

ON WAVEFRONT ABERRATIONS IN ASYMMETRIC AND MULTIPLE APERTURE OPTICAL SYSTEMS

Présentée le 26 février 2021

Faculté des sciences et techniques de l'ingénieur
Laboratoire de nanophotonique et métrologie
Programme doctoral en photonique

pour l'obtention du grade de Docteur ès Sciences

par

Alessandro GROSSO

Acceptée sur proposition du jury

Prof. Y. Bellouard, président du jury
Dr T. Scharf, directeur de thèse
Prof. J. Sasiàn, rapporteur
Prof. H. Gross, rapporteur
Prof. H. P. Herzig, rapporteur

Acknowledgements

I am extremely grateful to my supervisors, Toralf Scharf and Kurt Vonmetz, for having accepted my application to a Ph.D. fellowship in the NOLOSS project. Thanks to them, between 2017 and 2019 I had many stimulating experiences during numerous workshops and scientific training, I had the unique opportunity to attend the very expensive courses in Optical System Design using OpticStudio held in the UK and USA and to participate in several conferences around the world.

I want to thank all the colleagues and friends I had the privilege to meet during my stay in Datalogic IP Tech Srl in the fantastic scenario of Bologna, a city of innovation, art and culture.

I am grateful to Prof. Olivier Martin and to all the colleagues and friends I had the fortune to meet in the Nanophotonics and Metrology Laboratory (NAM) at EPFL in Lausanne.

I express my infinite gratitude to my beloved family for endless support and encouragement and I dedicate this work to my dear father who departed this life prematurely but will continue to live in me.

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 675745.

Abstract

The present work deals with monochromatic wavefront aberrations in optical systems without symmetries. The treatment begins with a class of systems characterized by misaligned spherical surfaces whose behavior is analyzed using the wavefront aberration expansion proposed in the framework of the Nodal Aberration Theory (NAT). It is derived the full field behavior of the Zernike polynomials in the Fringe indexing scheme for this class of systems. Then, the attention is focused on a more general class of asymmetric systems where the misaligned surfaces can be individually double-plane symmetric. In this case, considering aberrations up to the 4th order, it is shown that the field dependence of Zernike terms is described by general second-degree polynomials. The presence of double-plane symmetric optical surfaces induces additional perturbations to the magnitude of the field variation of primary aberrations for this class of systems. In particular, one observes that coma aberration acquires an elliptical conic shape in the field domain, while the full field variation of primary astigmatism magnitude is described by a class of surfaces that we define as “generalized Cassini surfaces” because these are more general than the standard Cassini surfaces describing the binodal behavior of astigmatism in the class of optical systems analyzed with NAT wavefront aberration expansion.

These considerations are preliminary to the discussion of the second part of this thesis whose scope is to analyze monochromatic wavefront aberrations in a further class of systems, namely optical systems characterized by multiple apertures. In this sense, it is first introduced a general description of the wavefront aberration function in the framework of Hamiltonian Optics. This consists of a full power series expansion in the ray coordinates that provides the most general representation of optical systems without symmetries. These introductory remarks are necessary to carry out the analysis of optical systems with many apertures. Such a class of systems is well represented by light field (or plenoptic) cameras. Their general structure consists of a main objective followed by an ensemble of apertures whose function is to divide the field of view into many partitions. Each aperture defines an optical channel. The partial overlap between adjacent field of view partitions serves to extract depth information from the scene in a similar manner to stereo cameras. The wavefront aberration analysis of this class of systems is primarily based on the definition of an ensemble of base-rays playing the role of reference axis for the various channels. The wavefront error for each optical channel is described with a general power series in the ray coordinates expanded about the inherent base-ray. Finally, different approaches are expounded to calculate and visualize the evolution of the aberration behavior of the various channels of this class of optical systems.

Keywords

Hamiltonian Optics, Nodal Aberration Theory, Asymmetric optical systems, Wavefront error, monochromatic aberrations, Zernike polynomials expansion, Multiple aperture optical systems, Light field camera, Plenoptic camera

Riassunto

Il presente lavoro di tesi si focalizza sulla teoria delle aberrazioni monocromatiche in sistemi ottici asimmetrici. La trattazione inizia da una classe di sistemi ottici composti da superfici sferiche lievemente disallineate la cui descrizione si basa sull'espansione delle aberrazioni del fronte d'onda proposta nell'ambito della Nodal Aberration Theory (NAT). In particolare, viene studiato il comportamento dei polinomi di Zernike in funzione delle coordinate del campo di vista di tali sistemi ottici. In seguito, viene considerata un'altra classe di sistemi ottici asimmetrici composti da superfici ottiche con due piani di simmetria ortogonali. In questo caso, limitando lo studio alle aberrazioni primarie, si dimostra che la dipendenza dei polinomi di Zernike dal campo di vista è ben descritta da polinomi di secondo grado. L'assenza di simmetria induce un effetto perturbativo sulla variazione delle aberrazioni primarie in funzione del campo di vista. Nel dettaglio, l'evoluzione dell'aberrazione di coma in dipendenza del campo di vista è descritta da superfici coniche ellittiche mentre quella dell'astigmatismo è descritta da superfici particolari, da noi definite come "superfici di Cassini generalizzate" in quanto più generali delle superfici di Cassini standard che, al contrario, ben descrivono il comportamento binodale dell'astigmatismo nella classe di sistemi ottici introdotta in precedenza.

Queste considerazioni sono preliminari alla discussione della seconda parte di questo lavoro di tesi il cui obiettivo è analizzare le aberrazioni monocromatiche di un'ulteriore classe di sistemi ottici caratterizzati da molteplici aperture. In tal senso, viene dapprima introdotto un metodo generale di analisi di questa categoria di sistemi nell'ambito dell'Ottica Hamiltoniana. Tale metodo è basato su un'espansione in serie di potenze nelle coordinate dei raggi luminosi che permette di descrivere nel modo più generale il comportamento di sistemi ottici privi di simmetria. Queste note introduttive sono indispensabili per l'analisi di sistemi ottici caratterizzati da molteplici aperture, ben rappresentati dalla telecamera plenottica. La struttura generale di questa classe di sistemi ottici consta di un obiettivo seguito da un insieme di aperture la cui funzione è quella di dividere il campo di vista in corrispettive partizioni. Ogni apertura definisce un canale ottico. La parziale sovrapposizione tra partizioni contigue consente di calcolare la distanza degli oggetti nella profondità di campo della scena inquadrata. L'analisi delle aberrazioni del fronte d'onda in tali sistemi si basa primariamente sulla definizione di un insieme di raggi base rispetto ai quali lo sviluppo in serie di potenze delle coordinate dei raggi luminosi è riferito. Infine, vengono introdotti diversi approcci per calcolare e visualizzare l'evoluzione del comportamento delle aberrazioni monocromatiche in funzione dei parametri geometrici dei diversi canali di un sistema ottico con molteplici aperture.

Parole chiave

Ottica Hamiltoniana, Nodal Aberration Theory, sistemi ottici asimmetrici, errore di fronte d'onda, aberrazioni monocromatiche, polinomi di Zernike, sistemi ottici con molteplici aperture, telecamera plenottica

Contents

Acknowledgements.....	i
Abstract	ii
Keywords	ii
Riassunto	iii
Parole chiave	iii
Contents	iv
List of Figures.....	vii
List of Tables.....	xii
List of Equations	xiii
Introduction.....	1
Chapter 1 Geometrical Optics.....	4
1.1 Finite raytracing.....	7
1.1.1 Refraction.....	7
1.1.2 Ray transfer.....	9
1.2 Wavefront aberration function	11
1.3 Perturbed optical systems	15
Chapter 2 Scalar analytical expressions for the field dependence of Zernike polynomials in asymmetric optical systems with circular symmetric surfaces ..	20
2.1 Introduction	20
2.2 Full-field wavefront aberration function in asymmetric optical systems with tilted and decentered circular symmetric surfaces	21
2.3 Full-field dependence of Zernike polynomials in asymmetric systems..	27
2.4 Validation of results	30
2.5 Digression on the origin of the binodal behavior of primary astigmatism magnitude	37
2.6 Full-field behavior of higher-order Zernike terms	40
2.7 Conclusions	44
Chapter 3 Full-field dependence of primary aberrations in perturbed double- plane symmetric systems with a circular pupil.....	47
3.1 Introduction	47
3.2 Full-field dependence of Zernike polynomials.....	48

3.3	Full-field primary aberrations	53
3.3.1	Full-field behavior of spherical aberration	54
3.3.2	Full-field behavior of Primary Coma	55
3.3.3	Full-field behavior of Primary Astigmatism	57
3.3.4	Full-field behavior of Field Curvature	61
3.4	Verification of results: an example with two aspheric lenses.....	63
3.4.1	Primary coma	65
3.4.2	Primary Astigmatism.....	66
3.4.3	Field curvature	68
3.5	Conclusions	69
Chapter 4 Asymmetric systems with decentred spherical surfaces in the framework of Hamiltonian Optics: a comparison with Nodal Aberration Theory		
	72
4.1	Introduction	72
4.2	Parabasal optics of systems without symmetries.....	73
4.3	General power series description of the full field wavefront aberration function	75
4.4	Comparison with the NAT wavefront aberration function	76
4.5	Study of the limits of applicability of NAT	79
4.6	Conclusions	83
Chapter 5 Wavefront aberration behavior in multiple aperture systems		85
5.1	Optical systems with multiple apertures.....	85
5.2	Parabasal behavior.....	86
5.3	Full-field wavefront aberration function in optical systems with many apertures.....	88
5.4	Wavefront aberration function expanded with a general power series	90
5.4.1	Wavefront aberration function expanded with full field-dependent Zernike polynomials	92
5.5	Application: aberrations behavior in a plenoptic camera.....	93
5.5.1	Transverse ray aberration display.....	96
5.5.2	Wavefront aberration coefficients display	99
5.5.3	Zernike polynomials expansion	105
5.6	Conclusions	110
Chapter 6 Conclusions.....		112

Bibliography	115
Curriculum Vitae	119

List of Figures

Figure 1.1: Principle of stationary action.....	5
Figure 1.2: Comparison of the optical distances between neighboring points.....	5
Figure 1.3: Illustration of refraction $n_2 > n_1$. I_1 and I_2 are the angles of incidence and refraction at the interface. The coordinate system is right-handed with the x axis oriented into the paper.	7
Figure 1.4: Ray transfer from P_1 to P_2	9
Figure 1.5: Illustration of wavefront and transverse ray aberrations.	13
Figure 1.6: Change of wavefront error due to a shift of the center of the reference sphere while keeping its vertex fixed.	16
Figure 1.7: Simple system with one single refractive surface S_1 studied to derive the effect of the change of the radius of the reference sphere on the aberration associated with a light ray. The distances are in mm. The aperture stop is placed on the surface S_2 that is the reference sphere whose distance t from S_1 is varied in the interval $0mm, 15mm$. In red and in green are shown the two limiting cases $S_2 = S_2'$ for $t = 0mm$ and $S_2 = S_2''$ for $t = 15mm$. In correspondence of these two positions the radius of curvature of the reference sphere is $R_2 = R_2'$ and $R_2 = R_2''$ respectively. Thus, the reference spheres S_2' and S_2'' are concentric.....	17
Figure 1.8: The variation of the wavefront aberration due to a change of radius of the reference sphere is a quadratic function of the angular ray aberration.	18
Figure 2.1: Reference system.	24
Figure 2.2: 3D layout and reference system of the optical system under study (the dimensions are in mm). The lenses are indicated respectively with L1 and L2. In red and green are highlighted the local coordinate systems of the curved surfaces in both plano-convex spherical lenses.	30
Figure 2.3: Field displacement vectors in the FOV plane.	32
Figure 2.4: Surface fitting of the field dependence of the first Zernike coefficient ($m = 0, n = 0$).....	33
Figure 2.5: Surface fitting of the field dependence of the second Zernike coefficient ($m = 1, n = 1$).....	33
Figure 2.6: Surface fitting of the field dependence of the third Zernike coefficient ($m = -1, n = 1$).....	33
Figure 2.7: Surface fitting of the field dependence of the fourth Zernike coefficient ($m = 0, n = 2$).....	34
Figure 2.8: Surface fitting of the field dependence of the fifth Zernike coefficient ($m = 2, n = 2$).....	35
Figure 2.9: Surface fitting of the field dependence of the sixth Zernike coefficient ($m = -2, n = 2$).....	35
Figure 2.10: Surface fitting of the field dependence of the seventh Zernike coefficient ($m = 1, n = 3$).....	36
Figure 2.11: Surface fitting of the field dependence of the eighth Zernike coefficient ($m = -1, n = 3$).....	36

Figure 2.12: Surface fitting of the field dependence of the ninth Zernike coefficient ($m = 0, n = 4$).	37
Figure 2.13: a) Tangential and sagittal image surfaces for the circular symmetric version of the optical system analyzed in section 2.4. The surface plot is cut along the tangential plane yz to emphasize that, along the y direction, the curvature of the tangential image surface is larger than that of the sagittal image surface. The height of the surfaces is measured in mm from the paraxial image surface. b) The top view emphasizes the coincidence of these two image surfaces at the origin of the FOV.	38
Figure 2.14: a) Tangential and sagittal image surfaces for the asymmetric optical system analyzed in the previous section. b) Top view of the tangential and sagittal image surfaces.	39
Figure 2.15: Contour plot of the full field-dependent magnitude of primary astigmatism overlapped to the tangential and sagittal focal surfaces. The black dots indicate the positions of the nodes of primary astigmatism.	39
Figure 2.16: Surface fitting of the field dependence of the tenth Zernike coefficient ($m = 3, n = 3$).	42
Figure 2.17: Surface fitting of the field dependence of the eleventh Zernike coefficient ($m = -3, n = 3$).	43
Figure 2.18: Surface fitting of the field dependence of the twelfth Zernike coefficient ($m = 2, n = 4$).	43
Figure 2.19: Surface fitting of the field dependence of the thirteenth Zernike coefficient ($m = -2, n = 4$).	43
Figure 2.20: Surface fitting of the field dependence of the fourteenth Zernike coefficient ($m = 1, n = 5$).	44
Figure 2.21: Surface fitting of the field dependence of the fifteenth Zernike coefficient ($m = -1, n = 5$).	44
Figure 3.1: a) Conic-shaped surface describing the full field magnitude of primary coma. b) Elliptical level curves of full-field primary coma magnitude.	56
Figure 3.2: Full-field phase of primary coma.	57
Figure 3.3: a) Cassini surface describing the full field behavior of primary astigmatism magnitude in NAT systems. b) Cassini ovals representing the level curves of Cassini surfaces. The lemniscate is emphasized with a bold dash-dot line. c) Generalized Cassini surfaces describing the full field behavior of astigmatism magnitude in perturbed double-plane symmetric systems. d) Rational bicircular quartic curves representing the level curves of generalized Cassini surfaces. In all figures, ax, ay and bx, by denote the coordinates of the nodes of primary astigmatism in the image plane.	59
Figure 3.4: Origin of the binodal behavior of primary astigmatism magnitude. The surfaces z_1 and z_2 (along with their zero level curves) are in red and green respectively. The intersection points define the nodes of primary astigmatism magnitude.	60
Figure 3.5: Full-field behavior of primary astigmatism phase.	61
Figure 3.6: a) Displaced and rotated double-symmetric surface describing the full field behavior of field curvature. b) The level curves are ellipses.	63

Figure 3.7: Layout and reference system of the optical system (the dimensions are in mm). The tangential and sagittal sections of the biconic surface are shown respectively in red and green. Two ray bundles are shown: in blue for the field point $Hx = 0, Hy = 0$ and in orange for the field point $Hx = 0.707, Hy = 0.707$	64
Figure 3.8: Conic-shaped surface fit to full-field primary coma magnitude.	65
Figure 3.9: Contour plot of the conic-shaped full-field primary coma magnitude data.	66
Figure 3.10: Surface model fitting the primary coma phase data.	66
Figure 3.11: Generalized Cassini surface fit to full-field primary astigmatism magnitude of perturbed double-plane symmetric systems.	67
Figure 3.12: Rational bicircular quartic curves define the contour plot of the generalized Cassini surface fitting the full field primary astigmatism magnitude data.	67
Figure 3.13: a) Surface model fitting the primary astigmatism phase. b) Contour plot of the surface model for primary astigmatism phase. The conic sections whose intersections locate the nodes of primary astigmatism are shown in red and green.	68
Figure 3.14: Fitting of full-field behavior of field curvature with the surface model in Equation 3.26.	68
Figure 3.15: Full-field behavior of field curvature: contour plot of the simulated data.	69
Figure 4.1: Example of an asymmetric optical system. The local reference systems at the exit pupil plane and the image plane are shown. They are located along the base-ray, in red. A wavefront surface is shown in yellow at the exit pupil plane and the reference sphere is shown in red. In green, several displaced optical surfaces are shown. This system is displaced from the mechanical axis (m. a.). ..	74
Figure 4.2: a) Simple optical system with two spherical lenses $L1$ and $L2$. While $L1$ is fixed, $L2$ is allowed to be decentered along the x and y directions of the global reference system xyz . In red and blue are indicated two configurations characterized by different transversal displacements $Dx2, Dy2$ of the second lens $L2$. b) Front view.	80
Figure 4.3: Variation of the spherical aberration coefficient $w0031$ with respect to the decentering parameters of the second lens constituting the optical system in Figure 4.2.....	81
Figure 4.4: Variation of the coma aberration coefficients $w1021$ and $w1003$ with respect to the decentering parameters of the second lens $Dx2$ and $Dy2$	82
Figure 4.5: Variation of astigmatism and field curvature aberration coefficients $w0211$ and $w1102$ with respect to the decentering parameters of the second lens $Dx2$ and $Dy2$	82
Figure 5.1: Multiple aperture system with 3x3 channels. The central channel is the only one with circular symmetry and its base-ray coincides with its axis of symmetry that, in turn, corresponds to the mechanical axis of the system as a whole.	86
Figure 5.2: Multi-aperture system with the respective base-rays (red), FOV partitions (in yellow it is shown the domain sampled by a rectangular grid of points), exit pupil planes (in red it is shown the domain sampled by a rectangular grid of points) and reference spheres centered at the intersection points between the base-rays and the sub-FOV. The reference axis of the exit pupil and sub-FOV	

planes are indicated only for the channel c . The parameters denoting the transversal displacement of the pupils from the mechanical axis are Dx and Dy	88
Figure 5.3: It is shown the overlap between the multiple aperture plane $DxDy$ and the image-side FOV plane $hxhy$. Each channel corresponds to a partition of the FOV whose local coordinates are the field variables Hx and Hy (ray coordinates). It is highlighted the correspondence between channels located in different quadrants of the FOV plane.	89
Figure 5.4: Layout of an example of a plenoptic camera: side view.	94
Figure 5.5: Layout of an example of a plenoptic camera: isometric view.	94
Figure 5.6: The object-side FOV is divided into multiple partitions, each one related to a particular decentered pupil in the multi-aperture plane. Adjacent partitions are partially overlapped; this means that the common area shared between them is “seen” from two different points of observation.	95
Figure 5.7: Front view of the system under study: for clarity, only 5 channels are shown along with their respective base-rays.	96
Figure 5.8: Multiple aperture (and pupil) plane with decentering parameters Dx and Dy overlapped to the corresponding sub-areas of the image-side FOV where transverse ray aberrations are displayed as parametric curves over grids of 11x11 field points. This grid of points has been arbitrarily chosen for visualization purposes and it is different from the 14x14 sampling grid mentioned above.	97
Figure 5.9: The parametric curves describing transverse ray aberrations in different field points in the channels $c2$, $c3$ and $c4$ can be obtained from the respective curves in the channel $c1$ through reflections across the axis and point reflection operations.	98
Figure 5.10: FOV partitions in the first quadrant. Grids of 11x11 parametric curves are displayed.	99
Figure 5.11: Wavefront aberration coefficients related to primary spherical aberration displayed as functions of the decentering parameters of the different channels Dx and Dy . Each blue dot is related to a specific channel.	101
Figure 5.12: Wavefront aberration coefficients related to primary coma.	102
Figure 5.13: Wavefront aberration coefficients related to field curvature and primary astigmatism.	103
Figure 5.14: Wavefront aberration coefficients related to distortion.	104
Figure 5.15: Field dependence of spherical aberration calculated from the coefficient of the Zernike term $n = 4, m = 0$ in the Fringe indexing scheme (Equation 5.5). The subplots' titles indicate the decentering parameters (in mm) of the inherent branch of the camera under investigation. The subplots' domain corresponds to the FOV-partitions, therefore the values are in mm.	106
Figure 5.16: Field dependence of coma aberration calculated with the coefficients of the Zernike terms $n = 3, m = 1$ and $n = 3, m = -1$ in the Fringe indexing scheme according to the second term in Equation 5.5.	107
Figure 5.17: Field dependence of primary astigmatism calculated with the coefficients of the Zernike terms $n = 2, m = 2$ and $n = 2, m = -2$ in the Fringe indexing scheme according to the fourth term in Equation 5.5. Because of the large variation of astigmatism in the various branches of the plenoptic camera under	

study, two different colormaps and colorbars are used for the most external channels $Dx = \pm 5mm$, $Dy = \pm 5mm$ and for the remaining channels $Dx \neq \pm 5mm$, $Dy \neq \pm 5mm$ respectively.....	108
Figure 5.18: Contour plot of the full field-dependent magnitude of primary astigmatism for the 5x5 channels of the system under study.	109
Figure 5.19: Field dependence of field curvature magnitude calculated with the coefficients of the Zernike polynomials reported in the third term of Equation 5.5.	110

List of Tables

Table 2.1: Simulation parameters	31
Table 2.2: Seidel coefficients and displacement vectors.....	31
Table 3.1: Simulation parameters of the optical system under study.....	64
Table 4.1: Spherical Aberration terms.....	77
Table 4.2: Coma Aberration terms	77
Table 4.3: Astigmatism and field curvature terms.....	78
Table 4.4: Distortion terms	78
Table 4.5: Simulation parameters	80
Table 5.1: System parameters.....	94

List of Equations

Equation 1.1.....	4
Equation 1.2.....	5
Equation 1.3.....	6
Equation 1.4.....	6
Equation 1.5.....	6
Equation 1.6.....	6
Equation 1.7.....	8
Equation 1.8.....	8
Equation 1.9.....	8
Equation 1.10.....	8
Equation 1.11.....	8
Equation 1.12.....	8
Equation 1.13.....	9
Equation 1.14.....	9
Equation 1.15.....	9
Equation 1.16.....	10
Equation 1.17.....	10
Equation 1.18.....	10
Equation 1.19.....	10
Equation 1.20.....	10
Equation 1.21.....	10
Equation 1.22.....	11
Equation 1.23.....	11
Equation 1.24.....	11
Equation 1.25.....	11
Equation 1.26.....	12
Equation 1.27.....	12
Equation 1.28.....	13
Equation 1.29.....	13
Equation 1.30.....	14
Equation 1.31.....	14
Equation 1.32.....	14
Equation 1.33.....	14
Equation 1.34.....	15
Equation 1.35.....	15
Equation 1.36.....	15
Equation 1.37.....	16
Equation 1.38.....	17
Equation 2.1.....	21
Equation 2.2.....	22
Equation 2.3.....	22
Equation 2.4.....	23
Equation 2.5.....	24

Equation 2.6.....	24
Equation 2.7.....	25
Equation 2.8.....	25
Equation 2.9.....	25
Equation 2.10.....	25
Equation 2.11.....	26
Equation 2.12.....	26
Equation 2.13.....	26
Equation 2.14.....	27
Equation 2.15.....	27
Equation 2.16.....	28
Equation 2.17.....	28
Equation 2.18.....	29
Equation 2.19.....	40
Equation 2.20.....	41
Equation 2.21.....	41
Equation 2.22.....	41
Equation 2.23.....	42
Equation 3.1.....	49
Equation 3.2.....	50
Equation 3.3.....	51
Equation 3.4.....	51
Equation 3.5.....	51
Equation 3.6.....	51
Equation 3.7.....	52
Equation 3.8.....	52
Equation 3.9.....	52
Equation 3.10.....	52
Equation 3.11.....	52
Equation 3.12.....	52
Equation 3.13.....	53
Equation 3.14.....	54
Equation 3.15.....	54
Equation 3.16.....	55
Equation 3.17.....	55
Equation 3.18.....	56
Equation 3.19.....	56
Equation 3.20.....	57
Equation 3.21.....	58
Equation 3.22.....	58
Equation 3.23.....	58
Equation 3.24.....	60
Equation 3.25.....	61
Equation 3.26.....	61
Equation 3.27.....	62
Equation 3.28.....	62

Equation 3.29.....67

Equation 4.1.....75

Equation 4.2.....75

Equation 5.1.....90

Equation 5.2.....91

Equation 5.3.....91

Equation 5.4.....92

Equation 5.5.....93

Introduction

The work presented in this thesis is the result of a collaboration between two institutions, Datalogic IP Tech Srl (DLIPTECH) and École Polytechnique Fédérale de Lausanne (EPFL), located respectively in Bologna (Italy) and in Lausanne (Switzerland). This collaboration occurred in the framework of the NOLOSS project (Lossless Photon management – Optical design for manufacture at different length scales) and received funding from a MSCA-ITN-EID (Marie-Sklodowska Curie Action - Innovative Training Network - European Industrial Doctorate). In this particular case, the NOLOSS project has promoted the collaboration between industry (DLIPTECH) and academia (EPFL) in the field of system optics with the following work theme: “Non-conventional multi-channel camera for Auto-ID applications”. This work has been carried out by myself under the supervision of Toralf Scharf (EPFL) and Kurt Vonmetz (DLIPTECH). I have been hired first by EPFL from 15.01.2017 to 31.08.2017, then by DLIPTECH from 01.09.2017 to 31.08.2019 and finally again by EPFL from 01.09.2019 to 31.12.2020. The main goal of this Ph.D. work has been the theoretical and experimental study and evaluation of a plenoptic camera in Galilean configuration for automatic identification applications. The experimental investigation resulted in the realization of a prototype of a plenoptic camera. This thesis work presents only the result of the theoretical investigation concerning the wavefront aberration behavior of this class of multi-channel optical systems.

This thesis investigates the behavior of monochromatic wavefront aberrations in optical systems characterized by the absence of symmetry with the ultimate goal to study the wavefront aberrations in multiple aperture optical systems like light field (or plenoptic) cameras. This class of optical systems is constituted by an ensemble of individual optical channels whose aberration properties are strongly affected by the absence of any specific symmetry due to the transversal displacement of the optical surfaces with respect to the mechanical axis of the system.

The first chapter of this thesis introduces the main concepts of geometrical optics in the framework of Hamiltonian Optics. The fundamental concepts of the characteristic function, wavefront aberration function and transverse ray aberration function are introduced along with the method of finite raytracing used to calculate them. In the second chapter, the full field dependence of Zernike Fringe coefficients in asymmetric optical systems with tilted and decentered spherical surfaces is analyzed. The wavefront aberration function is decomposed in terms of full field-dependent Zernike Fringe polynomials whose locations of the centers of symmetry in the field of view are calculated numerically and are compared with the centers of symmetry calculated analytically using the formulas proposed in the framework of the Nodal Aberration Theory. It is observed a mismatch between numerical and analytical calculations attributed to the presence of secondary aberrations in the optical system under study.

The third chapter discusses another class of optical systems defined as perturbed double-plane symmetric systems. These are essentially asymmetric optical systems characterized by weak geometrical perturbations (such as tilts and displacements) of the constituting optical surfaces that can be rotationally symmetric and double-plane symmetric. The aberrations (up to the 4th order) for this class of optical systems are decomposed again in terms of Zernike Fringe polynomials whose field dependence is more general than that

found in the second chapter. It is shown a method to calculate the full field dependence of primary aberrations sampling the Zernike Fringe coefficients in correspondence of few points in the field of view.

In the fourth chapter, asymmetric optical systems with more severe displacements of the optical components are considered. A comparison is carried out between two different representations of the associated full field wavefront aberration function: the first is based on the Nodal Aberration Theory (as proposed by Thompson), while the second is based on decomposing the aberration function with a general power series in the ray coordinates. This latter representation is best suited to the analysis of this class of systems due to its greater generality.

In the fifth chapter, the considerations done in the previous chapter are applied to the analysis and representation of the wavefront aberration function of the several channels constituting multiple aperture optical systems like plenoptic cameras. More in detail, general power series in the ray coordinates are used to decompose the wavefront function in numerous aberration terms deriving from the general absence of any specific symmetry in the channels of a plenoptic camera. The aberration coefficients are calculated fitting multivariate polynomials in the ray coordinates to raytracing data obtained for each channel. A method to visualize the evolution of such aberration coefficients as a function of the displacement parameters of the several channels is described. Finally, it is shown another method to visualize the behavior of the aberrations depending on the specific optical channel of interest using the Full Field Display available in programs for optical design such as Zemax OpticStudio.

Chapter 1 Geometrical Optics

The application of the Hamiltonian method of computation to problems in geometrical optics led to the birth of the theory of Hamiltonian Optics. This theory is based on the fundamental concept of characteristic function corresponding to the principal function introduced by Hamilton in the field of dynamics. Although the relevance of the characteristic function is mainly theoretical, it is nevertheless a fundamental building block in the construction of many concepts related to geometrical optics. The primary and fundamental assumption on which geometrical optics is constructed consists of neglecting the finiteness of the wavelength of light, $\lambda_0 \rightarrow 0$. The main implication of this assumption is that the phenomena of propagation of light in a medium, of refraction and reflection at the boundary between two media can be studied using geometrical constructions known as light rays. These are nothing but curves providing a geometrical description of the physical path followed by electromagnetic waves during their propagation and interaction with a medium when the phenomenon of diffraction is negligible.

In the framework of Hamiltonian Optics, the characteristic function $V(x, y, z, x', y', z')$ is introduced to describe quantitatively the propagation of light rays between two distinct points in space indicated as $P(x, y, z)$ and $P'(x', y', z')$ where (x, y, z) and (x', y', z') denote the coordinates of such points in two different reference systems. Generally, in imaging applications these two distinct points are qualified according to the space where they are situated, the object and image space respectively, and the role of the characteristic function (in reality being a functional) is to assign a measure to the propagation of light rays between P and P' . This measure is nothing but the optical path length traveled by a light ray from P to P' . More in detail, the characteristic function V accounts not only for the path length associated to the propagation but also for the phenomena of refraction and reflection occurring during the propagation of light rays between P and P' . In mathematical terms, the characteristic function V maps a set of six spatial coordinates (related to the initial and final positions of a light ray) to a scalar factor corresponding exactly to the optical path length traveled by a light ray during its propagation in a medium with refracting index $n = n(\vec{s})$. This latter, in general, depends on the position vector in the medium, $\vec{s} = \hat{x}\vec{i} + \hat{y}\vec{j} + \hat{z}\vec{k}$. This situation is described by the following equation

$$V(x, y, z, x', y', z') = \int_P^{P'} n(\vec{s}) \delta \vec{s}$$

Equation 1.1

where $\delta \vec{s} = (\delta x, \delta y, \delta z)$ expresses the length of the infinitesimal element of arc of the curve joining the points P and P' . Equation 1.1 states the equivalence between the characteristic function and the optical distance traveled by a light ray from the point P to the point P' . The expression in Equation 1.1 entails that the rays connecting the points P and P' are geometrical curves but it provides no information about the modality in which they propagate through space. In other words, Equation 1.1 does not explain how the geometrical construction of these curves occurs. To this end, it is in general assumed that the propagation of light rays is governed by the principle of stationary action. Such principle states that the trajectories of light rays are determined by the condition that the optical length associated with the propagation between two points P and P' is stationary. Stated another way, the action (in this context represented by the characteristic function) satisfies a

variational principle. This assumption is immediately expressed by the following Equation 1.2 which imposes a precise condition to be met for the geometrical construction of a ray connecting two points.

$$\delta V = 0$$

Equation 1.2

The following Figure 1.1 exemplifies this situation. In particular, it shows that, among different potential curves connecting two points, the one satisfying Equation 1.2 describes the actual path traveled by a light ray.

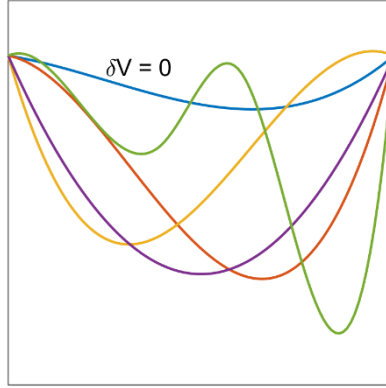


Figure 1.1: Principle of stationary action.

The basic equations of Hamiltonian Optics can be derived from the application of the variational principle (Equation 1.2) to the point characteristics functions (Equation 1.1) joining two couples of neighbor points. This is illustrated in Figure 1.2.

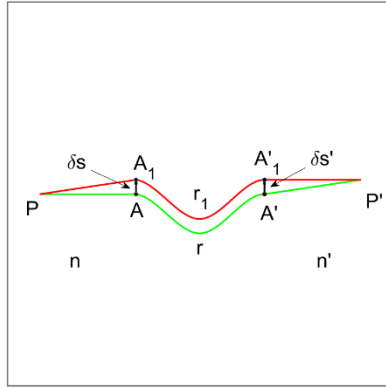


Figure 1.2: Comparison of the optical distances between neighboring points.

More in details, Equation 1.1 associates a value, the optical path, to a pair of arbitrary points, A and A' in Figure 1.2. One expects that, to the first order, another couple of points characterized by small displacements $\delta\vec{s} = (\delta x, \delta y, \delta z)$ and $\delta\vec{s}' = (\delta x', \delta y', \delta z')$ from the original pairs of points (denoted as A_1 and A'_1 in Figure 1.2) exhibit the same path length between their neighbors A and A' . This situation is expressed mathematically by Equation 1.3

$$\begin{aligned}\delta V &= V(A_1, A'_1) - V(A, A') = \frac{\partial V}{\partial x'} \delta x' + \frac{\partial V}{\partial y'} \delta y' + \frac{\partial V}{\partial z'} \delta z' + \frac{\partial V}{\partial x} \delta x + \frac{\partial V}{\partial y} \delta y + \frac{\partial V}{\partial z} \delta z \\ &= n'(\alpha' \delta x' + \beta' \delta y' + \gamma' \delta z') - n(\alpha \delta x + \beta \delta y + \gamma \delta z)\end{aligned}$$

Equation 1.3

where $\alpha, \beta, \gamma, \alpha', \beta'$ and γ' are the direction cosines associated with the tangents of the light rays at the points A and A' . Equating corresponding terms in Equation 1.3, the basic equations of Hamiltonian Optics are obtained

$$\begin{cases} n'\alpha' = \frac{\partial V}{\partial x'} & n'\beta' = \frac{\partial V}{\partial y'} & n'\gamma' = \frac{\partial V}{\partial z'} \\ n\alpha = -\frac{\partial V}{\partial x} & n\beta = -\frac{\partial V}{\partial y} & n\gamma = -\frac{\partial V}{\partial z} \end{cases}$$

Equation 1.4

The relevance of Equation 1.4 resides in the fact that, if the characteristic function V is known, then the direction cosines of a light ray passing through the points A and A' can be calculated. In general, when Equation 1.4 are applied to the propagation of light rays in an optical system, they can be simplified further considering that the object and image points are conjugate and in particular they belong respectively to the object and image base-planes characterized by the equations $z = 0$ and $z' = 0$. This further simplification implies that Equation 1.4 are reduced to four equations related to the x and y coordinates in the object and image spaces. Besides the point characteristic function, Hamilton introduced also the angle and mixed characteristic functions (not described in this context).

The relations between a particular characteristic function (being the point, angle or mixed characteristic function) and the ray coordinates can be expressed with the following general equation

$$p_i = \frac{\partial V}{\partial q_i} \quad (i = 1, \dots, 4)$$

Equation 1.5

where q_i are the ray-coordinates and p_i are their conjugates, represented by x, y, x', y' and $\alpha, \beta, \alpha', \beta'$ respectively in Equation 1.4. From Equation 1.5, the characteristic function V can be reformulated in differential form as follows

$$dV = \sum_{i=1}^4 \frac{\partial V}{\partial q_i} dq_i = \sum_{i=1}^4 p_i dq_i$$

Equation 1.6

In the previous considerations, it has been assumed that $n(\vec{s})$, the refractive index of the medium in which light rays travel, depends locally on the spatial coordinates identified by the vector \vec{s} . In the vast majority of cases of practical interest, the propagation of light rays is assumed to occur in an isotropic and homogeneous medium such that $n(\vec{s}) = n_0$. This assumption is of fundamental importance since its main implication is that light rays propagate along straight lines as shown in the following section where finite raytracing is discussed.

1.1 Finite raytracing

Raytracing consists of the iteration of two operations describing the propagation of light rays in homogeneous media characterized by different refractive indices. In particular, these operations are refraction (or reflection) at the intersection point with a surface and transfer between two surfaces. The first operation occurs as a consequence of the change of refractive index encountered by a light ray when passing from a medium with an index of refraction n_1 to another medium with an index of refraction n_2 (in the case of reflection it is sufficient to consider $n_2 = -n_1$). The second operation describes the physical propagation of a light ray between two optical surfaces separated by the same medium with a particular index of refraction.

1.1.1 Refraction

The application of the variational principle to the Hamilton point characteristic function can be used to derive the law of refraction. Substantially, this latter emerges from the coplanarity between three unit vectors associated with the propagation of a light ray. The first two unit vectors indicate respectively the directions of incidence and refraction of the light ray, the third unit vector is in the direction normal to the optical surface at the point of incidence with the light ray. In turn, the unit vectors (with their components) along the incident and refracted rays are denoted with $\vec{r}_1 = (\alpha_1, \beta_1, \gamma_1)$ and $\vec{r}_2 = (\alpha_2, \beta_2, \gamma_2)$ and the unit vector along the normal of the refracting surface at the point of incidence is indicated with $\vec{g} = (L, M, N)$. The following Figure 1.3 shows an example of refraction of a light ray at the interface between two media. The light ray propagates along straight lines considering that the media are homogeneous with indices of refraction $n_1(\vec{s}) = n_1$ and $n_2(\vec{s}) = n_2$.

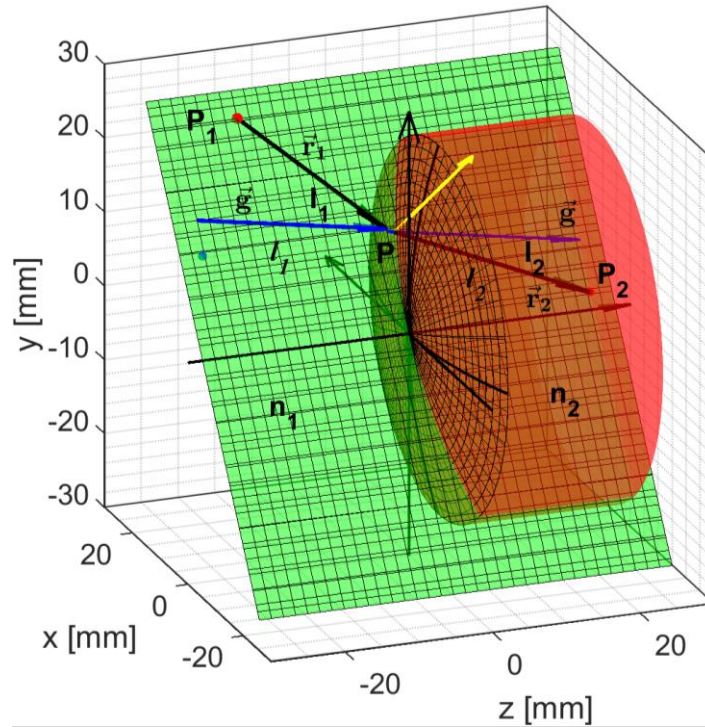


Figure 1.3: Illustration of refraction ($n_2 > n_1$). I_1 and I_2 are the angles of incidence and refraction at the interface. The coordinate system is right-handed with the x axis oriented into the paper.

The measure of the optical distance traveled by the light ray from the point P_1 to the point P_2 is quantified by the point characteristic function $V = V(P_1, P_2)$. This distance is the sum of two straight lines associated with the propagation of the light ray in two media with different refractive indices. Such straight lines are represented by the segments $\overline{P_1P} = l_1$ and $\overline{PP_2} = l_2$ multiplied by the refractive indices of the respective medium. This translates into the following measure of the optical distance, $V(P_1PP_2) = n_1 l_1 + n_2 l_2$. The application of the variational principle to this optical distance requires that the position of the point P is allowed to be slightly modified of a quantity $\delta \vec{l}$ while granting that it lies on the optical surface keeping fixed the geometrical distance $l_1 + l_2$. This results in the following equation

$$\delta V = (n_2 \vec{r}_2 - n_1 \vec{r}_1) \cdot \delta \vec{l} = [(n_2 \vec{r}_2 - n_1 \vec{r}_1) \times \vec{g}] \cdot \delta \vec{a} = 0$$

Equation 1.7

where one considers that the variation of the geometrical distance must be perpendicular to \vec{g} , that is $\delta \vec{l} = \vec{g} \times \delta \vec{a}$ and $\delta \vec{a}$ is an arbitrary vector. Equation 1.7 is satisfied by putting the term in the square bracket equal to zero. This gives Equation 1.8 which implies that the three aforementioned unit vectors \vec{r}_1 , \vec{r}_2 and \vec{g} are located in the same plane, known as plane of incidence, represented in green in Figure 1.3.

$$n_2 \vec{r}_2 - n_1 \vec{r}_1 = \sigma \vec{g}$$

Equation 1.8

In Equation 1.8, σ is a scalar term accounting for the different magnitudes of the terms on the left and right sides of the equation. The explicit value of σ is obtained doing a scalar multiplication of both sides of Equation 1.8 with \vec{g} , the vector normal to the optical surface at the point of incidence of the light ray. Doing this scalar multiplication, it is obtained the following expression for the quantity σ

$$\sigma = n_2 \cos I_2 - n_1 \cos I_1$$

Equation 1.9

and the final expression of the law of refraction can be formulated as follows

$$n_2 \vec{r}_2 - n_1 \vec{r}_1 = (n_2 \cos I_2 - n_1 \cos I_1) \vec{g}$$

Equation 1.10

The expression in Equation 1.10 can be alternatively obtained starting from another vector formula, most frequently encountered in practice, shown in Equation 1.11

$$n_2 (\vec{r}_2 \times \vec{g}) = n_1 (\vec{r}_1 \times \vec{g})$$

Equation 1.11

whose scalar version is

$$n_2 \sin I_2 = n_1 \sin I_1$$

Equation 1.12

Calculating the vector product of the left and right side of Equation 1.11 with the unit vector \vec{g} , one obtains

$$n_2 \vec{g} \times (\vec{r}_2 \times \vec{g}) = n_1 \vec{g} \times (\vec{r}_1 \times \vec{g})$$

Equation 1.13

and using the triple product expansion $\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$, Equation 1.13 transforms into the following Equation 1.14

$$n_2(\vec{r}_2 - (\vec{g} \cdot \vec{r}_2)\vec{g}) = n_1(\vec{r}_1 - (\vec{g} \cdot \vec{r}_1)\vec{g})$$

Equation 1.14

whose components are

$$\begin{cases} n_2 \alpha_2 - n_1 \alpha_1 = \sigma L \\ n_2 \beta_2 - n_1 \beta_1 = \sigma M \\ n_2 \gamma_2 - n_1 \gamma_1 = \sigma N \end{cases}$$

Equation 1.15

where $\sigma = n_2(\vec{g} \cdot \vec{r}_2) - n_1(\vec{g} \cdot \vec{r}_1) = n_2 \cos I_2 - n_1 \cos I_1$ in agreement with Equation 1.9 obtained before.

1.1.2 Ray transfer

The transfer of a light ray between two spherical surfaces is shown, for clarity, in Figure 1.4.

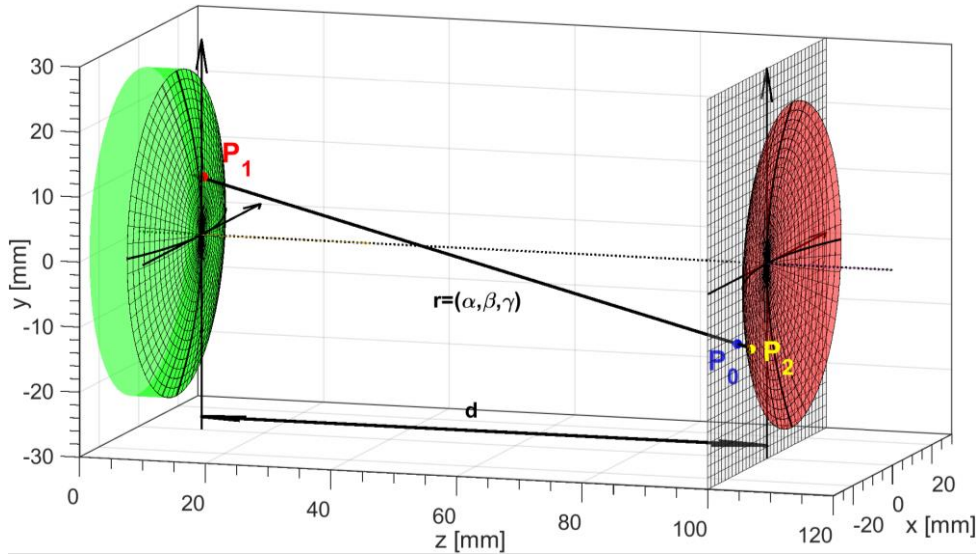


Figure 1.4: Ray transfer from P_1 to P_2 .

In this example, the transfer occurs from the point $P_1 = (x_1, y_1, z_1)$ on the first surface (green), to the point $P_2 = (x_2, y_2, z_2)$ on the second surface (red). The coordinates of the points P_1 and P_2 are referred to the local reference systems of the first and second surface. This second operation of raytracing consists, initially, in retrieving the coordinates of the intersection of the light ray in question with the plane tangent to the second surface at its vertex, namely at the point $P_0 = (x, y, 0)$. Considering as known the unit vector (or equivalently the direction cosines) of the light ray $r = (\alpha, \beta, \gamma)$ and the coordinates of the

intersection with the preceding surface (x_1, y_1, z_1) , the coordinates $(x, y, 0)$ are retrieved as follows

$$\begin{cases} x = x_1 + V(d - z_1) \\ y = y_1 + W(d - z_1) \end{cases}$$

Equation 1.16

where $V = \frac{\alpha}{\gamma}$ and $W = \frac{\beta}{\gamma}$ are the direction tangents of the light ray and d is the distance between the two surfaces in correspondence of their vertices. Afterward, the coordinates of the intersection with the spherical surface are found as follows

$$\begin{cases} x_2 = x + \alpha h \\ y_2 = y + \beta h \\ z_2 = \gamma h \end{cases}$$

Equation 1.17

where h corresponds to the distance between the points P_2 and P_0 , denoting respectively the intersection of the light ray with the second spherical surface and the intersection of the light ray with the plane tangent to this spherical surface at the vertex. The z_2 coordinate in Equation 1.17 coincides with the sag (or height) of the spherical surface in its local coordinates system (with the origin at the vertex). In fact, the point P_2 belongs to the second surface and its coordinates must satisfy the equation of this surface. Given that the sag of a spherical surface with curvature c is calculated as

$$z_2 = \frac{c}{2}(x_2^2 + y_2^2 + z_2^2)$$

Equation 1.18

putting Equation 1.17 into Equation 1.18, the distance h is found to be

$$h = \frac{c(x^2 + y^2)}{N - c(\alpha x + \beta y) + \sqrt{[\gamma - c(\alpha x + \beta y)]^2 - c^2(x^2 + y^2)}}$$

Equation 1.19

Substituting the value of h in Equation 1.17 allows finding the coordinates of the intersection of the light ray with the second spherical surface in the point P_2 .

