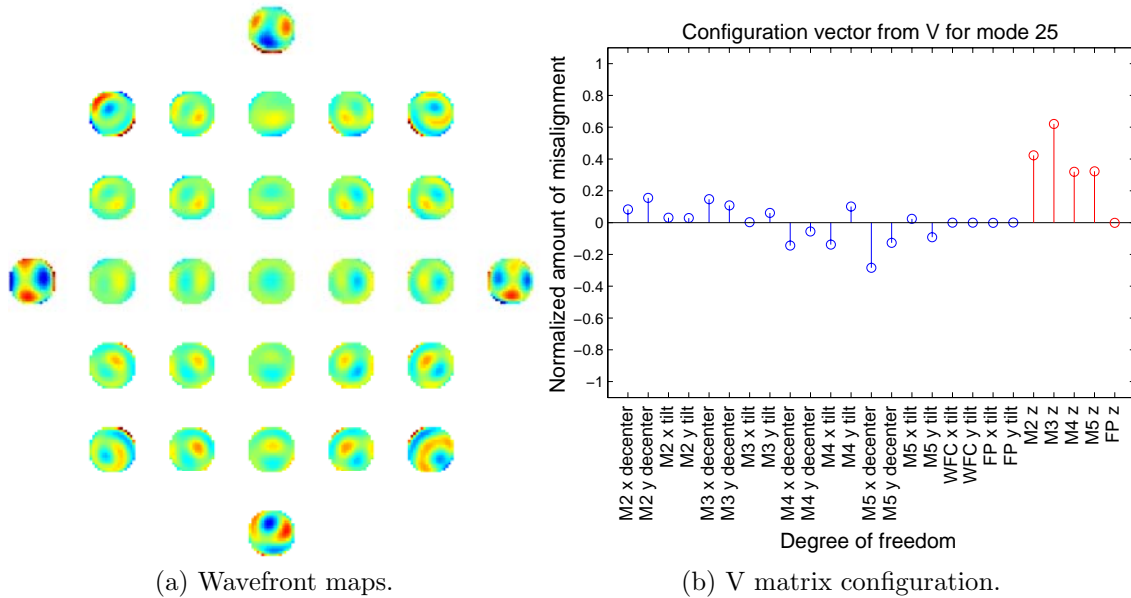


Figure O.25: HET SVD mode 25.

## REFERENCES

- Agurok, Il'ya P. 1998. Double expansion of wavefront deformation in Zernike polynomials over the pupil and the field of view of optical systems: lens design, testing, and alignment. Proc. SPIE, Vol. 3430, 80-87.
- Agurok, Il'ya P. 2000. Optimum position and minimum number of field-of-view points for optical system wavefront comprehension. Proc. SPIE, Vol. 4092, 38-47.
- Barrett, H. H., and Myers, K. J. 2004. *Foundations of Image Science*. Wiley-Interscience.
- Beckmann, Petr. 1973. *Orthogonal Polynomials for Engineers and Physicists*. Golem Press.
- Bhatia, R. K. 1995. Telescope alignment: is the zero-coma condition sufficient? Proc. SPIE, Vol. 2479, 354.
- Booth, John A., MacQueen, Phillip J., Good, John M., Wesley, Gordon L., Hill, Gary J., Palunas, Povilas, Segura, Pedro R., and Calder, Robert E. 2006. The wide field upgrade for the Hobby-Eberly Telescope. Proc. SPIE, Vol. 6267, 62673W.
- Born, Max, and Wolf, Emil. 1999. *Principles of Optics, seventh (expanded) edition*. Cambridge University Press, Cambridge.
- Buchdahl, H.A. 1968. *Optical Aberration Coefficients*. Dover Publications, Inc.
- Buchroeder, Richard Alfred. 1976. *Tilted component optical systems*. Ph.D. dissertation, University of Arizona, Tucson, Arizona.
- Burge, James H. 1993. *Advanced techniques for measuring primary mirrors for astronomical telescopes*. Ph.D. dissertation, University of Arizona, Tucson, Arizona.
- Burge, James H., Su, Peng, Zhao, Chunyu, and Zobrist, Tom. 2007. Use of a commercial laser tracker for optical alignment. Proc. SPIE, Vol. 6676, 66760E.
- Chapman, Henry N., and Sweeney, Donald W. 1998. Rigorous method for compensation selection and alignment of microlithographic optical systems. Proc. SPIE, Vol. 3331, 102-113.
- Clampin, Mark. 2008. Status of the James Webb Space Telescope (JWST). Proc. SPIE, Vol. 7010, 70100L.