Equation 1.19 is easily generalized to the case in which the second surface (towards which the ray transfers) is a quadric of revolution. The equation describing the sag of this type of surface is the following

$$z_2 = \frac{c}{2}(x_2^2 + y_2^2 + \varepsilon z_2^2)$$

Equation 1.20

where ε is a parameter accounting for the “degree” of asphericity of the quadric of revolution. In this case, Equation 1.19 turns into the following Equation 1.21

$$h = \frac{c(x^2 + y^2)}{\gamma - c(\alpha x + \beta y) + \sqrt{[\gamma - c(\alpha x + \beta y)]^2 - (1 + k\gamma^2)c^2(x^2 + y^2)}}$$

Equation 1.21

The parameter k is the conic constant of the surface in question and it is calculated as $k = \varepsilon - 1 = -e^2$ where e is another parameter known as eccentricity.

In the above derivations related to the refraction of a light ray with a spherical surface, it has not been mentioned how the direction cosines of the vector normal to the surface, $\vec{g} = (L, M, N)$, are calculated. For this purpose, it is necessary to express the equation of the optical surface in implicit form $S(x, y, z) = 0$ and to expand it as a Taylor series in the point $(x + \delta x, y + \delta y, z + \delta z)$. These operations give the following expression

$$S(x, y, z) + \frac{\partial S}{\partial x} \delta x + \frac{\partial S}{\partial y} \delta y + \frac{\partial S}{\partial z} \delta z + O(2) = 0$$

Equation 1.22

Considering $(\delta x, \delta y, \delta z)$ and $(\frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial z})$ as the components of two vectors, from Equation 1.22 it can be deduced that they are orthogonal and in particular, the latter points along the direction normal to the surface since the former is constrained to be on the surface. Consequently, one obtains the direction cosines of the normal to the surface at the point of intersection with a light ray as follows

$$\vec{g} = (L, M, N) = \frac{-\left(\frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial z}\right)}{\sqrt{\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right]}}$$

Equation 1.23

In the case of a spherical surface, recalling from Equation 1.18 that $S(x, y, z) = \frac{c}{2}(x^2 + y^2 + z^2) - z = 0$, the direction cosines of the normal have the following values

$$\vec{g} = (L, M, N) = (-cx, -cy, 1 - cz)$$

Equation 1.24

and, more in general, for quadric of revolution described by Equation 1.20, the direction cosines of the normal are given by Equation 1.25

$$\vec{g} = (L, M, N) = \frac{(-cx, -cy, 1 - cz)}{\sqrt{1 - 2ckz + c^2k(1 + k)z^2}}$$

Equation 1.25

Once the direction cosines of the unit vector normal to the optical surface are retrieved explicitly according to Equation 1.24-Equation 1.25, their expressions can be replaced into Equation 1.15 in order to find the relations between the direction cosines of the incident and refracted rays.

The points discussed so far are preliminary to the introduction of the fundamental concepts of wavefront aberration and transverse ray aberration functions. These quantities are practically computed by tracing finite rays through an optical system and measuring the associated optical path lengths.

1.2 Wavefront aberration function

Optical design for imaging applications is aimed at developing optical systems with image-forming capabilities whose quality and properties are defined within certain tolerances. In

other words, the design of an optical system is considered satisfactory if the required optical performances are met within defined tolerance intervals. The most common method to describe the optical performances of an image forming system relies on the concept of aberration function that is, essentially, an attempt to quantify the departure from the ideality of an optical system. In the present context, an optical system performs ideally if it forms an image that is a replica of the object, resized on a smaller or larger geometrical scale. This situation can be properly represented with an ideal characteristic function accounting for the description of the only ideal behavior of the optical system. The description of the real behavior of the system is encompassed in the aberration function that takes into account the real capability of the optical system itself to form an image as close as possible to the prescribed behavior. The total characteristic function V of an optical system is described with the following equation

$$V = V_i + W$$

Equation 1.26

where V_i is the ideal characteristic function, while W is the aberration function. The ideal characteristic function V_i accounts for the ideal image-forming capabilities of an optical system. In the general case of an optical system without symmetries, V_i describes the paraxial behavior of such a system. This implies that the ideal characteristic function contains terms up to the second degree in the ray coordinates, namely $V = V_i + O(3)$. Therefore, $W = O(3)$ in the ray coordinates. In the more specific case of a rotationally symmetric system, V_i describes its fundamental behavior as a projective transformation system meaning that the three-dimensional object space is mapped into a similar three-dimensional image space without deformations. Consequently, the geometrical modification taken into account by V_i is the resizing of the object occurring in the image space, that is the lateral magnification. Equivalently, one can say that V_i does not predict any transverse displacement of light rays with respect to their nominal intersection coordinates with the image plane (resulting from only terms of the second degree). On the other hand, potential displacements of the intersection coordinates of light rays with the image plane, resulting from higher-order effects, are taken into account by the aberration function W . This means that the aberration function supplements the description of the imaging behavior of an optical system encompassing the additional deviations occurring in the propagation of light rays within the optical system. In addition to this, W describes the departure from the ideal design of an optical system due to eventual manufacturing defects or assembly errors of the optical components constituting the system itself.

At this point, it is beneficial for a better understanding to remark that the exact dependence of the functions in Equation 1.26 on the ray coordinates is never sought because, in the majority of the cases of practical interest, such expressions are expected to be extremely complicated. For this simple and practical reason, they are approximated with expansions in power series depending on the ray coordinates. The possibility of approximating the characteristic functions with a power series is ensured by the property of regularity of an optical system jointly with a particular coordinate system. This is a consequence of the fact that the property of regularity is necessarily associated with a properly defined reference system for the ray coordinates that makes the power series convergent.

Therefore, the characteristic function V is expressed as follows

$$V = \sum_{n=1}^{\infty} V_n$$

Equation 1.27

The terms V_n are calculated as follows

$$V_n = \sum_{\lambda=0}^n \sum_{\mu=0}^{\lambda} \sum_{\nu=0}^{\mu} v_{n-\lambda, \lambda-\mu, \mu-\nu, \nu} q_1^{n-\lambda} q_2^{\lambda-\mu} q_3^{\mu-\nu} q_4^{\nu}$$

Equation 1.28

where q_1, q_2, q_3 and q_4 constitute a particular set of ray coordinates and $v_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}$ are coefficients weighting the contribution given by the corresponding monomial in the power series of Equation 1.27-Equation 1.28. Since the ideal characteristic function V_i is described by the truncation to the second order of Equation 1.27, it can be denoted with V_2 as follows

$$V_2 = \sum_{k,l=1}^2 v_{kl} q_k q_l$$

Equation 1.29

where v_{kl} are coefficients multiplying any combination of the four ray coordinates up to the second order.

Equations 1.27-1.29 are valid for any optical system without restrictions on the coefficients imposed by symmetry. In the next chapters, this power series representation will be extended to the wavefront aberration function W . In the remainder of this section, several basic concepts related to the aberration function are introduced.

Two alternative approaches are commonly employed to describe the aberrations of an optical system, the first is based on the calculation of the wavefront error at the exit pupil, the second relies on the calculation of the displacement of the ray coordinates at the image plane. The following considerations are aimed at exemplifying more in detail the connection between the definition of the wavefront error and the ray transverse displacement at the image plane. In Figure 1.5 it is shown with a green curve a cross-section of the reference sphere S_0 , centered at the image point $P'_0(H_{x0}, H_{y0})$ that is defined by the intersection of the chief ray r_c with the image plane. A red curve indicates a cross-section of the actual wavefront surface S generated by a congruence of light rays emerging from a point P in the object space (not indicated). A particular light ray r , belonging to the congruence mentioned above and identified by the direction cosines (α, β, γ) , intersects the reference sphere and the wavefront surface respectively in the points $Q_0(\rho_x, \rho_y, \rho_z)$ and Q and lands at the point $P'(H_x, H_y)$ in the image plane (whose equation is $H_z = 0$).

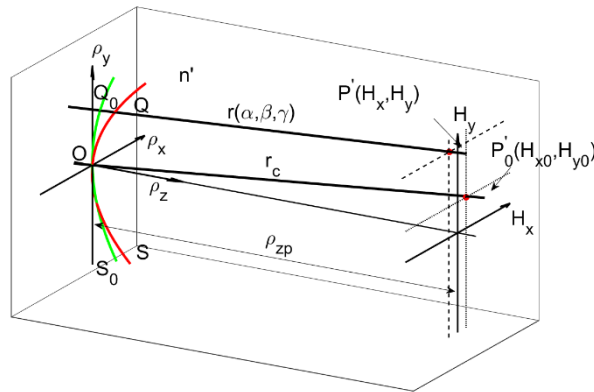


Figure 1.5: Illustration of wavefront and transverse ray aberrations.

The deviation of the light ray $r(\alpha, \beta, \gamma)$ from the chief ray position $P'_0(H_{x0}, H_{y0})$ at the image plane is known as transverse ray aberration. This is described with two parametric equations quantifying the deviation of the light ray r in the H_x and H_y directions. As will be shown later, the transverse ray aberration is linked to the wavefront error measured along the ray r , that is essentially the optical distance between the points Q and Q_0 . These two alternative approaches to describe the aberration behavior of a light ray and, in general, the aberration behavior of an optical system are linked by the characteristic function V . This fundamental point is developed more extensively in the following part.

As already remarked, the wavefront error is the optical path difference between two points defined by the intersections between a specific light ray with the reference sphere and the actual wavefront surface associated with the congruence of light rays emerging from a point in the object space. That is to say, along the ray $r(\alpha, \beta, \gamma)$, the optical path difference is $n'Q_0Q$ and this value is associated with a certain displacement of the ray coordinates (H_x, H_y) with respect to the point of intersection of the chief ray with the image plane. The optical distance from the object point P and the intersection points Q_0 and Q is given by the respective point characteristic functions as described by the following Equation 1.30

$$W(\rho_x, \rho_y) = V(P, Q) - V(P, Q_0) = V(P, O) - V(P, Q_0)$$

Equation 1.30

since the points Q and O are on the same wavefront, they have the same optical path length from the object point P . Differentiating Equation 1.30 with respect to the exit pupil coordinates, one obtains

$$\begin{cases} \frac{\partial W}{\partial \rho_x} = -\frac{\partial V}{\partial \rho_x} - \frac{\partial V}{\partial \rho_z} \frac{\partial \rho_z}{\partial \rho_x} \\ \frac{\partial W}{\partial \rho_y} = -\frac{\partial V}{\partial \rho_y} - \frac{\partial V}{\partial \rho_z} \frac{\partial \rho_z}{\partial \rho_y} \end{cases}$$

Equation 1.31

Recalling the basic equations of Hamiltonian optics (Equation 1.4) and considering that

$$\begin{cases} \frac{\partial \rho_z}{\partial \rho_x} = \frac{\rho_x - H_{x0}}{\rho_{zp} - \rho_z} \\ \frac{\partial \rho_z}{\partial \rho_y} = \frac{\rho_y - H_{y0}}{\rho_{zp} - \rho_z} \end{cases}$$

Equation 1.32

one obtains

$$\begin{cases} \frac{1}{n'} \frac{\partial W}{\partial \rho_x} = -\left[\alpha + \frac{\gamma(\rho_x - H_{x0})}{\rho_{zp} - \rho_z} \right] \\ \frac{1}{n'} \frac{\partial W}{\partial \rho_y} = -\left[\beta + \frac{\gamma(\rho_y - H_{y0})}{\rho_{zp} - \rho_z} \right] \end{cases}$$

Equation 1.33

Recalling that the directions cosines of the ray r are equal to

$$(\alpha, \beta, \gamma) = \frac{(H_{x0} - \rho_x, H_{y0} - \rho_y, \rho_{zp} - \rho_z)}{Q_0 P'}$$

Equation 1.34

one obtains the final relations between the wavefront error and the ray coordinates displacement in the image plane

$$\begin{cases} \delta H_x = -\frac{Q_0 P'}{n'} \frac{\partial W}{\partial \rho_x} \\ \delta H_y = -\frac{Q_0 P'}{n'} \frac{\partial W}{\partial \rho_y} \end{cases}$$

Equation 1.35

In practical calculations, the distance $Q_0 P'$ can be replaced with R , the radius of the reference sphere S_0 . Then, Equation 1.35 turns into

$$\begin{cases} \delta H_x = -\frac{R}{n'} \frac{\partial W}{\partial \rho_x} \\ \delta H_y = -\frac{R}{n'} \frac{\partial W}{\partial \rho_y} \end{cases}$$

Equation 1.36

In the previous considerations, the aberration function has been regarded as dependent on the only coordinates of intersection between the ray and the reference sphere (ρ_x, ρ_y) . In reality, the wavefront aberration function depends on the entire set of ray coordinates, namely $W = W(\rho_x, \rho_y, H_x, H_y)$. The field coordinates (H_x, H_y) are omitted when, as in the present case, the field dependence of the wavefront aberration function is implicitly embedded in the location of the point P in the object space from which the cone of light rays is emitted. The full field dependence of the wavefront aberration function will be treated throughout the rest of the thesis for different classes of optical systems.

1.3 Perturbed optical systems

As highlighted in the considerations above, the study of the wavefront aberration function of an optical system is based on an appropriate definition of the reference sphere. This is defined by specifying the radius of curvature R and the location of its center in the image plane while its vertex intersects the center of the exit pupil plane. Since the full expression of the wavefront aberration function depends not only on the exit pupil coordinates but also on the field coordinates, multiple reference spheres are defined, each one referenced to a specific point in the FOV. In fact, from each field point, it is traced a principal ray whose point of intersection with the image plane defines the center of the respective reference sphere. Therefore, it is desirable to give an adequate definition of the reference sphere to give consistency to the full field expression of the wavefront aberration function. In addition to these considerations, an optical system can be characterized by tilted and decentered surfaces with respect to the mechanical axis. The analysis of the aberration behavior of such optical systems has to take into account the change of the wavefront aberration function due to this type of geometrical perturbations. Since the definition of the wavefront aberration function depends on the definition of the reference sphere, in general, this latter is modified in order to minimize the Root Mean Square wavefront error that is the square root of the variance of the wavefront aberration. In particular, if the center of the reference sphere is slightly shifted from the point $P_0 = (H_{x0}, H_{y0}, 0)$ to the point $P_1 = (H_{x0} + \delta H_{x0}, H_{y0} +$

$\delta H_{y0}, \delta H_{z0}$) while its vertex is kept fixed at the origin of the exit pupil plane, the change of the wavefront aberration is given by the following Equation 1.37

$$\begin{aligned}\delta W &= -\frac{n'(\rho_x \delta H_{x0} + \rho_y \delta H_{y0} + \rho_z \delta H_{z0})}{\alpha'(\rho_x - H_{x0}) + \beta'(\rho_y - H_{y0}) + \gamma'(\rho_z - \rho_{zp})} \\ &= \frac{n'}{R}(\rho_x \delta H_{x,0} + \rho_y \delta H_{y,0} + \rho_z \delta H_{z,0})\end{aligned}$$

Equation 1.37

This situation is illustrated in the following Figure 1.6.

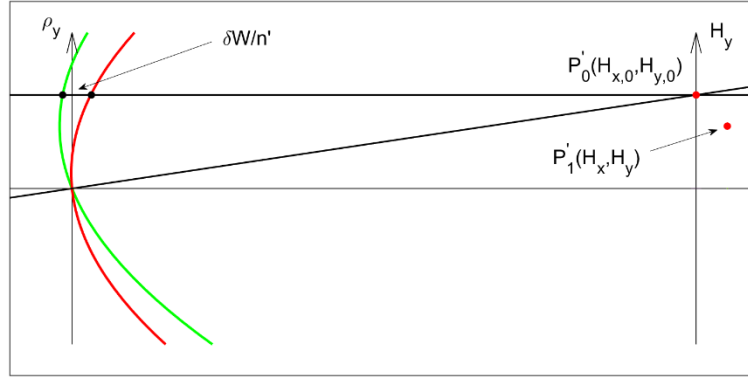


Figure 1.6: Change of wavefront error due to a shift of the center of the reference sphere while keeping its vertex fixed.

If the change of the reference sphere is more complex than a small shift of its center, the effect on the phase error associated with a light ray propagating through the optical system can be more severe. This is the case, for example, if the vertex of the reference sphere is shifted from the origin of the exit pupil plane or if its curvature is changed. These modifications of the reference sphere occur in the actual computation of the wavefront aberration function of an optical system. As a matter of fact, the numerical wavefront function, sampled by a grid of light rays traced through the exit pupil of an optical system, is commonly referenced to a modified reference sphere whose displacement and curvature are adjusted to minimize the variance of the wavefront function itself. These remarks emphasize further that the change of the reference sphere, either in terms of shift or in terms of change of the radius of curvature, has a crucial effect on the calculation of the wavefront aberration function.

Of particular interest is to analyze how the wavefront aberration of a particular ray is affected by the change of the radius of curvature of the reference sphere while keeping its center fixed at the point of intersection of the principal ray with the image plane. Since the aberrations are expressed with polynomials in the ray coordinates truncated at a certain order of interest, this type of investigation allows us to understand specifically at what order the aberrations start to be affected by a change of radius of curvature of the reference sphere. This analysis is done for a simple optical system constituted by a single refractive spherical surface S_1 illustrated in black in the following Figure 1.7 using Zemax OpticStudio. The entrance pupil diameter (ENPD) is equal to 20mm, the effective focal length (EFL) is equal to 96.4mm and the simulation wavelength is $\lambda_0 = 0.55\mu m$. The surface S_1 , made of N-BK7 glass and characterized by a radius of curvature $R_1 = 50mm$, is followed by another surface S_2 where it is set the aperture stop of the system. The exit pupil is also located in correspondence of S_2 and its distance from the image plane, known as exit pupil position (EXPP), is defined to be equal to the radius of the reference sphere, denoted with R_2 . For this reason, such surface S_2 can represent the reference sphere of the system under study provided that its radius of curvature R_2 is set equal to EXPP. The distance t between

S_1 and S_2 is varied in the range $[0\text{mm}, 15\text{mm}]$. The radius of curvature of the reference sphere R_2 changes accordingly in the range $[95.28\text{mm}, 80.28\text{mm}]$ and this is also the case for the EXPP, as explained above. The position of the image plane is automatically constrained to be at the paraxial focus bearing in mind that in correspondence with this location the height of the marginal ray is zero. While t varies, the calculation of the wavefront aberration is carried out for a light ray r whose normalized field and pupil coordinates are $(H_x = 0, H_y = 1, \rho_x = 0, \rho_y = 0.5)$. In Figure 1.7, in addition to the refractive surface S_1 , the reference sphere S_2 is shown in the two limiting positions $t = 0\text{mm}$ (in red $S_2 = S_2'$) and $t = 15\text{mm}$ (in green $S_2 = S_2''$). In these two cases, the radius of curvature of the reference sphere is equal to $R_2 = R_2' = \text{EXPP}'$ and $R_2 = R_2'' = \text{EXPP}''$ respectively. Therefore, varying t , the reference sphere is shifted and its curvature is modified.

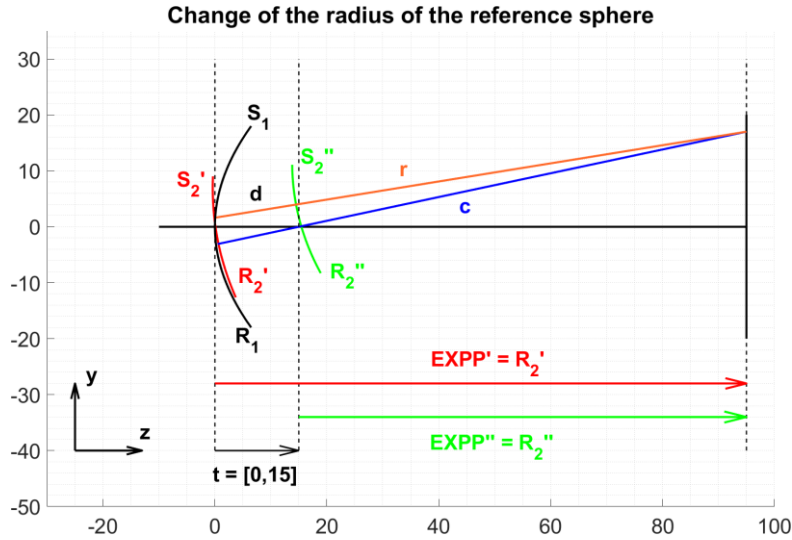


Figure 1.7: Simple system with one single refractive surface S_1 studied to derive the effect of the change of the radius of the reference sphere on the aberration associated with a light ray. The distances are in mm.

The aperture stop is placed on the surface S_2 that is the reference sphere whose distance t from S_1 is varied in the interval $[0\text{mm}, 15\text{mm}]$. In red and in green are shown the two limiting cases $S_2 = S_2'$ for $t = 0\text{mm}$ and $S_2 = S_2''$ for $t = 15\text{mm}$. In correspondence of these two positions the radius of curvature of the reference sphere is $R_2 = R_2'$ and $R_2 = R_2''$ respectively. Thus, the reference spheres S_2' and S_2'' are concentric.

The wavefront aberration associated with the ray r (in orange) is given by the optical path difference with respect to the chief ray c (in blue) in waves at the wavelength λ_0 . In this calculation, the chief ray is emitted from a field point with coordinates $(H_x = 0, H_y = 10^\circ)$ while the system is at infinite conjugates. The optical path lengths of the ray r and the chief ray c , respectively W_r and W_c , are affected by the shift of the reference sphere, denoted with $(\Delta x, \Delta y, \Delta z)$, and by the modification of its curvature, denoted with Δc . For the ray r , this quantity is given by

$$W_r = nd = n \left[\frac{(K\Delta x + L\Delta y + M\Delta z)}{\cos I} + \frac{\frac{1}{2}\Delta c(x^2 + y^2 + z^2)}{\cos I} \right]$$

Equation 1.38

where d is the distance between the new and the old sphere along the ray r , the parameters (K, L, M) and I are the components of the vector normal to the surface of incidence and the angle of incidence respectively. The optical path difference between the ray r and the chief ray c is given simply by $\delta W = W_c - W_r$. It is remarked again that δW

is a sort of perturbation of the wavefront aberration associated with the ray r deriving from a modification of the reference sphere. In the following Figure 1.8, it is shown that this change of the optical path δW due to a modification of the radius of the reference sphere varies with the square of the angular ray aberration α_y along the H_y direction. The angular ray aberration α_y for the ray r at the wavelength λ_0 is measured as the difference between the direction cosine of the ray r along H_y and the corresponding direction cosine of the chief ray c . It is again emphasised that such a change of the radius of the reference sphere is performed keeping its center of curvature fixed at the point of intersection of the chief ray with the image plane.

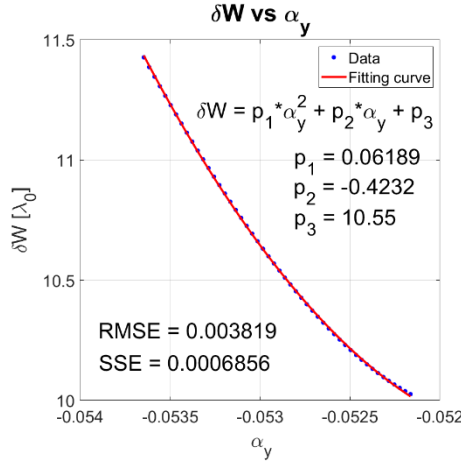


Figure 1.8: The variation of the wavefront aberration due to a change of radius of the reference sphere is a quadratic function of the angular ray aberration.

The fitting curve shown in Figure 1.8 puts in evidence that the variation of the optical path difference due to a change of curvature of the reference sphere varies quadratically with the angular ray aberration. This last quantity is linearly related to the transverse ray aberration, namely $\alpha_y \sim \delta H_y = O(3)$. Therefore, the change of the radius of curvature of the reference sphere R_2 determines a change of the wavefront aberration associated with a particular ray r that is $\delta W \sim \alpha_y^2 = O(6)$. This implies that such a change of the reference sphere does not affect primary wavefront aberrations since these are $O(4)$, but it affects secondary and higher order aberrations. This state of affairs is known as the summation theorem for primary aberrations [1].

Chapter 2 Scalar analytical expressions for the field dependence of Zernike polynomials in asymmetric optical systems with circular symmetric surfaces

In this chapter [2], we derive a series of scalar analytical expressions describing the full field dependence of Zernike polynomials in optical systems without symmetries. We consider the general case of optical systems constituted by arbitrarily tilted and decentered circular symmetric surfaces. The resulting analytical formulas are inferred from a modified version of the full field-dependent wavefront aberration function proposed in the Nodal Aberration Theory (NAT). Such formula is modified with the scope of solving few critical points arising when primary and higher-order aberrations are both presents in an optical system. It is shown that when secondary aberrations are taken into account in the wavefront aberration function, the final effect is a perturbation to the symmetry of the field dependence of the Zernike polynomials. In particular, the centers of symmetry of the Zernike polynomial field dependences are shifted with respect to the locations predicted using the NAT equations as a consequence of the presence of higher-order aberrations. The retrieved analytical expressions are verified through surface fitting to real raytracing data obtained for a simple optical system.

2.1 Introduction

The Nodal Aberration Theory (NAT) proposed by Thompson [3, 4] constitutes a fundamental step forward to the development of the wave theory of aberrations for non-circular symmetric optical systems characterized by tilted and decentered circular symmetric surfaces. NAT is built on the vectorial wavefront aberration formula introduced by Shack. Such vectorial expression is a reformulation of the scalar wavefront aberration formula introduced by H. H. Hopkins [5] for circular symmetric optical systems. NAT demonstrates that, in an optical system, the effect of geometrical perturbations such as tilt and decentering of optical surfaces, is to induce a particular nodal behavior in the field dependence of individual aberration types. This means that NAT provides a complete description of the aberration function of asymmetric optical systems as a function of both pupil and field coordinates.

An alternative description of the wavefront aberration function of an optical system is based on the use of Zernike polynomials. They represent a weighted polynomial expansion used to approximate the wavefront aberration function of an optical system with a circular pupil. The weights of individual Zernike polynomial terms in the wavefront decomposition are represented by multiplicative coefficients retrieved with a polynomial fitting routine applied to the wavefront function itself. The Zernike polynomials decomposition is normally applied to pupil-dependent wavefront functions whose field dependence is omitted because the wavefront function itself is referenced to a single field point in the field of view (FOV) plane. Nevertheless, such polynomial expansion can be extended to a full field-dependent wavefront aberration function, provided that the coefficients of the Zernike terms are turned into field-dependent functions. This approach opens up the possibility of decomposing the four-dimensional full field wavefront aberration function of an optical system in terms of analytical expressions consisting of the product between Zernike polynomials (depending on the pupil coordinates) and functions depending on the field coordinates.

In the design of optical systems where the full field behavior of aberrations is of concern, it is desirable to dispose of such analytical expressions. For example, the design of a multi-aperture optical system whose optical branches consist of decentered optical components can benefit from these formulas (as will be discussed in Chapter 5). Therefore, the main goal of the present work is to derive the full field dependence of Zernike polynomials for asymmetric optical systems constituted by tilted and decentered circular symmetric surfaces. For this purpose, we will use a slightly modified version of the full field wavefront aberration function described in the NAT theory and revisit partly the work done in [6]. Additionally, we aim at verifying numerically the expressions introduced in NAT to retrieve the location of the centers of symmetry in the FOV of primary aberrations in asymmetric optical systems. Finally, we validate the analytical expressions of the field dependence of Zernike polynomials with ray-trace-based calculation obtained for a simple optical system constituted by a sequence of two tilted and decentered spherical lenses.

2.2 Full-field wavefront aberration function in asymmetric optical systems with tilted and decentered circular symmetric surfaces

As said above, the Nodal Aberration theory was introduced to explain the occurrence of nodes in the full-field aberration function of optical systems characterized by tilted and decentered circular symmetric surfaces. The effect of surface tilt and decentering with respect to the mechanical axis of the system is modeled as a perturbation to the full field dependence of the aberration function of the optical system itself. The vectorial wavefront aberration function introduced by NAT is shown in Equation 2.1 [3]

$$W(\vec{H}, \vec{\rho}) = \sum_j \sum_p \sum_n \sum_m (W_{klm})_j \left((\vec{H} - \vec{\sigma}_j) \cdot (\vec{H} - \vec{\sigma}_j) \right)^p (\vec{\rho} \cdot \vec{\rho})^n \left((\vec{H} - \vec{\sigma}_j) \cdot \vec{\rho} \right)^m$$

Equation 2.1

where $\vec{\rho}$ is the pupil vector and \vec{H} is the field vector. It can be defined as an effective field vector $\vec{H}_{eff} = \vec{H} - \vec{\sigma}_j$ accounting for the perturbation induced by $\vec{\sigma}_j$, a vector in the FOV plane pointing to the center of symmetry of the aberration fields for individual tilted and decentered surfaces indexed with j . The $(W_{klm})_j$ are the aberration coefficients of the surface j , identified by the indexes m , $k = 2p + m$ and $l = 2n + m$. Formally, Equation 2.1 is the extension to asymmetric optical systems with circular symmetric surfaces of the formula introduced by Hopkins [5] to describe the aberrations in circular symmetric optical systems.

In [6], Equation 2.1 is expanded up to secondary aberrations retaining the summation over the surfaces indexed with j . This formulation is valid in the case when only primary aberrations are of concern, but it is no longer valid to describe the behavior of higher-order aberrations for asymmetric optical systems consisting of circular symmetric surfaces. In this regard, we believe it is important to mention the following considerations. Relatively to circular symmetric optical systems, it is known that the contributions of primary aberrations can be added surface by surface according to the summation theorem for primary aberrations [1]. This theorem entails that, when higher-order aberrations are taken into account, it is no longer possible to assume the validity of this summation. More precisely, the summation of aberration contributions from individual surfaces is valid under the assumption that the total wavefront aberration function of an optical system is not affected by a change of the radius of curvature of the reference sphere. This assumption is valid only for primary aberrations (being of 4th order) since the change of the radius of the reference sphere determines a modification to the 6th order of the wavefront error, thus having an impact on aberrations of order higher than the fourth [1,7]. Additionally, we mention that for higher-order aberrations it is useful to discriminate between intrinsic and extrinsic aberrations where the former relates to a contribution that can be effectively ascribed to a surface j , while the latter relates to the contributions generated by the surfaces $1, \dots, j-1$ preceding the surface j . Sasià in [8] addressed this problem retrieving a series of complex formulas to calculate secondary aberration coefficients in circular symmetric optical systems. These observations lead us to consider a modified version of Equation 2.1 relatively to asymmetric optical systems. Therefore, to avoid the aforementioned points of criticism, we remove from Equation 2.1 the summation \sum_j over the surfaces of the optical system. Nevertheless, for the main purpose of this chapter, that is the calculation of the field dependence of Zernike Equation 2.1 which points out the possibility to identify different centers of symmetry in the field dependence of different aberration terms. Consequently, modifying Equation 2.1 consistently with the previous considerations, we obtain the following Equation 2.2 that is the starting point of our further derivation of the field dependence of Zernike polynomials

$$W(\vec{H}, \vec{\rho}) = \sum_p \sum_n \sum_m W_{klm} \left((\vec{H} - \vec{a}_{klm}) \cdot (\vec{H} - \vec{a}_{klm}) \right)^p (\vec{\rho} \cdot \vec{\rho})^n \left((\vec{H} - \vec{a}_{klm}) \cdot \vec{\rho} \right)^m$$

Equation 2.2

The coefficients W_{klm} and the displacement vectors \vec{a}_{klm} denote respectively the net aberration coefficients and the net field displacement vectors of the total wavefront error of an asymmetric optical system. Equation 2.2 can be reduced to Equation 2.1 under the condition that secondary and higher-order aberrations are negligible with respect to primary aberrations, i.e. in optical systems with small apertures and small field angles. In this case, we can use the following Equations 2.3 to calculate the net primary aberration coefficients and displacement vectors as shown in [3, 9, 10, 11].

$$\begin{cases} W_{klm} = \sum_j W_{klm,j} \\ \vec{a}_{klm} = \frac{\sum_j W_{klm,j} \vec{\sigma}_j}{W_{klm}} \end{cases}$$

Equation 2.3

In Equations 2.3, the summation over the surfaces j is valid supposing that the optical system is characterized predominantly by primary aberrations, in fact, the coefficients W_{klm} are nothing but the Seidel sums [1]. On the other side, if we are concerned with higher-order aberrations, the calculation of the coefficients W_{klm} and displacement vectors \vec{a}_{klm} is

no longer accurate if Equations 2.3 are used. In this case, the displacement vectors \vec{a}_{klm} related to primary aberrations are affected by the contribution of higher-order aberrations and in addition to this, the coefficients W_{klm} cannot be simply calculated summing the aberration contributions of different surfaces in the optical system.

As mentioned above, in Equation 2.2 we retain the general structure of Equation 2.1 which highlights the possibility to identify a center of symmetry (located by the field displacement vector) in the field dependence of different aberration terms. Developing Equation 2.2 up to the 6th order, we obtain the following Equation 2.4 describing the wavefront aberration function constituted by primary and secondary aberrations for asymmetric optical systems with tilted and decentered circular symmetric surfaces.

$$\begin{aligned}
 W(\vec{H}, \vec{\rho}) = & W_{000} + W_{020}(\vec{\rho} \cdot \vec{\rho}) + W_{111}[(\vec{H} - \vec{a}_{111}) \cdot \vec{\rho}] \\
 & + W_{040}(\vec{\rho} \cdot \vec{\rho})^2 + W_{131}[(\vec{H} - \vec{a}_{131}) \cdot \vec{\rho}](\vec{\rho} \cdot \vec{\rho}) + W_{220}[(\vec{H} - \vec{a}_{220}) \cdot (\vec{H} - \vec{a}_{220})](\vec{\rho} \cdot \vec{\rho}) \\
 & + W_{222}[(\vec{H} - \vec{a}_{222}) \cdot \vec{\rho}]^2 + W_{311}[(\vec{H} - \vec{a}_{311}) \cdot (\vec{H} - \vec{a}_{311})][(\vec{H} - \vec{a}_{311}) \cdot \vec{\rho}] \\
 & + W_{060}(\vec{\rho} \cdot \vec{\rho})^3 + W_{151}[(\vec{H} - \vec{a}_{151}) \cdot \vec{\rho}](\vec{\rho} \cdot \vec{\rho})^2 + W_{240}[(\vec{H} - \vec{a}_{240}) \cdot (\vec{H} - \vec{a}_{240})](\vec{\rho} \cdot \vec{\rho})^2 \\
 & + W_{242}[(\vec{H} - \vec{a}_{242}) \cdot \vec{\rho}]^2(\vec{\rho} \cdot \vec{\rho}) + W_{331}[(\vec{H} - \vec{a}_{331}) \cdot (\vec{H} - \vec{a}_{331})][(\vec{H} - \vec{a}_{331}) \cdot \vec{\rho}](\vec{\rho} \cdot \vec{\rho}) \\
 & + W_{333}[(\vec{H} - \vec{a}_{333}) \cdot \vec{\rho}]^3 + W_{420}[(\vec{H} - \vec{a}_{420}) \cdot (\vec{H} - \vec{a}_{420})][(\vec{H} - \vec{a}_{420}) \cdot (\vec{H} - \vec{a}_{420})](\vec{\rho} \cdot \vec{\rho}) \\
 & + W_{422}[(\vec{H} - \vec{a}_{422}) \cdot (\vec{H} - \vec{a}_{422})][(\vec{H} - \vec{a}_{422}) \cdot \vec{\rho}]^2 \\
 & + W_{151}[(\vec{H} - \vec{a}_{151}) \cdot \vec{\rho}](\vec{\rho} \cdot \vec{\rho})^2
 \end{aligned}$$

Equation 2.4

In general, the aberration coefficients W_{klm} in Equation 2.4 are different from the aberration coefficients related to the wavefront aberration function of a corresponding circular symmetric optical system. As a matter of fact, expanding Equation 2.4, we obtain a power series in the ray coordinates containing terms of even and odd degree up to the 6th order while the wavefront aberration function of a circular symmetric system contains only terms of even order. It is worth noting that if secondary aberrations give a significant contribution to the total wavefront aberration function in Equation 2.4, then several lower-order terms (of both even and odd order) are generated and added to the terms attributable to primary aberrations. Therefore, in Equation 2.4 the displacement vectors locating the centers of symmetry of the field dependence of primary aberrations are influenced by the presence of higher-order aberrations. Such perturbative effect of higher-order aberrations on primary aberrations is treated in [10, 11] where effective aberration coefficients and effective field displacement vectors are introduced to account for the shift induced by secondary aberrations.

In Equation 2.4 each aberration term exhibits a center of symmetry that in general does not coincide with a node location for the specific aberration type. In particular, as regards binodal astigmatism, the vector \vec{a}_{222} locates the intermediate position in the FOV plane between the two nodes characterizing this aberration type. Equivalently, the vector \vec{a}_{222} indicates the center of symmetry of primary astigmatism in the FOV plane in a non-circular symmetric optical system. On the other hand, regarding primary coma the vector \vec{a}_{131} locates at the same time the center of symmetry of this aberration and the position of its single node in the FOV plane.

In order to retrieve the field dependence of Zernike coefficients in asymmetric optical systems, we proceed to convert Equation 2.4 into a scalar form. The reference system used in this work is shown in Figure 2.1.

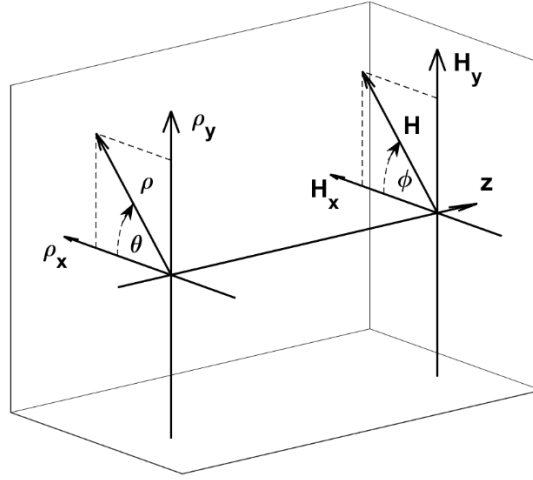


Figure 2.1: Reference system.

The vectors $\vec{\rho}$, \vec{H} and \vec{a}_{klm} are respectively

$$\begin{aligned}\vec{\rho} &= \begin{pmatrix} \rho_x \\ \rho_y \end{pmatrix} = \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \\ \vec{H} &= \begin{pmatrix} H_x \\ H_y \end{pmatrix} = \begin{pmatrix} H \cos \varphi \\ H \sin \varphi \end{pmatrix} \\ \vec{a}_{klm} &= \begin{pmatrix} a_{klm,x} \\ a_{klm,y} \end{pmatrix}\end{aligned}$$

Equation 2.5

Using the expression exposed by Thompson in [3, 4] to calculate the square of a generic vector $\vec{v} = v_x \vec{i} + v_y \vec{j}$ (where $\vec{v} = v e^{i\alpha} = v \cos \alpha \vec{i} + v \sin \alpha \vec{j}$)

$$(\vec{v})^2 = \begin{pmatrix} v_x \\ v_y \end{pmatrix}^2 = \begin{pmatrix} v \cos 2\alpha \\ v \sin 2\alpha \end{pmatrix} = \begin{pmatrix} v_x^2 - v_y^2 \\ 2v_x v_y \end{pmatrix}$$

Equation 2.6

one obtains the following series of equations from Equation 2.7 to Equation 2.15 for all primary aberrations and only a few secondary aberrations. Among them, we consider oblique spherical aberration W_{OSA} , field cubed coma W_{CC} , secondary field curvature W_{FC2} and secondary astigmatism W_{A2} . These secondary aberrations are explicitly considered because they give a more consistent contribution to the field dependence of Zernike coefficients shown below in the ray-trace-based example. The complete scalar full field wavefront aberration function $W(H_x, H_y, \rho, \vartheta)$ for asymmetric optical systems constituted by tilted and decentered circular symmetric surfaces is finally given by the sum of terms exposed in equations from Equation 2.7 to Equation 2.15.

We obtain for primary spherical aberration $W_{SA}(H_x, H_y, \rho, \vartheta)$ the well-known field independent expression in Equation 2.7.

$$W_{SA}(H_x, H_y, \rho, \vartheta) = W_{040} \left[\begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right]^2 = W_{040} (\rho^2 \cos^2 \vartheta + \rho^2 \sin^2 \vartheta)^2 = W_{040} \rho^4$$

Equation 2.7

For primary coma $W_C(H_x, H_y, \rho, \vartheta)$, one obtains Equation 2.8.

$$W_C(H_x, H_y, \rho, \vartheta) = W_{131} \left[\begin{pmatrix} H_x - a_{131,x} \\ H_y - a_{131,y} \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right] \left[\begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right] = W_{131} \left[(H_x - a_{131,x}) \rho \cos \vartheta + (H_y - a_{131,y}) \rho \sin \vartheta \right] (\rho^2 \cos^2 \vartheta + \rho^2 \sin^2 \vartheta) = W_{131} \left[(H_x - a_{131,x}) \rho^3 \cos \vartheta + (H_y - a_{131,y}) \rho^3 \sin \vartheta \right]$$

Equation 2.8

For primary astigmatism $W_A(H_x, H_y, \rho, \vartheta)$, one obtains Equation 2.9.

$$W_A(H_x, H_y, \rho, \vartheta) = \frac{1}{2} W_{222} \left[\begin{pmatrix} H_x - a_{222,x} \\ H_y - a_{222,y} \end{pmatrix}^2 \right]^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}^2 = \frac{1}{2} W_{222} \left[\left((H_x - a_{222,x})^2 - (H_y - a_{222,y})^2 \right) \rho^2 \cos 2\vartheta + 2(H_x - a_{222,x})(H_y - a_{222,y}) \rho^2 \sin 2\vartheta \right]$$

Equation 2.9

Relatively to field curvature $W_{FC}(H_x, H_y, \rho, \vartheta)$, one obtains the following Equation 2.10.

$$W_{FC}(H_x, H_y, \rho, \vartheta) = W_{220} \left[\begin{pmatrix} H_x - a_{220,x} \\ H_y - a_{220,y} \end{pmatrix}^T \cdot \begin{pmatrix} H_x - a_{220,x} \\ H_y - a_{220,y} \end{pmatrix} \right] \left[\begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right] = W_{220} \left[(H_x - a_{220,x})^2 + (H_y - a_{220,y})^2 \right] \rho^2$$

Equation 2.10

For Equation 2.9 and Equation 2.10 we have considered that field curvature and primary astigmatism are referred to the medial astigmatic surface setting $W_{220} = W_{220S} + \frac{W_{222}}{2}$ (where the subscript of the term W_{220S} stands for *Sagittal*). For distortion $W_D(H_x, H_y, \rho, \vartheta)$, we obtain Equation 2.11.

$$W_D(H_x, H_y, \rho, \vartheta) = W_{311} \left[\begin{pmatrix} H_x - a_{311,x} \\ H_y - a_{311,y} \end{pmatrix}^T \cdot \begin{pmatrix} H_x - a_{311,x} \\ H_y - a_{311,y} \end{pmatrix} \right] \left[\begin{pmatrix} H_x - a_{311,x} \\ H_y - a_{311,y} \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right] = W_{311} \left[(H_x - a_{311,x})^2 + (H_y - a_{311,y})^2 \right] \left[(H_x - a_{311,x}) \rho \cos \vartheta + (H_y - a_{311,y}) \rho \sin \vartheta \right]$$

Equation 2.11

For oblique spherical aberration $W_{OSA}(H_x, H_y, \rho, \vartheta)$, we obtain Equation 2.12 similar to Equation 2.10 for field curvature regarding the only field dependence.

$$W_{OSA}(H_x, H_y, \rho, \vartheta) = W_{240} \left[\begin{pmatrix} H_x - a_{240,x} \\ H_y - a_{240,y} \end{pmatrix}^T \cdot \begin{pmatrix} H_x - a_{240,x} \\ H_y - a_{240,y} \end{pmatrix} \right] \left[\begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right] = W_{240} \left[(H_x - a_{240,x})^2 + (H_y - a_{240,y})^2 \right] \rho^4$$

Equation 2.12

In Equation 2.12 the coefficient W_{240} is defined as $W_{240} = W_{240S} + \frac{W_{242}}{2}$ in order to reference it to the medial surface. For field cubed coma $W_{CC}(H_x, H_y, \rho, \vartheta)$, we obtain Equation 2.13 similar to Equation 2.11 obtained for distortion (concerning the only field behavior)

$$W_{CC}(H_x, H_y, \rho, \vartheta) = W_{331M} \left[\begin{pmatrix} H_x - a_{331,x} \\ H_y - a_{331,y} \end{pmatrix}^T \cdot \begin{pmatrix} H_x - a_{331,x} \\ H_y - a_{331,y} \end{pmatrix} \right] \left[\begin{pmatrix} H_x - a_{331,x} \\ H_y - a_{331,y} \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right] \left[\begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right] = W_{331M} \left[(H_x - a_{331,x})^2 + (H_y - a_{331,y})^2 \right] \left[(H_x - a_{331,x}) \rho \cos \vartheta + (H_y - a_{331,y}) \rho \sin \vartheta \right] \left[\rho^2 \cos^2 \vartheta + \rho^2 \sin^2 \vartheta \right] = W_{331M} \left[(H_x - a_{331,x})^2 + (H_y - a_{331,y})^2 \right] \left[(H_x - a_{331,x}) \rho^3 \cos \vartheta + (H_y - a_{331,y}) \rho^3 \sin \vartheta \right]$$

Equation 2.13

In Equation 2.13 the coefficient W_{331M} is defined as $W_{331M} = W_{331} + \frac{3}{4}W_{333}$. This implies that it contains part of trefoil aberration weighted by the coefficient W_{333} .

For secondary field curvature $W_{FC2}(H_x, H_y, \rho, \vartheta)$, we obtain Equation 2.14.

$$\begin{aligned}
 W_{FC2}(H_x, H_y, \rho, \vartheta) &= W_{420} \left[\begin{pmatrix} H_x - a_{420,x} \\ H_y - a_{420,y} \end{pmatrix}^T \cdot \begin{pmatrix} H_x - a_{420,x} \\ H_y - a_{420,y} \end{pmatrix} \right] \\
 &\left[\begin{pmatrix} H_x - a_{420,x} \\ H_y - a_{420,y} \end{pmatrix}^T \cdot \begin{pmatrix} H_x - a_{420,x} \\ H_y - a_{420,y} \end{pmatrix} \right] \left[\begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right] = \\
 &W_{420} \left[(H_x - a_{420,x})^2 + (H_y - a_{420,y})^2 \right]^2 \rho^2
 \end{aligned}$$

Equation 2.14

Finally, for secondary astigmatism $W_{A2}(H_x, H_y, \rho, \vartheta)$, we obtain Equation 2.15.

$$\begin{aligned}
 W_{A2}(H_x, H_y, \rho, \vartheta) &= \frac{1}{2} W_{422} \left[\begin{pmatrix} H_x - a_{422,x} \\ H_y - a_{422,y} \end{pmatrix}^T \cdot \begin{pmatrix} H_x - a_{422,x} \\ H_y - a_{422,y} \end{pmatrix} \right] \\
 &\left[\begin{pmatrix} (H_x - a_{422,x})^2 \\ (H_y - a_{422,y})^2 \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix}^2 \right] = \\
 &\frac{1}{2} W_{422} \left[(H_x - a_{422,x})^2 + (H_y - a_{422,y})^2 \right] \\
 &\left[\left((H_x - a_{422,x})^2 - (H_y - a_{422,y})^2 \right) \rho^2 \cos 2\vartheta + 2(H_x - a_{422,x})(H_y - a_{422,y}) \rho^2 \sin 2\vartheta \right]
 \end{aligned}$$

Equation 2.15

To be noted that in Equation 2.14 and Equation 2.15, similarly with what we have done previously for primary field curvature and primary astigmatism, we have set $W_{420} = W_{420S} + \frac{W_{422}}{2}$ in order to fix the medial astigmatic surface as a reference.

The total scalar full field-dependent and pupil dependent wavefront aberration function $W(H_x, H_y, \rho, \vartheta)$ results from the sum of individual aberration terms reported from Equation 2.7 to Equation 2.15. The remaining secondary aberration terms are not mentioned in this work because they are less relevant in the ray-trace data used for verification in the last paragraph. The scalar wavefront aberration function $W(H_x, H_y, \rho, \vartheta)$ is used in the following part to retrieve the field behavior of Zernike coefficients in asymmetric optical systems constituted by tilted and decentered surfaces.

2.3 Full-field dependence of Zernike polynomials in asymmetric systems

As said before, in this chapter we partly revisit the work done in [6] using a slightly different version of the full field wavefront aberration function proposed in NAT. Such reformulation of the wavefront aberration function is essentially designed to avoid contradictions with the summation theorem for primary aberrations. In this section, we derive the full field dependence of Zernike polynomials for asymmetric optical systems that contain circular symmetric surfaces. Further on, we provide a numerical verification of the expressions derived in the present section.

Zernike polynomials, indicated in this work as $C_n^m(\rho, \vartheta)$, are a complete set of polynomials depending on the radial and azimuthal coordinates of the pupil. They are orthonormal in a continuous fashion over the interior of a unit circle and are expressed as the product of a polynomial function $R_n^m(\rho)$ which depends only on the radial coordinate

and a trigonometric function $G^m(\vartheta)$ that depends only on the azimuthal coordinate. In detail, they are described as follows

$$C_n^m(\rho, \vartheta) = F_n^m R_n^m(\rho) G^m(\vartheta) = F_n^m Z_n^m(\rho, \vartheta)$$

Equation 2.16

where F_n^m are the Zernike coefficients and $Z_n^m(\rho, \vartheta)$ are pupil dependent functions resulting from the product between $R_n^m(\rho)$ and $G^m(\vartheta)$. We use the Fringe indexing scheme. Similarly to what has been shown by Gray et al. in [12] for the case of circular symmetric optical systems, the computation of the field dependence of Zernike coefficients in asymmetric optical systems is retrieved by projecting the scalar full field-dependent and pupil dependent wavefront aberration function $W(H_x, H_y, \rho, \vartheta)$ into the Zernike polynomial basis. Equivalently, the field-dependent coefficients of individual Zernike polynomials in asymmetric optical systems are obtained computing the following double integral with respect to the polar coordinates of the pupil shown in Equation 2.17

$$F_n^m(H_x, H_y) = \frac{1}{N_{nm}} \int_0^1 \int_0^{2\pi} W(H_x, H_y, \rho, \vartheta) Z_n^m(\rho, \vartheta) \rho d\rho d\vartheta$$

Equation 2.17

where N_{nm} is the norm of Zernike polynomials. The computation of Equation 2.17 results in a new expression (Equation 2.18) for the full field wavefront aberration function in asymmetric optical systems described as the sum of full field-dependent Zernike polynomials. Individual terms in Equation 2.18 are given by the product between the field-dependent coefficients $F_n^m(H_x, H_y)$ retrieved with Equation 2.17 and the respective Zernike polynomials $Z_n^m(\rho, \vartheta)$. The field dependent terms $F_n^m(H_x, H_y)$ in Equation 2.18 are determined by the contributions of various aberration types with different centers of symmetry and weighted by their respective wave aberration coefficients W_{klm} . The field dependent functions $F_n^m(H_x, H_y)$ are actually power series containing terms of even and odd order in the field coordinates originated by the absence of a well-defined symmetry in optical systems with displaced and tilted surfaces. Writing the result down leads to Equation 2.18 that reads

$$\begin{aligned}
 W(H_x, H_y, \rho, \vartheta) = & \left[\frac{W_{020}}{2} + \frac{W_{040}}{3} + \frac{W_{220}}{2} \left((H_x - a_{220,x})^2 + (H_y - a_{220,y})^2 \right) + \frac{W_{420}}{2} \left((H_x - a_{420,x})^2 + (H_y - a_{420,y})^2 \right)^2 \right] Z_0^0(\rho, \vartheta) + \\
 & \left[W_{111} (H_x - a_{111,x}) + \frac{2W_{131}}{3} (H_x - a_{131,x}) + \frac{2W_{331}}{3} \left((H_x - a_{331,x})^2 + (H_y - a_{331,y})^2 \right) (H_x - a_{331,x}) \right] Z_1^1(\rho, \vartheta) + \\
 & \left[W_{111} (H_y - a_{111,y}) + \frac{2W_{131}}{3} (H_y - a_{131,y}) + \frac{2W_{331}}{3} \left((H_x - a_{331,x})^2 + (H_y - a_{331,y})^2 \right) (H_y - a_{331,y}) \right] Z_1^{-1}(\rho, \vartheta) + \\
 & \left[\frac{W_{020}}{2} + \frac{W_{040}}{2} + \frac{W_{220}}{2} \left((H_x - a_{220,x})^2 + (H_y - a_{220,y})^2 \right) + \frac{W_{420}}{2} \left((H_x - a_{420,x})^2 + (H_y - a_{420,y})^2 \right)^2 \right] Z_2^0(\rho, \vartheta) + \\
 & \left[\frac{W_{222}}{2} \left((H_x - a_{222,x})^2 - (H_y - a_{222,y})^2 \right) + \frac{W_{422}}{2} \left((H_x - a_{422,x})^4 - (H_y - a_{422,y})^4 \right) \right] Z_2^2(\rho, \vartheta) + \\
 & \left[W_{222} (H_x - a_{222,x}) (H_y - a_{222,y}) + W_{422} \left(\left((H_x - a_{422,x})^2 + (H_y - a_{422,y})^2 \right) (H_x - a_{422,x}) (H_y - a_{422,y}) \right) \right] Z_2^{-2}(\rho, \vartheta) + \\
 & \left[\frac{W_{131}}{3} (H_x - a_{131,x}) + \frac{W_{331}}{3} \left((H_x - a_{331,x})^2 + (H_y - a_{331,y})^2 \right) (H_x - a_{331,x}) \right] Z_3^1(\rho, \vartheta) + \\
 & \left[\frac{W_{131}}{3} (H_y - a_{131,y}) + \frac{W_{331}}{3} \left((H_x - a_{331,x})^2 + (H_y - a_{331,y})^2 \right) (H_y - a_{331,y}) \right] Z_3^{-1}(\rho, \vartheta) + \\
 & \left[\frac{W_{040}}{6} + \frac{W_{240}}{6} \left((H_x - a_{240,x})^2 + (H_y - a_{240,y})^2 \right) \right] Z_4^0(\rho, \vartheta)
 \end{aligned}$$

Equation 2.18

The terms in Equation 2.18 correspond to the first nine Zernike polynomials using the Fringe indexing scheme. These terms are at most of the 4th order in their pupil coordinates dependence. Considering their field coordinates dependence, Equation 2.18 explicitly shows that in an asymmetric optical system the full field behavior of Zernike polynomials is described by a superposition of polynomial surfaces whose respective displacements in the FOV plane are defined by the centers of symmetry of the inherent aberration types contributing to the description of the Zernike terms in question. In fact, we see that considering explicitly the contributions from only the primary aberrations (setting to zero the wave coefficients W_{klm} related to secondary aberrations), the field dependence of Zernike polynomials in asymmetric optical systems is unchanged with respect to their field dependence in circular symmetric optical systems. The only difference being that in the former case the full field behavior of Zernike polynomials exhibits a new center of

symmetry displaced from the origin of the FOV plane. On the other hand, taking into consideration the contributions to the full field dependence of Zernike terms deriving jointly from primary and secondary aberrations, we can observe that in an asymmetric optical system the field dependence of each Zernike term no longer exhibits a well-defined symmetry because of the concomitant contribution of different order terms. In fact, it results from the superposition of field surfaces with not coincident centers of symmetry.

From the above derivation, Equation 2.18 demonstrates that the co-presence of primary and secondary aberrations modifies the symmetry of the field behavior of individual Zernike polynomials. In addition to this, when higher-order field surfaces (related to secondary aberrations) are overlaid to lower-order field surfaces (related to primary aberrations), the centers of symmetry of the latter can be displaced in the FOV plane. In optical systems where only primary aberrations are considered to contribute to the wavefront deformation, the aberration coefficients W_{klm} and the centers of symmetry of the field dependence of Zernike polynomials \vec{a}_{klm} can be calculated using Equation 2.3. In this specific case, using Equation 2.3 to calculate the total primary aberration coefficients W_{klm} , is equivalent to assume that tilts and displacements of the surfaces in the optical system under study induce weak perturbations to the original symmetry of the system. In the following section, we compare the analytical result obtained so far with ray tracing data calculated for a simple optical system. Additionally, we calculate the centers of symmetry of primary aberrations using Equation 2.3 (derived from NAT) in order to compare the prediction of NAT equations with real raytracing data.

2.4 Validation of results

To validate the analytical expressions describing the full field behavior of Zernike coefficients in asymmetric optical systems shown in the previous section, we compare them with real raytracing data obtained for a simple system constituted of a sequence of two tilted and decentered plano-convex spherical lenses. The 3D layout of the optical system under study is shown in Figure 2.2.

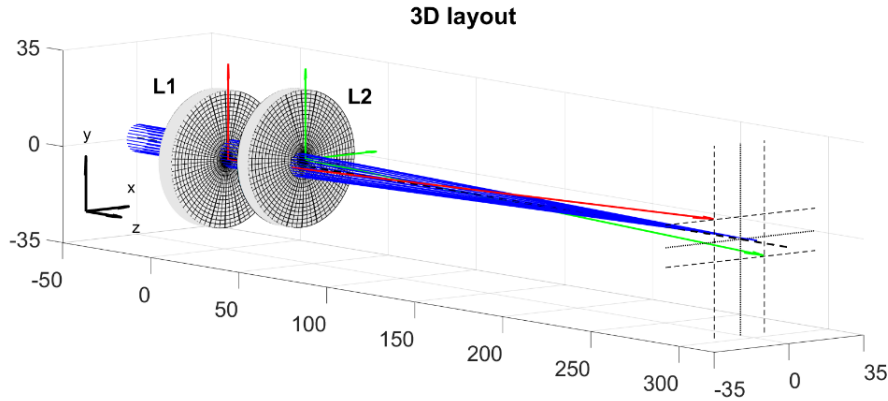


Figure 2.2: 3D layout and reference system of the optical system under study (the dimensions are in mm). The lenses are indicated respectively with L1 and L2. In red and green are highlighted the local coordinate systems of the curved surfaces in both plano-convex spherical lenses.

The system is at infinite conjugates. The aperture stop position is at the first surface (plane surface) of the first spherical lens L1 and the entrance pupil diameter is $ENPD = 10\text{mm}$. The simulation wavelength is $\lambda = 0.55\mu\text{m}$ and the FOV is $40^\circ \times 40^\circ$. It is worth pointing out that the field coordinates in the FOV plane can be equivalently described in terms of object height or image height. We represent the field coordinates H_x and H_y in terms of field angles in degrees measured with respect to the object space z axis and the paraxial entrance pupil position (in this example located in correspondence of the plane surface of L1 where it is also located the aperture stop of the optical system) [13]. The

spherical lenses are tilted and decentered with respect to the mechanical axis of the system as reported in Table 2.1. The distance between the two lenses is 20 mm. The image plane is at 292.4mm from the last optical surface. This distance has been chosen to minimize the RMS wavefront error and to constrain the defocus aberration to be zero for the field point ($H_x = 0^\circ, H_y = 0^\circ$). The parameters used for this example are summarized in Table 2.1. The simulation is performed with Zemax OpticStudio [13].

Table 2.1: Simulation parameters

	Parameter	Symbol	Value
L1	Diameter	D_1	50 mm
	Radius of curvature	R_1	-400 mm
	Thickness	t_1	7 mm
	Glass	g_1	N-BK7
	Conic constant	c_1	0
	x decentering	$D_{x,1}$	-2 mm
	y decentering	$D_{y,1}$	-2 mm
	x tilt	$\tau_{x,1}$	-2 °
	y tilt	$\tau_{y,1}$	-2 °
L2	Diameter	D_2	50 mm
	Radius of curvature	R_2	-250 mm
	Thickness	t_2	5 mm
	Glass	g_2	N-BK7
	Conic constant	c_2	0
	x decentering	$D_{x,2}$	2 mm
	y decentering	$D_{y,2}$	2 mm
	x tilt	$\tau_{x,2}$	2 °
	y tilt	$\tau_{y,2}$	2 °

In Table 2.2 we report the values of the Seidel aberration coefficients in waves for the circular symmetric version of the optical system under test obtained by setting to zero the parameters related to tilt and displacement of the two lenses L1 and L2.

Table 2.2: Seidel coefficients and displacement vectors

	$W_{131}(\lambda)$	$W_{222}(\lambda)$	$W_{220,S}(\lambda)$	$\sigma_x(^{\circ})$	$\sigma_y(^{\circ})$
L1	-0.2956	8.2344	6.9259	-1.55	2.44
L2	-0.9095	7.2874	8.1377	1.55	-3.25
TOT	-1.2051	15.5217	15.0636		

The coefficients $W_{220,S}$ reported in Table 2.2 are the sagittal field curvature Seidel terms. The Seidel coefficients related to spherical aberration are not reported because this aberration is independent of the field angle, while those related to distortion are not reported because in the following we neglect this aberration type (the reason for this will be clarified later). The calculation of the vectors $\vec{\sigma} = (\sigma_x, \sigma_y)$ is performed using the real raytracing method exposed in [14]. In this example, the $\vec{\sigma}$ vectors are calculated for the lenses L1 and L2. In Figure 2.3 we show the $\vec{\sigma}$ vectors describing the field perturbations due to decentering (in blue) and tilt (in red) of the lenses L1 (in the upper part) and L2 (in the bottom part of the graph). The final displacement vectors (in green) result from the cumulative effect of decentering and tilt respectively for L1 and L2.

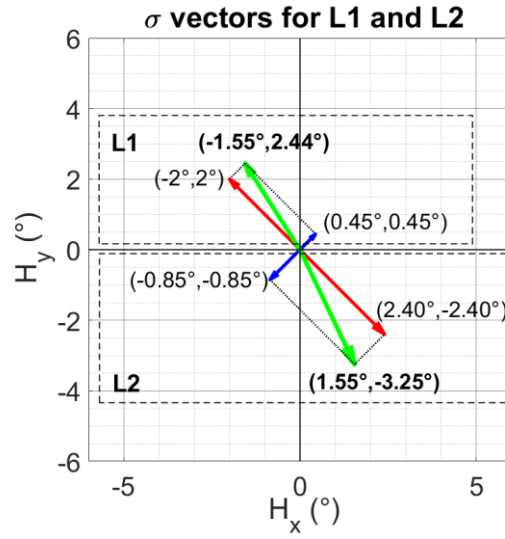


Figure 2.3: Field displacement vectors in the FOV plane.

In what follows, we propose a quantitative comparison between numerical data simulated with raytracing and the surface models shown in Equation 2.18 corresponding to the field behavior of the first nine Zernike coefficients $F_n^m(H_x, H_y)$ in the Fringe indexing scheme. The verification of the proposed analytical expressions is carried out fitting the surface models $F_n^m(H_x, H_y)$ to the ray-trace data. The RMSE (Root Mean Squared Error) and SSE (Sum of Squares due to Error) of the fitting process are indicated to report on the goodness of the fit along with the values of the parameters in the equations of the model surfaces for the full field dependence of the Zernike coefficients. The ray-trace data shown in the following figures from Figure 2.4 to Figure 2.12 result from the computation of the respective Zernike coefficient over a grid of 101×101 field points. For each of these field points, a grid of 128×128 rays is traced through the pupil of the optical system. The resulting wavefront matrix (with 128×128 elements) contains the optical path difference of the rays with respect to the reference sphere in correspondence with their intersection points with it. From the retrieved wavefront data, the mean value is subtracted and it is calculated the square root of its variance. Therefore, the data contained in the wavefront aberration matrix for each field point correspond to the RMS wavefront error referenced to its mean [7, 13]. In the calculation of the optical path difference of light rays, the reference sphere is centered on the point of intersection of the chief ray with the image plane and its radius is equal to the distance between the exit pupil and the image plane of the optical system (in this example the exit pupil is located 322mm before the image plane).

The surface models $F_n^m(H_x, H_y)$ used to fit the ray-trace data and indicated in the text box on the right side of every figure contain only the most relevant contributions from the constituting aberration types. More in detail, considering the example in Figure 2.4 and recalling from Equation 2.18 that the complete expression of $F_0^0(H_x, H_y)$ contains also the contributions of defocus (W_{020}) and primary spherical aberrations (W_{040}) in addition to those of primary field curvature (W_{220}) and secondary spherical aberration (W_{420}), in the surface fitting model $F_0^0(H_x, H_y)$ we have omitted the coefficients W_{020} and W_{040} because, as said before, the last surface to image distance of the optical system has been optimized to minimize the RMS wavefront error and the defocus term in correspondence to the center of the FOV (this implies that the coefficient W_{020} is always zero in the fitting models and the contribution of the coefficient W_{040} is encompassed in the contribution of other terms).

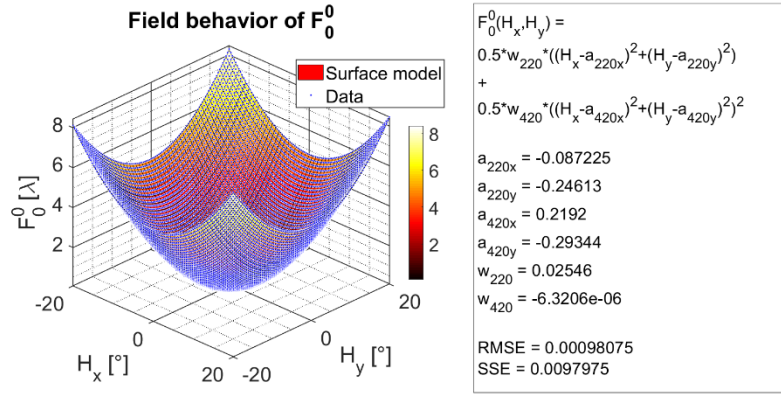


Figure 2.4: Surface fitting of the field dependence of the first Zernike coefficient ($m = 0, n = 0$).

In Figure 2.4 the full field behavior of the first Zernike coefficient is given by the superposition of a quadratic and a quartic surface respectively represented by primary field curvature and secondary spherical aberration. The contribution of oblique spherical aberration (W_{240}) is omitted in this case because such term exhibits the same quadratic dependence on the field coordinates as primary field curvature. These two overlapped polynomial surfaces exhibit different centers of symmetry. The contribution of primary field curvature to the field dependence of this Zernike polynomial is dominant for small FOV. When a wider FOV is required, the contribution from oblique spherical aberration becomes more and more relevant.

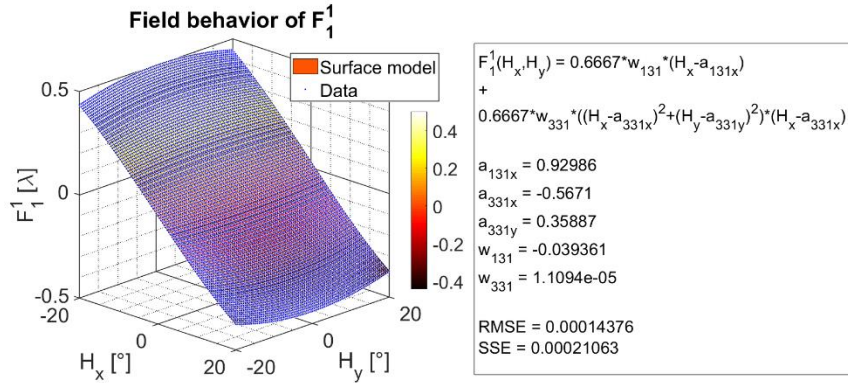


Figure 2.5: Surface fitting of the field dependence of the second Zernike coefficient ($m = 1, n = 1$).

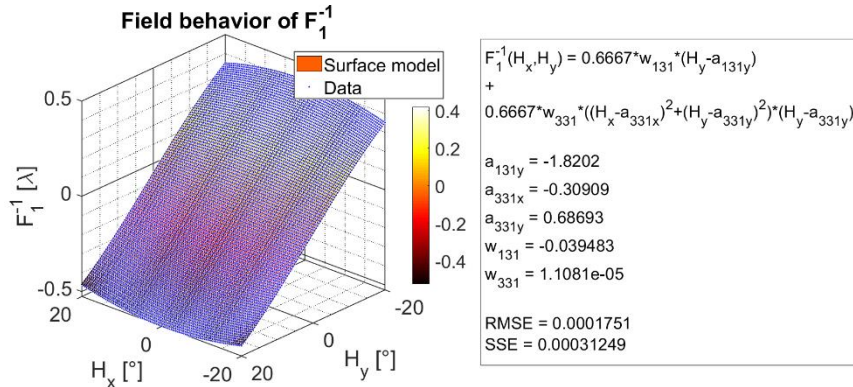


Figure 2.6: Surface fitting of the field dependence of the third Zernike coefficient ($m = -1, n = 1$).

In Figure 2.5 and Figure 2.6 the surface fitting of the second and third Zernike coefficients full-field behavior is carried out neglecting the contribution of distortion $W_D(H_x, H_y, \rho, \vartheta)$ because this aberration type does not degrade the image quality formed by an optical system. In fact, distortion is an image shape aberration measured differently from other aberrations. In particular, it is quantified tracing a grid of chief rays through the optical system and measuring the difference between the predicted and the real landing coordinates of such chief rays in the image plane. On the other side, the RMS wavefront error calculated for a specific field point is the optical path length difference of light rays over the pupil with respect to the chief ray. Therefore, the calculation of the RMS wavefront error does not account for distortion because the reference sphere is centered on the chief ray. From the fitting results in Figure 2.5 and Figure 2.6, we can observe that the centers of symmetry of the field dependence of the second and third Zernike coefficients, describing a displacement along the H_x (Figure 2.5) and H_y (Figure 2.6) directions, is located in the field point $\begin{pmatrix} a_{131,x} \\ a_{131,y} \end{pmatrix} = \begin{pmatrix} 0.92986^\circ \\ -1.8202^\circ \end{pmatrix}$ that represents the center of symmetry and also the node location of primary coma. This is also a consequence of the fact that the fitting models F_1^1 and F_1^{-1} do not include the linear displacement aberration W_{111} (boresight error) because its contribution is subtracted from the calculation of the wavefront phase. This is justified because W_{111} shifts the image location but otherwise has no effect on image quality. Thus, the same result obtained for F_1^1 and F_1^{-1} will be obtained in the case of higher order Zernike terms F_3^1 and F_3^{-1} related to primary coma. Notice that the coefficients W_{131} and W_{331} almost coincide in Figure 2.5 and Figure 2.6. These values are retrieved later as concerns the field dependence of Zernike polynomials F_3^1 and F_3^{-1} because, as explained above, from the calculation of F_1^1 and F_1^{-1} has been removed the contribution given by the field displacement term locating the center of the shifted image plane. The coefficients W_{131} and W_{331} are the wavefront aberration coefficients weighting the magnitude of primary coma and field-cubed secondary coma. The calculation of the primary coma node location using Equation 2.3 gives the following result $\begin{pmatrix} a_{131,x} \\ a_{131,y} \end{pmatrix} = \begin{pmatrix} 0.7877^\circ \\ -1.8568^\circ \end{pmatrix}$. The slight difference with the surface fitting result is probably due to the relevant contribution of higher order aberration terms such as field-cubed coma W_{331} that can shift the node position of a lower order aberration term such as primary coma as discussed in [10].

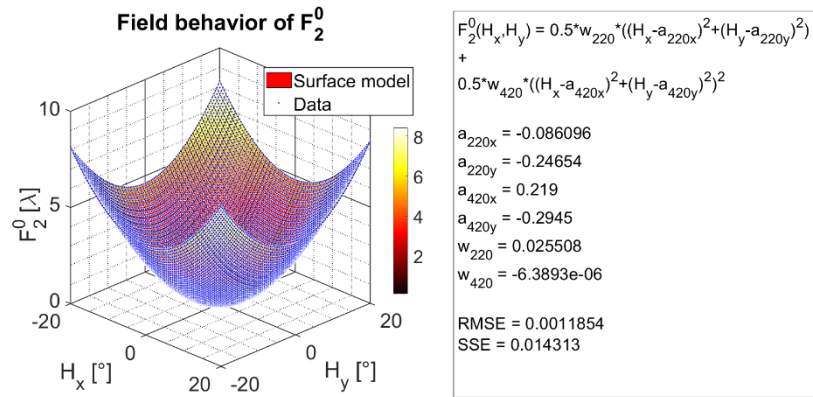


Figure 2.7: Surface fitting of the field dependence of the fourth Zernike coefficient ($m = 0, n = 2$).

The surface fitting result in Figure 2.7, related to the field dependence of the fourth Zernike Fringe coefficient, coincides with the result previously obtained for the first Zernike Fringe coefficient. This is due to the fact that both Zernike coefficients (the first and the fourth) exhibit the same dependence on the field coordinates but different behavior with respect to the pupil coordinates; in fact, the first one is a piston-like term while the fourth is a defocus term. The location of primary field curvature node, calculated using Equation 2.3, is $\begin{pmatrix} a_{220,x} \\ a_{220,y} \end{pmatrix} = \begin{pmatrix} 0.0503^\circ \\ -0.5021^\circ \end{pmatrix}$. This calculation is referenced to the medial focal

surface according to [3] using the data in Table 2.2. The mismatch with the raytracing data $\begin{pmatrix} a_{220,x} \\ a_{220,y} \end{pmatrix} = \begin{pmatrix} -0.086096^\circ \\ -0.24654^\circ \end{pmatrix}$ is attributed to the significance of secondary aberration due to the wide FOV considered in this example.

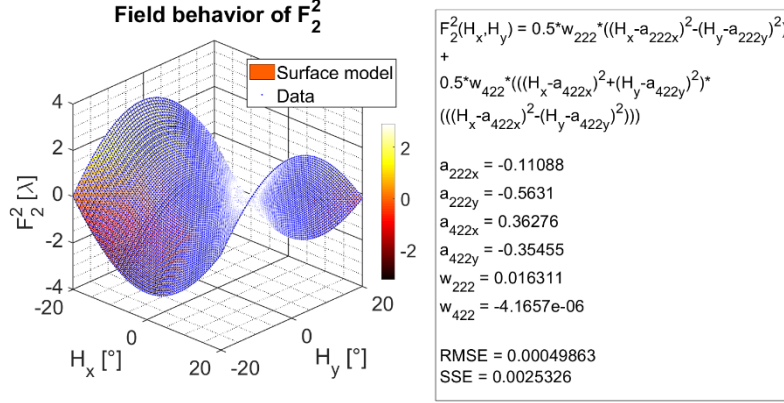


Figure 2.8: Surface fitting of the field dependence of the fifth Zernike coefficient ($m = 2, n = 2$).

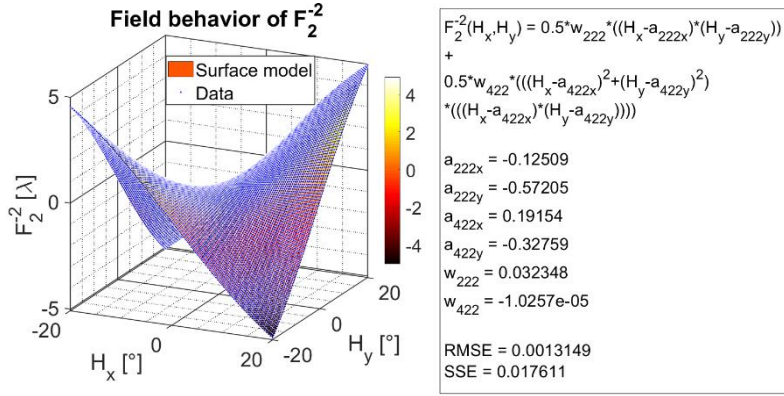


Figure 2.9: Surface fitting of the field dependence of the sixth Zernike coefficient ($m = -2, n = 2$).

Regarding the surface fitting results in Figure 2.8 and Figure 2.9 for the field behavior of Zernike terms related to astigmatism, we can observe that the center of symmetry related to the field dependence of primary astigmatism is approximately the same in the two figures, being $\begin{pmatrix} a_{222,x} \\ a_{222,y} \end{pmatrix} = \begin{pmatrix} -0.11088^\circ \\ -0.5631^\circ \end{pmatrix}$ for Figure 2.8 and $\begin{pmatrix} a_{222,x} \\ a_{222,y} \end{pmatrix} = \begin{pmatrix} -0.12509^\circ \\ -0.57205^\circ \end{pmatrix}$ for Figure 2.9. The calculation of the center of symmetry of binodal astigmatism with Equation 2.3 gives the following result $\begin{pmatrix} a_{222,x} \\ a_{222,y} \end{pmatrix} = \begin{pmatrix} 0.0135^\circ \\ -0.4347^\circ \end{pmatrix}$. The values of the aberration coefficients used for this calculation are relative to the tangential image surface. In practice, $W_{220T,j} = W_{220S,j} + W_{222,j}$ where the values of $W_{220S,j}$ and $W_{222,j}$ are those reported in Table 2.2 (with $j = L1, L2$) and $W_{220T,j}$ are the field curvature aberration coefficients referenced to the tangential image surface. We believe that this mismatch with the raytracing data is due to the important role played by secondary aberrations. In both examples, for larger FOV the contribution of secondary astigmatism (characterized by its center of symmetry not co-located with the center of symmetry of primary astigmatism) becomes more relevant. This can be observed in Figure 2.9 where at the corner of the FOV ($H_x = 20^\circ, H_y = 20^\circ$) the field behavior of the sixth Zernike polynomial tends to be slightly curved.

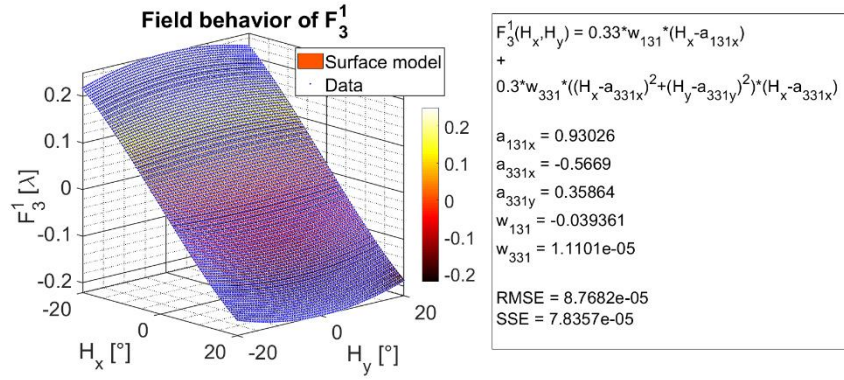


Figure 2.10: Surface fitting of the field dependence of the seventh Zernike coefficient ($m = 1, n = 3$).

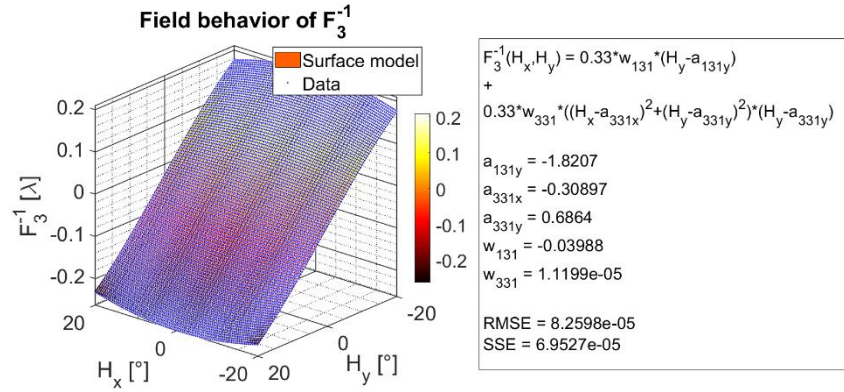


Figure 2.11: Surface fitting of the field dependence of the eighth Zernike coefficient ($m = -1, n = 3$).

The surface fitting results shown in Figure 2.10 and Figure 2.11 for primary coma replicate what has been previously obtained for the second and third Zernike Fringe coefficients related to tilt along the H_x and H_y directions. The location of the center of symmetry in the FOV plane for primary coma is found to be $\begin{pmatrix} a_{131,x} \\ a_{131,y} \end{pmatrix} = \begin{pmatrix} 0.93026^\circ \\ -1.8207^\circ \end{pmatrix}$. For this type of aberration, as mentioned above, the location of its node in the FOV plane coincides with the center of symmetry of the field dependence of the inherent Zernike polynomials. In particular, the coordinate along the H_x direction is determined by the displacement related to the seventh Zernike term and the coordinate along the H_y direction is due to the eighth Zernike polynomial. Observing Figure 2.10 and Figure 2.11, we can see that for larger values of the FOV the surfaces tend to be slightly “distorted”. This behavior is due to the third-order contribution (in the field coordinates) provided by field-cubed coma that has indeed the same field dependence of distortion so that the surfaces exhibit a barrel distortion-like deformation. Nonetheless, we point out that, even though such behavior is reminiscent of distortion, it cannot be ascribed to this type of aberration because, in line with the remarks above, distortion is not calculated with the raytracing method presented here.

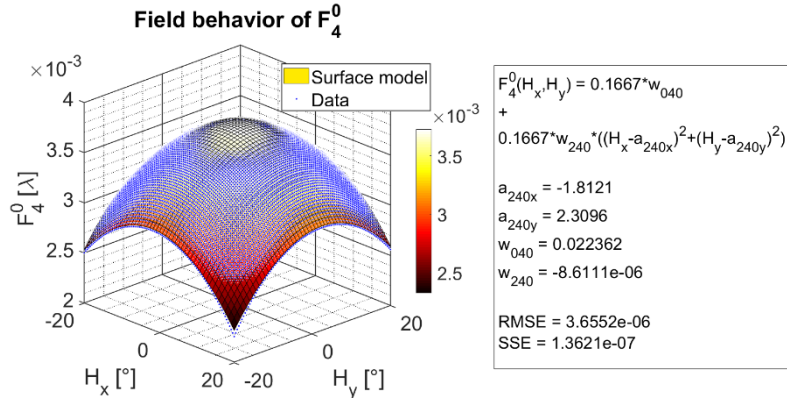


Figure 2.12: Surface fitting of the field dependence of the ninth Zernike coefficient ($m = 0, n = 4$).

Finally, in Figure 2.12 it is shown the surface fitting result obtained for the field behavior of the ninth Zernike Fringe coefficient dominated by the contribution of primary spherical aberration and oblique spherical aberration. In this case, we can observe that the center of symmetry of oblique spherical aberration is retrieved in the field point located in $\begin{pmatrix} a_{240,x} \\ a_{240,y} \end{pmatrix} = \begin{pmatrix} -1.8121^\circ \\ 2.3096^\circ \end{pmatrix}$.

2.5 Digression on the origin of the binodal behavior of primary astigmatism magnitude

The fifth ($m = 2, n = 2$) and sixth ($m = -2, n = 2$) Zernike Fringe terms are involved in the calculation of primary astigmatism. Since it results from the combination of these two terms, it is a vectorial aberration described by a magnitude and a phase. It is well known that the full field dependence of the magnitude of primary astigmatism in symmetric optical systems is described by a paraboloid of revolution. On the opposite side, the full field behavior of the magnitude of primary astigmatism in optical systems without symmetry (described by NAT [3]) exhibits a binodal behavior. This behavior consists of the occurrence of two nodes in the FOV. The nodes are specific locations in the FOV where the magnitude of primary astigmatism vanishes. The binodal behavior of primary astigmatism is not only characteristic of asymmetric systems consisting of tilted and decentered circular symmetric surfaces, but also of asymmetric systems containing tilted and decentered surfaces characterized by double-plane symmetry. This last class of optical systems is the subject of investigation in the following chapter where a mathematical description of this phenomenon is given. In this context, we suggest an intuitive explanation of the occurrence of this behavior.

The origin of the binodal behavior of primary astigmatism magnitude can be understood by studying the full field behavior of the tangential and sagittal field curvature surfaces associated with an asymmetric optical system. In this context, the tangential and sagittal directions are defined to lie along the y and x axis of the FOV plane. In particular, the tangential plane is the yz plane and the sagittal plane is orthogonal to the yz plane and intersects the center of the entrance pupil [13]. A short digression on the roles of astigmatism and field curvature in circular symmetric systems is a preliminary and necessary step to understand better the following considerations related to asymmetric optical systems.

Studying the aberration behavior of an optical system, field curvature and astigmatism are coupled together giving rise to tangential and sagittal curved image surfaces. These two surfaces are distinct if primary astigmatism does not vanish and, in this case, the separation between them in the meridional and the sagittal planes increases quadratically with the field coordinates. In a circular symmetric optical system, the study of aberrations is restricted to the meridional plane and consequently, the tangential and sagittal focal surfaces are

commonly represented as two different focal lines, whose curvatures depend jointly on the aberration coefficients of astigmatism and field curvature. These surfaces (or lines if the analysis is restricted to the meridional plane) are characterized by the fact that the distance between the tangential surface and the Petzval surface is three times the distance between the sagittal surface and the Petzval surface. Additionally, the intermediate surface between the sagittal and tangential surfaces is defined as the medial focal surface. Considering the asymmetric optical system of the previous section (Figure 2.2), the corresponding circular symmetric system is obtained setting to zero the tilt ($\tau_{x,1} = 0, \tau_{y,1} = 0, \tau_{x,2} = 0, \tau_{y,2} = 0$) and displacement parameters ($D_{x,1} = 0, D_{y,1} = 0, D_{x,2} = 0, D_{y,2} = 0$) of the lenses L1 and L2. The calculation of the tangential and sagittal focal surfaces in the optical system thus obtained, can be carried out by retrieving the distance between the paraxial image position and the image plane in the tangential plane and the sagittal plane respectively. The paraxial image position is located where the marginal ray height is equal to zero. Reiterating the calculations of these two focal distances with respect to a grid of field points in the FOV, the results shown in Figure 2.13 are generated.

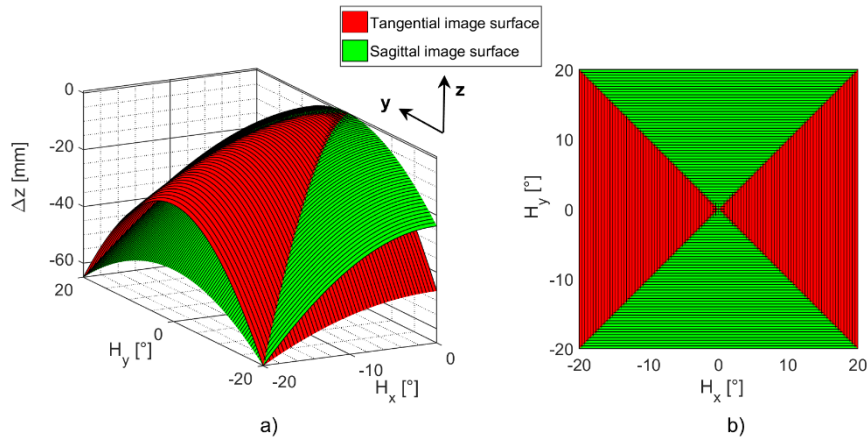


Figure 2.13: a) Tangential and sagittal image surfaces for the circular symmetric version of the optical system analyzed in section 2.4. The surface plot is cut along the tangential plane (yz) to emphasize that, along the y direction, the curvature of the tangential image surface is larger than that of the sagittal image surface. The height of the surfaces is measured in mm from the paraxial image surface. b) The top view emphasizes the coincidence of these two image surfaces at the origin of the FOV.

In correspondence to the origin of the FOV, there is no separation between the two focal surfaces because primary astigmatism vanishes on-axis in symmetric systems while considering any other off-axis field point, the separation between the sagittal and tangential image surfaces is proportional to the magnitude of astigmatism. The coincidence between these two surfaces in the origin of the FOV is evident if we observe Figure 2.13b while in Figure 2.13a it is put in evidence the separation between these two surfaces for any off-axis field point.

On the other hand, asymmetric optical systems with displaced rotationally symmetric surfaces are characterized by a separation between the sagittal and tangential image surfaces even in correspondence with the origin of the FOV. This situation is due to the fact that in asymmetric optical systems the field dependence of field curvature and astigmatism are perturbed by different field displacement vectors (indicated with $\vec{\sigma}$ in NAT). These vectors shift differently the sagittal and tangential image surfaces and the nodes of primary astigmatism magnitude are to be found along the curves defined by the intersection of the sagittal and tangential image surfaces themselves. In the points of intersection of these two image surfaces, primary astigmatism vanishes by definition. The nodes of primary astigmatism magnitude thus coincide with the field points where the distance between the paraxial image surface and the sagittal focal surface equals the distance between the paraxial image surface and the tangential focal surface. In the following Figure 2.14, it is

shown the overlap between the perturbed tangential and sagittal focal surfaces in the case of the asymmetric optical system shown in the example in Figure 2.2 in the previous section.

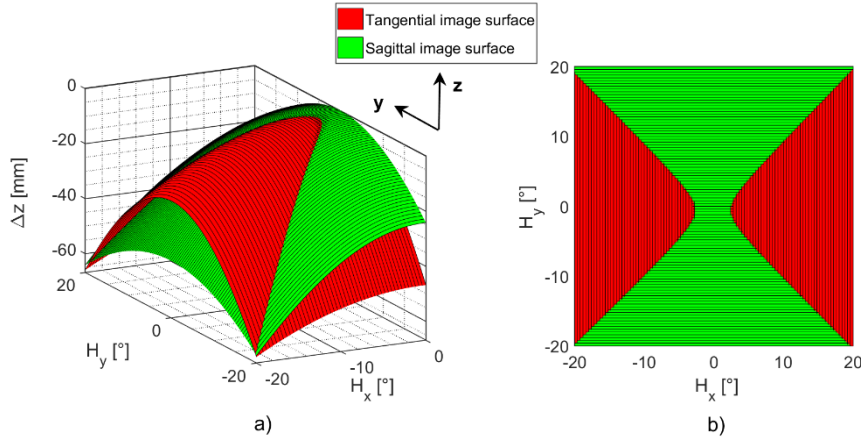


Figure 2.14: a) Tangential and sagittal image surfaces for the asymmetric optical system analyzed in the previous section. b) Top view of the tangential and sagittal image surfaces.

Figure 2.14a shows again that in the yz plane the curvature of the tangential image surface is larger than that of the sagittal image surface but these two surfaces are displaced due to the lack of symmetry. Furthermore, in Figure 2.14b it is emphasized that there is no longer coincidence between the tangential and sagittal focal surfaces at the origin of the FOV. In the following Figure 2.15, we show the field-dependent magnitude of primary astigmatism (for the asymmetric system of the previous section) overlapped to the tangential and sagittal focal surfaces to show more clearly that the location of the nodes of primary astigmatism is in correspondence to the intersection between the tangential and sagittal focal surfaces.

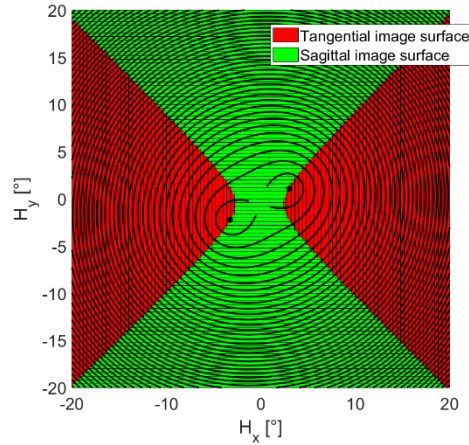


Figure 2.15: Contour plot of the full field-dependent magnitude of primary astigmatism overlapped to the tangential and sagittal focal surfaces. The black dots indicate the positions of the nodes of primary astigmatism.

In the next chapter, we will provide a mathematical model to describe such behavior in double-plane symmetric optical systems characterized by perturbations of the optical surfaces (such model is more general and adaptable to the class of optical systems dealt with in the present chapter).