- Claver, Charles F., Sweeney, Donald W., Tyson, John A., Althouse, Bryan, Axelrod, Timothy S., Cook, Kem H., Daggert, Larry G., Kantor, Jeffrey C., Kahn, Steven M., Krabbendam, Victor L., Pinto, Philip, Sebag, Jacques, Stubbs, C., and Wolff, Sidney C. 2004. Project status of the 8.4-m LSST. *Proc. SPIE*, Vol. 5489, 705–716.
- CODE V. 2009. *CODE V 10.1 Reference Manual*.
- Descour, M.R., Willer, M.R., Clarke, D.S., and Volin, C.E. 2000. Misalignment modes in high-performance optical systems. *Optical Engineering*, **39**, 1737–1747.
- Didkovsky, Leonid V., Kuhn, Jeff R., and Goode, Philip R. 2004. Optical design for a New off-axis 1.7 m Solar Telescope (NST) at Big Bear. *Proc. SPIE*, Vol. 5171, 333–343.
- Fehniger, M. J. 1980. Alignment of a full aperture system test of a Cassegrain telescope. *Proc. SPIE*, Vol. 251, 21–28.
- Hopkins, H. H. 1950. *The Wave Theory of Aberrations*. Oxford on Clarendon.
- Hvisc, Anastacia M., and Burge, James H. 2008. Alignment analysis of four-mirror spherical aberration correctors. *Proc. SPIE*, Vol. 7018, 701819.
- Johns, Matt. 2008. Progress on the GMT. *Proc. SPIE*, Vol. 7012, 70121B.
- Krabbendam, Victor L., Sebring, Thomas A., Ray, Frank B., and Fowler, James R. 1998. Development and performance of Hobby-Eberly Telescope 11-m segmented mirror. *Proc. SPIE*, Vol. 3352, 436–445.
- Kwee, I. W., and Braat, J. J. M. 1993. Double Zernike expansion of the optical aberration function. *Pure and Applied Optics: Journal of the European Optical Society Part A*, **2**, 21–32.
- Lundström, L., and Unsbo, P. 2007. Transformation of Zernike coefficients: scaled, translated, and rotated wavefronts with circular and elliptical pupils. *Journal of the Optical Society of America A*, **24**, 3, 569–577.
- Mahajan, V. N. 1983. Strehl ratio for primary aberrations in terms of their aberration variance. *Journal of the Optical Society of America*, **73**, 6, 860–861.
- Mahajan, V. N. 1993. Strehl ratio for primary aberrations: some analytical results for circular and annular pupils: erratum. *Journal of the Optical Society of America A*, **10**, 9, 2092–2092.
- Mahajan, V. N. 1998. *Optical Imaging and Aberrations, Part I, Ray Geometrical Optics*. SPIE Press.



- Mahajan, V. N. 2001. *Optical Imaging and Aberrations: Part II, Wave Diffraction Optics*. SPIE Press.
- Mahajan, Virendra N. 1981. Zernike annular polynomials for imaging systems with annular pupils. *Journal of the Optical Society of America*, **71**, 1, 75–85.
- Mahajan, Virendra N. 1982. Strehl ratio for primary aberrations: some analytical results for circular and annular pupils. *Journal of the Optical Society of America*, **72**, 9, 1258–1266.
- Mahajan, Virendra N., and Dai, Guang-ming. 2006. Orthonormal polynomials for hexagonal pupils. *Optics Letters*, **31**, 16, 2462–2464.
- Malacara, D. 2007. *Optical Shop Testing*. Wiley-Interscience.
- Manuel, Anastacia M., and Burge, James H. 2009. Alignment aberrations of the New Solar Telescope. Proc. SPIE, Vol. 7433, 74330A.
- Martin, H. M., Burge, J. H., Miller, S. M., Smith, B. K., Zehnder, R., and Zhao, C. 2006. Manufacture of a 1.7 m prototype of the GMT primary mirror segments. Proc. SPIE, Vol. 6273, 62730G.
- McLeod, B. A. 1996. Collimation of fast wide-field telescopes. *Publications of the Astronomical Society of the Pacific*, **108**, 217–219.
- Moore, Lori B., Hvisc, Anastacia M., and Sasian, José. 2008. Aberration fields of a combination of plane symmetric systems. *Optics Express*, **16**, 20, 15655–15670.
- Nienhuis, K., and Nijboer, B. R. A. 1949. The diffraction theory of optical aberrations. Part III: General formulae for small aberrations: experimental verification of the theoretical results. *Physica*, **14**, 590–608.
- Nijboer, B. R. A. 1943. The diffraction theory of optical aberrations. Part I: General discussion of the geometrical aberrations. *Physica*, **10**, 679–692.
- Nijboer, B. R. A. 1947. The diffraction theory of optical aberrations. Part II: Diffraction pattern in the presence of small aberrations. *Physica*, **13**, 10, 605–620.
- Noethe, L., and Guisard, S. 2000. Analytical expressions for field astigmatism in decentered two mirror telescopes and application to the collimation of the ESO VLT. *Astronomy and Astrophysics Supplement Series*, **144**, 1, 157–167.
- Noll, R. J. 1976. Zernike polynomials and atmospheric turbulence. *Journal of the Optical Society of America*, **66**, 3, 207–211.
- O'Donoghue, Darragh, and Swat, Arek. 2002. Spherical aberration corrector of the Southern African Large Telescope (SALT). Proc. SPIE, Vol. 4411, 72–78.