2.6 Full-field behavior of higher-order Zernike terms

The analysis of the full field behavior of the Zernike terms from the tenth to the fifteenth order (in the Fringe indexing scheme) requires first to obtain their respective expressions in scalar form starting from Equation 2.4. In particular, the field behavior of these Zernike terms depends on secondary aberration terms such as elliptical coma $W_{EC}(H_x, H_y, \rho, \vartheta)$, secondary astigmatism $W_{OSA2}(H_x, H_y, \rho, \vartheta)$ and field linear fifth order coma $W_{C5}(H_x, H_y, \rho, \vartheta)$ in addition to the aberration terms already introduced in section 2.3. Elliptical coma (also known as trefoil) has the following full-field expression

$$W_{EC}(H_x, H_y, \rho, \vartheta) = W_{333} \left[(\vec{H} - \vec{a}_{333}) \cdot \vec{\rho} \right]^3 = W_{333} \left[\begin{pmatrix} H_x - a_{333,x} \\ H_y - a_{333,y} \end{pmatrix}^T \cdot \begin{pmatrix} \rho \cos \vartheta \\ \rho \sin \vartheta \end{pmatrix} \right]^3 =$$

$$W_{333} \left[(H_x - a_{333,x}) \rho \cos \vartheta + (H_y - a_{333,y}) \rho \sin \vartheta \right]^3 =$$

$$W_{333} \left\{ \begin{aligned} & \left[\frac{(H_x - a_{333,x})^3}{4} - \frac{3}{4} (H_x - a_{333,x}) (H_y - a_{333,y})^2 \right] \rho^3 \cos(3\vartheta) + \\ & \left[-\frac{(H_y - a_{333,y})^3}{4} + \frac{3}{4} (H_y - a_{333,y}) (H_x - a_{333,x})^2 \right] \rho^3 \sin(3\vartheta) + \\ & \left[\frac{3}{4} (H_x - a_{333,x})^3 + \frac{3}{4} (H_x - a_{333,x}) (H_y - a_{333,y})^2 \right] \rho^3 \cos \vartheta + \\ & \left[\frac{3}{4} (H_y - a_{333,y})^3 + \frac{3}{4} (H_y - a_{333,y}) (H_x - a_{333,x})^2 \right] \rho^3 \sin \vartheta \end{aligned} \right\}$$

Equation 2.19

One can observe that the final expression in Equation 2.19 (in curly bracket) is composed of two pairs of terms characterized by two different behaviors with respect to the pupil coordinates. The first pair of terms depends on $\rho^3 \cos 3\vartheta$ and $\rho^3 \sin 3\vartheta$, while the second pair of terms depends on $\rho^3 \cos \vartheta$ and $\rho^3 \sin \vartheta$. Therefore, the second couple of terms exhibits the same pupil dependence characteristic of field cubed coma W_{CC} and, for this reason, in section 2.3 the coefficient for medial field cubed coma has been defined as $W_{331M} = W_{331} + \frac{3}{4} W_{333}$, namely the sum of two contributions, of which the first strictly derives from field cubed coma, while the second is due to elliptical coma. On the other side, the other two terms in the final expression of Equation 2.19 (those depending on $\rho^3 \cos 3\vartheta$ and $\rho^3 \sin 3\vartheta$) are to be considered as purely trefoil terms that are involved in the calculation of the field dependence of the tenth and eleventh Zernike coefficient in the Fringe indexing scheme (Figure 2.16-Figure 2.17).

The full field expression of secondary oblique spherical aberration is shown in Equation 2.20.

$$\begin{aligned}
 W_{OSA2}(H_x, H_y, \rho, \vartheta) = & \\
 W_{242} \left[\left(\vec{H} - \vec{a}_{242} \right) \cdot \vec{\rho} \right]^2 (\vec{\rho} \cdot \vec{\rho}) = & \\
 W_{242} \rho^4 \left\{ \frac{(H_x - a_{242,x})^2 + (H_y - a_{242,y})^2}{2} + \right. & \\
 \left. \frac{1}{2} \left[(H_x - a_{242,x})^2 - (H_y - a_{242,y})^2 \right] \cos(2\vartheta) + \right. & \\
 \left. (H_x - a_{242,x})(H_y - a_{242,y}) \sin(2\vartheta) \right\} &
 \end{aligned}$$

Equation 2.20

The full field expression of linear field fifth order coma is the following

$$\begin{aligned}
 W_{C5}(H_x, H_y, \rho, \vartheta) = & \\
 W_{151} \left[\left(\vec{H} - \vec{a}_{151} \right) \cdot \vec{\rho} \right] \rho^4 = & \\
 W_{151} \left[(H_x - a_{151,x}) \rho^5 \cos \vartheta + (H_y - a_{151,y}) \rho^5 \sin \vartheta \right] &
 \end{aligned}$$

Equation 2.21

Finally, it is also reported in Equation 2.22 the full field expression for an aberration term of 8th order, namely tertiary oblique spherical aberration whose 4th order field dependence is suited to describe the full field behavior of the twelfth and thirteenth Zernike Fringe coefficients.

$$\begin{aligned}
 W_{OSA3}(H_x, H_y, \rho, \vartheta) = & \\
 \frac{W_{442}}{2} \left[(H_x - a_{442,x})^2 + (H_y - a_{442,y})^2 \right] & \\
 \left\{ \left[(H_x - a_{442,x})^2 - (H_y - a_{442,y})^2 \right] \rho^4 \cos(2\vartheta) + \right. & \\
 \left. 2(H_x - a_{442,x})(H_y - a_{442,y}) \rho^4 \sin(2\vartheta) \right\} &
 \end{aligned}$$

Equation 2.22

Like what has been shown in section 2.3, the full field dependence of the Zernike Fringe coefficients (from the 10th to the 15th) can be calculated again using Equation 2.17 where the complete scalar expression of the wavefront aberration function (also including the expressions from Equation 2.19 to Equation 2.22) is projected onto the Zernike basis functions. Thus, the full field expression of these six Zernike terms is made explicit in the following Equation 2.23 that is the continuation of Equation 2.18 in section 2.3.

$$\begin{aligned}
 W_{10-15}(H_x, H_y, \rho, \vartheta) = & \\
 & W_{333} \left[\frac{(H_x - a_{333,x})^3}{4} - \frac{3}{4}(H_x - a_{333,x})(H_y - a_{333,y})^2 \right] Z_3^3(\rho, \vartheta) + \\
 & W_{333} \left[-\frac{(H_y - a_{333,y})^3}{4} + \frac{3}{4}(H_y - a_{333,y})(H_x - a_{333,x})^2 \right] Z_3^{-3}(\rho, \vartheta) + \\
 & \left\{ \frac{W_{442}}{2} \left[(H_x - a_{442,x})^2 + (H_y - a_{442,y})^2 \right] \left[(H_x - a_{442,x})^2 - (H_y - a_{442,y})^2 \right] + \right. \\
 & \left. \frac{W_{242}}{2} \left[(H_x - a_{242,x})^2 - (H_y - a_{242,y})^2 \right] \right\} Z_4^2(\rho, \vartheta) + \\
 & \left\{ W_{442} \left[(H_x - a_{442,x})^2 + (H_y - a_{442,y})^2 \right] \left[(H_x - a_{442,x})(H_y - a_{442,y}) \right] + \right. \\
 & \left. W_{242} \left[(H_x - a_{242,x})(H_y - a_{242,y}) \right] \right\} Z_4^{-2}(\rho, \vartheta) + \\
 & W_{151}(H_x - a_{151,x}) Z_5^1(\rho, \vartheta) + W_{151}(H_y - a_{151,y}) Z_5^{-1}(\rho, \vartheta)
 \end{aligned}$$

Equation 2.23

The verification of these expressions can be again carried out fitting them to raytracing data calculated for the example of section 2.4. These results are reported in the following figures from Figure 2.16 to Figure 2.18.

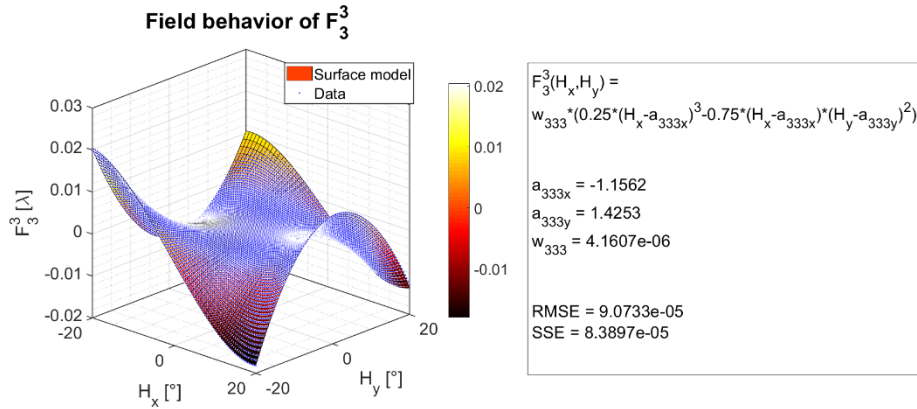


Figure 2.16: Surface fitting of the field dependence of the tenth Zernike coefficient ($m = 3, n = 3$).

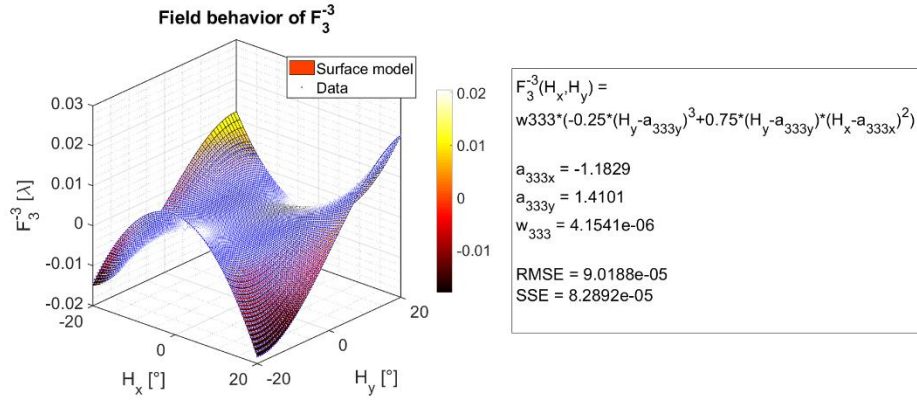


Figure 2.17: Surface fitting of the field dependence of the eleventh Zernike coefficient ($m = -3, n = 3$).

A proper combination of the full field dependences of the tenth and eleventh Zernike coefficients gives rise to the vectorial expression of trefoil aberration. This type of aberrations is characterized by the presence of three nodes over the FOV in an asymmetric optical system.

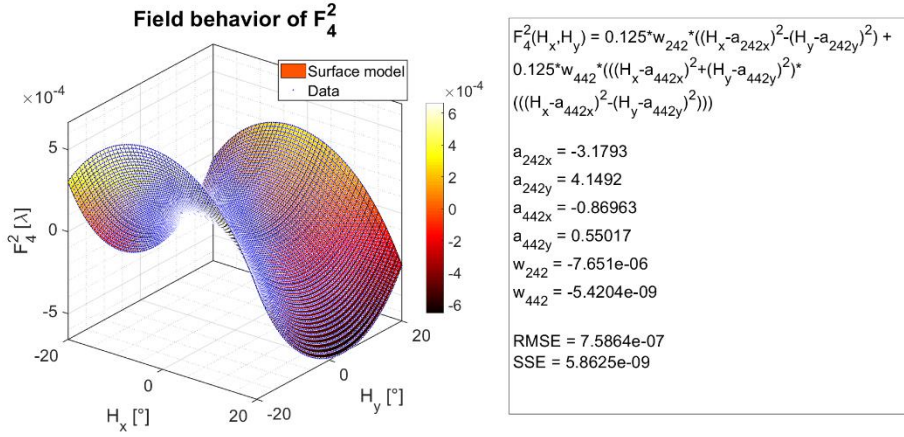


Figure 2.18: Surface fitting of the field dependence of the twelfth Zernike coefficient ($m = 2, n = 4$).

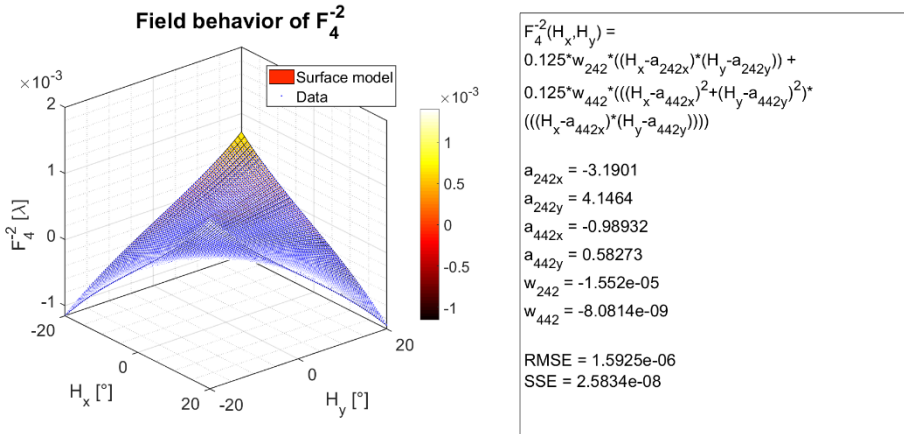


Figure 2.19: Surface fitting of the field dependence of the thirteenth Zernike coefficient ($m = -2, n = 4$).

The combination of the field dependences of these two last Zernike coefficients results in another vectorial aberration whose magnitude behaves similarly in magnitude as primary

astigmatism since, also in this case, two nodes occur in the FOV of an asymmetric optical system.

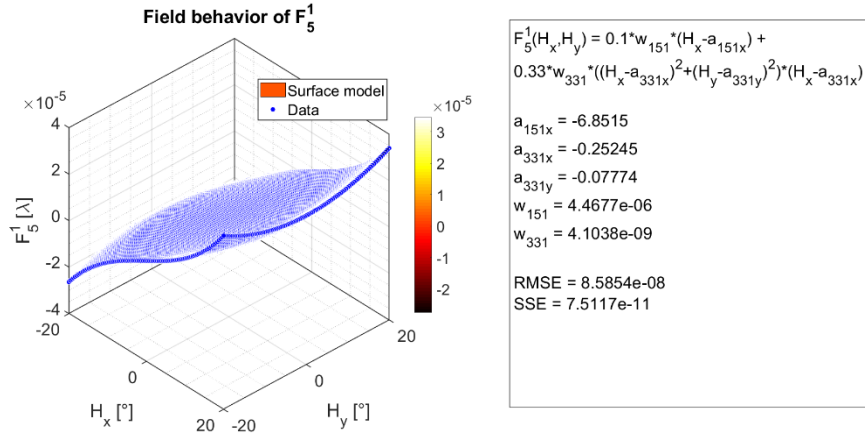


Figure 2.20: Surface fitting of the field dependence of the fourteenth Zernike coefficient ($m = 1, n = 5$).

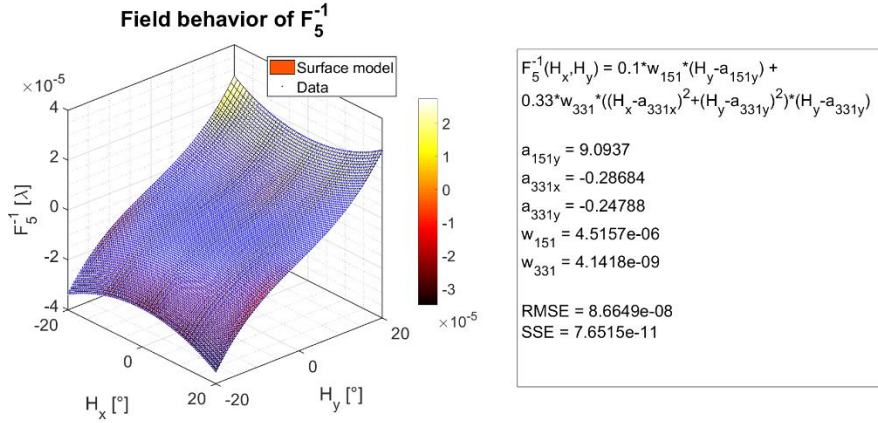


Figure 2.21: Surface fitting of the field dependence of the fifteenth Zernike coefficient ($m = -1, n = 5$).

Finally, the combination of the field dependences of the fourteenth and fifteenth Zernike terms originates a vectorial aberration whose field dependence features one single node.

The field behavior of the Zernike coefficients (in the Fringe indexing scheme) shown up to this point is useful to determine the occurrence of nodes for specific aberration types in asymmetric optical systems constituted by arbitrarily tilted and decentered surfaces. The equations of NAT [3] are applicable to locate the nodes of aberrations in the case only primary aberrations are of concern. If higher-order aberrations are relevant in an asymmetric optical system, the positions of their inherent nodes can be found with real raytracing computing the respective Zernike coefficients in correspondence of a grid of points over the FOV.

2.7 Conclusions

In this chapter, we presented a series of scalar analytical formulas describing the full field dependence of Zernike polynomials that derives from the field behavior of their inherent coefficients in asymmetric optical systems characterized by tilted and decentered circular symmetric surfaces. The starting point of such derivation is a modified version of the vectorial wavefront aberration function proposed in the Nodal Aberration theory. Such function is first transformed into its scalar counterpart, then it is projected onto the basis of Zernike polynomials in the Fringe indexing scheme in order to retrieve the full field

behavior of individual Zernike coefficients. It is highlighted the presence of centers of symmetry for the field dependence of individual Zernike terms when they are described by the contribution of the only primary aberrations. Furthermore, it is emphasized that secondary aberrations change the symmetry of the field behavior of Zernike coefficients adding higher-order overlay field surfaces with different centers of symmetry.

Chapter 3 Full-field dependence of primary aberrations in perturbed double-plane symmetric systems with a circular pupil

This chapter [15] deals with the theory of primary aberrations for perturbed double-plane symmetric optical systems consisting of a combination of tilted and decentered surfaces and a circular pupil. First, the analytical expressions describing the full field behavior of Zernike polynomials are derived from the fourth-order wavefront aberration function for this class of optical systems. Then, such expressions are combined to retrieve the full field dependence of primary coma, primary astigmatism, and field curvature. They are described by an elliptical conic-shaped surface with a variable apex location over the field of view, by a binodal surface with two nodes over the field of view, and by a general elliptical surface with one node. The proposed analytical expressions provide a better understanding of the primary aberration behavior for these systems and can be of great use in their optical design and aberration correction. An optical system constituted by a pair of tilted and decentered biconic lenses is studied to validate the proposed expressions.

3.1 Introduction

Over the past century, numerous authors have investigated the aberrations of optical systems emphasizing their intrinsic connection with the symmetry properties of the systems themselves. In this regard, the contributions of Hopkins, Buchdahl, Burfoot, Wynne, Sands [5, 16, 17, 18, 19] are of outstanding relevance. More recently, the introduction of a vectorial formulation of the wavefront aberration function due to Shack, initially for the only circular symmetric systems, has later paved the way to the investigation of optical systems without symmetry. Of particular relevance are the Nodal Aberration Theory (NAT) [3, 4] describing asymmetric optical systems with tilted and decentered circular symmetric surfaces, the study of bilateral symmetric and anamorphic systems [20, 21, 22, 23] and the investigation of optical systems with freeform surfaces [24, 25].

The goal of the present paper is to study the primary aberrations behavior in perturbed double-plane symmetric systems. This class of optical systems is characterized by a full double-plane symmetry in the absence of geometrical perturbations, such as tilts and displacements of the optical components with respect to the mechanical axis of the system itself. The presence of misalignments of the optical surfaces reduces the general double-plane symmetry of the system inducing characteristic modifications to the field dependence of primary aberrations, particularly evident in the displacement of the location of the node in the FOV. We derive a series of analytical expressions (with effective coefficients) describing the full field behavior of primary aberrations for this category of optical systems with geometrical perturbations. Our treatment starts from the general expression of the full field wavefront aberration function for double-plane symmetric systems deriving from the

theory of Hamiltonian optics. This expression is first generalized to include the description of perturbed double-plane symmetric systems, then it is projected onto the basis of Zernike polynomials to retrieve the latter's full field dependence for the optical systems of our concern. Finally, the full field Zernike polynomials are properly combined to elaborate a more advantageous analytical formulation (with effective coefficients) of the primary aberrations in perturbed double-plane symmetric systems. Formally, these expressions are analogous to the corresponding expression of the wavefront aberration function for a circular symmetric system in the Hopkins notation where the coefficients of the aberration terms are represented by the Seidel sums [5]. Such formal similarity between the primary aberration functions of circular symmetric and perturbed double-plane symmetric systems fosters a better understanding of the aberration behavior of the latter. The retrieved analytical expressions can be applied fruitfully in the optical design and optimization of this class of systems providing a better insight about the general behavior of their primary aberrations also in relation to the geometrical perturbations of their optical components. Moreover, it is shown that the effective aberration coefficients, present in the proposed analytical expressions, can be easily calculated by sampling the Zernike coefficients in correspondence with few field points of interest in the FOV. This is possible only in the case higher-order aberrations are negligible.

We show that in perturbed double-plane symmetric systems the full field magnitude of primary astigmatism exhibits a binodal behavior described by a surface model similar to a Cassini surface (that on the contrary well describes the binodal behavior of primary astigmatism in asymmetric optical systems with tilted and decentered surfaces as shown in NAT). On the other side, primary coma magnitude and field curvature full-field dependence are described by an elliptic conic-shaped surface and by a general ellipsoidal surface respectively. Distortion is neglected because it does not determine image blur, being an image shape aberration. In fact, its effect is to change the shape of the image introducing an error mapping between points in the object plane and corresponding points in the image plane. In the final section, an optical system consisting of a combination of biconic lenses (both tilted and decentered) is studied in order to verify the primary full field wavefront aberration function proposed in this paper. It is worth noting that this class of optical systems can be alternatively studied in the framework of the generalized nodal aberration theory [25] as shown in [26] relatively to an optical system constituted by double-plane symmetric surfaces. The results of [25] are related to freeform systems and consequently, no assumptions are made on the underlying symmetry of the optical systems under study. In the present paper, we follow a different approach for the analysis of the optical systems of our interest. In fact, we start our analysis by making specific assumptions on the symmetry properties of such systems and additionally we make use of the principles exposed in the theory of Hamiltonian optics [16] to introduce their general wavefront aberration function. We believe that our proposal, to describe the full field dependence of primary aberrations with functions with new effective coefficients obtained from recombination of Zernike polynomials, facilitates the understanding of the class of optical systems of our interest and can be beneficial for their design and optimization.

3.2 Full-field dependence of Zernike polynomials

According to the theory of Hamiltonian Optics [16], the 4th order wavefront aberration function of a double-plane symmetric system (supposing that lower orders are absent) is obtained combining the six reflection invariants $\rho_x^2, \rho_x H_x, H_x^2, \rho_y^2, \rho_y H_y, H_y^2$. The resulting equation is constituted by 21 terms (two of which are redundant and for this reason are omitted) weighted by proper coefficients with the notation w_{klpq} where the subscripts k, l, p and q indicate respectively the power to which the ray coordinates H_x, H_y, ρ_x and ρ_y are raised. Since we are concerned with 4th order (or primary) aberrations, the values of the subscripts k, l, p and q are constrained by the condition $k + l + p + q = 4$. Such equation is written as follows

$$\begin{aligned}
 W(H_x, H_y, \rho_x, \rho_y) = & w_{0040}\rho_x^4 + w_{0004}\rho_y^4 + w_{0022}\rho_x^2\rho_y^2 + \\
 & w_{1030}H_x\rho_x^3 + w_{0121}H_y\rho_x^2\rho_y + w_{1012}H_x\rho_x\rho_y^2 + w_{0103}H_y\rho_y^3 + \\
 & w_{2020}H_x^2\rho_x^2 + w_{0202}H_y^2\rho_y^2 + w_{0220}H_y^2\rho_x^2 + w_{2002}H_x^2\rho_y^2 + w_{1111}H_xH_y\rho_x\rho_y + \\
 & w_{3010}H_x^3\rho_x + w_{0301}H_y^3\rho_y + w_{1210}H_xH_y^2\rho_x + \\
 & w_{2101}H_x^2H_y\rho_y + w_{4000}H_x^4 + w_{2200}H_x^2H_y^2 + w_{0400}H_y^4
 \end{aligned}$$

Equation 3.1

The last three coefficients w_{4000} , w_{2200} and w_{0400} are set to zero since they correspond to quartic piston terms. The remaining 16 coefficients w_{klpq} (with $k = 0,1,2,3$, $l = 0,1,2,3$, $p = 0,1,2,3,4$ and $q = 0,1,2,3,4$) can be collected according to their dependence on the pupil and field coordinates in order to clarify their respective meaning. In detail, the first three coefficients describe spherical aberrations, the coefficients from the fourth to the seventh indicate coma aberration, those from the eighth to the twelfth describe a combination of astigmatism and field curvature, and finally, the coefficients from the thirteenth to the sixteenth quantify distortion.

We are interested in the analysis and description of double-plane symmetric systems whose symmetry properties are perturbed by internal misalignments such as eventual tilts and displacements of the optical components. In the most general case, when the components of a symmetric optical system are arbitrarily perturbed, the optical system itself is converted into an asymmetric system. In this context, we assume that this situation can be described as a perturbation to the FOV of the optical system. In other words, we assume that the presence of geometrical perturbations to the components of an otherwise double-plane symmetric system induces displacements to the FOV (or to the image plane). Upon closer inspection, this assumption is very similar to the one on which the NAT is founded [3]. According to NAT, an asymmetric optical system with tilted and decentered circular symmetric surfaces can be studied summing (over the surfaces of the system) the net contributions to the full field-dependent wavefront aberration function. Thus, the resulting aberration function is the sum of circular symmetric contributions oriented along their axis of symmetries whose intersections with the image plane define different nodes' positions. The shift of these nodes from the origin of the FOV is described with an equivalent displacement (perturbation) of the image plane. Despite this, in the optical systems of our concern, we do not address the problem of finding the individual contribution of each surface to the total wavefront aberration function, but we focus our attention directly on the total full field wavefront aberration function of the optical system. Bearing in mind these considerations, in agreement with the principles of Hamiltonian optics [16], the optical systems of our interest can be described generalizing Equation 3.1 including odd and even lower order terms to account for the net perturbations occurring to the image plane of the system. As a matter of fact, the occurrence of geometrical perturbations (such as tilt and decentering of surfaces constituting the optical system) generates several terms in the polynomial expansion of the wavefront error (Equation 3.1) containing lower orders in the field coordinates H_x and H_y . Such lower order terms, properly weighted by their respective coefficients w_{klpq} (where $k + l + p + q \leq 3$), reveal that geometrical perturbations to the optical surfaces are at the origin of corresponding perturbations to the system FOV. In practice, the field of view (or equivalently the image plane) of such an optical system is displaced as a consequence of the reduction of symmetry determined by perturbed surfaces in the system itself. This state of affairs translates into the following modification of Equation 3.1 where terms of order lower than 4 are included in the wavefront aberration expansion as shown in Equation 3.2

$$\begin{aligned}
W(H_x, H_y, \rho_x, \rho_y) = & w_{0040}\rho_x^4 + w_{0004}\rho_y^4 + w_{0022}\rho_x^2\rho_y^2 + \\
& (w_{1030}H_x\rho_x^3 + \mathbf{w}_{0030}\rho_x^3) + (w_{0121}H_y\rho_x^2\rho_y + \mathbf{w}_{0021}\rho_x^2\rho_y) + \\
& (w_{1012}H_x\rho_x\rho_y^2 + \mathbf{w}_{0012}\rho_x\rho_y^2) + (w_{0103}H_y\rho_y^3 + \mathbf{w}_{0003}\rho_y^3) + \\
& (w_{2020}H_x^2\rho_x^2 + \mathbf{w}_{1020}H_x\rho_x^2 + \mathbf{w}_{0020}\rho_x^2) + \\
& (w_{0202}H_y^2\rho_y^2 + \mathbf{w}_{0102}H_y\rho_y^2 + \mathbf{w}_{0002}\rho_y^2) + \\
& (w_{0220}H_y^2\rho_x^2 + \mathbf{w}_{0120}H_y\rho_x^2 + \mathbf{w}_{0020}\rho_x^2) + \\
& (w_{2002}H_x^2\rho_y^2 + \mathbf{w}_{1002}H_x\rho_y^2 + \mathbf{w}_{0002}\rho_y^2) + \\
& (w_{1111}H_xH_y\rho_x\rho_y + \mathbf{w}_{1011}H_x\rho_x\rho_y + \mathbf{w}_{0111}H_y\rho_x\rho_y + \mathbf{w}_{0011}\rho_x\rho_y) + \\
& (w_{3010}H_x^3\rho_x + \mathbf{w}_{2010}H_x^2\rho_x + \mathbf{w}_{1010}H_x\rho_x + \mathbf{w}_{0010}\rho_x) + \\
& (w_{0301}H_y^3\rho_y + \mathbf{w}_{0201}H_y^2\rho_y + \mathbf{w}_{0101}H_y\rho_y + \mathbf{w}_{0001}\rho_y) + \\
& (w_{1210}H_xH_y^2\rho_x + \mathbf{w}_{1110}H_xH_y\rho_x + \mathbf{w}_{1010}H_x\rho_x + \mathbf{w}_{0210}H_y^2\rho_x + \mathbf{w}_{0110}H_y\rho_x + \mathbf{w}_{0010}\rho_x) + \\
& (w_{2101}H_x^2H_y\rho_y + \mathbf{w}_{2001}H_x^2\rho_y + \mathbf{w}_{1101}H_xH_y\rho_y + \mathbf{w}_{1001}H_x\rho_y + \mathbf{w}_{0101}H_y\rho_y + \mathbf{w}_{0001}\rho_y)
\end{aligned}$$

Equation 3.2

Equation 3.2 contains 4th order terms ($k + l + p + q = 4$) collected in round brackets with the inherent lower order terms ($k + l + p + q \leq 3$) highlighted with bolded characters. The lower order terms derive essentially from a polynomial expansion in the only field coordinates H_x and H_y of the 4th order terms present in Equation 3.1. It is worth noting that in general the coefficients of the 4th order terms present in Equation 3.2 are different from those of the respective terms in Equation 3.1. They differ, essentially, because Equation 3.1 describes the 4th order aberration function of a purely double-plane symmetric system, while Equation 3.2 describes the 4th order aberration function of an asymmetric optical system (that we have defined here as perturbed double-plane symmetric) whose lower order terms arise from the perturbation induced to the original double-plane symmetry. For the sake of clarity and at the risk of being redundant, it is important to remark again that the inclusion of these lower order terms in the wavefront aberration function is to some extent similar to what is done in NAT where the displacement of the field coordinates (due to misalignments of the circular symmetric surfaces constituting an optical system) results exactly in a polynomial expansion of the aberration terms with respect to the field coordinates H_x and H_y [3].

Equation 3.2 is the starting point for the derivation of the field behavior of Zernike polynomials. In this work, we use the Fringe indexing scheme. These analytical expressions are necessary to retrieve the full field formulas describing primary aberrations for the category of optical systems about which we are concerned. The full field dependence of Zernike polynomials in perturbed double-plane symmetric systems characterized by the presence of misaligned components can be retrieved projecting the wavefront aberration function in Equation 3.2 (transformed in polar coordinates) onto the basis functions constituted by the polynomials themselves analogously to what has been done in the previous chapter for the field dependence of Zernike polynomials in asymmetric systems with rotationally symmetric surfaces. It is worth mentioning that, to our knowledge, the idea of expanding the wavefront deformation in Zernike polynomials over the pupil and the FOV was proposed for the first time in [27]. Assuming that the optical systems of our interest are characterized by a circular stop, the coordinates ρ_x and ρ_y denoting the intersection point of a ray with the aperture stop, can be written equivalently in polar coordinates as $\rho_x = \rho \cos \vartheta$ and $\rho_y = \rho \sin \vartheta$ where the angle ϑ is measured counter clockwise from the axis ρ_x . The basis of the Zernike polynomials in the Fringe indexing scheme is expressed as follows

$$C_n^m(\rho, \vartheta) = F_n^m R_n^m(\rho) G^m(\vartheta) = F_n^m Z_n^m(\rho, \vartheta)$$

Equation 3.3

$Z_n^m(\rho, \vartheta)$ results from the product between $R_n^m(\rho)$ and $G^m(\vartheta)$ where the former is a polynomial function depending on the only radial coordinate ρ of the pupil and the latter is a trigonometric function depending on the azimuthal coordinate ϑ , while F_n^m denotes the Zernike coefficients that are multiplicative factors weighting the respective polynomials. The full field dependence of such coefficients $F_n^m(H_x, H_y)$ can be retrieved calculating the following integral expression

$$F_n^m(H_x, H_y) = \frac{1}{N_{nm}} \int_0^{2\pi} \int_0^1 W(H_x, H_y, \rho, \vartheta) Z_n^m(\rho, \vartheta) \rho d\rho d\vartheta$$

Equation 3.4

where the factor N_{nm} represents the norm of the respective Zernike polynomials. The result of the integral Equation 3.4 is shown in the following series of equations from Equation 3.5 to Equation 3.13 where the coefficients $F_n^m(H_x, H_y)$ are written in full to emphasize the dependence on the field coordinates H_x and H_y . It is fundamental to note that in the following series of equations (Equation 3.5 - Equation 3.13), the field dependence of Zernike coefficients is obtained from the field dependence of several polynomial terms constituting the wavefront aberration function $W(H_x, H_y, \rho, \vartheta)$ up to the 4th order shown in Equation 3.2.

$$C_0^0(H_x, H_y, \rho, \vartheta) = \left(\begin{aligned} &\frac{w_{0002}}{2} + \frac{w_{0004}}{8} + \frac{w_{0020}}{2} + \frac{w_{0022}}{24} + \frac{w_{0040}}{8} + \\ &\frac{w_{0102}H_y}{4} + \frac{w_{0120}H_y}{4} + \frac{w_{0202}H_y^2}{4} + \frac{w_{0220}H_y^2}{4} + \\ &\frac{w_{1002}H_x}{4} + \frac{w_{1020}H_x}{4} + \frac{w_{2002}H_x^2}{4} + \frac{w_{2020}H_x^2}{4} \end{aligned} \right) Z_0^0(\rho, \vartheta)$$

Equation 3.5

$$C_1^1(H_x, H_y, \rho, \vartheta) = \left(\begin{aligned} &2w_{0010} + \frac{w_{0012}}{6} + \frac{w_{0030}}{2} + w_{0110}H_y + \\ &w_{0210}H_y^2 + 2w_{1010}H_x + \frac{w_{1012}H_x}{6} + \frac{w_{1030}H_x}{2} + \\ &w_{1110}H_xH_y + w_{1210}H_xH_y^2 + w_{2010}H_x^2 + w_{3010}H_x^3 \end{aligned} \right) Z_1^1(\rho, \vartheta)$$

Equation 3.6

$$C_1^{-1}(H_x, H_y, \rho, \vartheta) = \begin{pmatrix} 2w_{0001} + \frac{w_{0003}}{2} + \frac{w_{0021}}{6} + w_{1001}H_x + \\ 2w_{0101}H_y + \frac{w_{0103}H_y}{2} + \frac{w_{0121}H_y}{6} + w_{2001}H_x^2 + \\ w_{0201}H_y^2 + w_{0301}H_y^3 + w_{1101}H_xH_y + w_{2101}H_x^2H_y \end{pmatrix} Z_1^{-1}(\rho, \vartheta)$$

Equation 3.7

$$C_2^0(H_x, H_y, \rho, \vartheta) = \begin{pmatrix} \frac{w_{0002}}{2} + \frac{3w_{0004}}{16} + \frac{w_{0020}}{2} + \frac{w_{0022}}{16} + \frac{3w_{0040}}{16} + \\ \frac{w_{1002}H_x}{4} + \frac{w_{1020}H_x}{4} + \frac{w_{0202}H_y^2}{4} + \frac{w_{0220}H_y^2}{4} + \\ \frac{w_{0102}H_y}{4} + \frac{w_{0120}H_y}{4} + \frac{w_{2002}H_x^2}{4} + \frac{w_{2020}H_x^2}{4} \end{pmatrix} Z_2^0(\rho, \vartheta)$$

Equation 3.8

$$C_2^2(H_x, H_y, \rho, \vartheta) = \begin{pmatrix} -w_{0002} - \frac{3w_{0004}}{8} + w_{0020} + \frac{3w_{0040}}{8} - \frac{w_{1002}H_x}{2} + \\ \frac{w_{1020}H_x}{2} - \frac{w_{0202}H_y^2}{2} + \frac{w_{0220}H_y^2}{2} - \frac{w_{0102}H_y}{2} + \\ \frac{w_{0120}H_y}{2} - \frac{w_{2002}H_x^2}{2} + \frac{w_{2020}H_x^2}{2} \end{pmatrix} Z_2^2(\rho, \vartheta)$$

Equation 3.9

$$C_2^{-2}(H_x, H_y, \rho, \vartheta) = \left(\frac{w_{0011}}{2} + \frac{w_{1011}H_x}{2} + \frac{w_{0111}H_y}{2} + \frac{w_{1111}H_xH_y}{2} \right) Z_2^{-2}(\rho, \vartheta)$$

Equation 3.10

$$C_3^1(H_x, H_y, \rho, \vartheta) = \left(\frac{w_{0012}}{12} + \frac{w_{0030}}{4} + \frac{w_{1012}H_x}{12} + \frac{w_{1030}H_x}{4} \right) Z_3^1(\rho, \vartheta)$$

Equation 3.11

$$C_3^{-1}(H_x, H_y, \rho, \vartheta) = \left(\frac{w_{0003}}{4} + \frac{w_{0021}}{12} + \frac{w_{0103}H_y}{4} + \frac{w_{0121}H_y}{12} \right) Z_3^{-1}(\rho, \vartheta)$$

Equation 3.12

$$C_4^0(H_x, H_y, \rho, \vartheta) = \left(\frac{w_{0004}}{16} + \frac{w_{0022}}{48} + \frac{w_{0040}}{16} \right) Z_4^0(\rho, \vartheta)$$

Equation 3.13

Equations from Equation 3.5 to Equation 3.13 result from the product between a field-dependent function and the respective pupil-dependent Zernike polynomial. The field-dependent functions $F_n^m(H_x, H_y)$ constitute analytical expressions describing the field behavior of the first 9 Zernike Fringe polynomials for perturbed double-plane symmetric optical systems. This result is functional to the main goal of this chapter, namely the investigation of primary aberrations in perturbed double-plane symmetric optical systems consisting of tilted and decentered surfaces and circular pupils. It is worth making two important remarks at the end of this section to clarify further what has been done so far. The first important aspect to remark is that this method to derive the field dependence of Zernike polynomials can be easily extended to the study of more general systems without imposing the condition of double-plane symmetry. This can be done using a wavefront aberration function $W(H_x, H_y, \rho_x, \rho_y)$ more general than that in Equation 3.2 containing all the possible terms of a full polynomial expansion in the ray coordinates H_x, H_y, ρ_x and ρ_y . The second important aspect to remark is that this method works also for systems characterized by large geometrical perturbations provided that the wavefront aberration function in Equation 3.2 is expanded up to higher orders (for example up to the 6th order if secondary aberrations are of concern or up to the 8th order if tertiary aberrations are relevant). In the present case, we intentionally focus our attention on primary aberrations. This implies that the derived formulas can be applied to optical systems characterized by “weak” geometrical perturbations. In the following section, the full field-dependent Zernike polynomials shown in Equation 3.5 – Equation 3.13 will be combined in order to retrieve the full field behavior of primary aberrations for the optical systems of our concern in this chapter.

3.3 Full-field primary aberrations

The computation of primary aberrations of an optical system can be carried out properly combining the first 9 Zernike polynomials. In particular, it is well known that, in this way, the full field magnitude and phase of vectorial aberrations such as primary astigmatism and primary coma can be retrieved. In the following Equation 3.14, the Zernike polynomials in the Fringe indexing scheme are recombined according to the formulation presented in Equation 61 in [28] to find the full field dependence of primary aberrations in the optical systems of our interest. The Zernike polynomials are written with their full field and pupil dependence. To be noted that in the following Equation 3.14 the pupil dependent parts of the Zernike polynomials (in Equation 3.5 to Equation 3.13 indicated with $Z_n^m(\rho, \vartheta)$) are written explicitly, while the field dependent parts are indicated as previously with $F_n^m(H_x, H_y)$. Additionally, to limit the complexity of Equation 3.14, we have omitted the contribution of lower order Zernike terms corresponding to piston (Z_0^0), tilt along ρ_x (Z_1^1) and tilt along ρ_y (Z_1^{-1}). This choice is justified by the fact that, in the actual computation of the Zernike terms describing the wavefront aberration function of an optical system, the wavefront function itself can be referenced to a properly shifted and tilted reference sphere with the scope of minimizing the RMS wavefront error. In practice, this means that the contribution of tilt along ρ_x and ρ_y can be subtracted out from the wavefront aberration function referencing the wavefront error (or equivalently the optical path difference of light rays) to the chief ray.

$$\begin{aligned}
 W(H_x, H_y, \rho, \vartheta) = & 6F_4^0(H_x, H_y)\rho^4 + \\
 & 3\sqrt{\left(F_3^1(H_x, H_y)\right)^2 + \left(F_3^{-1}(H_x, H_y)\right)^2}\rho^3\cos\left[\vartheta - \tan^{-1}\left(\frac{F_3^{-1}(H_x, H_y)}{F_3^1(H_x, H_y)}\right)\right] + \\
 & \left[2F_2^0(H_x, H_y) - 6F_4^0(H_x, H_y) - \sqrt{\left(F_2^2(H_x, H_y)\right)^2 + \left(F_2^{-2}(H_x, H_y)\right)^2}\right]\rho^2 + \\
 & 2\sqrt{\left(F_2^2(H_x, H_y)\right)^2 + \left(F_2^{-2}(H_x, H_y)\right)^2}\rho^2\cos^2\left[\vartheta - 0.5\tan^{-1}\left(\frac{F_2^{-2}(H_x, H_y)}{F_2^2(H_x, H_y)}\right)\right]
 \end{aligned}$$

Equation 3.14

Individual terms of Equation 3.14 represent primary aberrations. We emphasize that the full field variation of individual primary aberration terms in Equation 3.14 is embedded in the field dependence of the Zernike polynomials involved in the calculation of the primary aberrations themselves. As previously remarked, the reformulation of primary aberrations (resulting from the combination of Zernike polynomials) exposed in Equation 3.14 is functional to the scope of the present chapter, that is the investigation of the full field behavior of primary aberrations in perturbed double-plane symmetric optical systems. In addition to this, it is important to note the formal similarity between Equation 3.14 and the corresponding wavefront aberration function for circular symmetric systems in the Hopkins notation [5], with the difference that in the latter the field dependence is considered in the only meridional plane (because it is related to circular symmetric systems), while in Equation 3.14 the full two-dimensional field dependence of aberrations is taken into account. Moreover, the second and the fourth term of Equation 3.14, corresponding respectively to primary coma and primary astigmatism, are described also by a phase term that is dependent only on the field coordinates. For this reason, primary coma and astigmatism can be represented as vectors and, in this work, we refer to them as vectorial aberrations.

In the following sections, we focus our analysis on the only field dependence of individual terms shown in Equation 3.14. In particular, we indicate the magnitude and phase of the full field behavior of vectorial aberrations with $M_i(H_x, H_y)$ and $P_i(H_x, H_y)$ respectively where the subscript i can be any of the following strings $i = \{s, c, a, fc\}$ standing for spherical aberration, coma, astigmatism and field curvature respectively. We remark that spherical aberration and field curvature do not have a phase term and consequently only their magnitudes will be discussed in the following part. The field-dependent part of the Zernike polynomials contained in the expressions of magnitude and phase of primary aberrations are simplified combining the coefficients w_{klpq} in Equation 3.2 with the same field coordinates dependence. As we will see, this simplification of the polynomials expressions describing the field dependence of Zernike coefficients (Equation 3.5 to Equation 3.13) derives from the pupil dependence of the Zernike polynomials themselves. The simplification of these expressions is aimed at obtaining new expressions with new effective coefficients whose application can be advantageous in the design and optimization of the optical systems under study.

3.3.1 Full-field behavior of spherical aberration

Spherical aberration is described by the first term in Equation 3.14. Writing the explicit expression of F_4^0 , we obtain the following Equation 3.15

$$M_s(H_x, H_y) = 6F_4^0(H_x, H_y) = 6\left(\frac{w_{0004}}{16} + \frac{w_{0022}}{48} + \frac{w_{0040}}{16}\right)$$

Equation 3.15

whose inspection immediately reveals that spherical aberration is field independent.

3.3.2 Full-field behavior of Primary Coma

The full field primary coma magnitude M_c and phase P_c are contained in the second term of Equation 3.14. They result from the combination of the field-dependent parts of the Zernike Fringe terms C_3^1 and C_3^{-1} as follows

$$M_c(H_x, H_y) = 3\sqrt{\left(F_3^1(H_x, H_y)\right)^2 + \left(F_3^{-1}(H_x, H_y)\right)^2} = 3\sqrt{(s_{00} + s_{10}H_x)^2 + (r_{00} + r_{01}H_y)^2}$$

Equation 3.16

In Equation 3.16 the expressions of F_3^1 and F_3^{-1} (deriving from Equation 3.11 and Equation 3.12) are simplified combining the coefficients w_{klpq} with the same field dependence in order to obtain new effective coefficients indicated with s_{kl} and r_{kl} . In this specific case, we obtain respectively

$$\begin{aligned} \frac{w_{0012}}{12} + \frac{w_{0030}}{4} &= s_{00} & \frac{w_{0003}}{4} + \frac{w_{0021}}{12} &= r_{00} \\ \frac{w_{1012}}{12} + \frac{w_{1030}}{4} &= s_{10} & \frac{w_{0103}}{4} + \frac{w_{0121}}{12} &= r_{01} \end{aligned}$$

Equation 3.17

This simplification is justified by the nature of the Zernike polynomials in question whose aperture dependence is described in rectangular coordinates by the following expressions $Z_3^1(\rho_x, \rho_y) = 3\rho_x^3 + 3\rho_x\rho_y^2 - 2\rho_x$ and $Z_3^{-1}(\rho_x, \rho_y) = 3\rho_y^3 + 3\rho_x^2\rho_y - 2\rho_y$. Therefore, the coefficients w_{klpq} in the expression of F_3^1 with $pq = \{3,0\}$ and $pq = \{1,2\}$ and with the same subscripts kl are combined. For example in Equation 3.17, w_{0012} is combined with w_{0030} to give s_{00} because both of them share the same dependence on the field coordinates (i.e. they have the same subscripts $k = 0$ and $l = 0$) and in addition to this, their pupil coordinates dependence is relative to the monomials contained in the expression of $Z_3^1(\rho_x, \rho_y)$ reported above, namely ρ_x^3 and $\rho_x\rho_y^2$. To summarize, regarding F_3^1 , the coefficients w_{kl30} and w_{kl12} are combined in a new effective coefficient s_{kl} , while the coefficients w_{kl03} and w_{kl21} are combined in a new effective coefficient r_{kl} for the case of F_3^{-1} . The introduction of these effective coefficients s_{kl} and r_{kl} is of great aid in the optimization of the optical systems under study. More specifically, from Equation 3.16 we infer that s_{kl} and r_{kl} can be calculated evaluating the respective Zernike coefficients in two field points of reference, for example at the origin ($H_x = 0, H_y = 0$) and at the corner ($H_x = 0.707, H_y = 0.707$) of the field of view (here the field is expressed in normalized coordinates). We obtain

$$\begin{aligned}
 s_{00} &= F_3^1(H_x = 0, H_y = 0) \\
 s_{10} &= \frac{F_3^1(H_x = 0.707, H_y = 0.707) - s_{00}}{0.707} \\
 r_{00} &= F_3^{-1}(H_x = 0, H_y = 0) \\
 r_{01} &= \frac{F_3^{-1}(H_x = 0.707, H_y = 0.707) - r_{00}}{0.707}
 \end{aligned}$$

Equation 3.18

Equation 3.18 shows that the effective coefficients s_{kl} and r_{kl} can be easily obtained calculating the respective Zernike coefficients F_3^1 and F_3^{-1} (in this case in the Fringe indexing scheme) in correspondence of two field points of interest, the first is the origin of the FOV while the choice of the second point is at the discretion of the optical designer. Here, we choose the upper right point at the corner of the FOV in normalized coordinates. The expression of the full field primary coma magnitude $M_c(H_x, H_y)$ in Equation 3.16 describes an elliptical conic surface whose vertex is displaced in the FOV in a point corresponding to its node. This entails that the magnitude of primary coma exhibits a linear dependence on the coordinates of the field starting from its node that is located in the point $(c_x = -\frac{s_{00}}{s_{10}}, c_y = -\frac{r_{00}}{r_{01}})$, calculated using the effective coefficients s_{kl} and r_{kl} . In addition to this, such a linear dependence is different in the tangential and sagittal directions due to the underlying double-plane symmetry of the optical system. In fact, the coefficients s_{10} and r_{01} (indicating the slopes of F_3^1 and F_3^{-1}) are different. Figure 3.1a shows this behavior and Figure 3.1b shows that the level curves are ellipses.