- Phillion, Donald W., Olivier, Scot S., Baker, Kevin, Seppala, Lynn, and Hvisc, Stacie. 2006. Tomographic wavefront correction for the LSST. *Proc. SPIE*, Vol. 6272, 627213.
- Ramsey, Lawrence W., Adams, Mark T., III, Thomas G. Barnes, Booth, John A., Cornell, Mark E., Fowler, James R., Gaffney, Niall I., Glaspey, John W., Good, John M., Hill, Gary J., Kelton, Philip W., Krabbendam, Victor L., Long, Larry E., MacQueen, Phillip J., Ray, Frank B., Ricklefs, Randall L., Sage, J., Sebring, Thomas A., Spiesman, William J., and Steiner, M. 1998. Early performance and present status of the Hobby-Eberly Telescope. *Proc. SPIE*, Vol. 3352, 34–42.
- Rimmer, M. 1970. Analysis of perturbed lens systems. *Applied Optics*, **9**, 3, 533–537.
- Sasian, Jose. 2009. *Theory of sixth-order wave aberrations*. Submitted to *Applied Optics*.
- Schmid, Tobias, Thompson, Kevin, and Rolland, Jannick. 2008. Alignment of two mirror astronomical telescopes (the astigmatic component). *Proc. SPIE*, Vol. 7017, 70170C.
- Schroeder, D. J. 1999. *Astronomical Optics*. Academic Press.
- Seidel, L. 1865. *About the third order expansion that describes the path of a light beam outside the plane of the axis through an optical system of refracting elements*. translated from German by R. Zehnder on behalf of J. Sasian, available from jose.sasian@optics.arizona.edu (2007), Originally published in *Astronomische Nachrichten*, 1027, 1028, 1029.
- Shack, Roland. 2005. *Opti514: Aberration theory*. University of Arizona, OSC Class notes.
- Shack, Roland V., and Thompson, Kevin. 1980. Influence of alignment errors of a telescope system on its aberration field. *Proc. SPIE*, Vol. 251, 146–153.
- Shiri, Ron, Aronstein, David L., Smith, Jeffery S., Dean, Bruce H., and Sabatke, Erin M. 2007. Wavefront control toolbox for James Webb Space Telescope testbed. *Proc. SPIE*, Vol. 6711, 67110P.
- Tessieres, Régis. 2003. *Analysis for alignment of optical systems*. M.S. thesis, University of Arizona, Tucson, Arizona.
- Tessieres, Régis. 2004. *Alignment strategy for the LSST*. LSST Observatory Report.
- Thompson, Kevin P. 1980. *Aberration fields in nonsymmetric optical systems*. Ph.D. dissertation, University of Arizona, Tucson, Arizona.

- Thompson, Kevin P. 2005. Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry. *Journal of the Optical Society of America A*, **22**, 7, 1389–1401.
- Thompson, Kevin P. 2009. Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: spherical aberration. *Journal of the Optical Society of America A*, **26**, 5, 1090–1100.
- Thompson, Kevin P., Schmid, Tobias, and Rolland, Jannick P. 2008. The misalignment induced aberrations of TMA telescopes. *Optics Express*, **16**, 25, 20345–20353.
- Thompson, Kevin P., Schmid, Tobias, Cakmakci, Ozan, and Rolland, Jannick P. 2009. Real-ray-based method for locating individual surface aberration field centers in imaging optical systems without rotational symmetry. *Journal of the Optical Society of America A*, **26**, 6, 1503–1517.
- Turner Jr., Theodore S. 1992. Vector aberration theory on a spreadsheet: analysis of tilted and decentered systems. Proc. SPIE, Vol. 1752, 184.
- Upton, Robert, and Ellerbroek, Brent. 2004. Gram-Schmidt orthogonalization of the Zernike polynomials on apertures of arbitrary shape. *Optics Letters*, **29**, 24, 2840–2842.
- Welford, W. T. 1974. *Aberrations of the symmetrical optical system*. Academic Press.
- Welford, W. T. 1986. *Aberrations of optical systems*. Taylor & Francis.
- Wilson, R. N. 1996. *Reflecting Telescope Optics I: Basic Design Theory and its Historical Development*. Springer.
- Wilson, R. N., and Delabre, B. 1997. Concerning the alignment of modern telescopes: theory, practice, and tolerances illustrated by the ESO NTT. *Publications of the Astronomical Society of the Pacific*, **109**, 53.
- ZEMAX. 2009. *ZEMAX Optical Design Program User's Guide*.
- Zernike, F. 1934. Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. *Mon. Not. R. Astron. Soc*, **94**, 377–384.
- Zernike, F. 2002. Diffraction theory of the knife-edge test and its improved form, the phase-contrast method. *Journal of Microlithography, Microfabrication, and Microsystems*, **1**, 87–94.

- Zhao, Chunyu, and Burge, James H. 2002. Conditions for correction of linear and quadratic field-dependent aberrations in plane-symmetric optical systems. *Journal of the Optical Society of America A*, **19**, 12, 2467–2472.
- Zhou, Ping. 2009. *Error analysis and data reduction for interferometric surface measurements*. Ph.D. dissertation, University of Arizona, Tucson, Arizona.