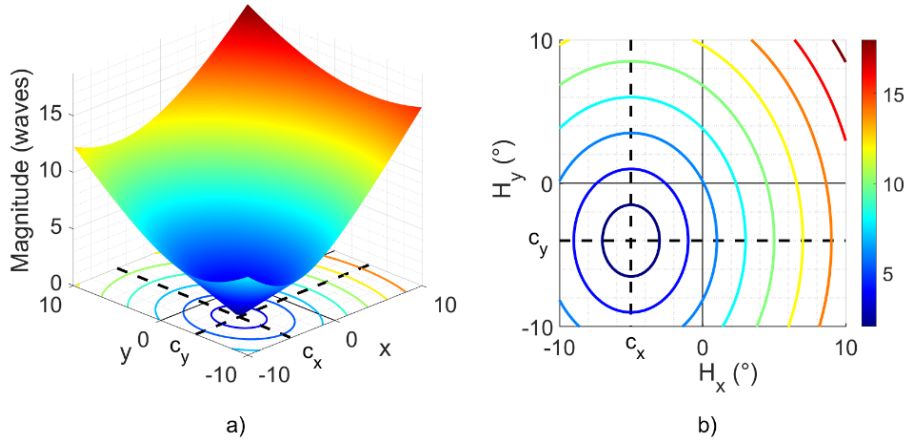


Figure 3.1: a) Conic-shaped surface describing the full field magnitude of primary coma. b) Elliptical level curves of full-field primary coma magnitude.

The full field behavior of primary coma phase P_c is described with the following expression deriving from the second term in Equation 3.14

$$\begin{aligned}
 P_c(H_x, H_y) &= \tan^{-1} \left(\frac{F_3^{-1}(H_x, H_y)}{F_3^1(H_x, H_y)} \right) = \\
 &\quad \tan^{-1} \left(\frac{r_{00} + r_{01}H_y}{s_{00} + s_{10}H_x} \right)
 \end{aligned}$$

Equation 3.19

where the same considerations as before apply. This behavior is shown in Figure 3.2 for the sake of completeness. In correspondence with each point (H_x, H_y) , the phase indicates the azimuthal orientation of primary coma aberration in the FOV plane.

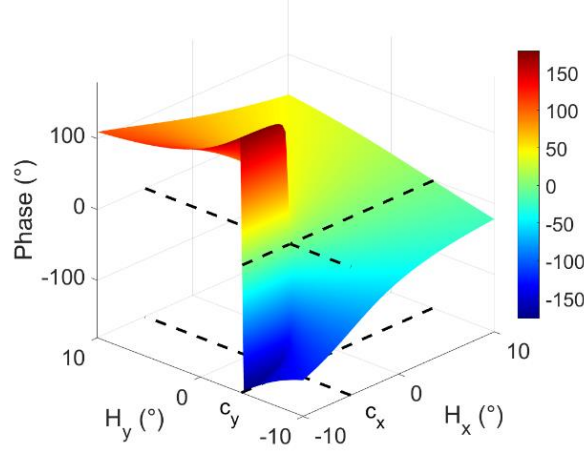


Figure 3.2: Full-field phase of primary coma.

In general, the magnitude and phase of vectorial aberrations are represented jointly in the so-called Full Field Display [13] where specific icons, characterized by certain sizes and orientations and located over a grid of points in the FOV, are used to symbolize the inherent vectorial aberrations (with their magnitude and phase values).

Equation 3.16 and Equation 3.19 can be conveniently used in the design of this category of optical systems. The optimization process can be properly controlled by building the merit function with the effective aberration coefficients s_{kl} and r_{kl} obtained computing the respective Zernike coefficients in correspondence of two field points of interest. This understanding can be of great aid in the definition and control of the optimization of this class of optical systems.

3.3.3 Full-field behavior of Primary Astigmatism

Full-field primary astigmatism is calculated with the fourth term in Equation 3.14. Its magnitude M_a results from the combination of the Zernike coefficients F_2^2 and F_2^{-2} as follows

$$M_a(H_x, H_y) = 2\sqrt{\left[F_2^2(H_x, H_y)\right]^2 + \left[F_2^{-2}(H_x, H_y)\right]^2} =$$

$$2\sqrt{\left[u_{00} + u_{10}H_x + u_{01}H_y + u_{20}H_x^2 + u_{02}H_y^2\right]^2 + \left[v_{00} + v_{10}H_x + v_{01}H_y + v_{11}H_xH_y\right]^2}$$

Equation 3.20

Analogously to the case of primary coma, the coefficients w_{klpq} exhibiting the same field dependence (but different pupil coordinates dependence) are combined. This is again justified by the pupil dependent expressions of the involved Zernike polynomials, namely $Z_2^2(\rho_x, \rho_y) = \rho_x^2 - \rho_y^2$ and $Z_2^{-2}(\rho_x, \rho_y) = 2\rho_x\rho_y$. More in detail, regarding the expression of F_2^2 (Equation 3.9), the coefficients w_{klpq} with the same dependence on the field (i.e. with the same subscripts kl) but with different dependence on the pupil coordinates (i.e. with different subscripts pq) are combined. The following sets of coefficients are summed to obtain a series of equivalent effective coefficients u_{kl} shown in Equation 3.21:

$$\begin{aligned}
 -w_{0002} - \frac{3w_{0004}}{8} + w_{0020} + \frac{3w_{0040}}{8} &= u_{00} \\
 -\frac{w_{1002}}{2} + \frac{w_{1020}}{2} &= u_{10} \\
 -\frac{w_{0102}}{2} + \frac{w_{0120}}{2} &= u_{01} \\
 \frac{w_{2002}}{2} + \frac{w_{2020}}{2} &= u_{20} \\
 -\frac{w_{0202}}{2} + \frac{w_{0220}}{2} &= u_{02}
 \end{aligned}$$

Equation 3.21

Regarding F_2^{-2} , the following coefficients (Equation 3.22) are simplified with the notation v_{kl} retaining only their field dependence:

$$\begin{aligned}
 \frac{w_{0011}}{2} &= v_{00} \\
 \frac{w_{1011}}{2} &= v_{10} \\
 \frac{w_{0111}}{2} &= v_{01} \\
 \frac{w_{1111}}{2} &= v_{11}
 \end{aligned}$$

Equation 3.22

Similarly with the considerations done in the previous section relatively to primary coma, the effective coefficients u_{kl} and v_{kl} can be used advantageously in the optimization of this category of systems. In particular, they can be calculated from the involved Zernike coefficients F_2^2 and F_2^{-2} as shown in Equation 3.23

$$\begin{aligned}
 u_{00} &= F_2^2(H_x = 0, H_y = 0) \\
 u_{00} + u_{10} + u_{20} &= F_2^2(H_x = 1, H_y = 0) \\
 u_{00} - u_{10} + u_{20} &= F_2^2(H_x = -1, H_y = 0) \\
 u_{00} + u_{01} + u_{02} &= F_2^2(H_x = 0, H_y = 1) \\
 u_{00} - u_{01} + u_{02} &= F_2^2(H_x = 0, H_y = -1) \\
 v_{00} &= F_2^{-2}(H_x = 0, H_y = 0) \\
 v_{10} &= F_2^{-2}(H_x = 1, H_y = 0) - v_{00} \\
 v_{01} &= F_2^{-2}(H_x = 0, H_y = 1) - v_{00} \\
 v_{11} &= \frac{F_2^{-2}(H_x = 0.707, H_y = 0.707) - v_{00} - 0.707v_{10} - 0.707v_{01}}{0.707^2}
 \end{aligned}$$

Equation 3.23

In Equation 3.23 the field points of interest used to compute the effective coefficients u_{kl} and v_{kl} are again the origin of the FOV ($H_x = 0, H_y = 0$) and five additional points at the corners of the FOV in normalized coordinates, namely ($H_x = 1, H_y = 0$),

($H_x = -1, H_y = 0$), ($H_x = 0, H_y = 1$), ($H_x = 0, H_y = -1$) and ($H_x = 0.707, H_y = 0.707$). In particular, the effective coefficients u_{10} and u_{20} can be calculated solving a system of two linear equations with two unknowns where the two linear equations are the second and third in Equation 3.23. The effective coefficients u_{01} and u_{02} can be calculated solving another system of two linear equations with two unknowns where the two linear equations are the fourth and fifth in Equation 3.23. The calculation of the effective coefficients v_{kl} is easier because it does not require to solve systems of equations but it can be done retrieving directly the values of the inherent Zernike coefficient in correspondence of the field points of interest as shown in Equation 3.23. The surface $M_a(H_x, H_y)$ described in Equation 3.20 exhibits a binodal behavior similar to the one characterizing the full field dependence of primary astigmatism in asymmetric systems described by NAT. More in detail, the binodal behavior of primary astigmatism in NAT optical systems is well described with Cassini surfaces whose level curves are Cassini ovals. On the other side, the binodal behavior of primary astigmatism in perturbed double-plane symmetric optical systems, described by $M_a(H_x, H_y)$ in Equation 3.20, is more general than Cassini surfaces (Figure 3.3c). As a matter of fact, the level curves of this type of surface are rational bicircular quartics (Figure 3.3d), a more general family of curves to which Cassini ovals belong. For this reason, in this chapter, we define the category of surfaces described by $M_a(H_x, H_y)$ in Equation 3.20 as generalized Cassini surfaces. These two behaviors are shown for comparison in Figure 3.3.

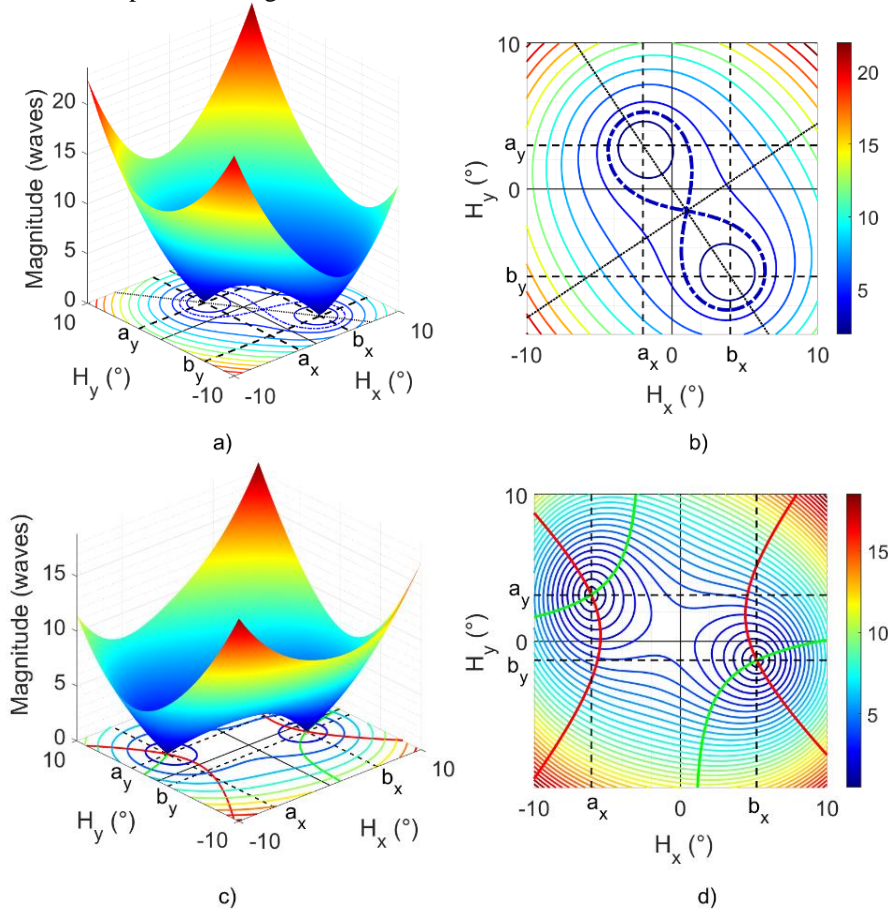


Figure 3.3: a) Cassini surface describing the full field behavior of primary astigmatism magnitude in NAT systems. b) Cassini ovals representing the level curves of Cassini surfaces. The lemniscate is emphasized with a bold dash-dot line. c) Generalized Cassini surfaces describing the full field behavior of astigmatism magnitude in perturbed double-plane symmetric systems. d) Rational bicircular quartic curves representing the level curves of generalized Cassini surfaces. In all figures, (a_x, a_y) and (b_x, b_y) denote the coordinates of the nodes of primary astigmatism in the image plane.

In Figure 3.3d it is emphasized the fact that the location of the node in the FOV of the systems of our interest is defined by the intersection of two conic sections (in red and green) respectively described by the equations $F_2^2(H_x, H_y) = 0$ and $F_2^{-2}(H_x, H_y) = 0$. These two equations are nothing but the zero level curves of the equations of the surfaces described by $z_1 = F_2^2(H_x, H_y)$ and $z_2 = F_2^{-2}(H_x, H_y)$ respectively. As a consequence, the nodes of primary astigmatism can be located solving the system of two second-degree equations $F_2^2(H_x, H_y) = 0$ and $F_2^{-2}(H_x, H_y) = 0$. This observation is particularly relevant because it contributes to clarifying the origin of the two nodes visible in the magnitude of primary astigmatism. This fact is emphasized further in the following Figure 3.4 where the surfaces $z_1 = F_2^2(H_x, H_y)$ and $z_2 = F_2^{-2}(H_x, H_y)$ are overlapped and the origin of nodes of primary astigmatism in correspondence of the intersections between their respective zero level curves $z_1 = 0$ and $z_2 = 0$ is highlighted.

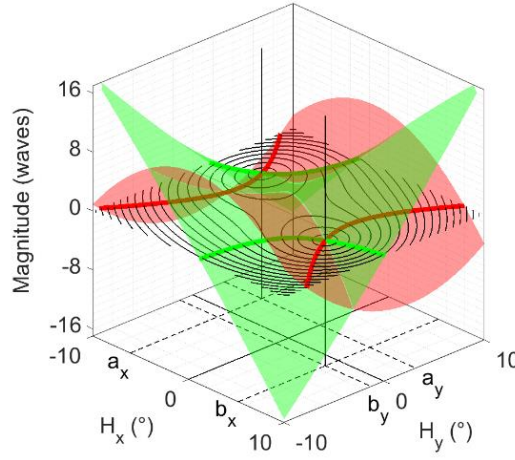


Figure 3.4: Origin of the binodal behavior of primary astigmatism magnitude. The surfaces z_1 and z_2 (along with their zero level curves) are in red and green respectively. The intersection points define the nodes of primary astigmatism magnitude.

The full field behavior of primary astigmatism phase $P_a(H_x, H_y)$ is described by an expression as follows deriving from the fourth term in Equation 3.14.

$$P_a(H_x, H_y) = 0.5 \tan^{-1} \left(\frac{F_2^{-2}(H_x, H_y)}{F_2^2(H_x, H_y)} \right) = 0.5 \tan^{-1} \left(\frac{v_{00} + v_{10}H_x + v_{01}H_y + v_{11}H_xH_y}{u_{00} + u_{10}H_x + u_{01}H_y + u_{20}H_x^2 + u_{02}H_y^2} \right)$$

Equation 3.24

The previous considerations relative to the simplification of the coefficients w_{klpq} remain valid. This behavior is shown in Figure 3.5.

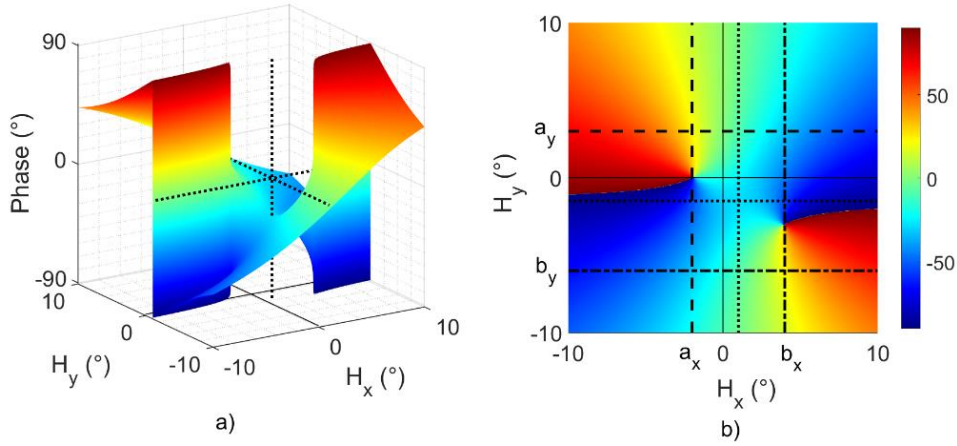


Figure 3.5: Full-field behavior of primary astigmatism phase.

The magnitude and phase of primary astigmatism are in general represented jointly in the Full Field Display with icons represented by bars with variable sizes and orientations corresponding respectively to the magnitude and phase of this aberration type.

Similarly with what we have previously said for primary coma, in this case, the values of the effective coefficients u_{kl} and v_{kl} can be properly targeted in customized merit functions defined in the optimization routines during the optical design of this class of systems.

3.3.4 Full-field behavior of Field Curvature

Field curvature M_{fc} as a function of field coordinates is calculated as follows (from the third term in Equation 3.14)

$$M_{fc}(H_x, H_y) = 2F_2^0(H_x, H_y) - 6F_4^0(H_x, H_y) - \sqrt{\left(F_2^2(H_x, H_y)\right)^2 + \left(F_2^{-2}(H_x, H_y)\right)^2}$$

Equation 3.25

Field curvature aberration is zero for on-axis field points and exhibits a quadratic dependence on the field coordinate for circular symmetric systems. Regarding optical systems described by NAT, the behavior is substantially the same with the additional possibility for a displacement of the node location in the FOV. In perturbed double-plane symmetric optical systems, the behavior of $M_{fc}(H_x, H_y)$ is more general than the previous two cases. In fact, the variation of field curvature with the field coordinates is generally described by an elliptical surface displaced in the FOV. This situation is expressed by the following Equation 3.26 obtained from Equation 3.25 replacing the terms F_n^m with their explicit expressions (deriving from Equation 3.8, Equation 3.9, Equation 3.10 and Equation 3.13) after simplifying the coefficients w_{klpq} as it is shown below

$$M_{fc}(H_x, H_y) = t_{00} + t_{10}H_x + t_{01}H_y + t_{20}H_x^2 + t_{02}H_y^2 - \sqrt{\left(F_2^2(H_x, H_y)\right)^2 + \left(F_2^{-2}(H_x, H_y)\right)^2}$$

Equation 3.26

The effective coefficients t_{kl} are obtained in Equation 3.27.

$$\begin{aligned}
 2\left(\frac{w_{0002}}{2} + \frac{3w_{0004}}{16} + \frac{w_{0020}}{2} + \frac{w_{0022}}{16} + \frac{3w_{0040}}{16}\right) - 6\left(\frac{w_{0004}}{16} + \frac{w_{0022}}{48} + \frac{w_{0040}}{16}\right) &= t_{00} \\
 2\left(\frac{w_{1002}}{4} + \frac{w_{1020}}{4}\right) &= t_{10} \\
 2\left(\frac{w_{0102}}{4} + \frac{w_{0120}}{4}\right) &= t_{01} \\
 2\left(\frac{w_{2002}}{4} + \frac{w_{2020}}{4}\right) &= t_{20} \\
 2\left(\frac{w_{0202}}{4} + \frac{w_{0220}}{4}\right) &= t_{02}
 \end{aligned}$$

Equation 3.27

The coefficient t_{00} contains the additional field independent contribution given by spherical aberration. Equation 3.20, relative to primary astigmatism magnitude and contained in Equation 3.26, has been discussed in the previous sub-section and for this reason, it is not expressed explicitly in Equation 3.26. It is important to note that the presence of this term in the expression for field curvature intrinsically results from the recombination of the inherent Zernike polynomials as shown in Equation 3.14. In fact, the full field Zernike polynomials C_2^2 and C_2^{-2} quantify the joint contributions of primary astigmatism and defocus to the total wavefront aberration function. The new effective coefficients t_{kl} can be again retrieved calculating the involved Zernike coefficients in correspondence of specific field points of interest. These are the origin of the FOV and four additional points at the corners of the FOV in normalized coordinates, namely $(H_x = 1, H_y = 0)$, $(H_x = -1, H_y = 0)$, $(H_x = 0, H_y = 1)$ and $(H_x = 0, H_y = -1)$. More in detail, the effective coefficients are found as shown in Equation 3.28.

$$\begin{aligned}
 t_{00} &= 2F_2^0(H_x = 0, H_y = 0) - 6F_4^0(H_x = 0, H_y = 0) \\
 t_{00} + t_{10} + t_{20} &= 2F_2^0(H_x = 1, H_y = 0) - 6F_4^0(H_x = 1, H_y = 0) \\
 t_{00} - t_{10} + t_{20} &= 2F_2^0(H_x = -1, H_y = 0) - 6F_4^0(H_x = -1, H_y = 0) \\
 t_{00} + t_{01} + t_{02} &= 2F_2^0(H_x = 0, H_y = 1) - 6F_4^0(H_x = 0, H_y = 1) \\
 t_{00} - t_{01} + t_{02} &= 2F_2^0(H_x = 0, H_y = -1) - 6F_4^0(H_x = 0, H_y = -1)
 \end{aligned}$$

Equation 3.28

The effective coefficients t_{10} and t_{20} are obtained solving a system of two linear equations (the second and the third in Equation 3.28) with two unknowns, while the effective coefficients t_{01} and t_{02} are retrieved solving another system of two linear equations (the fourth and the fifth in Equation 3.28) with two unknowns. Equation 3.28 along with Equation 3.23, previously discussed relative to the effective coefficients of astigmatism and computed in the same field points of interest, allows to completely determine the surface describing the field behavior of field curvature expressed with Equation 3.26. After the effective coefficients u_{kl} , v_{kl} and t_{kl} are retrieved with Equation 3.23 and Equation 3.28 respectively, the full field surface $M_{fc}(H_x, H_y)$ in Equation 3.26 or alternatively its zero level curves $M_{fc}(H_x, H_y) = 0$ can be easily calculated. The full expression of $M_{fc}(H_x, H_y)$ describes in general an elliptical surface with a sort of deformations in correspondence of the nodes of primary astigmatism. In Figure 3.6a it is shown the full field behavior described with an elliptical surface without the deformation induced by the nodes of primary astigmatism. In the last section related to the simulation example, we will show the effect of considering the additional contribution of the

magnitude of primary astigmatism to the total expression of field curvature. In Figure 3.6b the associated zero level curves are shown to be ellipses decentered and rotated in the FOV. This observation is useful to find the coordinates of the node of field curvature that coincides with the center of symmetry of such a family of ellipses. More specifically, the ellipsoidal surface contained in Equation 3.26 is described by the part of the equation containing the effective coefficients t_{kl} , namely $E(H_x, H_y) = t_{00} + t_{10}H_x + t_{01}H_y + t_{20}H_x^2 + t_{02}H_y^2$. The inherent level curves are found setting $E(H_x, H_y) = 0$, that is a second-degree polynomial equation in the field coordinates, describing precisely an ellipse. For this reason, the coordinates (d_x, d_y) of the node of field curvature can be calculated as $d_x = \frac{-0.5t_{02}t_{10}}{\delta}$ and $d_y = \frac{-0.5t_{20}t_{01}}{\delta}$, where $\delta = \det \begin{bmatrix} t_{20} & 0 \\ 0 & t_{02} \end{bmatrix}$ is the determinant of the matrix of the quadratic form associated with the equation $E(H_x, H_y) = 0$.

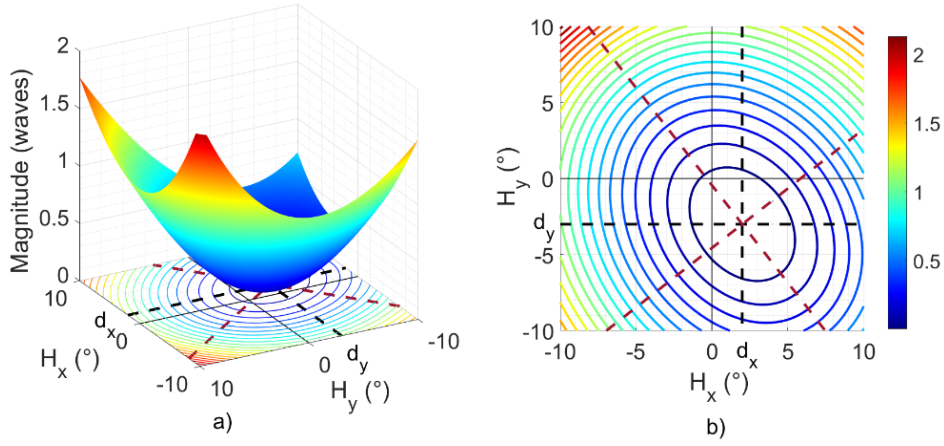


Figure 3.6: a) Displaced and rotated double-symmetric surface describing the full field behavior of field curvature. b) The level curves are ellipses.

So far, we have explicitly restricted our analysis to only primary aberrations. If higher-order aberrations are present, it is necessary to consider the previous Equation 3.2 expanded up to the order of interest (for example 6th order) and again project the Zernike polynomials basis using Equation 3.4. The retrieved expressions of $F_n^m(H_x, H_y)$ will consist of polynomials of order higher than the second (for example of 4th order). In this case, in general, it is not possible to determine such full-field expressions sampling the respective Zernike coefficients in correspondence of few field points, therefore different approaches must be put in place to simplify the study of this category of optical systems.

3.4 Verification of results: an example with two aspheric lenses

In order to demonstrate the full field behavior of primary aberrations (except distortion) in perturbed double-plane symmetric optical systems with circular pupil, we consider an optical system constituted by a combination of two aspheric lenses whose first surface is planar and the second one is biconic. Both lenses are tilted and decentered with respect to the mechanical axis of the system. A biconic surface is characterized by different conic constants ($\varepsilon_y \neq \varepsilon_x$) and different radii of curvatures ($R_y \neq R_x$) along the tangential and sagittal directions respectively. As a consequence, a biconic surface is intrinsically double-plane symmetric in both x and y sections.

The layout of the optical system under study and the reference system are shown in Figure 3.7.

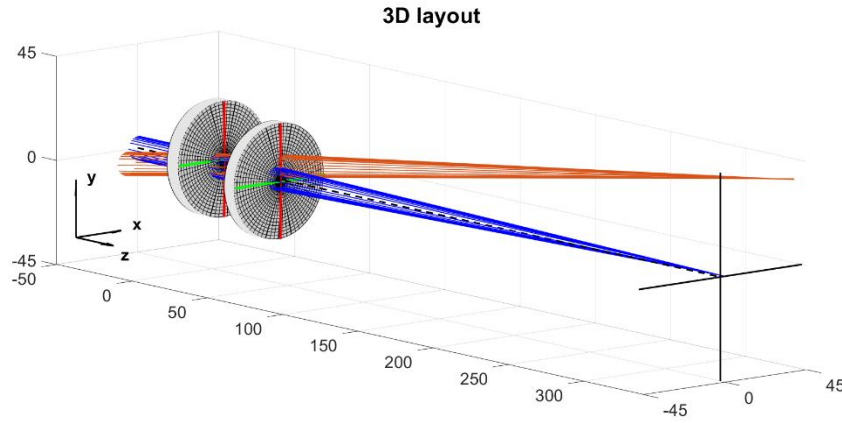


Figure 3.7: Layout and reference system of the optical system (the dimensions are in mm). The tangential and sagittal sections of the biconic surface are shown respectively in red and green. Two ray bundles are shown: in blue for the field point ($H_x = 0, H_y = 0$) and in orange for the field point ($H_x = 0.707, H_y = 0.707$).

The system is set at infinite conjugates with an Entrance Pupil Diameter equal to $ENPD = 10\text{mm}$. The simulation wavelength is $\lambda = 0.55\mu\text{m}$ and the FOV is $7^\circ \times 7^\circ$. The distance between the two lenses is 20mm. The stop position is at the first plane surface and the distance between the last surface and the image plane is optimized in order to minimize the RMS wavefront error.

The parameters used for this example are summarized in Table 3.1. The parameters of the initial surfaces of the first and second biconic lenses are omitted because both surfaces are planar. The ray-tracing simulation is performed in OpticStudio [13].

Table 3.1: Simulation parameters of the optical system under study

	Parameter	Symbol	Value
First Biconic Lens	Diameter	D_1	50mm
	Sagittal Radius	R_x	-402 mm
	Tangential Radius	R_y	-400 mm
	Thickness	t_1	7 mm
	Glass	g_1	N-BK7
	Conic constant x	ε_x	24
	Conic constant y	ε_y	-8
	x decentering	$D_{x,1}$	3.7mm
	y decentering	$D_{y,1}$	2.3mm
	x tilt	$\tau_{x,1}$	2.6°
	y tilt	$\tau_{y,1}$	-1.8°
Second Biconic Lens	Diameter	D_2	50mm
	Sagittal Radius	R_x	-250mm
	Tangential Radius	R_y	-249mm
	Thickness	t_2	5 mm
	Glass	g_2	N-BK7
	Conic constant x	ε_x	3
	Conic constant y	ε_y	-2
	x decentering	$D_{x,2}$	2.4mm
	y decentering	$D_{y,2}$	-1.5mm

	x tilt	$\tau_{x,2}$	-3.8°
	y tilt	$\tau_{y,2}$	2.6°

In the following sections, we show the simulation results related to the full field behavior of individual aberration types for the optical system under study and we validate the mathematical models presented in the previous sections through surface fitting. The quality of the fit is quantified by calculating the corresponding RMSE (Root Mean Squared Error) and SSE (Sum of Squares due to Error).

3.4.1 Primary coma

Primary coma magnitude varies linearly with the distance from its node that can be displaced from the center of the FOV depending on the symmetry properties of the optical system. As previously said, the general model describing such behavior is an elliptic conic-shaped surface with variable apex position in the FOV. This is demonstrated by fitting the model surface in Equation 3.16 to the simulation results. Figure 3.8 and Figure 3.9 show the fitting result.

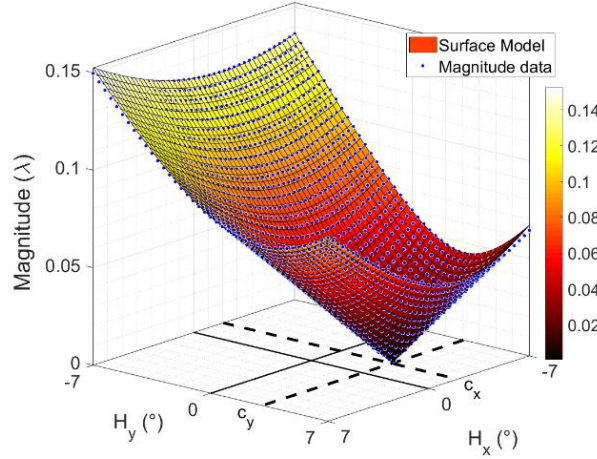


Figure 3.8: Conic-shaped surface fit to full-field primary coma magnitude.

Figure 3.9 shows that the level curves of the fitted model (Equation 3.16) to the primary coma magnitude data are ellipses whose displacement in the FOV corresponds to the node location. The coefficients in the box on the right side of Figure 3.9 result from the fitting process. Their values coincide with those of the effective coefficients s_{kl} and r_{kl} calculated with Equation 3.18 reported in the previous section. The computation of the effective coefficients with Equation 3.18 is carried out sampling the inherent Zernike coefficients in only two points in the FOV, the first of which is the origin of the FOV while the second is chosen by us to be at the corner of the FOV, namely $(H_x = 0.707, H_y = 0.707)$. This result is particularly useful considering that the full field behavior of primary coma can be completely determined in this category of optical systems by the direct computation of the effective coefficients s_{kl} and r_{kl} in two predefined field points of interest.

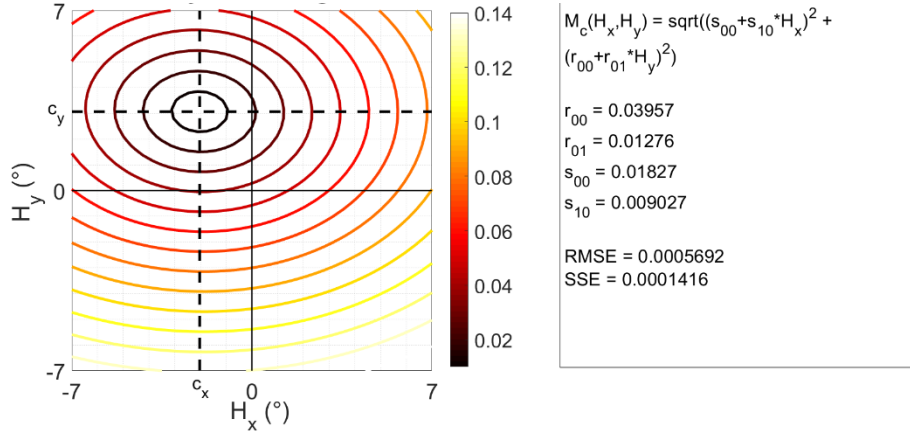


Figure 3.9: Contour plot of the conic-shaped full-field primary coma magnitude data.

The node of primary coma in the FOV is located in $(c_x = -\frac{s_{00}}{s_{10}} = -2.0228^\circ, c_y = -\frac{r_{00}}{r_{01}} = 3.0984^\circ)$. The phase of primary coma is modeled with a surface described by Equation 3.19. Figure 3.10 shows the fit result for completeness. To be noticed that the calculation of the phase is done using the four-quadrant inverse tangent.

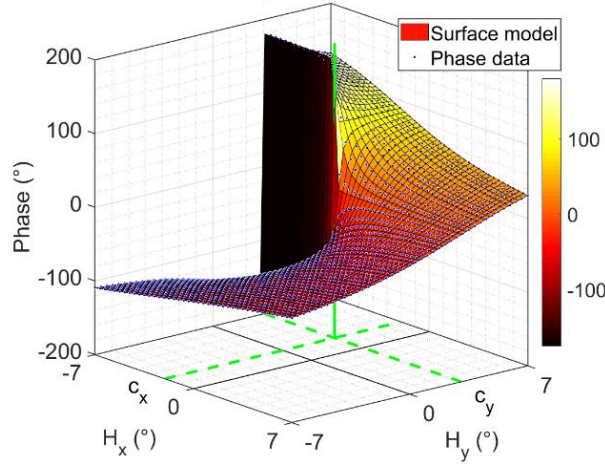


Figure 3.10: Surface model fitting the primary coma phase data.

3.4.2 Primary Astigmatism

The field dependence of primary astigmatism magnitude exhibits a generalized Cassinian behavior. This is demonstrated by fitting the surface model (Equation 3.20) to the simulated field-dependent primary astigmatism magnitude as shown in Figure 3.11 and Figure 3.12.

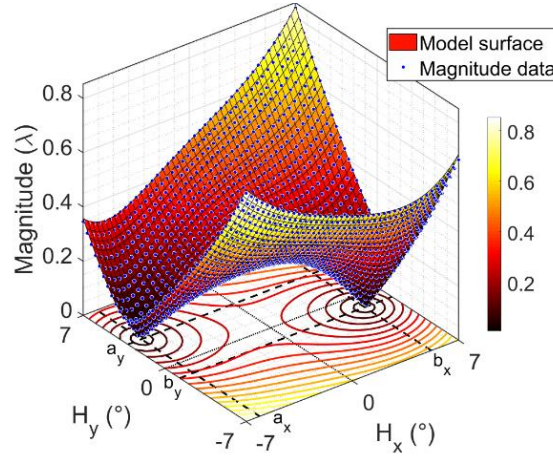


Figure 3.11: Generalized Cassini surface fit to full-field primary astigmatism magnitude of perturbed double-plane symmetric systems.

Figure 3.12 shows the level curves of the fitted model (Equation 3.20) to the simulation data. As previously remarked, such curves are rational bicircular quartic, a family of curves including also the Cassini ovals.

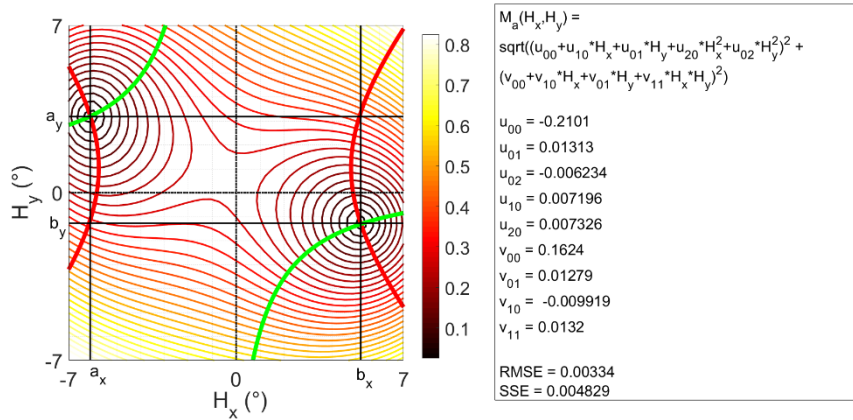


Figure 3.12: Rational bicircular quartic curves define the contour plot of the generalized Cassini surface fitting the full field primary astigmatism magnitude data.

The values of the effective coefficients u_{kl} and v_{kl} reported in the text box on the right side of Figure 3.12 are obtained fitting the surface model for primary astigmatism magnitude to the simulation data. Analogously to the above observations about the effective coefficients of primary coma, the values of the effective coefficients of astigmatism u_{kl} and v_{kl} can be alternatively found using Equation 3.23 described in the previous section. With this equation, only a few field points of interest are sufficient to retrieve the values of the effective coefficients and consequently the complete behavior of Equation 3.20. The nodes of primary astigmatism are located by the intersection points between the conics defined by the equations $F_2^2(H_x, H_y) = 0$ and $F_2^{-2}(H_x, H_y) = 0$ given explicitly in Equation 3.29

$$\begin{aligned} u_{00} + u_{10}H_x + u_{01}H_y + u_{20}H_x^2 + u_{02}H_y^2 &= 0 \\ v_{00} + v_{10}H_x + v_{01}H_y + v_{11}H_xH_y &= 0 \end{aligned}$$

Equation 3.29

In Figure 3.12 they are represented in red and green respectively. The intersection points between these curves can be found solving the system of two second-degree equations in two variables shown in Equation 3.29. In this example, we obtain $(a_x = -6.146^\circ, a_y = 3.164^\circ)$ and $(b_x = 5.249^\circ, b_y = -1.296^\circ)$. We point out again that Equation 3.29 defines the zero level curves of the respective surfaces $z_1 = F_2^2(H_x, H_y)$ and $z_2 = F_2^{-2}(H_x, H_y)$. Such zero level curves can be alternatively referred to as nodal lines as it is done in [29]. The surface model of Equation 3.24 fitting the phase of primary astigmatism is displayed in Figure 3.13.

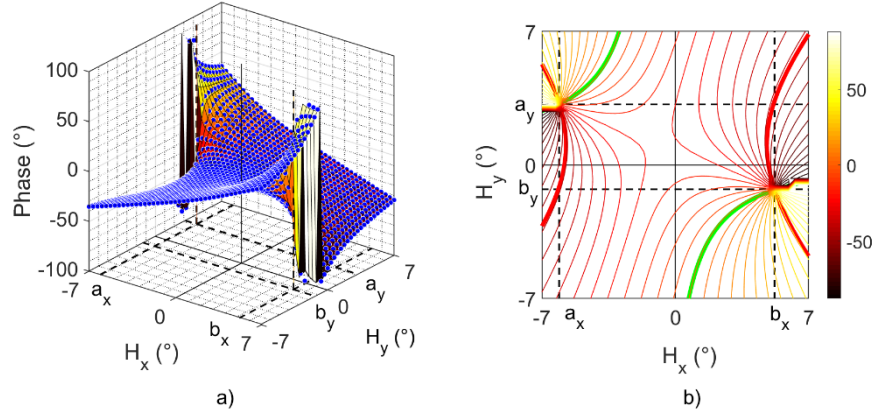


Figure 3.13: a) Surface model fitting the primary astigmatism phase. b) Contour plot of the surface model for primary astigmatism phase. The conic sections whose intersections locate the nodes of primary astigmatism are shown in red and green.

Figure 3.13b shows again the conic sections described by Equation 3.29, in this case, overlapped to the contour plot of primary astigmatism phase.

3.4.3 Field curvature

The full field dependence of field curvature is modeled by Equation 3.25 and Equation 3.26. Figure 3.14 and Figure 3.15 display the fitting result.

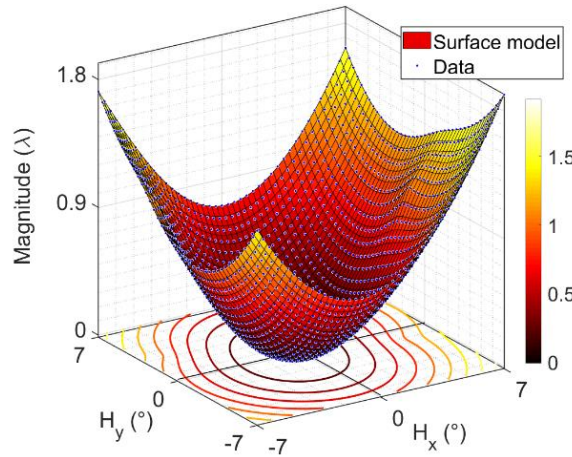


Figure 3.14: Fitting of full-field behavior of field curvature with the surface model in Equation 3.26.

The level curves are elliptical except close to the nodes of primary astigmatism (indicated with two green dots in Figure 3.15). These are the areas where the nodes of primary astigmatism are located.

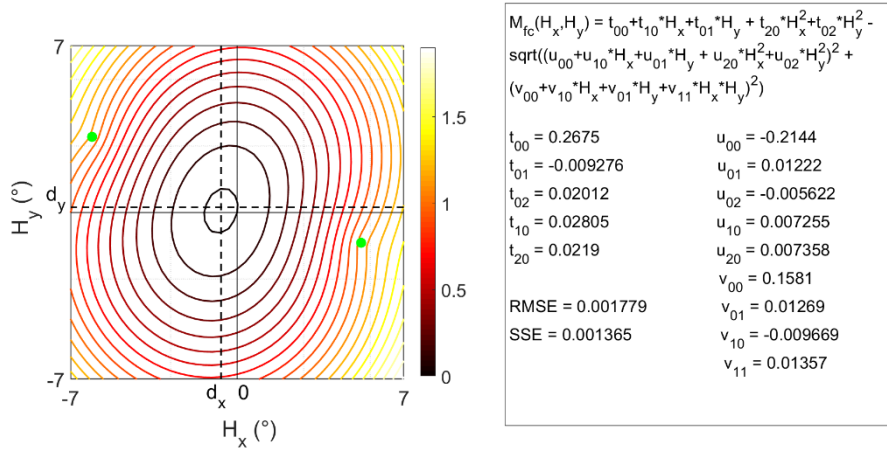


Figure 3.15: Full-field behavior of field curvature: contour plot of the simulated data.

The fitting result relative to field curvature contains the same coefficients previously reported in the surface model for primary astigmatism indicated with u_{kl} and v_{kl} . These effective coefficients have almost the same values retrieved in the previous section regarding primary astigmatism magnitude. The effective coefficients t_{kl} can be alternatively obtained using Equation 3.28 where the involved Zernike coefficients, F_2^0 and F_4^0 , are sampled in correspondence with five field points of interest. In this particular example, we can observe that the contribution of primary astigmatism (with effective coefficients u_{kl} and v_{kl}) to field curvature causes a sort of “distortion” to the elliptical level curves described by the part of Equation 3.26 with effective coefficients t_{kl} (shown also in the text box in Figure 3.15). The node of field curvature coincides with the center of symmetry of the family of ellipses described by the part of Equation 3.26 containing the effective coefficients t_{kl} , namely $E(H_x, H_y) = t_{00} + t_{10}H_x + t_{01}H_y + t_{20}H_x^2 + t_{02}H_y^2$. Bearing in mind the considerations done in the previous section, the coordinates of the node of field curvature are found to be $(d_x = -0.6788^\circ, d_y = 0.2209^\circ)$ in this example.

The computation of the effective coefficients for the various types of primary aberrations is performed evaluating the inherent Zernike coefficients in correspondence with few field points of interest. And as shown for coma, astigmatism and field curvature, the location of the respective nodes can be carried out algebraically directly from the introduced formulas containing effective coefficients. This fact is of particular importance because it allows determining the complete full-field behavior of primary aberration in perturbed double-plane symmetric optical systems in a form similar to the wavefront aberration function in the Hopkins notation for circular symmetric optical systems. We emphasize especially that this formulation promotes a better understanding of the general behavior of this category of optical systems and can be fruitfully used for their optimization during optical design, for example building proper merit functions containing the desired values of the effective coefficients.

3.5 Conclusions

In this chapter, we have presented a series of analytical formulas describing the full field dependence of Zernike polynomials for perturbed double-plane symmetric optical systems characterized by misalignments of the internal components. Such expressions are derived projecting the general full field wavefront aberration function of double-plane symmetric systems up to the 4th order onto the basis of pupil-dependent Zernike polynomials in the Fringe indexing scheme. In turn, such expressions are combined to define full field-dependent and pupil-dependent analytical formulas with effective coefficients for primary aberrations (excluding distortion) for perturbed double-plane symmetric systems with a circular pupil. The proposed analytical expressions account for both magnitude and phase

of vectorial aberrations such as primary coma and primary astigmatism. In particular, the full field behavior of the magnitudes of primary astigmatism, primary coma and field curvature are described respectively with a special binodal surface (more general than the Cassini surface), with an elliptical conic-shaped surface and with a general elliptical surface with double-plane symmetry that is slightly distorted in correspondence of the astigmatism nodes. The location of nodes of individual aberrations over the FOV is accounted for by the displacement and rotation of the respective surfaces in the FOV. The application of these expressions with new effective coefficients can be particularly beneficial to the design and optimization of optical systems characterized in general by double-plane symmetry. The proposed analytical expressions are verified with a numerical example relative to an optical system consisting of misaligned biconic surfaces.

Chapter 4 Asymmetric systems with decentred spherical surfaces in the framework of Hamiltonian Optics: a comparison with Nodal Aberration Theory

The goal of the present chapter is to carry out a comparison between two different descriptions of asymmetric optical systems with decentered spherical surfaces. The first description is provided in the framework of Hamiltonian Optics, the second is given with the formulation of the Nodal Aberration Theory. These two theories are simply compared based on their respective full field wavefront aberration expansions. Particular aberration terms in the general power series expansion of Hamiltonian Optics are not included in the NAT formulation. The computation of these specific aberration terms in an asymmetric system with decentered spherical surfaces allows us to define what are the actual limits of applicability of NAT to the analysis of this category of optical systems. Such computations consist of fitting the general multivariate polynomial model foreseen by Hamiltonian Optics to the wavefront aberration data obtained with real raytracing. The ultimate scope of the considerations presented here is to justify the use of the Hamilton power series expansion rather than the NAT formula to describe this category of optical systems. In this sense, the present chapter is ancillary to the next one where this computational approach is extensively used to investigate the full field aberration behavior of optical systems with multiple apertures. Finally, this chapter is also intended to pave the way to an integration of the Nodal Aberration Theory [3] into the more general and complete theory of Hamiltonian Optics [16].

4.1 Introduction

As already remarked in the previous chapters, the relevance of Hamiltonian Optics comes from the possibility to predict the structure of the characteristic function of an optical system taking into consideration its symmetry properties. In the most general case of an optical system without symmetry, the characteristic function is expected to be expanded with a polynomial in the ray coordinates whose all terms are potentially different from zero. On the opposite side, a circular symmetric system exhibits a simplified characteristic function (if compared to the aforementioned case), resulting from certain combinations of the only rotational invariants. In fact, this class of optical systems is characterized by the co-presence of axial symmetry and plane symmetry and consequently, most of the coefficients multiplying the terms of the polynomial expansion are identically zero. Between these two opposite cases, different typologies of optical systems with an intermediate degree of

symmetry can be devised. For example, in [16], it is described a class of systems, named semi-symmetric optical systems, characterized by only axial symmetry (while plane-symmetry is absent). Optical systems with only plane-symmetry were addressed in [30]. In [20], bilateral symmetric optical systems are introduced as plane-symmetric systems where the optical surfaces can be additionally tilted in the plane of symmetry. The several classes of optical systems mentioned up to now are characterized by a common point, that is the possibility to be described, in the framework of Hamiltonian Optics, with a power series expansion in the ray coordinates whose aberration terms reflect the underlying symmetry.

The Nodal Aberration theory (NAT) deals with asymmetric optical systems constituted by tilted and decentered rotationally symmetric surfaces reformulating the scalar wavefront aberration expansion proposed by Hopkins for circular symmetric system into a more general vectorial form. Such vectorial reformulation has proved to be beneficial for the treatment of general asymmetric systems with rotationally symmetric surfaces (as demonstrated by Thompson). Nevertheless, in such a vast theoretical development it does not seem to have been emphasized sufficiently the underlying derivation of NAT [3] from the theory of Hamiltonian optics. In fact, from a careful examination of both theories, it emerges clearly that the former is already contained in the latter. More specifically, the vectorial full-field aberration function introduced in NAT is encompassed in the characteristic function of a general asymmetric system as foreseen by Hamiltonian Optics. This can be demonstrated considering that, as mentioned above, the theory of Hamiltonian Optics prescribes a polynomial description of the characteristic function where, in the most general case, all the coefficients multiplying the polynomial terms are different from zero. On the other hand, the formula expressing the wavefront aberration function introduced in NAT is less general than the corresponding formula envisaged by Hamiltonian Optics. This state of things is clarified in this chapter where the polynomial expansions of the wavefront aberration proposed in the framework of these two theories are compared to highlight their main differences. In the following section, we give an overview of the paraxial behavior of asymmetric optical systems, necessary to introduce the concept of base-ray. In the last section, we analyze the full field wavefront aberration function of a simplified optical system in various geometrical configurations. These configurations are distinguished from one another by different values of the decentering parameters of the second lens composing the system. The choice to analyze the variation of the aberrations depending on these specific geometrical parameters is functional to the topic dealt with in the next chapter, which is the study of optical systems with multiple apertures. The present analysis indeed lends itself to being adapted to the study of this last class of optical systems since they can be classified based on specific decentering parameters, as will be clarified later.

4.2 Paraxial optics of systems without symmetries

As already said in the introduction of this thesis, the characteristic function quantifies the optical path length between any pairs of points belonging to two well-defined base-planes in an optical system. This implies that it is of primary importance to fix the anterior and the posterior base-planes constituting the starting and arrival planes of the rays passing through the system. At the same level, in order to study the properties of an asymmetric optical system, it is fundamental to make an adequate choice of the base-ray. This last one plays the role of reference axis in a general system without symmetries. In fact, the calculation of the optical path length of any light ray connecting two points lying on the anterior and posterior base-planes is referenced to the optical path traced by the base-ray. The application of the point characteristic function to the calculation of these optical distances is consistent only if the property of regularity is ensured for the optical system under study. This is the possibility to express the characteristic function (and hence the wavefront aberration function) of an optical system with a polynomial expansion. In the case of an asymmetric optical system consisting of decentered surfaces, this property is provided by appropriately defining the reference systems of the ray coordinates to lie along the base-ray at the point of intersection with the anterior and the posterior base-planes respectively. This point clarifies the reason why the definition of such base-ray is of primary importance,

certainly comparable to that of the optical axis in a rotationally symmetric system. In addition to this, the base-ray is defined to pass through the center of the aperture stop of an optical system. Thus, tracing the base-ray through the system also serves to fix the origin of the local reference systems of the different optical surfaces at the point of intersection with the base-ray itself. Moreover, such local reference systems have their z axis lying along the base-ray. Last but not least, the base-ray specifies the local reference system at the exit pupil plane of an optical system. In practical calculations, the distance between the exit pupil plane and the image plane is defined to be the radius of the reference sphere to which the wavefront aberration function is referenced. Therefore, once a base-ray is properly defined in a system without symmetries, the optical aberration associated with an arbitrary ray (belonging to a pencil of light rays) can be described as the optical path length difference between the light ray itself and the base-ray. Alternatively, the optical aberration of an arbitrary ray can be also defined as the optical distance between the reference sphere (centered at the point of intersection between the base-ray and the image plane) and the wavefront of the pencil along the light ray in question. This situation is illustrated in the following Figure 4.1 where it is traced the base-ray of an optical system without symmetries along with the exit pupil plane and the reference sphere.

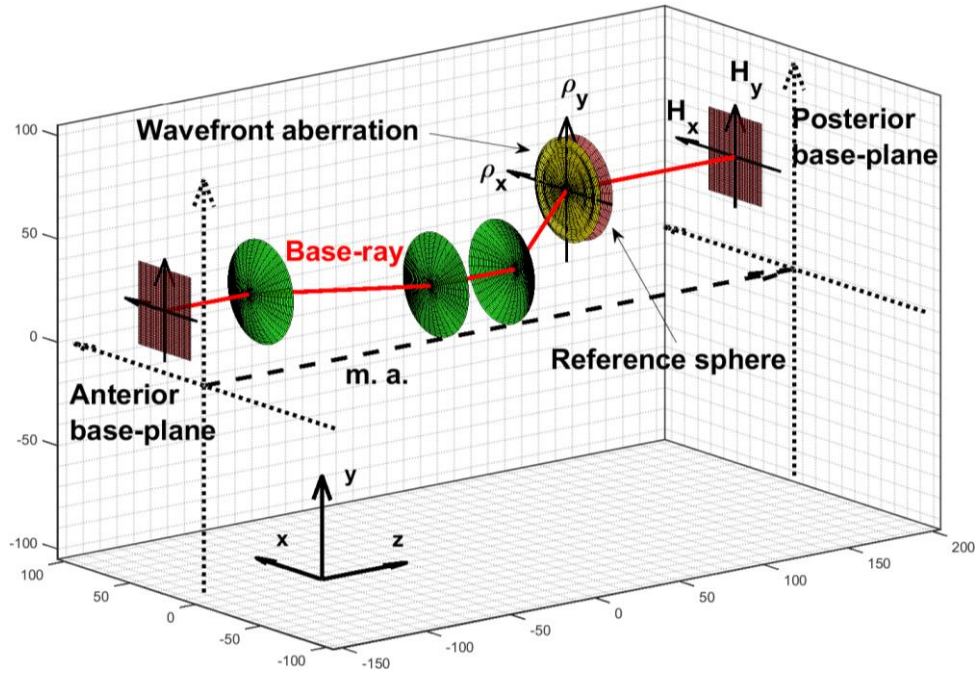


Figure 4.1: Example of an asymmetric optical system. The local reference systems at the exit pupil plane and the image plane are shown. They are located along the base-ray, in red. A wavefront surface is shown in yellow at the exit pupil plane and the reference sphere is shown in red. In green, several displaced optical surfaces are shown. This system is displaced from the mechanical axis (m. a.).

Once the base-ray has been traced, the paraxial behavior of an asymmetric optical system can be investigated. This clarifies further that the result of a paraxial analysis is strictly dependent on the particular definition of the base-ray.

Analogous to paraxial optics, paraxial optics deals with the behavior of light rays near the base-ray. As mentioned in Chapter 1 of this thesis, the behavior of paraxial rays is described by the characteristic function truncated to the second degree shown in Equation 4.1

$$\begin{aligned}
 V(H_x, H_y, \rho_x, \rho_y) = & v_{0000} + v_{2000}H_x^2 + v_{1100}H_xH_y + \\
 & v_{1010}H_x\rho_x + v_{0101}H_y\rho_y + v_{0200}H_y^2 + v_{0110}H_y\rho_x + \\
 & v_{0101}H_y\rho_y + v_{0020}\rho_x^2 + v_{0011}\rho_x\rho_y + v_{0002}\rho_y^2
 \end{aligned}$$

Equation 4.1

Applying the basic equations of Hamiltonian Optics (shown in Chapter 1) to Equation 4.1, one obtains four equations describing the parabal relations between the ray coordinates and their conjugates.

The practical computation of such parabal relations is done by tracing few parabal rays through the optical system under study and retrieving the final values of the ray coordinates at the points of intersection with the posterior base-plane. Such parabal rays are conveniently defined fixing their initial coordinates at the anterior base-plane. The relations obtained in this way are linear transformations from the object space coordinates to the image-space coordinates.

4.3 General power series description of the full field wavefront aberration function

In the present section, it is introduced the general full field wavefront aberration function within the framework of the theory of Hamiltonian Optics. The importance of this theory relies on its rigorous and exhaustive description of the imagery of optical systems classified according to their symmetry properties. As illustrated in Figure 4.1, the wavefront aberration function associated with a light ray is expressed as the difference between the point characteristic function at the origin of the exit pupil plane and the point characteristic function at the point of intersection of the light ray with the wavefront itself. This entails that the considerations done in the introduction relatively to the expansion in power series of the point characteristic function can be extended to the wavefront aberration function. Consequently, taking for granted again that a general optical system is regular with respect to a chosen reference system, its general full field wavefront aberration function $W = W(H_x, H_y, \rho_x, \rho_y)$ can be expanded as a power series in the ray coordinates H_x, H_y, ρ_x and ρ_y containing terms of even and odd degree up to the order n_c of concern. Thus, the full field wavefront aberration function is expressed as follows

$$W(H_x, H_y, \rho_x, \rho_y) = \sum_{n=0}^{n_c} \sum_{\lambda=0}^n \sum_{\mu=0}^{\lambda} \sum_{\nu=0}^{\mu} w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu} H_x^{n-\lambda} H_y^{\lambda-\mu} \rho_x^{\mu-\nu} \rho_y^{\nu}$$

Equation 4.2

where $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}$ are coefficients weighting the respective monomials. The maximum number of coefficients in the power series up to the order n_c is given by the binomial coefficient $\binom{n_c+4}{n_c} = \frac{(n_c+4)!}{n_c!4!}$. In the most general case of an optical system without any symmetry, all the coefficients $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}$ can be different from zero. This observation emphasizes the fundamental role played by symmetry in the theory of aberrations. In fact, whenever an optical system exhibits a particular symmetry, then the number of aberration coefficients calculated with Equation 4.2 is reduced accordingly.

The computation of the coefficients $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}$ in the general power series in Equation 4.2 can be done through real raytracing. Families of rays, traced through the optical system, are identified by their intersection points with the field plane and the exit pupil plane. The phase of rays is measured at the intersection point with the reference sphere. Such phase is the optical path difference of a ray with respect to the chief ray in correspondence of the intersection point with the reference sphere. Therefore, these phase

data are nothing but the wavefront aberration function related to a specific field point in the FOV plane. Reiterating this calculation over a grid of field points allows obtaining the full field wavefront aberration function. Then, the coefficients $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}$ can be retrieved fitting (in the least-squares sense) the multivariate polynomial model in Equation 4.2 (up to the desired order n_c) to the ray-trace data. In section 4.5, this approach is used to calculate the value of few particular aberration coefficients of the power series in Equation 4.2 for different geometrical configurations of the optical system that will be analyzed. In Chapter 5 of the thesis, this same approach is used to calculate the actual coefficients $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}$ for the various branches of an optical system with multiple apertures.

In the following section, a comparison is made between the more general wavefront aberration function of Equation 4.2 provided by Hamiltonian Optics and that less general introduced by NAT. This allows delineating more clearly the limits of applicability of the latter.

4.4 Comparison with the NAT wavefront aberration function

In this section, some important considerations are done regarding the nature of the wavefront aberration function in the NAT formulation. In particular, the main goal is to make a comparison between the aberration terms present in the general power series shown in Equation 4.2 and the respective aberration terms of the NAT wavefront aberration formula. This comparison is done between terms up to the 4th degree. It is important to point out that, relatively to NAT, it is considered that version of the wavefront aberration function without the summation of the aberration contributions over the surfaces of the optical system, that is Equation 2.4 in Chapter 2. This is done, again, to circumvent the intrinsic criticality of Equation 2.2 in Chapter 2 in describing optical systems characterized by significant secondary aberrations. In fact, in this scenario, the summation theorem for primary aberrations is inapplicable and it is more convenient to use a formulation characterized by the absence of any summation over the optical surfaces, as already explained in Chapter 2.

The power series in the ray coordinates shown in Equation 4.2 provides a very general description of the wavefront aberration behavior of an optical system. This description is not limited to asymmetric systems consisting of only tilted and decentered spherical surfaces, but it is also extended to asymmetric optical systems constituted by surfaces with arbitrary shapes. Such surfaces must meet the essential condition of regularity, that is the possibility to be locally approximated with a polynomial. This observation underlines the greater generality of Equation 4.2 compared to the corresponding formulation provided by NAT. As a matter of fact, NAT is concerned with optical systems made of arbitrarily tilted and displaced only rotationally symmetric surfaces. To clarify further this state of things, in the following it is carried out a comparison between the aberration terms of both formulas with the additional aim to provide a better understanding of the meaning of the numerous terms contained in the power series in Equation 4.2. In fact, the juxtaposition of the aberration terms in Equation 4.2 with their counterparts in the NAT wavefront aberration formula gives insight into their respective meaning. Furthermore, the aberration terms that are not accounted for in the NAT formulation, are immediately put in evidence. This comparison is preliminary to the following section where some of the aberration coefficients $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}$ listed in the tables, will be calculated for a very simple optical system in different geometrical configurations. An attempt will be made to define the limits of applicability of the NAT wavefront aberration formula to the description of the aberration behavior of asymmetric systems with decentered spherical surfaces.

The following tables from 4.1 to 4.4 point out the correspondence between the aberration terms of the two formulations. In order of appearance, Tables from Table 4.1 to Table 4.4 show a comparison of the aberration terms of spherical aberration, coma, astigmatism and field curvature and distortion.

Table 4.1: Spherical Aberration terms

Power series expansion	NAT expansion	
$w_{0040}\rho_x^4$	$W_{040}(\rho_x^2 + \rho_y^2)^2$	4 th order
$w_{0022}\rho_x^2\rho_y^2$		
$w_{0004}\rho_y^4$		
$w_{0031}\rho_x^3\rho_y$	-	
$w_{0013}\rho_x\rho_y^3$	-	

Among the five primary spherical aberration terms foreseen by the Hamilton wavefront aberration function, the first three are collected together in a single aberration term in the NAT formula, while the last two terms are not foreseen. As will be explained in the next section, the study of the dependence of these spherical aberration coefficients, w_{0031} and w_{0013} , on the parameters defining the various geometrical configurations of the system that will be analyzed, helps to recognize when NAT is unsuitable for such description.

Table 4.2 compares the primary coma aberration coefficients in the general Hamilton formula and in the NAT representation.

Table 4.2: Coma Aberration terms

Power series expansion	NAT expansion	
$w_{1030}H_x\rho_x^3$	$W_{131}H_x\rho_x(\rho_x^2 + \rho_y^2)$	4 th order
$w_{1012}H_x\rho_x\rho_y^2$		
$w_{1003}H_x\rho_y^3$	-	
$w_{1021}H_x\rho_x^2\rho_y$	-	
$w_{0103}H_y\rho_y^3$	$W_{131}H_y\rho_y(\rho_x^2 + \rho_y^2)$	
$w_{0121}H_y\rho_x^2\rho_y$		
$w_{0130}H_y\rho_x^3$	-	
$w_{0112}H_y\rho_x\rho_y^2$	-	
$w_{0030}\rho_x^3$	$-W_{131}a_{131,x}\rho_x(\rho_x^2 + \rho_y^2)$	3 rd order
$w_{0012}\rho_x\rho_y^2$		
$w_{0003}\rho_y^3$	$-W_{131}a_{131,y}\rho_y(\rho_x^2 + \rho_y^2)$	
$w_{0021}\rho_x^2\rho_y$		

Among the eight terms describing primary coma in the general power series in Equation 4.2, four are not accounted for in the NAT formulation. Moreover, the four remaining terms are collected in two pairs of terms with the same aberration coefficient. Similarly, the 3rd order comatic terms are collected in two pairs of terms with the same aberration coefficient.

Table 4.3 deals with the aberration coefficients of primary astigmatism and field curvature. To prevent any misunderstanding, it needs to be clarified that the 3rd and 2nd order terms, listed in Table 4.3 below, are considered to be part of astigmatism and field curvature in the sense that, according to the NAT formulation (Equation 2.4, Chapter 2), they are generated from a “perturbation” of the field dependence of these 4th order aberration terms. Therefore, the 3rd order aberration coefficients w_{1011} and w_{0111} can be defined as linear astigmatism terms and those with coefficients w_{1020} , w_{1002} , w_{0120} and w_{0102} can be defined as field tilt terms. In addition to this, the 2nd order terms with aberration coefficients w_{0020} , w_{0011} and w_{0002} can be defined as constant astigmatism terms. These 3rd and 2nd order terms (generated by a field perturbation) will be added to other aberration terms of the same order like defocus (that is of 2nd order) if present. On the other side, in the general power series expansion of Equation 4.2, terms of 2nd and 3rd order are not considered to derive from an

eventual field perturbation of 4th or higher-order terms, but rather they are intrinsically present in the expansion due to the lack of symmetry of an optical system.

Table 4.3: Astigmatism and field curvature terms

Power series expansion	NAT expansion	
$w_{2020}H_x^2\rho_x^2$	$\left(W_{220} + \frac{1}{2}W_{222}\right)H_x^2\rho_x^2$	4 th order
$w_{2011}H_x^2\rho_x\rho_y$	-	
$w_{2002}H_x^2\rho_y^2$	$\left(W_{220} - \frac{1}{2}W_{222}\right)H_x^2\rho_y^2$	
$w_{1120}H_xH_y\rho_x^2$	-	
$w_{1111}H_xH_y\rho_x\rho_y$	$W_{222}H_xH_y\rho_x\rho_y$	
$w_{1102}H_xH_y\rho_y^2$	-	
$w_{0220}H_y^2\rho_x^2$	$\left(W_{220} - \frac{1}{2}W_{222}\right)H_y^2\rho_x^2$	
$w_{0211}H_y^2\rho_x\rho_y$	-	
$w_{0202}H_y^2\rho_y^2$	$\left(W_{220} + \frac{1}{2}W_{222}\right)H_y^2\rho_y^2$	
$w_{1020}H_x\rho_x^2$	$\left(2W_{220}a_{220,x} + \frac{1}{2}W_{222}a_{222,x}\right)H_x\rho_x^2$	3 rd order
$w_{1011}H_x\rho_x\rho_y$	$-2W_{222}a_{222,y}H_x\rho_x\rho_y$	
$w_{1002}H_x\rho_y^2$	$\left(-2W_{220}a_{220,x} + W_{222}a_{222,x}\right)H_x\rho_y^2$	
$w_{0120}H_y\rho_x^2$	$\left(-2W_{220}a_{220,y} + W_{222}a_{222,y}\right)H_y\rho_x^2$	
$w_{0111}H_y\rho_x\rho_y$	$-W_{222}a_{222,y}H_y\rho_x\rho_y$	
$w_{0102}H_y\rho_y^2$	$\left(-2W_{220}a_{220,y} - W_{222}a_{222,y}\right)H_y\rho_y^2$	
$w_{0020}\rho_x^2$	$\left[W_{220}(a_{220,x}^2 + a_{220,y}^2) + \frac{1}{2}W_{222}(a_{222,x}^2 - a_{222,y}^2)\right]\rho_x^2$	2 nd order
$w_{0002}\rho_y^2$	$\left[W_{220}(a_{220,x}^2 + a_{220,y}^2) - \frac{1}{2}W_{222}(a_{222,x}^2 - a_{222,y}^2)\right]\rho_y^2$	
$w_{0011}\rho_x\rho_y$	$2W_{222}a_{222,x}a_{222,y}\rho_x\rho_y$	

Within the nine coefficients of 4th order related to astigmatism and field curvature in the power series expansion, there are four that have no equivalent in the NAT formulation. Finally, Table 4.4 makes a comparison between the distortion terms in the two formulations.

Table 4.4: Distortion terms

Power series expansion	NAT expansion	
$w_{3010}H_x^3\rho_x$	$W_{311}H_x^3\rho_x$	4 th order

$w_{3001}H_x^3\rho_y$	-	
$w_{2110}H_x^2H_y\rho_x$	-	
$w_{2101}H_x^2H_y\rho_y$	$W_{311}H_x^2H_y\rho_y$	
$w_{1210}H_xH_y^2\rho_x$	$W_{311}H_xH_y^2\rho_x$	
$w_{1201}H_xH_y^2\rho_y$	-	
$w_{0310}H_y^3\rho_x$	-	
$w_{0301}H_y^3\rho_y$	$W_{311}H_y^3\rho_y$	
$w_{2010}H_x^2\rho_x$	$-3W_{311}a_{311,x}H_x^2\rho_x$	3 rd order
$w_{2001}H_x^2\rho_y$	$-W_{311}a_{311,y}H_x^2\rho_y$	
$w_{1110}H_xH_y\rho_x$	$-2W_{311}a_{311,y}H_xH_y\rho_x$	
$w_{1101}H_xH_y\rho_y$	$-2W_{311}a_{311,x}H_xH_y\rho_y$	
$w_{0210}H_y^2\rho_x$	$-W_{311}a_{311,x}H_y^2\rho_x$	
$w_{0201}H_y^2\rho_y$	$-3W_{311}a_{311,y}H_y^2\rho_y$	
$w_{1010}H_x\rho_x$	$W_{311}(3a_{311,x}^2 + a_{311,y}^2)H_x\rho_x$	2 nd order
$w_{1001}H_x\rho_y$	$2W_{311}a_{311,x}a_{311,y}H_x\rho_y$	
$w_{0110}H_y\rho_x$	$2W_{311}a_{311,x}a_{311,y}H_y\rho_x$	
$w_{0101}H_y\rho_y$	$W_{311}(a_{311,x}^2 + 3a_{311,y}^2)H_y\rho_y$	
$w_{0010}\rho_x$	$-W_{311}a_{311,x}(a_{311,x}^2 + a_{311,y}^2)\rho_x$	1 st order
$w_{0001}\rho_y$	$-W_{311}a_{311,y}(a_{311,x}^2 + a_{311,y}^2)\rho_y$	

Considerations similar to those done for the previous aberration types remain valid for 4th order distortion terms. 3rd, 2nd and 1st order aberration terms are respectively defined as quadratic distortion terms, anamorphism terms and field displacement terms.

4.5 Study of the limits of applicability of NAT

The strategy put in place in this section to define the limits of applicability of the NAT version of the wavefront aberration function is based on the practical study of the aberration behavior of a simple optical system constituted by the sequence of two spherical lenses of which the second is allowed to be displaced transversally at different distances from the mechanical axis of the system. The second lens is smaller in size with respect to the first one and in addition to this, it defines also the aperture stop of the system. In practice, every couple of values of the decentering parameters of the second lens defines a specific configuration of the optical system under study. De facto, such a system can be studied as a multi-configurations optical system. The study carried out in this section is, to some extent, an anticipation of the topic discussed in the following chapter where the several optical configurations (mentioned here) will be replaced by different optical branches (or channels) of a multiple aperture optical system. The simple system under study is represented in the following Figure 4.2

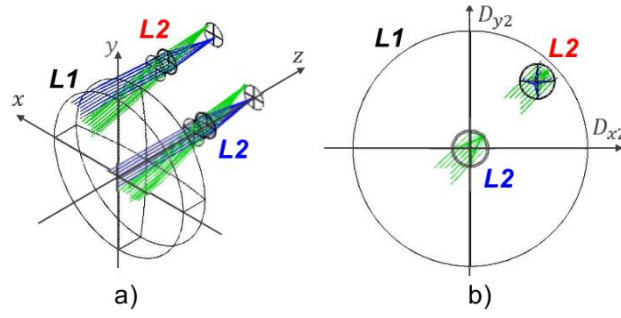


Figure 4.2: a) Simple optical system with two spherical lenses $L1$ and $L2$. While $L1$ is fixed, $L2$ is allowed to be decentered along the x and y directions of the global reference system (xyz) . In red and blue are indicated two configurations characterized by different transversal displacements (D_{x2}, D_{y2}) of the second lens $L2$. b) Front view.

In Figure 4.2 two different geometrical configurations are displayed. They differ for the decentering parameters of the second lens (D_{x2}, D_{y2}) . The following Table 4.5 reports the parameters of the optical system shown in Figure 4.2.

Table 4.5: Simulation parameters

	Parameter	Symbol	Value
$L1$	Diameter	D_1	100 mm
	Radius of curvature	R_1	-200 mm
	Thickness	t_1	15 mm
	Glass	g_1	N-BK7
	Conic constant	c_1	0
$L2$	Diameter	D_2	10 mm
	Radius of curvature	R_2	-15 mm
	Thickness	t_2	7 mm
	Glass	g_2	N-BK7
	Conic constant	c_2	0
	x decentering	D_{x2}	Variable
	y decentering	D_{y2}	Variable
	x tilt	$\tau_{x,2}$	0 °
	y tilt	$\tau_{y,2}$	0 °

Each configuration can be imagined to be a particular channel of a composite optical system with multiple apertures. The objective of this section is to make an analysis of the different configurations of the system shown in Figure 4.2 comparing the respective full field wavefront aberration functions (Equation 4.2) obtained with the approach described in section 4.3. The scope of this comparison is to ascertain the conditions in which the NAT full field wavefront aberration function fails to describe the effective wavefront aberrations of the various configurations of the optical system under study. Because of the large number of aberration terms in the general power series expressed by Equation 4.2, the investigation made here is focused on the behavior of only a few terms that are more relevant for the analysis. The terms in question are those not included in the NAT wavefront expansion. As remarked in the previous section, these aberration terms result from certain combinations of the ray coordinates that are not considered in NAT because the terms of this latter theory contain combinations of the pupil coordinates, ρ_x and ρ_y , deriving from the only rotational invariants $(\rho_x^2 + \rho_y^2)$ and $(H_x \rho_x + H_y \rho_y)$. For example, regarding primary spherical aberration, the terms of election to carry out the present analysis are respectively $w_{0031} \rho_x^3 \rho_y$ and $w_{0013} \rho_x \rho_y^3$. Considering primary coma, the terms $w_{1003} H_x \rho_y^3$, $w_{1021} H_x \rho_x^2 \rho_y$,

$w_{0130}H_y\rho_x^3$ and $w_{0112}H_y\rho_x\rho_y^2$ are of interest. Continuing in the same way, for astigmatism and field curvature the elected terms are $w_{2011}H_x^2\rho_x\rho_y$, $w_{1120}H_xH_y\rho_x^2$, $w_{1102}H_xH_y\rho_y^2$ and $w_{0211}H_y^2\rho_x\rho_y$. Therefore, calculating the value of the listed aberration terms for each geometrical configuration of the optical system under study, allows to determine exactly those configurations that are not properly described by the NAT wavefront aberration function. In particular, if the magnitude of one of the listed aberration terms is different from zero for a specific configuration, the consequence is that this same configuration is not eligible to be described with the NAT wavefront aberration formula. Alternatively, one can say that the eventual description with the NAT formula will introduce an error. It is recalled here that such configurations differ from each other for the decentering parameters of the second lens. This analysis, if extended to the rest of the wavefront aberration terms, allows to completely define the limits of applicability of NAT to this category of systems. Furthermore, it provides an insight into the general variation of the wavefront aberration function in optical systems whose “degree of symmetry” can be conveniently related to the value of some geometrical parameters of interest. In this particular case, the geometrical parameters involved in the definition of the “degree of symmetry” of any specific configuration of the system in Figure 4.2 are, as mentioned above, the transversal displacements of the second spherical lens along the x and y directions in the global reference system denoted with D_{x2} and D_{y2} .

Starting with the coefficients of spherical aberration (relevant for the scope of this study), they are respectively w_{0031} and w_{0013} . Since both these coefficients exhibit the same variation over the decentering parameters D_{x2} and D_{y2} , in the following Figure 4.3, it is shown the behavior of the first of them, $w_{0031}(D_{x2}, D_{y2})$. The height of the surface in correspondence with a grid of points in the domain $0\text{mm} \leq D_{x2} \leq 20\text{mm}$ and $0\text{mm} \leq D_{y2} \leq 20\text{mm}$ quantifies the magnitude of the spherical aberration coefficient $w_{0031}(D_{x2}, D_{y2})$ in waves ($\lambda_0 = 0.55\mu\text{m}$). Each point (D_{x2}, D_{y2}) indicates a particular geometrical configuration of the optical system under study, characterized exactly by the corresponding values of the decentering parameters of the second lens.

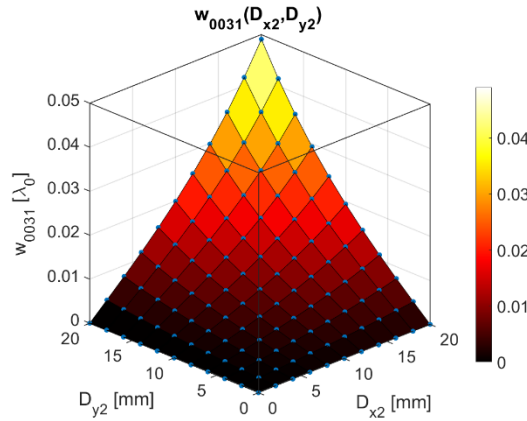


Figure 4.3: Variation of the spherical aberration coefficient w_{0031} with respect to the decentering parameters of the second lens constituting the optical system in Figure 4.2.

Figure 4.3 demonstrates that the spherical aberration coefficient in question (not included in the NAT wavefront aberration expansion) is nonzero for $(D_{x2} \neq 0, D_{y2} \neq 0)$ in the analyzed domain. This suggests that the geometrical configurations whose decentering parameters meet the condition above are not properly described using the NAT wavefront aberration function since this last formula does not include the particular spherical aberration coefficients shown above (it is reminded that the coefficient $w_{0013}(D_{x2}, D_{y2})$ is not shown because it exhibits the same behavior as the coefficient w_{0031}). Following the same path, Figure 4.4 shows the behavior of two primary coma coefficients relevant to this

analysis. These are respectively $w_{1021}(D_{x2}, D_{y2})$ and $w_{1003}(D_{x2}, D_{y2})$, while the remaining aberration coefficients, $w_{0112}(D_{x2}, D_{y2})$ and $w_{0130}(D_{x2}, D_{y2})$, are omitted because they behave exactly as the formers.

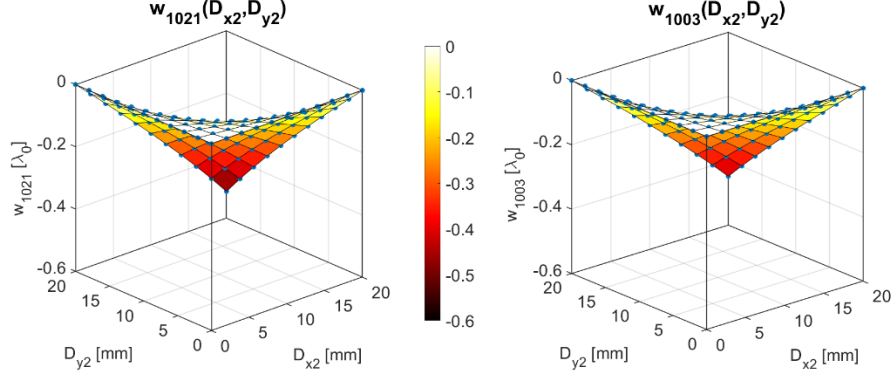


Figure 4.4: Variation of the coma aberration coefficients w_{1021} and w_{1003} with respect to the decentring parameters of the second lens D_{x2} and D_{y2} .

Similar to the case of spherical aberration, the coma coefficients in Figure 4.4 are different from zero if the condition ($D_{x2} \neq 0, D_{y2} \neq 0$) is met. This situation supports further the idea that an increasing numerical error is introduced using the NAT wavefront aberration formula to describe geometrical configurations with increasing values of the decentring parameters.

Extending the considerations done before to astigmatism and field curvature, in this case, the aberration coefficients of interest for the present investigation are $w_{0211}(D_{x2}, D_{y2})$ and $w_{1102}(D_{x2}, D_{y2})$, while the other two coefficients, namely $w_{2011}(D_{x2}, D_{y2})$ and $w_{1120}(D_{x2}, D_{y2})$, are again neglected given that they vary exactly like the other two coefficients. The following Figure 4.5 illustrates the behavior of both aberration coefficients.

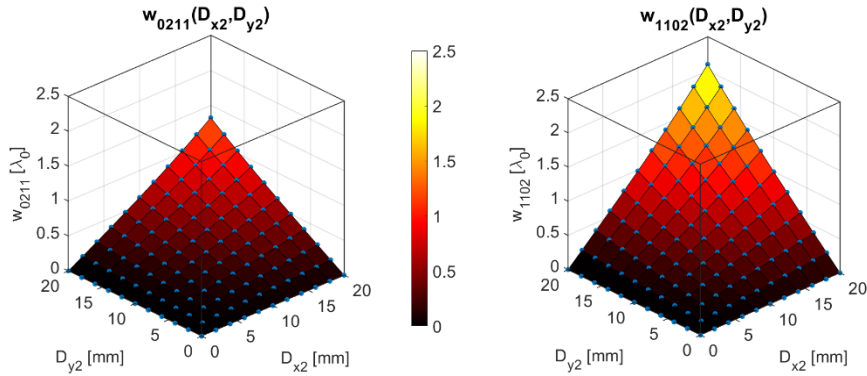


Figure 4.5: Variation of astigmatism and field curvature aberration coefficients w_{0211} and w_{1102} with respect to the decentring parameters of the second lens D_{x2} and D_{y2} .

This last example confirms further the previous considerations regarding the increasing error done in the case the NAT wavefront aberration function is used to describe geometrical configurations with larger decentring parameters.

The series of examples in this section and, in general, this chapter are propaedeutic to the next chapter where it is carried out a study about optical systems with multiple apertures. Substantially, the observations done up to now seem to suggest unambiguously that the use of the general power series formula is more promising than the NAT formula to describe

the aberration behavior of asymmetric systems characterized by decentered spherical lenses.

4.6 Conclusions

The general goal of this chapter is to show that, substantially, the wavefront aberration function introduced by NAT relatively to asymmetric optical systems is nothing but a reformulation of the more general wavefront aberration function described in the theory of Hamiltonian Optics for optical systems without symmetry. In addition to this, due to the greater generality of the Hamilton formulation compared to that of NAT, the former appears to be best suited to the analysis of the aberration behavior of asymmetric systems characterized by more pronounced decentering parameters of the spherical surfaces. This analysis has been carried out for a simplified optical system consisting of two spherical lenses, the latter of which is allowed to be decentered from the z axis of the global reference system. Such analysis has been focused on the only transversal displacement of the second spherical lens, intentionally neglecting the possibility to be tilted with respect to the reference axis. The approach presented in this chapter is prone to be applied to the analysis of the full field aberration behavior of optical systems with multiple apertures, which is the subject of the next chapter of this thesis. In particular, the full field wavefront aberration function expressed by Equation 4.2 will be computed for the several branches of an optical system with many apertures since, as pointed out in the present chapter, this formulation is more appropriate for this purpose.

Chapter 5 Wavefront aberration behavior in multiple aperture systems

In this chapter, it is proposed a general approach to describe the full field aberration behavior in optical systems with many apertures. The formulas presented in the previous chapters, properly adapted to account for the presence of several apertures in the optical system, are used for this purpose. The additional goal of the present treatment is to emphasize the possibility to exploit the occurrence of symmetries in the channels of a multi-aperture system to acquire a better understanding of its image-forming properties and its aberrations. In the last section, it is shown a practical application of the considerations done throughout the chapter. A light field camera is studied as an exemplary realization of a multi-aperture optical system. In this case, it is shown the possibility to predict the aberration behavior of the whole optical system restricting the analysis to the only channels spanning the first quadrant of the field of view.

5.1 Optical systems with multiple apertures

The main feature of the class of optical systems of interest in this chapter is the presence of multiple apertures whose primary function is to perform a spatial channeling of the incoming light disturbance. Such a channeling, more precisely, consists of the spatial sampling of the incoming light. This function is, in practice, realized by providing an optical system with special optical components characterized by many apertures. An alternative way to devise and implement this category of optical systems consists of providing them simply with many single-aperture optical components properly decentered with respect to the mechanical axis of the whole system. Not surprisingly, a multiple-aperture optical component can be conceived as an ensemble of many single-aperture optical components properly arranged in a grid structure. In this case, the grid arrangement of the single-aperture optical components can be designed arbitrarily according to the functions of the multiple aperture system. The category of optical systems under study is well represented by light field cameras, also known as plenoptic cameras. These cameras are typically constituted by a main objective followed by a microlens array (MLA) that is a monolithic optical component consisting of many lenses whose diameters commonly range from $100\mu\text{m}$ to 1mm . In the present treatment, there is no need to restrict the diameters of such multiple apertures to be in that range. In fact, it is assumed that the diameters of the multiple apertures can be chosen arbitrarily. This assumption affects strongly the final number of channels that a multi-aperture system exhibits. In the last section, it is analyzed an example of a multiple-aperture optical system with 5×5 channels with diameters equal to 2.5mm . It is worth observing that, in this context, the terms optical channel and optical branch are used interchangeably.

As already mentioned above, in the class of optical systems under study, multiple apertures ensure a spatial sampling or partitioning of the incoming light. This concept is further clarified if one considers that the light disturbance emitted from (or reflected by) an object plane, located in front of the optical system, undergoes spatial partitioning when traveling through the optical system itself as a direct consequence of the spatial arrangement of the many apertures. Such spatial partitions are then projected onto the image plane that is thereby nothing but a collection of resized sub-areas of the object plane. Depending on

the particular architecture of these optical systems and on the parameters of the individual optical components, a multiple-aperture system can be designed to sample the object plane with a prescribed overlap between adjacent partitions. This means that adjacent channels of the system sample a common sub-area of the object plane, corresponding to the region of overlap between adjacent partitions. In practice, the overlap of adjacent sub-areas of the object plane determines a certain degree of redundancy of the final image at the image plane. As a matter of fact, the final image appears as a collection of portions of the object plane captured from different points of observation. This effect of the shift of the point of observation is exploited in post-processing algorithms that, typically, allow to retrieve a depth map of the observed scene.

The goal of the present chapter is to explore the opportunities offered by symmetry to facilitate the study of optical systems with many apertures. For the sake of clarity and to dispel any doubt, it is important to remark again that this category of optical systems can be conveniently conceived as an ensemble of independent optical sub-systems, each one characterized by peculiar symmetry properties. Depending on the particular optical system under study, it may be possible to identify groups of channels characterized by the same symmetry properties. In this case, the study of the whole optical system is greatly facilitated.

5.2 Parabasal behavior

The application of the principles of Hamiltonian Optics to the study of multi-aperture optical systems relies on the definition of an ensemble of base-rays, one per optical channel, playing the role of reference axis to which the respective characteristic functions are referenced. If the symmetry properties of the several branches composing a multi-aperture optical system are known in advance, a proper definition of the aforementioned ensemble of base-rays is easier. For example, the base-ray of a plane-symmetric system can be suitably defined in the plane of symmetry itself. In the more general case of an optical system without symmetry, an appropriate definition of the base-ray is not immediate. In this case, depending on the characteristics of the optical system under study, a base-ray can be conveniently defined as the ray that connects the center of the aperture stop with the center of the FOV on the object and image sides. In the case of a multi-channel optical system, the definition of a base-ray for each channel is to some extent simplified by the prescribed displacement of the optical components from the mechanical axis of the system. In fact, individual channels are characterized by a specific shift of the aperture stop with respect to the mechanical axis of the optical system itself. Therefore, the definition of a base-ray for each channel is somehow eased by the specific displacement of the respective pupils. In the following Figure 5.1, this concept is illustrated for a multi-aperture system characterized by 3x3 channels arranged in a rectangular grid.

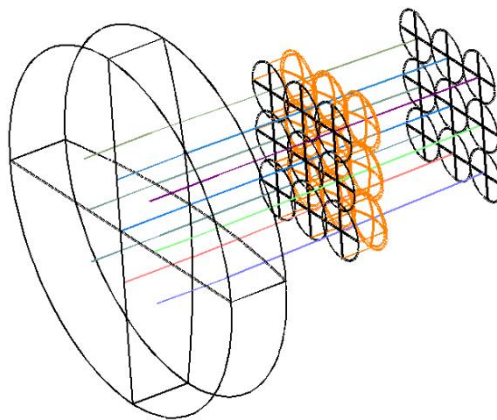


Figure 5.1: Multiple aperture system with 3x3 channels. The central channel is the only one with circular symmetry and its base-ray coincides with its axis of symmetry that, in turn, corresponds to the mechanical axis of the system as a whole.

As remarked above, the definition of such an ensemble of base-rays is of primary importance since each base-ray plays the role of reference axis for its inherent optical channel that, in general, can be a system without any predefined symmetry. In the case the system has a circular symmetric channel, the base-ray coincides with its axis of symmetry that, in turn, corresponds to its mechanical axis.

Tracing the base-rays requires first the definition of a global coordinates system for the whole multi-aperture optical system. Such a global reference system is conveniently defined with the origin lying in the point of intersection between the mechanical axis of the multi-aperture system and an arbitrary surface preceding the multiple pupils plane. Additionally, the global reference system is defined to have the z -axis along the mechanical axis itself. Once a global reference system is established, the tracing of an ensemble of base-rays, one per each channel, is driven by the particular displacements of the aperture stop surfaces of the channels themselves with respect to the z -axis of the global reference system. Substantially, this means that, in the raytracing algorithm, the base-rays are aimed at the predefined locations of the decentered pupils. A base-ray, traversing a particular channel c , defines a series of points of intersection with the surfaces constituting the channel itself. Such points of intersection, in turn, define the origin of two local coordinate systems for each surface of the channel c , indexed with j ; the first coordinate system is related to rays before refraction at the surface j , while the second is related to rays after refraction. Both local reference systems are characterized by the fact that their respective z -axis lie along the base-ray at the intersection with the surface j . In line with the considerations done in the previous chapter, the base-ray of each channel defines also the local coordinates systems associated with the object plane, image plane and pupil plane. This point deserves to be emphasized further in order to acquire a better understanding of the information contained in the full field wavefront aberration function of any optical channel discussed in the next section. More in details, the coordinates of the local reference system in the object plane (or equivalently in the image plane) and those of the local reference system in the exit pupil plane, are the variables used to define the characteristic function associated with any light ray propagating through an optical channel. Therefore, these are also the variables on which depends the full field wavefront aberration function of each channel. These variables have been denoted as H_x , H_y , ρ_x and ρ_y throughout this thesis. Moreover, the various reference spheres centered at various field points of an optical channel can be conveniently defined to have a common vertex at the point of intersection of the base-ray with the exit pupil plane. For the sake of clarity, the local coordinates systems in the image plane and the exit pupil plane are illustrated for a multi-aperture optical system with 9 channels in Figure 5.2. The reference spheres centered at the intersection between the base-ray and the image plane are depicted in green.

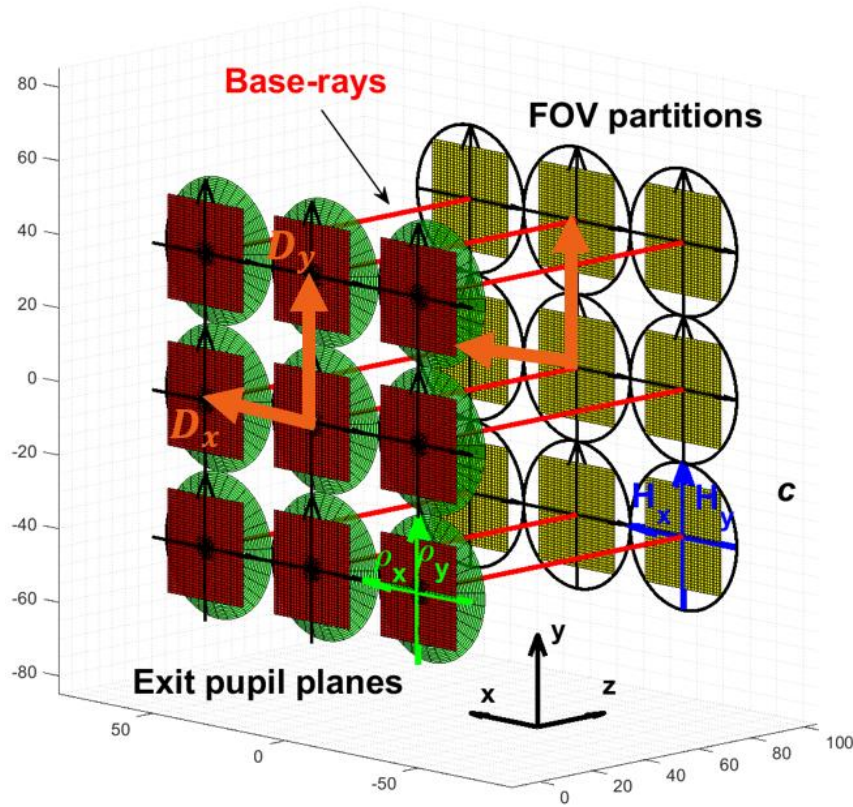


Figure 5.2: Multi-aperture system with the respective base-rays (red), FOV partitions (in yellow it is shown the domain sampled by a rectangular grid of points), exit pupil planes (in red it is shown the domain sampled by a rectangular grid of points) and reference spheres centered at the intersection points between the base-rays and the sub-FOV. The reference axis of the exit pupil and sub-FOV planes are indicated only for the channel *c*. The parameters denoting the transversal displacement of the pupils from the mechanical axis are D_x and D_y .

Once the base-rays and the reference axis of the local exit pupil and FOV planes of each branch of the system are defined, the paraxial behavior of the optical channels can be studied by tracing paraxial rays similar to the study of the paraxial behavior of circular symmetric optical systems. Paraxial rays are traced in a neighborhood of the base-rays.

In the next section, few considerations are done relative to the possibility to simplify the study of the aberration behavior of this class of systems exploiting the symmetry of different FOV partitions in corresponding channels. Then, different methods of visualization of the aberration behavior of the various channels are described.

5.3 Full-field wavefront aberration function in optical systems with many apertures

Different strategies can be devised to describe the aberration behavior of optical systems with many apertures. One potential route to analyze the aberration behavior of these optical systems relies on the possibility to leverage the symmetry of the various channels constituting the system as a whole. In fact, in the category of multi-channel optical systems considered in this chapter, it is typically possible to make specific assumptions on the class of symmetry to which different groups of optical channels belong. These assumptions can be made based on the specific geometry of the various channels constituting the optical system. For instance, considering a multi-channel system like a light field camera, this consists of the sequence of two sub-systems, of which the first is the main objective while the second is the multi-aperture system, namely an array of microlenses. This latter sub-

system's function is to provide a channeling of the incoming light and for such reason, different microlenses can be regarded as defining different channels of the optical system. Specifically, the various optical channels can be classified depending on the peculiar geometrical displacement of the inherent microlens with respect to the mechanical axis of the system. Consequently, each channel projects a particular partition of the object-side FOV onto a corresponding partition of the image-side FOV located behind the microlens. Each image-side partition of the FOV has the same area as a microlens. Thus, the area of the entire image-side FOV is equal to the area of the multi-aperture plane in the present treatment. These observations pave the way to the application of the concepts of Hamiltonian Optics to the study of the aberration behavior of different optical channels classified according to the displacement of the respective exit pupils in the multi-aperture plane or equivalently, according to the displacement of the respective partitions of the image-side FOV.

The following Figure 5.3 is aimed at illustrating the correspondence between the optical channels $c1$, $c2$, $c3$ and $c4$ whose pupils are located in corresponding positions in different quadrants of the image-side FOV.

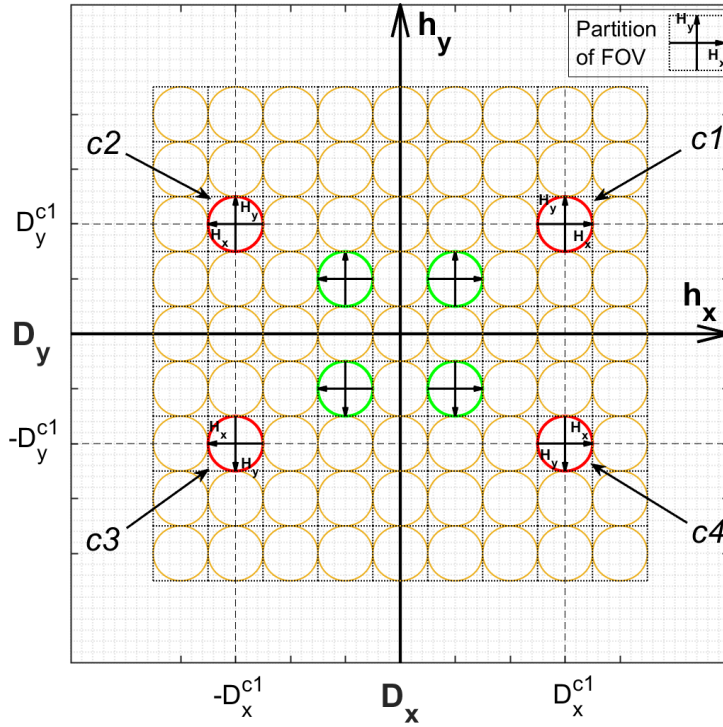


Figure 5.3: It is shown the overlap between the multiple aperture plane $D_x D_y$ and the image-side FOV plane $h_x h_y$. Each channel corresponds to a partition of the FOV whose local coordinates are the field variables H_x and H_y (ray coordinates). It is highlighted the correspondence between channels located in different quadrants of the FOV plane.

The optical channel $c1$ characterized by a specific displacement of the pupil ($D_x = D_x^{c1} > 0, D_y = D_y^{c1} > 0$) with respect to the mechanical axis of the system, can be expected to have a behavior that is symmetric about the origin with respect to that of the corresponding optical channel $c3$ whose specific displacement of the pupil is ($D_x = D_x^{c3}, D_y = D_y^{c3}$) where ($D_x^{c3} = -D_x^{c1}, D_y^{c3} = -D_y^{c1}$). The same considerations are valid for the channels $c2$ and $c4$, symmetric about the h_y and h_x axis with respect to the channel $c1$, provided that their displacement coordinates are ($D_x^{c2} = -D_x^{c1}, D_y^{c2} = D_y^{c1}$) and ($D_x^{c4} = D_x^{c1}, D_y^{c4} = -D_y^{c1}$). For the sake of simplicity, in the present treatment, multi-aperture optical systems constituted by only spherical surfaces are taken into account. This

facilitates further assumptions on the symmetry of optical channels characterized by a similar displacement of the pupil.

Analyzing Figure 5.3, it can be observed that the field variables (H_x, H_y) of each channel can be identified with the coordinates of the local reference system in the partition of the image-side FOV behind the channel itself. In the light of this observation, it can therefore be deduced that an advantageous approach to analyze optical systems with many apertures can be based on the study of the aberration behavior of optical channels grouped in the same quadrant of the plane $h_x h_y$ that is the image-side FOV of the optical system as a whole. The advantage of this approach will become evident in section 5.5 where the parametric curves describing the transverse ray aberrations will be displayed for channels in different quadrants of an example of a plenoptic camera. For now, it is sufficient to observe that, as a consequence of the previous considerations, the full field wavefront aberration function $W^{c3}(H_x, H_y, \rho_x, \rho_y)$ of the channel $c3$, located in the third quadrant, can be easily retrieved from the full field wavefront aberration function $W^{c1}(H_x, H_y, \rho_x, \rho_y)$ of the corresponding channel $c1$, located in the first quadrant of the plane $h_x h_y$, just inverting the sign of the field coordinates of the wavefront aberration function itself, namely $H_x \rightarrow -H_x$ and $H_y \rightarrow -H_y$. The following Equation 5.1 simply describes this situation for the channels mentioned above and illustrated in Figure 5.3.

$$\begin{cases} W^{c2}(H_x, H_y, \rho_x, \rho_y) = W^{c1}(-H_x, H_y, \rho_x, \rho_y) \\ W^{c3}(H_x, H_y, \rho_x, \rho_y) = W^{c1}(-H_x, -H_y, \rho_x, \rho_y) \\ W^{c4}(H_x, H_y, \rho_x, \rho_y) = W^{c1}(H_x, -H_y, \rho_x, \rho_y) \end{cases}$$

Equation 5.1

These observations facilitate substantially the study of the full field aberration behavior of multi-aperture systems like a light field camera. In fact, once the aberration behavior of the optical channels located in one particular quadrant of the FOV has been obtained, the extension of this result to optical channels belonging to the remaining quadrants is straightforward. Bearing in mind these considerations, in the following sections, two different approaches are presented to analyze the aberration behavior of a multi-aperture optical system. The first is based on the analysis of the aberration coefficients depending on the decentering parameters of the channels constituting the optical system itself. This first approach has been already introduced in Chapter 4 of this thesis to investigate the dependence of the aberration coefficients on the decentering parameters of the second lens constituting the simple multi-configuration system studied in that context. The second approach is based on a combination of Zernike Fringe coefficients to obtain full-field expressions for different types of aberrations represented again as functions of the decentering parameters. This second approach has been introduced throughout Chapters 2 and 3. These two different descriptions of the full field wavefront aberration function of a system with multiple apertures can be restricted to the channels spanning the only first quadrant of the image-side FOV since the behavior of the channels in the remaining quadrants can be found with symmetry operations about the axis and the origin of the FOV plane.

5.4 Wavefront aberration function expanded with a general power series

Equation 4.2 from Chapter 4 can be used to describe the general behavior of the various branches of a multi-aperture optical system. For the sake of clarity, this equation is reported here as Equation 5.2 where the simple addition of a superscript c to the wavefront function W^c and to the inherent aberration coefficients w^c is useful to denote the channel to which this expression is related.

$$W^c(H_x, H_y, \rho_x, \rho_y) = \sum_{n=0}^{n_c} \sum_{\lambda=0}^n \sum_{\mu=0}^{\lambda} \sum_{\nu=0}^{\mu} w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}^c H_x^{n-\lambda} H_y^{\lambda-\mu} \rho_x^{\mu-\nu} \rho_y^{\nu}$$

Equation 5.2

The superscript c is assumed to be $c = 1, \dots, N_c$ where N_c is the total number of channels present in the multi-aperture optical system. The coefficients $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}^c$ weight the respective terms in the polynomial expansion in Equation 5.2. The subscripts $n - \lambda, \lambda - \mu, \mu - \nu, \nu$ indicate the power to which the ray coordinates H_x, H_y, ρ_x and ρ_y are raised. In particular, the subscript n , indicating the total power of the inherent term, is such that $n = 1, \dots, n_c$ where n_c is the maximum order of interest in the wavefront aberration expansion. For example, if we are concerned with primary or secondary aberrations (retaining the same nomenclature used for rotationally symmetric systems), then $n_c = 4$ or $n_c = 6$ respectively. Different channels exhibit, in general, different symmetry properties. This situation is reflected in the values of the aberration coefficients that can be displayed as functions of the decentering parameters of different channels of a multi-aperture system. This representation is similar to what has been discussed in Chapter 4. The main difference is that, in the present context, the possible displacements (D_x, D_y) of the different channels are constrained to be in a grid of points defined by the multiple aperture plane structure, while in Chapter 4 such constraint was absent.

In agreement again with the considerations done in Chapter 4, in general in Equation 5.2, there are 35 terms of 4th order for each optical channel. If a particular channel in the optical system exhibits a specific symmetry behavior, the number of 4th order terms in Equation 5.2 can be reduced following the specific symmetry. Equivalently, it can be assumed that the terms that are absent due to symmetry in the power series of Equation 5.2, have the respective coefficients equal to zero. For example, in a multiple aperture system with 5x5 optical branches, the variation of primary spherical aberration over the channels can be visualized displaying the inherent 4th order coefficients as functions of the channels decentering coordinates D_x and D_y . More in detail, for primary spherical aberration, the five inherent 4th order terms $\rho_x^4, \rho_x^3\rho_y, \rho_x^2\rho_y^2, \rho_x\rho_y^3$ and ρ_y^4 are weighted by the coefficients $w_{0040}^c, w_{0013}^c, w_{0022}^c, w_{0031}^c$ and w_{0004}^c respectively. As said above, these coefficients can be displayed as functions of the individual optical channel decentering parameters in analogy to what has been done in the previous chapter.

As an alternative to the full-field wavefront aberration function, it can be equivalently investigated the transverse ray aberration behavior for each branch of the whole system. This behavior can be obtained deriving the equations that describe the ray coordinates displacements along the H_x and H_y directions in the image plane. In Equation 1.35- Equation 1.36 in Chapter 1, these quantities have been indicated with δH_x and δH_y respectively. To avoid misunderstanding, in this chapter they are denoted respectively with $\varepsilon_x^c(H_x, H_y, \rho_x, \rho_y)$ and $\varepsilon_y^c(H_x, H_y, \rho_x, \rho_y)$ where the full dependence on the ray coordinates serves to point out the general nature of these two functions as will be clarified later. Additionally, the superscript c specifies again the associated channel. Thus, the transverse ray coordinates ε_x^c and ε_y^c are computed as follows

$$\begin{cases} \varepsilon_x^c(H_x, H_y, \rho_x, \rho_y) = -\frac{R}{n'} \frac{\partial W^c(H_x, H_y, \rho_x, \rho_y)}{\partial \rho_x} \\ \varepsilon_y^c(H_x, H_y, \rho_x, \rho_y) = -\frac{R}{n'} \frac{\partial W^c(H_x, H_y, \rho_x, \rho_y)}{\partial \rho_y} \end{cases}$$

Equation 5.3

where n' is the refractive index in the image space and R is the radius of curvature of the reference sphere for the optical channel c . In principle, the radius of curvature of the reference sphere changes for different field points in the same optical channel, but in

practical calculations, it is conveniently considered to be equal to the exit pupil position, corresponding to the distance between the exit pupil and the image plane. Additionally, here it is supposed that the radius of curvature of the reference sphere is the same for all the optical channels constituting the whole optical system. This assumption is necessary to give consistency to the forthcoming comparison of the full field aberration behavior of different optical channels. As a matter of fact, in the absence of this assumption, such a comparison would be based on computations performed with respect to a different reference sphere for each optical channel and this situation would lack consistency. Recalling that the term W^c (contained in Equation 5.3) is a power series expansion in the ray coordinates, it is immediately evident that the transverse ray coordinates in Equation 5.3 are also described by two power series expansions in the ray coordinates H_x , H_y , ρ_x and ρ_y whose order is reduced by one.

After that, considering any channel of the whole system, the transverse ray aberrations along the inherent partition of FOV can be plotted fixing the field coordinates in correspondence with a field point of interest ($H_x = H_{x0}$, $H_y = H_{y0}$) and replacing the pupil coordinates ρ_x and ρ_y with new parametrized coordinates $\rho_x(t) = \rho_0 \cos t$ and $\rho_y(t) = \rho_0 \sin t$ where t is a parameter and ρ_0 is the radius of the circle determined by the intersection of the congruence of light rays emitted from the point (H_{x0}, H_{y0}) with the decentered exit pupil of the optical channel under study. As a consequence of this, the transverse ray aberrations have the following parametrized coordinates along the H_x and H_y directions of the FOV partition identified by the channel c

$$\begin{cases} \varepsilon_x^c(H_{x0}, H_{y0}, \rho_0 \cos t, \rho_0 \sin t) = \varepsilon_x^c(t) \\ \varepsilon_y^c(H_{x0}, H_{y0}, \rho_0 \cos t, \rho_0 \sin t) = \varepsilon_y^c(t) \end{cases}$$

Equation 5.4

If normalized pupil coordinates are used, setting $\rho_0 = 1$ in Equation 5.4 implies that the transverse ray coordinates $\varepsilon_x^c(t)$ and $\varepsilon_y^c(t)$ in the image plane are related to a congruence of light rays grazing the rim of the exit pupil.

In section 5.5, the wavefront aberration coefficients w^c and the parametric curves $\varepsilon_x^c(t)$ and $\varepsilon_y^c(t)$ describing transverse ray aberrations are calculated for each channel of an example of a plenoptic camera.

5.4.1 Wavefront aberration function expanded with full field-dependent Zernike polynomials

In this section, it is shown a generalization of the formulas presented in Chapters 2 and 3 to multiple-aperture optical systems. This generalization consists simply of allowing the field dependence of Zernike coefficients to be described by general power series in the field coordinates up to the 4th order. In this way, the peculiar nodal behavior developed by different types of aberrations in the FOV of the channels of a multiple aperture system can be easily described. Recalling Equation 3.14 from Chapter 3, it has been used to describe the full field behavior of primary aberrations in double-plane symmetric optical systems with perturbations due to tilts and displacements of the optical components. In that case, the full field expressions of the Zernike coefficients have been obtained projecting the full field wavefront aberration function of perturbed double-plane symmetric systems up to the 4th order onto the basis of the pupil dependent Zernike polynomials in the Fringe indexing scheme. The resulting expressions were general power series in the field coordinates containing even and odd terms up to the 2nd degree since we were concerned with only primary aberrations. For the present treatment, these expressions are further expanded up to the 4th order to correctly describe the influence of higher-order aberrations on the behavior of the optical channels characterized by a more pronounced decentering from the mechanical axis of the system. In line with the treatment of Chapter 2, different types of

vectorial aberrations are obtained from a proper combination of the field-dependent expressions of the Zernike coefficients. It is worth recalling that vectorial aberrations are represented similarly as vectors since they are characterized by magnitude and phase. For the sake of clarity, Equation 3.14 from Chapter 3 is reported in full below with the addition of terms describing secondary aberrations. In particular, one obtains the following full field wavefront aberration expression $W^c(H_x, H_y, \rho, \vartheta)$ where the superscript c is again necessary to specify each channel of the multiple aperture system.

$$\begin{aligned}
 W^c(H_x, H_y, \rho, \vartheta) = & 6F_4^0(H_x, H_y)\rho^4 + \\
 & 3\sqrt{[F_3^1(H_x, H_y)]^2 + [F_3^{-1}(H_x, H_y)]^2}\rho^3 \cos\left[\vartheta - \tan^{-1}\left(\frac{F_3^{-1}(H_x, H_y)}{F_3^1(H_x, H_y)}\right)\right] + \\
 & \left[2F_2^0(H_x, H_y) - 6F_4^0(H_x, H_y) - \sqrt{[F_2^2(H_x, H_y)]^2 + [F_2^{-2}(H_x, H_y)]^2}\right]\rho^2 + \\
 & 2\sqrt{[F_2^2(H_x, H_y)]^2 + [F_2^{-2}(H_x, H_y)]^2}\rho^2 \cos^2\left[\vartheta - 0.5 \tan^{-1}\left(\frac{F_2^{-2}(H_x, H_y)}{F_2^2(H_x, H_y)}\right)\right] + \\
 & \sqrt{[F_3^3(H_x, H_y)]^2 + [F_3^{-3}(H_x, H_y)]^2}\rho^3 \cos^3\left[\vartheta - 0.33 \tan^{-1}\left(\frac{F_3^{-3}(H_x, H_y)}{F_3^3(H_x, H_y)}\right)\right] + \\
 & 4\sqrt{[F_4^2(H_x, H_y)]^2 + [F_4^{-2}(H_x, H_y)]^2}\rho^4 \cos^2\left[\vartheta - 0.5 \tan^{-1}\left(\frac{F_4^{-2}(H_x, H_y)}{F_4^2(H_x, H_y)}\right)\right] + \\
 & 10\sqrt{[F_5^1(H_x, H_y)]^2 + [F_5^{-1}(H_x, H_y)]^2}\rho^5 \cos\left[\vartheta - \tan^{-1}\left(\frac{F_5^{-1}(H_x, H_y)}{F_5^1(H_x, H_y)}\right)\right]
 \end{aligned}$$

Equation 5.5

In Equation 5.5, the complete expressions of the Zernike field dependences, denoted with $F_n^m(H_x, H_y)$, are power series in the field coordinates expanded up to the 4th order to account for both primary and secondary aberration contributions. It is important to remark that, since such expressions are used here to describe the field dependence of asymmetric systems with arbitrarily decentered surfaces, they contain terms of both odd and even order in the coordinates H_x and H_y . In this way, a more general description is given to the possible field behaviors of primary and secondary aberrations for the optical channels of a multi-aperture system using the Zernike polynomials expansion. The computation of Equation 5.5, based on raytracing, allows obtaining the locations of the nodes of different aberration types in the portions of FOV inherent to the different optical channels identified with the superscript c . In the following section, the magnitude of the full field dependence of the various aberration types described in Equation 5.5 is calculated for the 5x5 channels of an example of a plenoptic camera.

5.5 Application: aberrations behavior in a plenoptic camera

In this section, it is analyzed the general behavior of a multi-aperture optical system like a plenoptic camera. In order to acquire a better understanding of the general structure of this class of optical systems, it is shown the optical layout of an example of this type of camera in the following Figure 5.4-Figure 5.5. The main objective is a Cooke triplet, while the multi-aperture component is constituted by plano-convex lenses with a radius of curvature $R_c = -8.36\text{mm}$, diameter $D = 2.5\text{mm}$ and working F-number $WFNO = 5$. The entrance pupil diameter of each channel is equal to $ENPD = 4.25\text{mm}$ and the effective

focal length is equal to $EFFL = 19mm$. The object plane is set at $z_o = 250mm$ from the system. The simulation wavelength is $\lambda_0 = 0.55\mu m$. The system under study is characterized by 5×5 optical branches (or optical channels). The aperture stop for each branch of the system is placed in correspondence with the last surface before the image plane, that is the second surface of each plano-convex lens. The following results are obtained with Zemax OpticStudio [13].

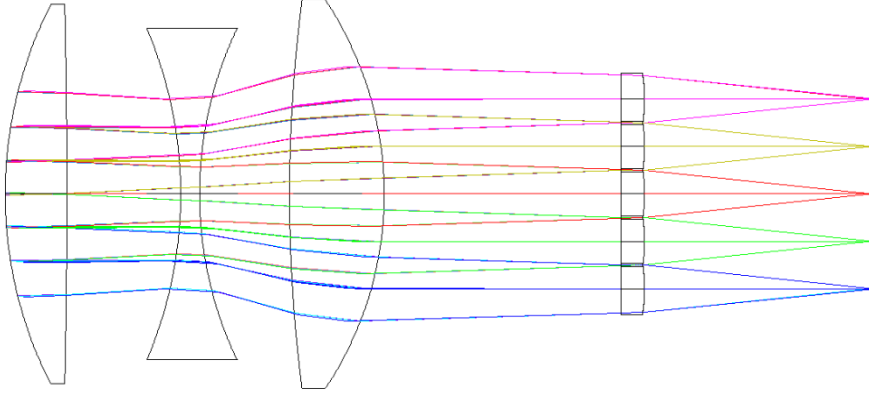


Figure 5.4: Layout of an example of a plenoptic camera: side view.

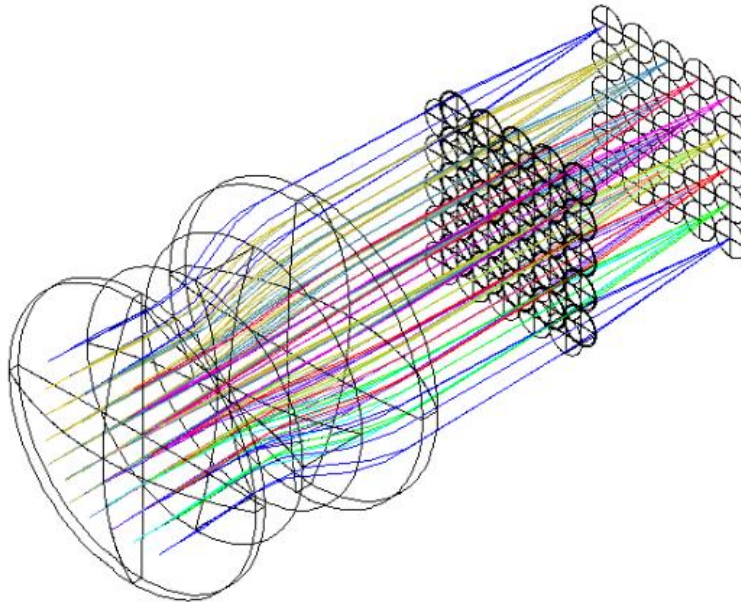


Figure 5.5: Layout of an example of a plenoptic camera: isometric view.

The optical parameters of this particular example are listed in Table 5.1, reported below.

Table 5.1: System parameters

Component	Surface	Radius (mm)	Thickness (mm)	Material	D_x (mm)	D_y (mm)
Cooke Triplet	1	22	3.25	SK16	-	-
	2	-436	6		-	-

	3	-22.25	1	F2	-	-
	4	20.3	4.76		-	-
	5	80	5	SK16	-	-
	6	-18.4	12.50		-	-
Decentred lenses	7	Inf	1.25	Fused Silica	[-5, -2.5, 0, 2.5, 5]	
	8 (STOP)	-8.36	12.2		[-5, -2.5, 0, 2.5, 5]	

The scope of a plenoptic camera is to project onto the image plane contiguous and partially overlapped portions of the object-side FOV. The degree of overlap between these adjacent sub-areas of the FOV plane plays a fundamental role in the realization of this category of optical systems that effectively behave as multi-stereo imaging systems. In the following Figure 5.6, this concept is exemplified to acquire a better understanding of the working principle of this type of optical system. It is shown the overlap between contiguous partitions of the object-side FOV.

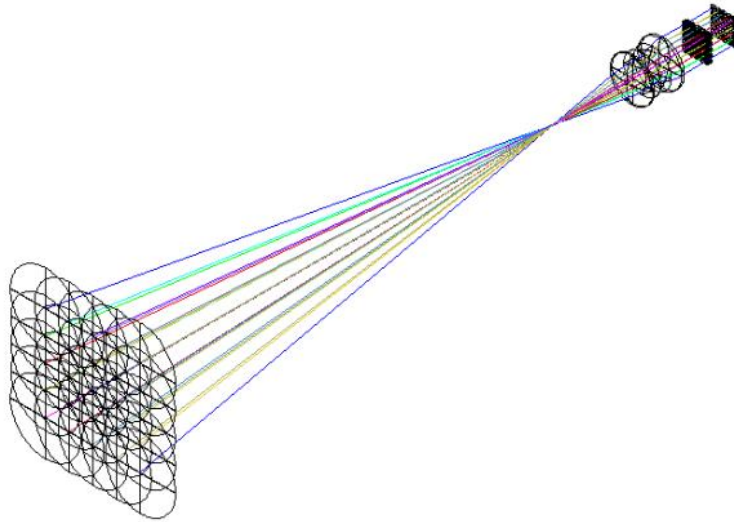


Figure 5.6: The object-side FOV is divided into multiple partitions, each one related to a particular decentered pupil in the multi-aperture plane. Adjacent partitions are partially overlapped; this means that the common area shared between them is “seen” from two different points of observation.

In Figure 5.6 the centers of the several partitions of FOV ascribed to individual branches of the multiple aperture camera are located by their respective base-rays. For the only central channel (that is rotationally symmetric), the base-ray coincides with its axis of symmetry that, in turn, coincides with the mechanical axis of the whole optical system. In the following Figure 5.7, it is shown a front view of the optical system under study to emphasize better the overlap between the adjacent portions of the FOV and the fundamental role played by the base-rays of individual optical branches. It is evidenced that the base-ray of each optical branch connects the center of the respective portion of the FOV in the object space with the center of the corresponding portion of FOV in the image space while passing through the center of the respective aperture stop. These two planes (the object-side FOV and the image-side FOV) can be equivalently defined as posterior and anterior base-planes respectively, using the same terminology related to paraxial optics.

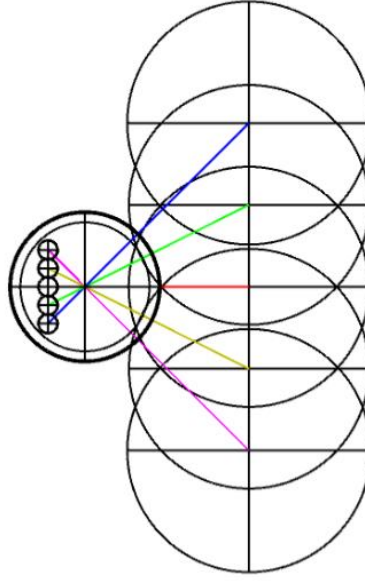


Figure 5.7: Front view of the system under study: for clarity, only 5 channels are shown along with their respective base-rays.

As explained in the previous section, the individual channels of a multi-aperture imaging system are characterized by specific symmetries that can be studied by analyzing the behavior of the aberration coefficients as a function of the particular geometry of the optical channels. In what follows, it is first shown an analysis display for the visualization of the transverse ray aberrations in the FOV partitions of the system under study. Later on, two additional methods to analyze the wavefront aberration coefficients for the various channels are shown, the first based on a general power series expansion, the second based on a Zernike polynomials expansion.

5.5.1 Transverse ray aberration display

Relatively to the system under investigation, different channels are properly classified with the decentering parameters D_x and D_y of the plano-convex lenses in the multi-aperture plane. The aperture stop is placed in correspondence with the last surface before the image plane, therefore, for every channel, the exit pupil is located in this same position. Thus, a multiple pupil plane containing the exit pupils for the several channels coincides with the multi-aperture plane of the optical system under study. The parameters D_x and D_y are measured from the mechanical axis of the system that coincides with the optical axis of the central axially symmetric channel. As a consequence, the central channel is identified by $D_x = 0mm$ and $D_y = 0mm$, while the remaining channels are characterized by parameters $D_x \neq 0mm$ and $D_y \neq 0mm$. For the sake of clarity, the multiple aperture (and pupil) plane is shown in the following Figure 5.8 to highlight the respective displacements D_x and D_y . Bearing in mind the remarks done before and recalling the previous Figure 5.4-Figure 5.5, the partitions of the image-side FOV associated with each channel are located just after the respective pupils. For this reason, in Figure 5.8 it is shown a grid of parametric curves (in correspondence with a grid of 11×11 field points) overlapped to the sub-areas of the exit pupils. The H_x and H_y coordinates of such parametric curves, indicated with $\varepsilon_x^c(t)$ and $\varepsilon_y^c(t)$, represent the transverse ray aberrations for the various optical channels constituting the plenoptic camera under study. In the particular case of Figure 5.8, ε_x^c and ε_y^c are calculated tracing $N_H^2 N_\rho = 14^2 * 50 = 9800$ rays through each channel of the system under study. Rays are traced from field points arranged in a square grid with $N_H = 14$ points per side and intersect the exit pupil in $N_\rho = 50$ points arranged in a circle with radius $\rho_0 =$

1 in normalized coordinates. The coordinates of the intersection of each ray with the several FOV partitions are measured with respect to chief rays, thus obtaining the transverse ray displacements. These data are then fitted (in the least-squares sense) with general power series up to the order n_c of interest (in this case $n_c = 5$). An alternative way to retrieve the parametric curves in Figure 5.8 is based on the calculation of Equation 5.3-Equation 5.4 setting the parameter ρ_0 equal to 1. Of course, this second method requires first the calculation of W^c expanded with a general power series.

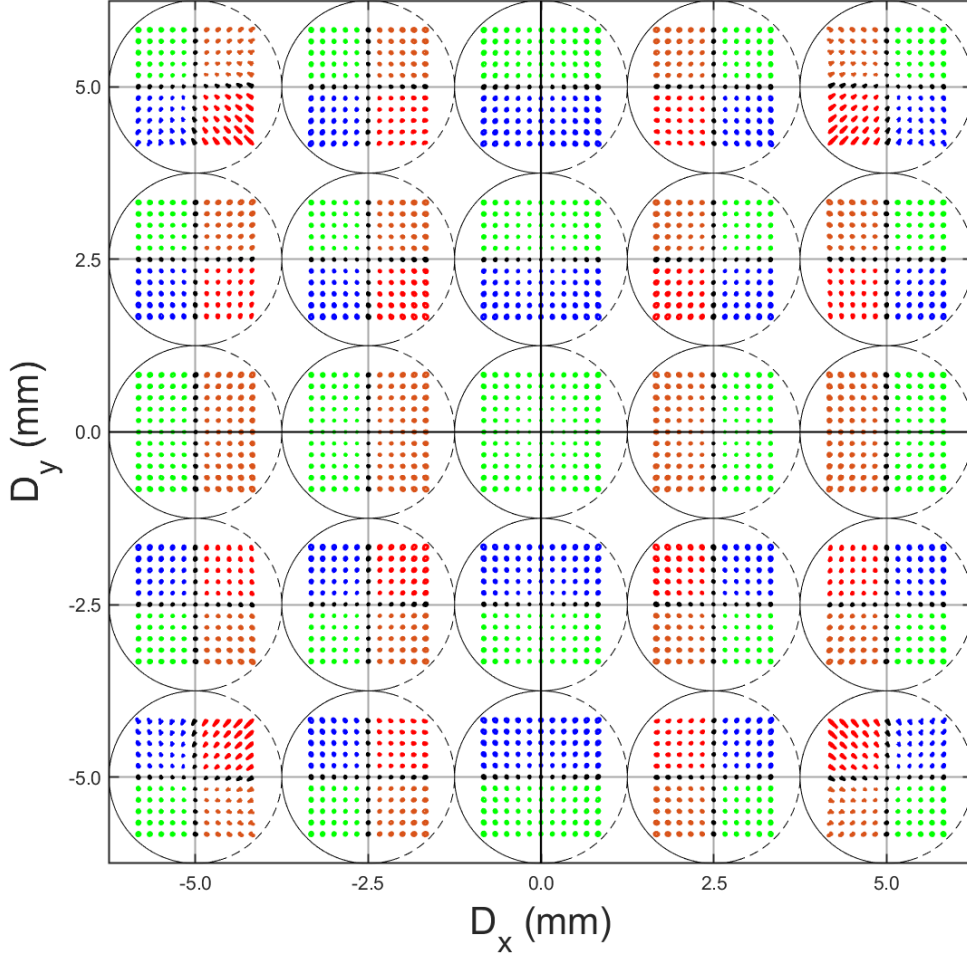


Figure 5.8: Multiple aperture (and pupil) plane with decentering parameters D_x and D_y overlapped to the corresponding sub-areas of the image-side FOV where transverse ray aberrations are displayed as parametric curves over grids of 11x11 field points. This grid of points has been arbitrarily chosen for visualization purposes and it is different from the 14x14 sampling grid mentioned above.

The decentering parameters of the plano-convex lenses in the multi-aperture plane play a substantial role in the identification of “symmetric behaviors” of the aberrations in different optical channels in agreement with the considerations done in the previous sections of this chapter. The same shapes of the parametric curves visible in a particular FOV partition are found also in the other corresponding partitions but relocated to different field points. To highlight this simple situation, parametric curves belonging to corresponding sub-quadrants of different FOV positions are colored in the same way. This important point can be understood better considering that the main objective (in this case a Cooke triplet) is rotationally symmetric and the decentering parameters of the plano-convex lenses in the multi-aperture plane act effectively as perturbations to the symmetry of the main objective. For this reason, parametric curves in corresponding FOV partitions have the same shape but different orientations depending on the peculiar displacements of the

pupils. Therefore, parametric curves in corresponding channels located in different quadrants of the FOV can be retrieved from each other performing symmetry operations of the FOV quadrants such as reflections about the vertical axis, about the horizontal axis and the origin. As a consequence, the analysis of the multi-aperture optical system shown in Figure 5.4-Figure 5.5 can be restricted to a single quadrant of the multi-aperture plane. This situation is better illustrated in the following Figure 5.9 where corresponding parametric curves, located in different partitions of the FOV (belonging to different quadrants), are plotted with the same color and labeled with the same number. In particular, the sub-FOV considered in this figure are at the corners of the quadrants of the full-FOV, in correspondence with the positions $(D_x = 5mm, D_y = 5mm)$, $(D_x = -5mm, D_y = 5mm)$, $(D_x = -5mm, D_y = -5mm)$ and $(D_x = 5mm, D_y = -5mm)$. For the sake of clarity, in each partition of the FOV are shown parametric curves over a grid of 3x3 field points. As said above, such parametric curves are general power series depending on the ray coordinates. In this case, these curves describe the transverse ray aberrations up to the 5th order for different optical channels constituting the optical system as a whole.

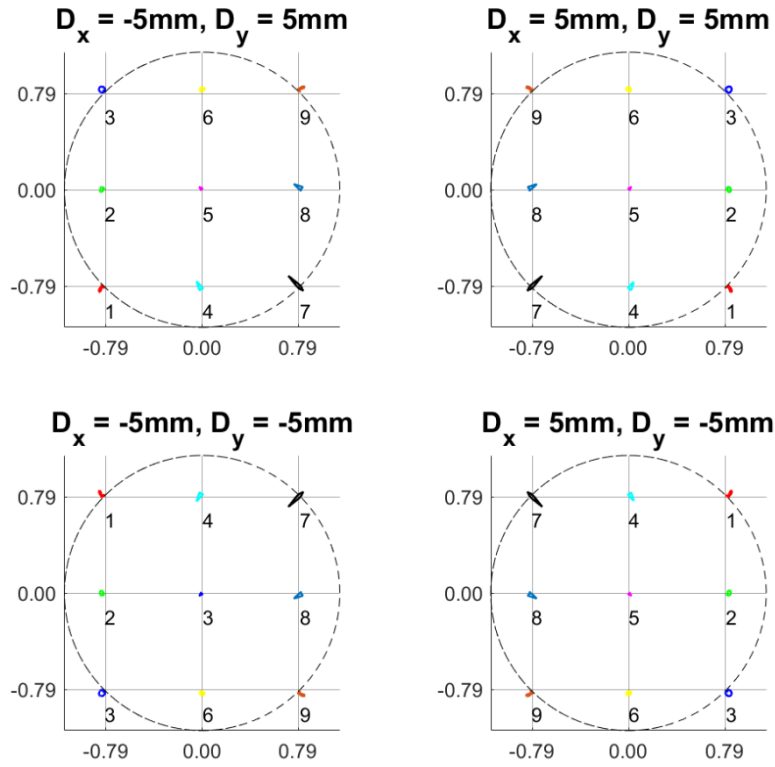


Figure 5.9: The parametric curves describing transverse ray aberrations in different field points in the channels c_2 , c_3 and c_4 can be obtained from the respective aberrations in the channel c_1 through reflections across the axis and point reflection operations.

Figure 5.9 is aimed at emphasizing the similarity of the aberration behavior of different optical channels that can be obtained by reflection symmetry operations. On close observation, it can be noticed that the parametric curve labeled with the number “7” (plotted in black) is located approximately in $(H_x = -0.79mm, H_y = -0.79mm)$ in the sub-FOV whose displacement is $(D_x = 5mm, D_y = 5mm)$ in the first quadrant. The corresponding curve in the sub-FOV of the channel with displacement $(D_x = -5mm, D_y = -5mm)$ located in the third quadrant, is found approximately in $(H_x = 0.79mm, H_y = 0.79mm)$. This example clearly illustrates that the analysis of this class of optical systems can be effectively simplified taking into account that, corresponding optical branches of the multi-aperture system as a whole, can be characterized by symmetric behaviors and consequently,

their investigation can be conveniently restricted to the optical branches located in the same quadrant of the full-FOV. In the following Figure 5.10, it is shown a blow-up of the first quadrant of the full-FOV (previously shown in Figure 5.8). The parametric curves are again plotted over a grid of 11x11 field points for each sub-FOV.

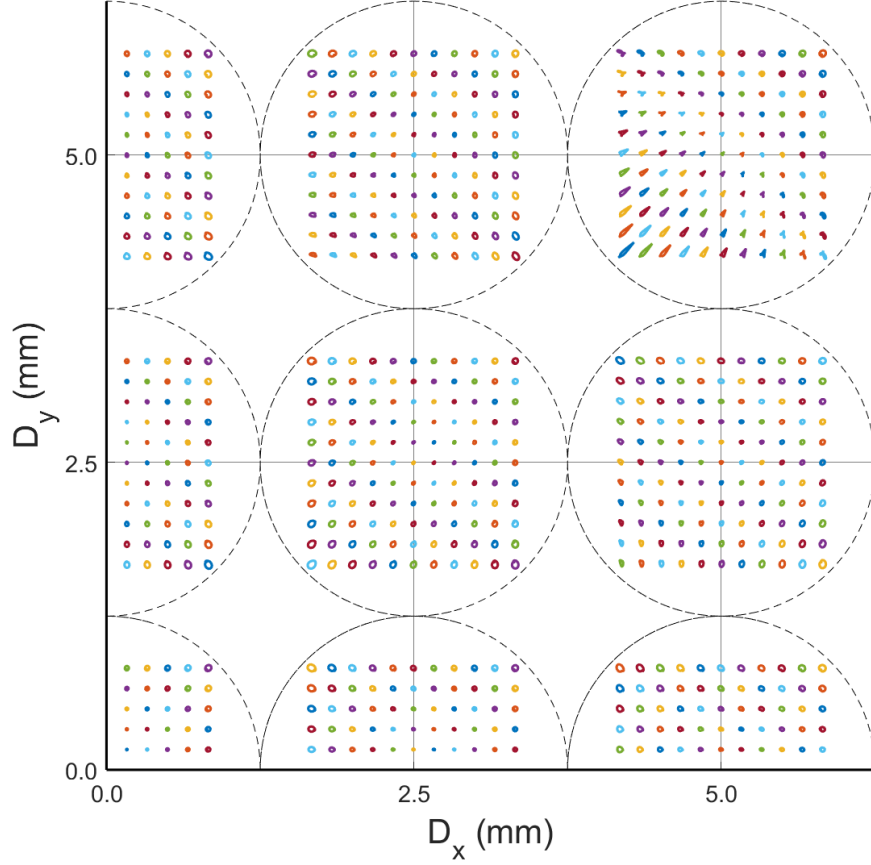


Figure 5.10: FOV partitions in the first quadrant. Grids of 11x11 parametric curves are displayed.

The complete transverse ray aberration behavior over the multiple channels constituting the multi-aperture optical system is obtained reflecting the parametric curves in Figure 5.10, as explained before, with respect to the vertical axis, to the horizontal axis and the origin of the FOV.

As explained up to now, the visualization of the parametric curves over a grid of field points for each optical channel of a multiple-aperture system can be beneficial to acquire a general overview of the aberration behavior of the different channels over the full-FOV. In this case, such parametric curves contain the contributions of all the aberrations up to the order of interest. If only a particular aberration type is of interest, rather than plotting the full power series constituting the parametric curves of the transverse ray aberrations, it is possible to plot only the corresponding term obtaining, in this case, the deviation of the light rays due to the contribution of the only aberration type taken into account.

5.5.2 Wavefront aberration coefficients display

An alternative way to visualize the aberration behavior of the channels of the light field camera considered in this example consists of displaying the magnitude of the wavefront aberration coefficients as a function of the displacements (D_x, D_y) of the channels themselves. As explained in the previous section, Equation 5.2 can be used for this purpose. In the following series of figures, the aberration coefficients $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}^c$ in Equation 5.2

are described as $w^c = w(D_x, D_y)$, namely as surfaces whose height, in correspondence of the coordinates (D_x, D_y) , represents the magnitude of the associated aberration type for the channel whose decentering parameters are exactly (D_x, D_y) . Therefore, the domain of these surfaces is clearly constituted by a grid of discrete points representing the decentering values of the respective optical channels with respect to the mechanical axis of the optical system or the displacements of the sub-FOV of the respective channels. The values of the aberration coefficients are retrieved with the method described in Chapter 4, namely fitting (in the least-square sense) the multivariate model (up to the 6th order) represented by Equation 5.2 to the raytracing data. In this case, for each channel of the light field camera under study is traced a number of rays equal to $N_H^2 N_\rho^2 = 160000$ where $N_H = 20$ and $N_\rho = 20$ are the number of sampling points along the directions of two uniform square grids in the FOV and exit pupil planes respectively. While tracing rays whose variables H_x, H_y, ρ_x and ρ_y are allowed to have values equal to the points of the sampling grids, it is measured the optical path difference with respect to the chief rays in waves at the wavelength $\lambda_0 = 0.55\mu\text{m}$. The rays traced from the FOV grid and passing through the center of the pupil grid correspond to an ensemble of chief rays to which the calculation of the optical path differences of other rays are referenced. In the figures below are shown the fitting results between the multivariate model in Equation 5.2 and the raw optical path differences of traced rays with respect to chief rays, that is the wavefront error associated with individual rays traced through the system. In these calculations, the mean value of the wavefront error of rays through the pupil is not subtracted out from the wavefront error of individual rays [13].

Focusing on primary (or 4th order) aberrations and recalling that, in general, an optical branch of the multiple-aperture system is asymmetric, the total number of 4th order terms in the power series of Equation 5.2 is 35 (this value is given by the difference between the binomial coefficients $\binom{n_1+4}{n_1} - \binom{n_2+4}{n_2} = 70 - 35 = 35$ where $n_1 = 4$ and $n_2 = 3$ in this case, or calculating directly the binomial coefficient $\binom{n_1+3}{n_1} = 35$ for a homogeneous polynomial of 4th order with four ray coordinates). More in detail, in Equation 5.2 the order n equals $n_c = 4$ and for a better understanding, the coefficients $w_{n-\lambda, \lambda-\mu, \mu-\nu, \nu}^c$ are indicated as w_{klpq} where the superscript c has been removed because the current optical channel is already identified by the decentering parameters D_x and D_y (being the axis of the domain) and $n - \lambda, \lambda - \mu, \mu - \nu$ and ν are replaced respectively with k, l, p and q that are the exponents to which the ray coordinates H_x, H_y, ρ_x and ρ_y are raised. In principle, the analysis of the aberrations of asymmetric optical systems should start considering terms of 3rd order since these give the first relevant contribution to the imagery of an asymmetric optical system beyond the paraxial behavior. Nevertheless, in the present treatment, we intentionally neglect this fact because our main concern is the description of the method of analysis.

The following figures are intended to give an overview of the behavior of the 4th order aberration coefficients for the 5x5 channels of the light field camera under study. This method of visualization of the aberration coefficients $w_{klpq} = w_{klpq}(D_x, D_y)$, namely as functions of the decentering parameters D_x and D_y , can be useful to understand the “structural behavior” of aberrations. Multiple aperture systems, like plenoptic cameras, can be conceived as an ensemble of many individual optical sub-systems differing from each other for some structural parameter, therefore the visualization of the aberration coefficients depending on such parameters can be insightful in this sense. In the following Figure 5.11, the behaviors of the five coefficients related to spherical aberration, $w_{0004}, w_{0013}, w_{0031}, w_{0040}$ and w_{0022} are visualized depending on the decentering parameters $-5\text{mm} \leq D_x \leq 5\text{mm}$ and $-5\text{mm} \leq D_y \leq 5\text{mm}$.

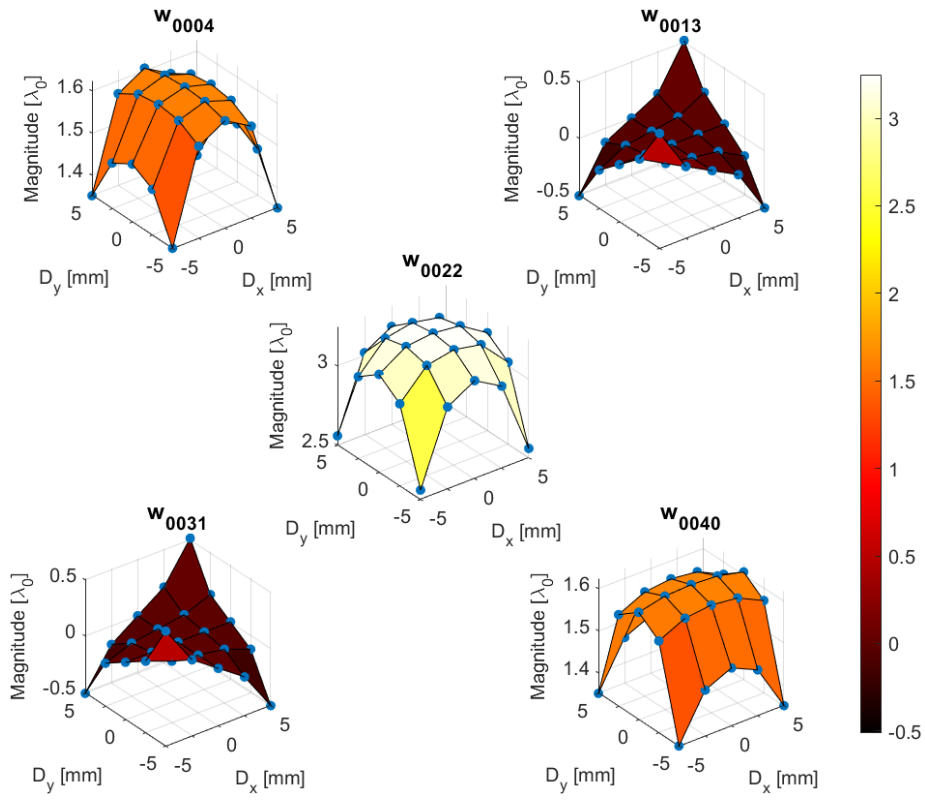


Figure 5.11: Wavefront aberration coefficients related to primary spherical aberration displayed as functions of the decentering parameters of the different channels D_x and D_y . Each blue dot is related to a specific channel.

The eight primary coma coefficients w_{0103} , w_{0112} , w_{0121} , w_{0130} , w_{1003} , w_{1012} , w_{1021} and w_{1030} are displayed in the following as a function of the parameters D_x and D_y .

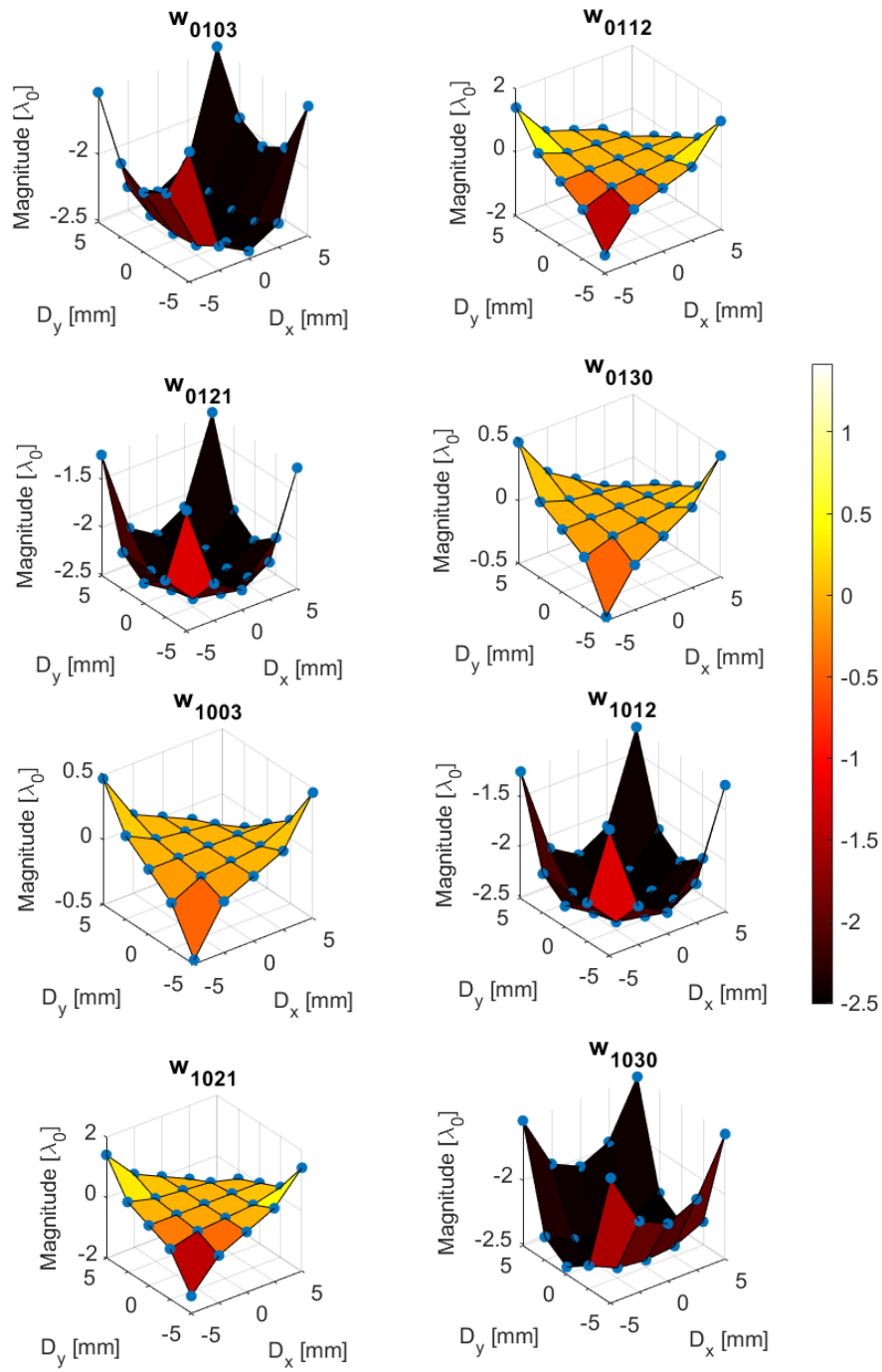


Figure 5.12: Wavefront aberration coefficients related to primary coma

Relatively to field curvature and primary astigmatism, there are in total nine coefficients whose dependence on the decentering parameters of the optical channels is displayed in the following Figure 5.13.

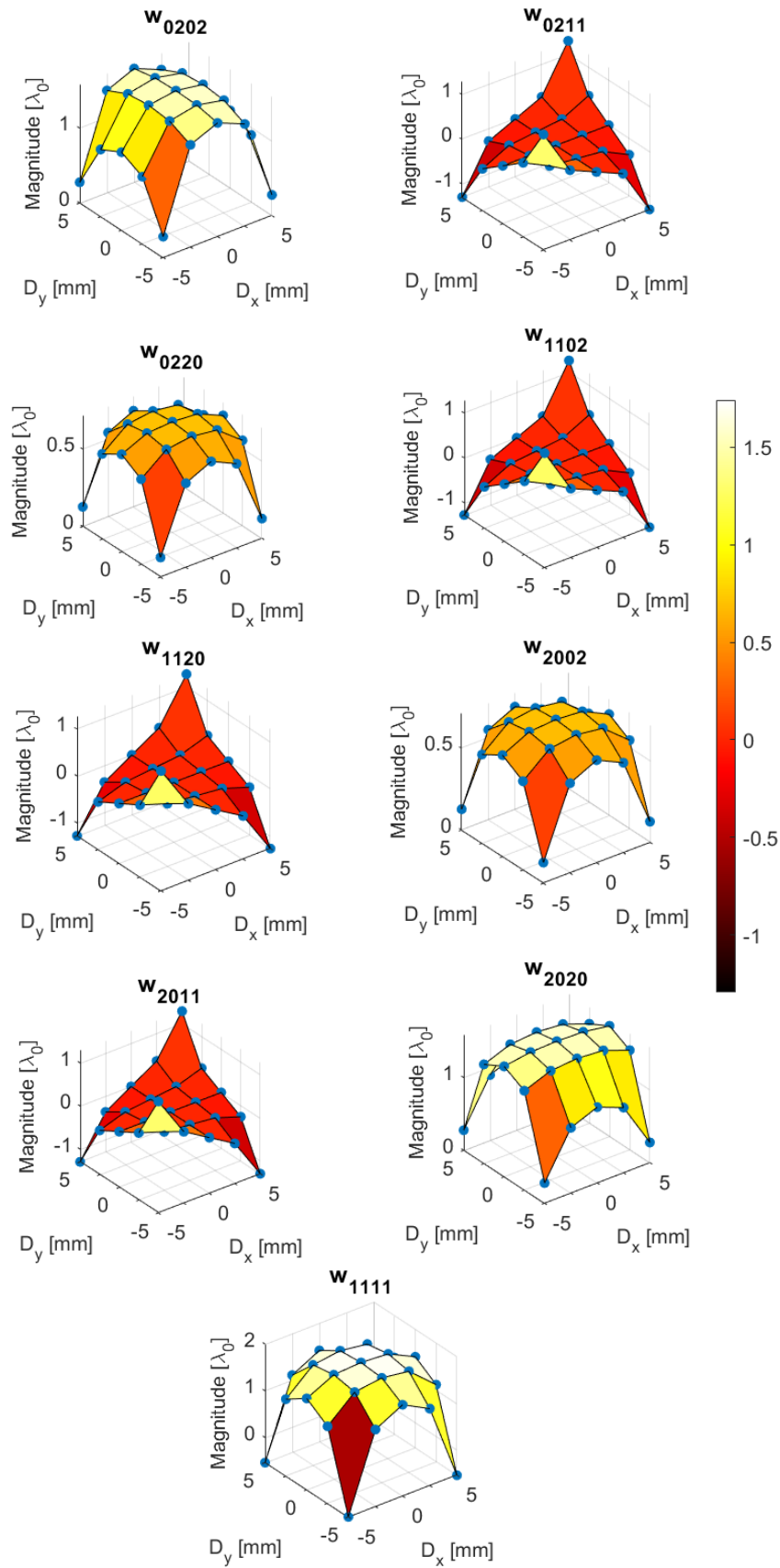


Figure 5.13: Wavefront aberration coefficients related to field curvature and primary astigmatism.

The remaining 4th order terms are represented by distortion terms (characterized by a linear dependence on the pupil coordinates ρ_x and ρ_y and by a cubic dependence on the field coordinates H_x and H_y).

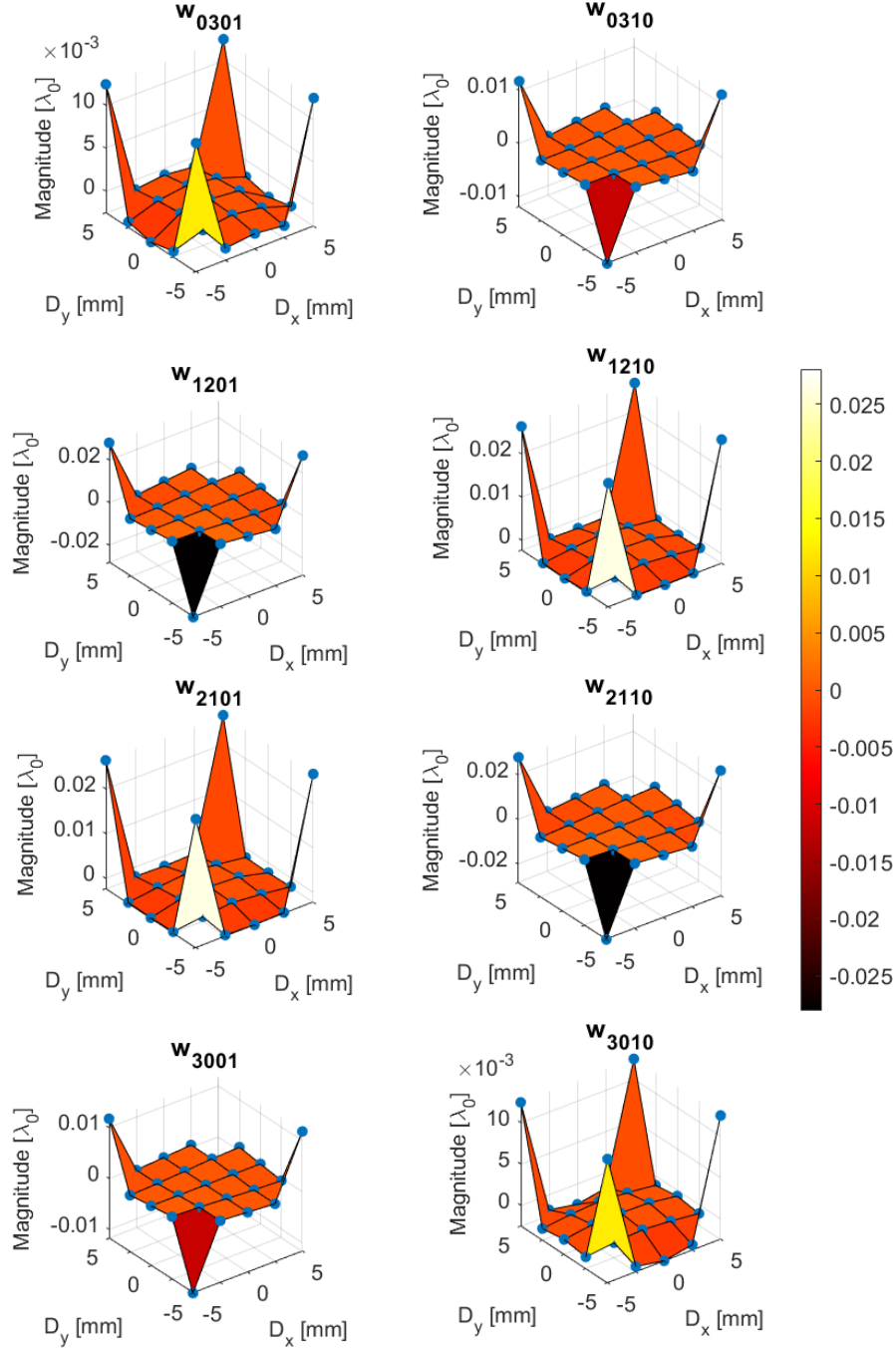


Figure 5.14: Wavefront aberration coefficients related to distortion.

The group of coefficients related to quartic piston terms (whose definition derives from the quartic dependence on the field coordinates) is omitted since they represent uniform phase terms over the pupil of an optical system.

The previous figures (from Figure 5.11 to Figure 5.14) are to be intended primarily as useful tools for the visualization of the aberrations behavior of different channels constituting an optical system with many apertures. As explained at the beginning of this chapter, the realization of a light field camera is not only based on the multiple aperture structure extensively discussed so far, but it also relies on post-processing algorithms for the elaboration of the captured raw images and extraction of useful information from them. In this sense, the visualization tools presented up to now can be of aid to design the necessary computational algorithms for the elaboration of the raw images providing, for example, a general understanding of the degradation of the image quality captured in correspondence of different partitions of the FOV.

5.5.3 Zernike polynomials expansion

In this last section, an alternative approach to represent the full field aberration behavior of a multi-aperture optical system is based on the formula of Equation 5.5 in section 5.4.1 where different aberration types are calculated properly combining the full field-dependent expressions of Zernike coefficients in the Fringe indexing scheme. The results are illustrated in the following series of figures. Of particular relevance is the behavior of primary astigmatism magnitude, shown in Figure 5.17-Figure 5.18. In this case, the optical branches with decentering parameters $D_x = \pm 2.5mm, \pm 5mm$ and $D_y = \pm 2.5mm, \pm 5mm$ exhibit an unconventional behavior for primary astigmatism magnitude.

In the following Figure 5.15, it is reported the full field behavior of spherical aberration (calculated with the first term of Equation 5.5) for the various channels of the light field camera under study. The domain of the subplots is represented by the size in millimeters of the images behind the plano-convex lenses. The colorbar is referred to all the subplots in Figure 5.15. It is worth observing that spherical aberration is actually characterized by a certain field dependence, evidenced in particular in the external channels. This situation is the result, in this example, of the increasing relevance of secondary aberrations with the channel displacement from to the mechanical axis of the system. In fact, as soon as the decentering coordinates (D_x, D_y) increase, the field dependence of spherical aberration is affected by a more relevant contribution of secondary aberrations. This situation has repercussions on the field dependence of the Zernike Fringe term $F_4^0(H_x, H_y)$ involved in the calculation of spherical aberration that is described, as shown in Chapter 2, by a power series involving terms of second order.

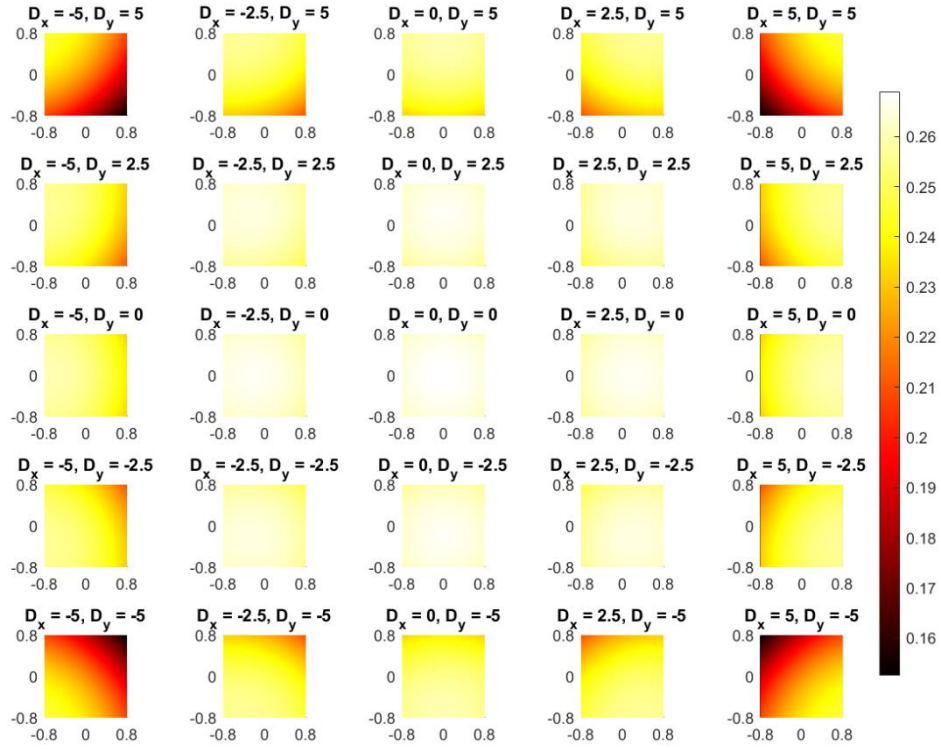


Figure 5.15: Field dependence of spherical aberration calculated from the coefficient of the Zernike term ($n = 4, m = 0$) in the Fringe indexing scheme (Equation 5.5). The subplots' titles indicate the decentering parameters (in mm) of the inherent branch of the camera under investigation. The subplots' domain corresponds to the FOV-partitions, therefore the values are in mm.

Regarding primary coma magnitude, for each optical branch, a single node in the FOV partition is visible and displaced with respect to the origin (except for the central channel whose node is at the center of the FOV due to its rotational symmetry). It is also evident that the magnitude of primary coma is properly described by a circular conic surface only for few central channels of the multiple aperture system, while in channels with larger decentering parameters, it is better described by elliptical conic surfaces whose apex is displaced over the sub-FOV. This situation is particularly accentuated in most external channels.

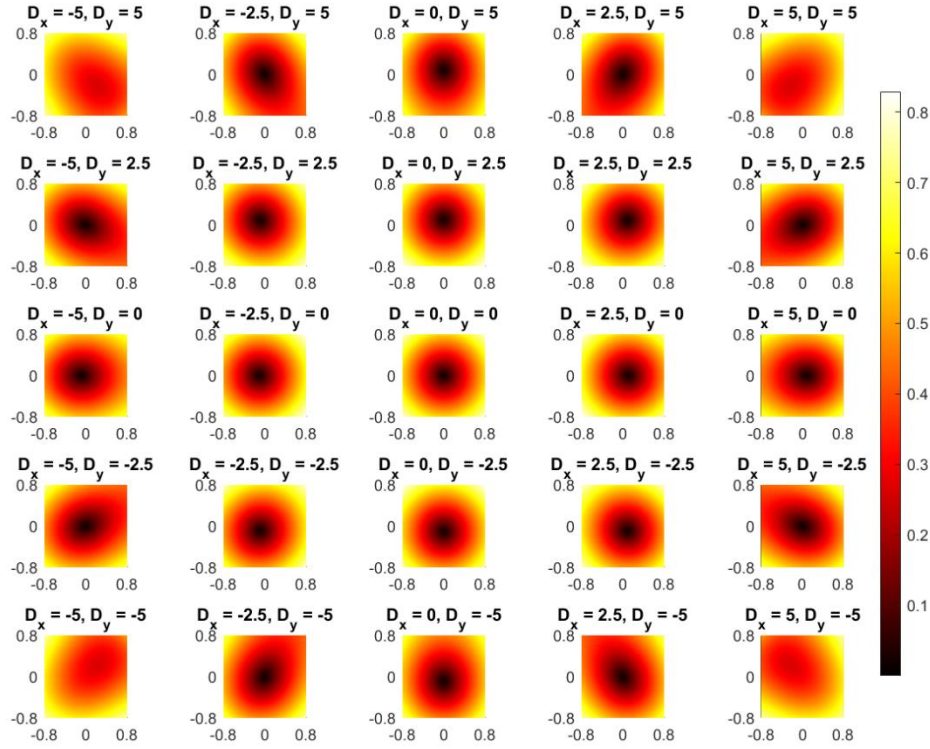


Figure 5.16: Field dependence of coma aberration calculated with the coefficients of the Zernike terms ($n = 3, m = 1$) and ($n = 3, m = -1$) in the Fringe indexing scheme according to the second term in Equation 5.5.

Such elliptical shape of the magnitude of the full field dependence of primary coma is not foreseen in the standard NAT [3] since this theory describes the aberration behavior in asymmetric optical systems (with decentered circular symmetric surfaces) characterized by only primary aberrations and by weak perturbations (weak decentering and tilt parameters of the surfaces constituting the optical system). The elliptical shape of the magnitude of the full field dependence of primary coma has been introduced in Chapter 3 relatively to a class of optical systems defined there as “perturbed” double-plane symmetric optical systems. In that context, the elliptical shape was due to the intrinsic double-plane symmetry of the investigated systems. In the present context, this behavior originates from the stronger perturbation induced by a more pronounced geometrical displacement of the external channels.

The field dependence of primary astigmatism for various branches of the multi-aperture system is particularly interesting because of the development of a well-known binodal behavior and a novel trinodal behavior over the FOV partitions of different channels. In particular, the optical channels characterized by decentering parameters $D_x = \pm 2.5\text{mm}, \pm 5\text{mm}$ and $D_y = \pm 2.5\text{mm}, \pm 5\text{mm}$ exhibit three nodes over the FOV, while the remaining channels exhibit two nodes, except the central channel whose only node is located at the center of its sub-FOV (due to its rotational symmetry). Due to the strong variation of the magnitude of primary astigmatism over the different channels, in the following Figure 5.17, the surface plots related to the fourth term of Equation 5.5 are represented with different colorbars. In particular, the four optical channels at the corners of the full-FOV with decentering parameters ($D_x = \pm 5\text{mm}, D_y = \pm 5\text{mm}$), exhibit the largest variation in the magnitude of the field dependence of primary astigmatism. For this reason, their respective colorbars are shown just nearby. On the other side, the magnitude of primary astigmatism for the remaining channels with decentering parameters ($D_x \neq \pm 5\text{mm}, D_y \neq \pm 5\text{mm}$), is referenced to the colorbar underneath Figure 5.17.

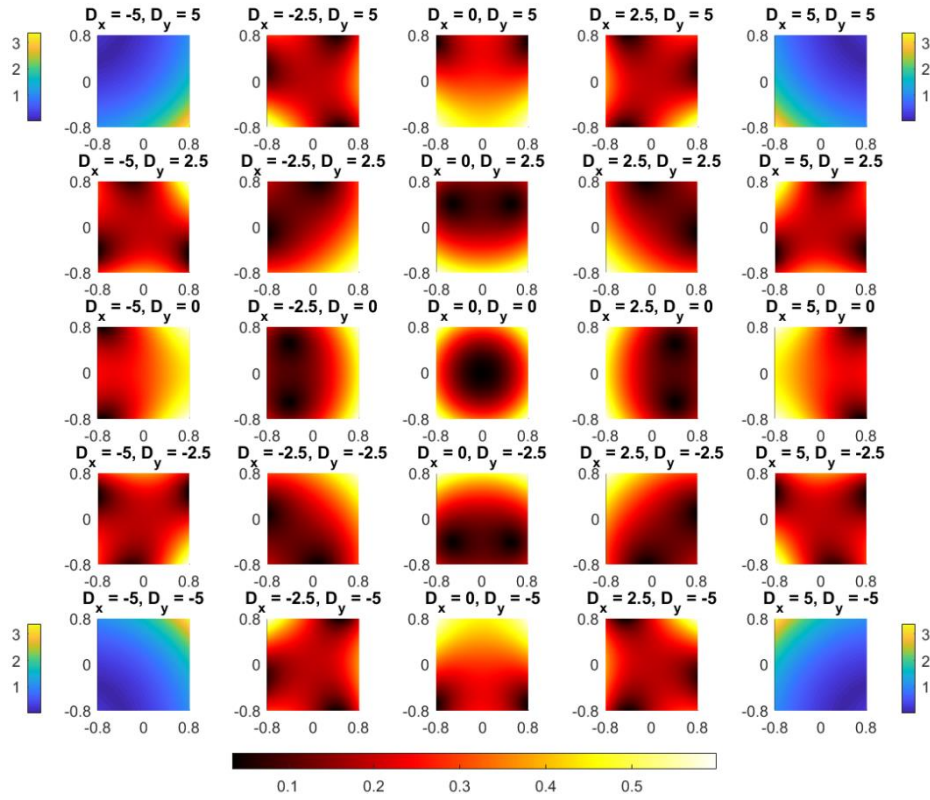


Figure 5.17: Field dependence of primary astigmatism calculated with the coefficients of the Zernike terms ($n = 2, m = 2$) and ($n = 2, m = -2$) in the Fringe indexing scheme according to the fourth term in Equation 5.5. Because of the large variation of astigmatism in the various branches of the plenoptic camera under study, two different colormaps and colorbars are used for the most external channels ($D_x = \pm 5\text{mm}, D_y = \pm 5\text{mm}$) and for the remaining channels ($D_x \neq \pm 5\text{mm}, D_y \neq \pm 5\text{mm}$) respectively.

In Figure 5.18 it is shown a contour plot of the data already shown in the previous Figure 5.17 in order to highlight more clearly the occurrence of nodes in the sub-FOV of the various channels.

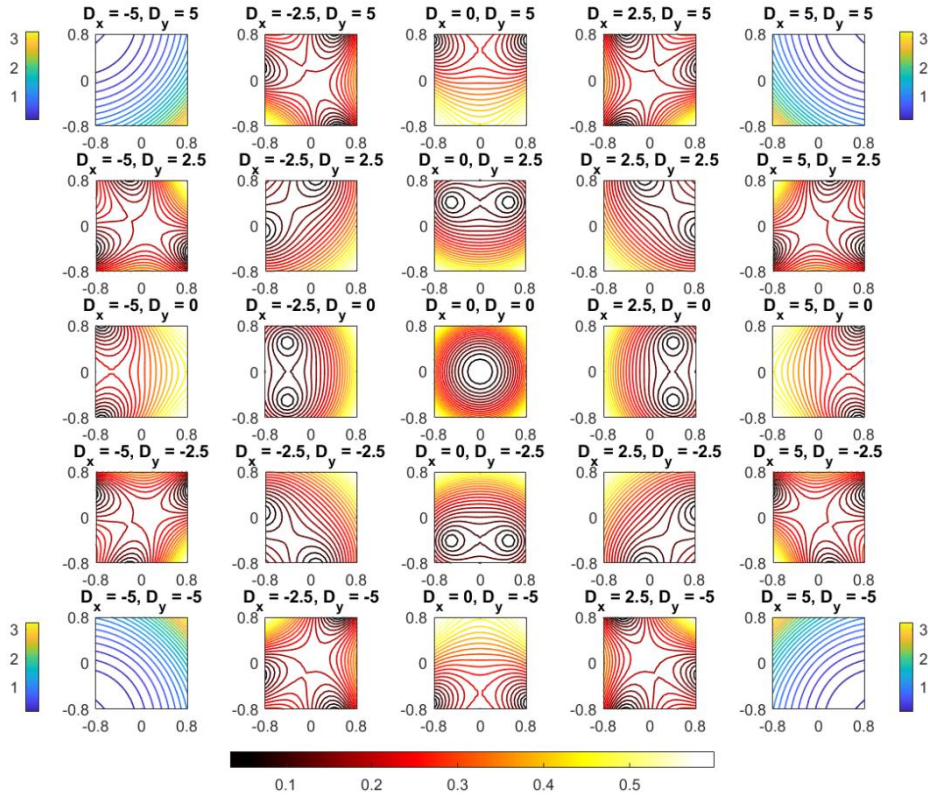


Figure 5.18: Contour plot of the full field-dependent magnitude of primary astigmatism for the 5x5 channels of the system under study.

The occurrence of three nodes in the magnitude of primary astigmatism is a consequence of the fact that the field dependence of the inherent Zernike coefficients $F_2^2(H_x, H_y)$ and $F_2^{-2}(H_x, H_y)$ is described by polynomials of order higher than 2. Therefore, the increasing impact of higher-order aberrations on the external channels manifests itself inducing a more complex variation of the aberrations over the FOV, like the trinodal behavior shown in Figure 5.17-Figure 5.18.

Regarding the full field dependence of field curvature, it is calculated with the third term of Equation 5.5 and it is shown in Figure 5.19. Due to the large variation of its magnitude, similar to the previous case, the channels at the corners of the full FOV have their colormaps (and colorbars), while the remaining channels are described with another colormap referenced to the colorbar underneath Figure 5.19.

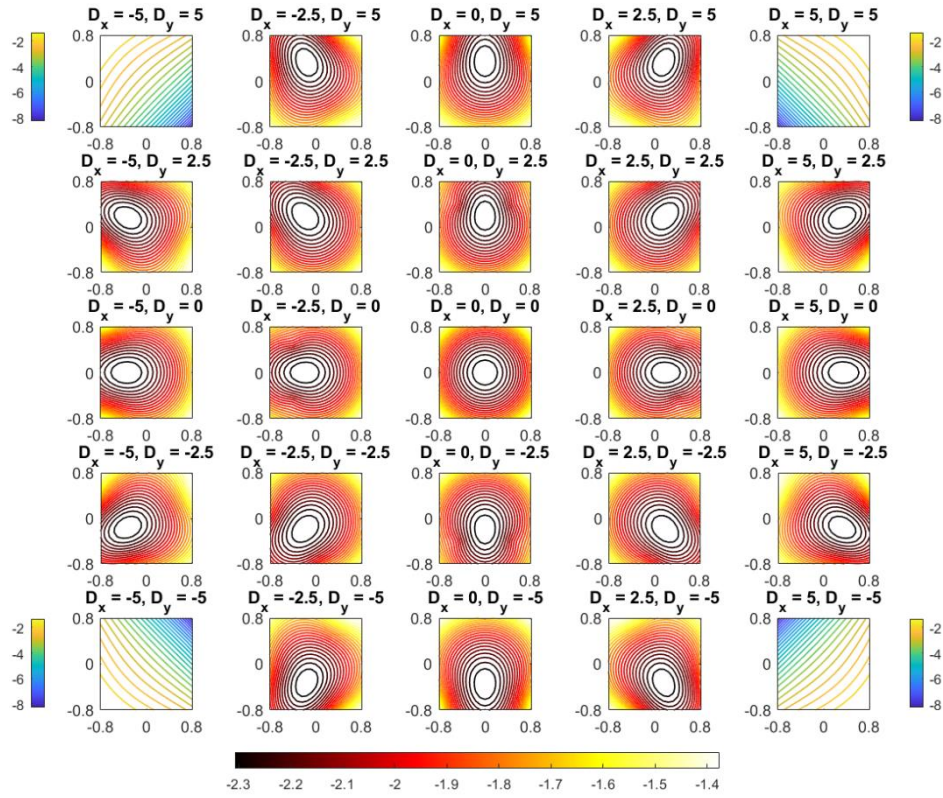


Figure 5.19: Field dependence of field curvature magnitude calculated with the coefficients of the Zernike polynomials reported in the third term of Equation 5.5.

5.6 Conclusions

In this last chapter of the thesis the wavefront aberration behavior of a class of optical systems with multiple apertures, represented by light field cameras, has been described using two different polynomial expansions. The first is a general power series expansion in the ray coordinates introduced in the framework of Hamiltonian Optics, the second is a full field Zernike polynomials-based expansion. Since this class of systems can be conceived as an ensemble of similar optical sub-systems differing from each other for the decentering parameters of the aperture stop surfaces, the concepts of paraxial optics are used at the beginning of the chapter to define an ensemble of base-rays, one for each channel. Such base-rays play the role of reference axis for the different branches of the systems under study, therefore they enable a proper choice of the ray variables on which the general power series expansion depends. Using this description, the wavefront aberration and the transverse ray aberration behaviors of a plenoptic camera are analyzed. It is discussed the possibility to restrict the study of these systems to the channels spanning the only first quadrant of the full FOV since, the underlying symmetry of the various channels, allows to easily retrieve the behavior of those in the remaining quadrants of the FOV performing reflection symmetry operations on the ray transverse aberration curves. The second description, based on proper combinations of Zernike polynomials in the Fringe indexing scheme, allows visualizing the nodes of different aberration types over the FOV partitions of the several channels. Additionally, this description of the wavefront aberrations allows us to visualize the variation of the magnitude of different aberration types due to the stronger perturbative effect of the geometrical displacement of the channels. The consequent enhancement of higher-order aberrations is visible in a more complex variation of the field dependence of Zernike polynomials that generates, for example, a particular trinodal behavior of primary astigmatism for the most decentered channels.

Chapter 6 Conclusions

This thesis deals with the study and representation of monochromatic wavefront aberrations in optical systems without symmetry and in multi-aperture optical systems. In the first chapter, the fundamental concepts of geometrical optics are introduced starting from Hamilton's characteristic function. In particular, the method of raytracing is described along with the concepts of wavefront aberration function and transverse ray aberration function. A brief demonstration of the summation theorem for primary aberrations is provided. The second chapter discusses monochromatic wavefront aberrations in optical systems without symmetry characterized by tilted and decentered spherical surfaces. In particular, it is shown that the presence of secondary aberrations has a perturbative effect on the location of the nodes of the field dependence of primary aberrations predicted by NAT. The third chapter discusses a class of systems defined as perturbed double-plane symmetric optical systems. Primary aberrations are analyzed combining full field-dependent Zernike Fringe polynomials. In particular, it is described a method to retrieve the field behavior of primary aberrations sampling the Zernike coefficients in correspondence with few field points of interest. The fourth chapter compares two different methods to represent the wavefront aberration function in asymmetric optical systems characterized by severe displacements of the constituting optical surfaces. The first method is based on the NAT wavefront aberration expansion. The second method uses a general power series expansion in the ray coordinates to describe the wavefront aberration function of this class of systems. The fourth chapter serves to introduce the topic of the fifth chapter that is the investigation of monochromatic wavefront aberrations in different channels of optical systems with many apertures. This investigation is carried out calculating numerically the coefficients of general power series in the ray coordinates describing the wavefront aberration functions of individual channels in a plenoptic camera. This approach is necessary because the channels of an optical system with multiple apertures like a plenoptic camera are, in general, asymmetric and are characterized by severe decentering parameters of the apertures of the different optical channels. It is shown a convenient method to visualize the evolution of the aberration coefficients for the various channels of a plenoptic camera. For completeness, it is also described a more common method of analyzing the aberration of asymmetric systems based on the combination of full field-dependent Zernike polynomials. This method emphasizes the presence of nodes in the FOV of the various channels of a light field camera. Of particular interest is the trinodal behavior exhibited by the field dependence of primary astigmatism magnitude in specific channels.

The numerical approach shown in this work to calculate the aberration coefficients of a general power series in the ray coordinates describing the wavefront aberration functions of the several branches of a plenoptic camera applies to the analysis of any system without symmetry. The only condition to be fulfilled is that the surfaces in the optical system must be developable in polynomial terms in the coordinates of a properly defined reference system. The main limitation of this method consists in the fact that the calculated coefficients describe the net aberrations of the whole system.

A further improvement of the presented method would be the calculation of the aberration coefficients contributed by the individual surfaces constituting the optical system

itself. In this case, this knowledge would be extremely useful to the optical designer to understand the individual contributions given by each surface to aberration terms of different order in the general power series development in the ray coordinates. Furthermore, this additional understanding would be helpful in the optimization process guiding the development of correct strategies to compensate for the aberrations.

Bibliography

- [1] W. T. Welford, *Aberrations of Optical Systems*, Adam Hilger Ltd, 1986.
- [2] A. Grosso and T. Scharf, “Scalar analytical expressions for the field dependence of Zernike polynomials in asymmetric optical systems with circular symmetric surfaces,” *OSA Continuum*, vol. 3, no. 10, pp. 2749-2765, 2020.
- [3] K. P. Thompson, “Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry,” *J. Opt. Soc. Am. A*, vol. 22, no. 7, p. 1389–1401, 2005.
- [4] K. P. Thompson, “Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry: errata,” *J. Opt. Soc. Am. A*, vol. 26, no. 3, pp. 699-699, 2009.
- [5] H. H. Hopkins, *The Wave Theory of Aberrations*, Oxford, UK: Oxford on Clarendon Press, 1950.
- [6] R. W. Gray and J. P. Rolland, “Wavefront aberration function in terms of R. V. Shack’s vector product and Zernike polynomial vectors,” *J. Opt. Soc. Am. A*, vol. 32, no. 10, pp. 1836-1847, 2015.
- [7] M. Rimmer, “Analysis of Perturbed Lens Systems,” *Appl. Opt.*, vol. 9, no. 3, pp. 533-537, 1970.
- [8] J. Sasián, “Theory of sixth-order wave aberrations,” *Appl. Opt.*, vol. 49, no. 16, p. D69–D95, 2010.
- [9] K. P. Thompson, “Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: spherical aberration,” *J. Opt. Soc. Am. A*, vol. 26, no. 5, p. 1090–1100, 2009.
- [10] K. P. Thompson, “Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: the comatic aberrations,” *J. Opt. Soc. Am. A*, vol. 27, no. 6, p. 1490–1504, 2010.
- [11] K. P. Thompson, “Multinodal fifth-order optical aberrations of optical systems without rotational symmetry: the astigmatic aberrations,” *J. Opt. Soc. Am. A*, vol. 28, no. 5, pp. 821-836, 2011.
- [12] R. W. Gray, C. Dunn, K. P. Thompson and J. P. Rolland, “An analytic expression for the field dependence of Zernike polynomials in rotationally

symmetric optical systems,” *Opt. Express*, vol. 20, no. 15, pp. 16436-16449, 2012.

- [13] ZEMAX user guide, July 2018.
- [14] K. P. Thompson, T. Schmid, O. Cakmakci and J. P. Rolland, “A real-ray based method for locating individual surface aberration field centers in imaging optical systems without rotational symmetry,” *J. Opt. Soc. Am. A*, vol. 26, no. 6, p. 1503–1517, 2009.
- [15] A. Grosso and T. Scharf, “Full field dependence of primary aberrations in perturbed double-plane symmetric systems with a circular pupil,” *J. Opt. Soc. Am. A*, vol. 37, no. 12, pp. 1999-2013, 2020.
- [16] H. A. Buchdahl, *An Introduction to Hamiltonian Optics*, Cambridge University, 1970.
- [17] J. C. Burfoot, “Third-order aberrations of doubly symmetric systems,” *Proc. Phys. Soc. B*, vol. 67, no. 7, p. 523–528, 1954.
- [18] C. G. Wynne, “The primary aberrations of anamorphic lens systems,” *Proc. Phys. Soc. B*, vol. 67, no. 7, p. 529–537, 1954.
- [19] P. J. Sands, “Aberration coefficients of double-plane-symmetric systems,” *J. Opt. Soc. Am.*, vol. 63, no. 4, pp. 425-430, 1973.
- [20] J. Sasiàn, “How to approach the design of a bilateral symmetric optical system,” *Opt. Eng.*, vol. 33, no. 6, pp. 2045-2061, 1994.
- [21] J. Sasiàn and S. Yuan, “Aberrations of anamorphic optical systems. I: the first-order foundation and method for deriving the anamorphic primary aberration coefficients,” *Appl. Opt.*, vol. 48, no. 13, pp. 2574-2584, 2009.
- [22] J. Sasiàn and S. Yuan, “Aberrations of anamorphic optical systems. II. Primary aberration theory for cylindrical anamorphic systems,” *Appl. Opt.*, vol. 48, no. 15, pp. 2836-2841, 2009.
- [23] J. Sasiàn and S. Yuan, “Aberrations of anamorphic optical systems III: the primary aberration theory for toroidal anamorphic systems,” *Appl. Opt.*, vol. 49, no. 35, pp. 6802-6807, 2010.
- [24] K. Fuerschbach, J. P. Rolland and K. P. Thompson, “A new family of optical systems employing ϕ -polynomial surfaces,” *Opt. Express*, vol. 19, no. 22, pp. 21919-21928, 2011.
- [25] K. Fuerschbach, J. P. Rolland and K. P. Thompson, “Theory of aberration fields for general optical systems with freeform surfaces,” *Opt. Express*, vol. 22, no. 22, p. 26585–26606, 2014.

- [26] H. Gross and Y. Zhong, "Vectorial aberrations of biconic surfaces," *J. Opt. Soc. Am. A*, vol. 35, no. 8, pp. 1385-1392, 2018.
- [27] I. Agurok, "Double expansion of wavefront deformation in Zernike polynomials over the pupil and the field of view of optical systems: lens design, testing, and alignment," in *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, San Diego, CA, United States, 1998.
- [28] J. C. Wyant and K. Creath, "Basic wavefront aberration theory for optical metrology," 1992, pp. 1-53.
- [29] D. Ochse, "Aberration fields of anamorphic systems," in *SPIE Optical Systems Design*, Frankfurt, Germany, 2018.
- [30] P. J. Sands, "Aberration Coefficients of Plane Symmetric Systems," *J. Opt. Soc. Am.*, vol. 62, no. 10, pp. 1211-1220, 1972.
- [31] A. Grosso, K. Vonmetz and T. Scharf, "Aberrations analysis of a focused plenoptic camera," in *SPIE Photonics Europe*, Strasbourg, France, 2018.
- [32] A. Grosso, "Depth of field extension through polarization aberrations," in *SPIE Optical Engineering + Applications*, San Diego, California, United States, 2019.
- [33] A. Grosso, "Depth of field extension through polarization aberrations," poster presentation at *SPIE Optical Engineering + Applications*, San Diego, California, USA, 2019.
- [34] A. Grosso and T. Scharf, "Plano-Concave mini-lens array for plenoptic imaging applications," poster presentation at *MOC2019, 24th MicroOptics Conference*, Toyama, Japan, 2019.
- [35] A. Grosso, K. Vonmetz and T. Scharf, "Design of fast hybrid microlens array," poster presentation at *ODF conference (Optics-Photonics Design & Fabrication)*, Hiroshima, Japan, 2018.
- [36] A. Grosso, K. Vonmetz and T. Scharf, "Wigner distribution representation of plenoptic imaging," poster presentation at *European Optical Society Biennial Meeting*, Delft, The Netherlands, 2018.
- [37] A. Grosso, K. Vonmetz and T. Scharf, "Aberrations analysis of a focused plenoptic camera," poster presentation at *SPIE Photonics Europe conference*, Strasbourg, France, 2018.

Curriculum Vitae

ALESSANDRO GROSSO, 18.08.1988, SIENA, ITALY

alessandro.grosso@alumni.epfl.ch



EDUCATION

- High School: Liceo Classico Ruggero Bonghi, Lucera, Italy
- University: Bachelor of Science and Master of Science in Engineering Physics at Politecnico di Milano, Milano, Italy

EMPLOYMENT HISTORY

- 04.03.2014 - 15.07.2016, R&D Engineer at Positech Srl, Milano, Italy
- 18.07.2016 - 31.12.2016, R&D Engineer at Fondazione CIFE (Centro Italiano di Fotonica per l'energia), Milano, Italy
- 15.01.2017 – 31.08.2017, Ph.D. Fellow at EPFL, Lausanne, Switzerland
- 01.09.2017 – 31.08.2019, Optical Engineer at Datalogic IP Tech Srl, Bologna, Italy
- 01.09.2019 – 31.12.2020, Ph.D. Fellow at EPFL, Lausanne, Switzerland

PUBLICATIONS

1. PEER-REVIEWED ARTICLES

- A. Grosso and T. Scharf, "Scalar analytical expressions for the field dependence of Zernike polynomials in asymmetric optical systems with circular symmetric surfaces," *OSA Continuum*, vol. 3, no. 10, pp. 2749-2765, 2020.
- A. Grosso and T. Scharf, "Full-field dependence of primary aberrations in perturbed double-plane symmetric systems with a circular pupil," *J. Opt. Soc. Am. A*, vol. 37, no. 12, pp. 1999-2013, 2020.

2. PROCEEDINGS

- A. Grosso, K. Vonmetz and T. Scharf, "Aberrations analysis of a focused plenoptic camera," in *SPIE Photonics Europe*, Strasbourg, France, 2018.
- A. Grosso, "Depth of field extension through polarization aberrations," in *SPIE Optical Engineering + Applications*, San Diego, California, United States, 2019.

3. CONFERENCE PRESENTATIONS

- A. Grosso, "Depth of field extension through polarization aberrations," poster presentation at *SPIE Optical Engineering + Applications*, San Diego, California, USA, 2019.
- A. Grosso and T. Scharf, "Plano-Concave mini-lens array for plenoptic imaging applications," poster presentation in *MOC2019, 24th MicroOptics Conference*, Toyama, Japan, 2019.
- A. Grosso, K. Vonmetz and T. Scharf, "Design of fast hybrid microlens array," poster presentation in *ODF conference (Optics-Photonics Design & Fabrication)*, Hiroshima, Japan, 2018.
- A. Grosso, K. Vonmetz and T. Scharf, "Wigner distribution representation of plenoptic imaging," poster presentation in *European Optical Society Biennial Meeting*, Delft, The Netherlands, 2018.
- A. Grosso, K. Vonmetz and T. Scharf, "Aberrations analysis of a focused plenoptic camera," poster presentation in *SPIE Photonics Europe conference*, Strasbourg, France, 2018.