

TOPICS IN MODERN LENS DESIGN

by

Dmitry Reshidko

Copyright © Dmitry Reshidko 2016

A Dissertation Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Dmitry Reshidko, titled Topics in Modern Lens Design and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Jose Sasian Date: (12/01/16)

Rongguang Liang Date: (12/01/16)

James Schwiegerling Date: (12/01/16)

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

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

Dissertation Director: Jose Sasian Date: (12/01/16)

STATEMENT BY AUTHOR

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

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Dmitry Reshidko

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Prof. Jose Sasian for the advice, support, patient guidance, and cheerful encouragement that allowed me to pursue research on topics for which I am truly passionate. I highly appreciate his time and efforts spent on me; he has been a wonderful mentor.

I would like to extend my appreciation to Prof. Roger Angel, with whom I've had the great privilege to collaborate. His drive and passion has been a great influence on me, has inspired me to work smarter and harder, and to constantly look for technological innovations. I would also like to thank Prof. Rongguang Liang and Prof. Jim Schwiegerling for serving on my defense committee.

It has been a great pleasure working with the faculty, staff, and students at the University of Arizona, particularly Eric Herman, Chia-Ling Li, Oscar Martinez, Yufeng Yan, Weibo Cheng, Shaun Pacheco, Brian Wheelwright, Jen Glass, and many others.

I gratefully acknowledge the Optical Design and Engineering Scholarship by SPIE and TRIF Imaging Fellowship Program, which have offered me funding to pursue new ideas. I also acknowledge the generous support of the College of Optical Sciences and my donors Romeo I. and Erlinda Mercado.

A special gratitude goes to my family whose love and unfailing support has enabled me to complete this work.

DEDICATION

This dissertation is dedicated to my father, Pavel Reshidko, who first taught me the value of education and critical thought. His enthusiasm, ambition, and encouragement have motivated me to constantly challenge myself and have inspired me to succeed.

TABLE OF CONTENTS

- LIST OF FIGURES 9
- LIST OF TABLES 16
- ABSTRACT 19
- CHAPTER I: INTRODUCTION..... 20
 - References..... 23
- CHAPTER II: ANALYSIS AND EVALUATION TOOLS IN MODERN LENS DESIGN 24
 - 2.1 Introduction 24
 - 2.2 Lens Evaluation From Power Distribution And Symmetry Principles 26
 - 2.3 Evaluation Of The Ray Invariant Product..... 30
 - 2.4 Evaluation Of Limiting Aberrations In A Lens 31
 - 2.5 Evaluation Of Higher Order Aberration Content..... 36
 - 2.6 Analysis Example: Mobile Phone Camera..... 42
 - 2.6.1 First-Order Properties 43
 - 2.6.2 Angles Of Refraction 45
 - 2.6.3 Spherical Aberration Correction Analysis 46
 - 2.6.4 Field Curvature Correction Analysis 49
 - 2.6.5 Image Quality..... 51
 - 2.6.6 Tolerance Analysis..... 52
 - 2.6.7 Total Length..... 53
 - 2.7 Conclusion..... 54
 - References..... 55
- CHAPTER III: GEOMETRICAL IRRADIANCE CHANGES IN A SYMMETRIC OPTICAL SYSTEM..... 57
 - 3.1 Introduction 57
 - 3.2 Radiative Transfer In An Optical System 61
 - 3.3 Irradiance Function Of A Pinhole Camera..... 63
 - 3.4 Irradiance On The Image Plane..... 64
 - 3.4 Irradiance On The Exit Pupil Plane..... 66
 - 3.5 Coefficient Relationship..... 68

3.6	Combination Of Irradiance Coefficients	69
3.7	Irradiance Coefficients And Choice Of Coordinates	70
3.8	Coefficients Verification	73
3.9	Relative Illumination.....	76
3.9.1	Historical Background	77
3.9.2	Relative Illumination Coefficient Summary.....	79
3.9.3	Example: Mobile Phone Camera Lens	82
3.9.4	Example: Wide-Angle Lens.....	85
3.9.5	Example: Lens Optimization For A Target Relative Illumination	88
3.9.6	Relative Illumination: Conclusion	90
3.9	Conclusion.....	91
	References.....	93
CHAPTER IV: RAY TRACING METHODS FOR CORRECTING CHROMATIC ABERRATIONS IN IMAGING SYSTEMS.....		95
4.1	Introduction	95
4.2	Forward And Reverse Ray Tracing.....	97
4.3	Aberration Subtraction Method	101
4.4	Guidelines Of Application Of The Methods	103
4.4	Example Of Optimization Using The Methods.....	104
4.4.1	Canon Telephoto Lens	104
4.4.2	Swir Zoom Lens.....	109
4.5	Aberration Subtraction Method For Athermalization Of Imaging Systems.....	114
4.5	Conclusion.....	119
	References.....	121
CHAPTER V: USEFUL ASPHERIC/FREEFORM SURFACE DESCRIPTIONS FOR IMAGING AND NON-IMAGING PROBLEMS		123
5.1	Introduction.....	123
5.2	Pedal Curve To The Ellipse	125
5.3	Off-Axis Conic Surface	131
5.4	Extended Cartesian Oval Surface	137
5.5	Conclusion	143
	References.....	144
CHAPTER VI: Conclusion.....		145

APPENDIX A: Lens Evaluation From Power Distribution And Symmetry Macro For Zemax	149
APPENDIX B: Evaluation Of The Ray Invariant Product Macro For Zemax.....	151
APPENDIX C: Evaluation Of Higher Order Aberration Content Macro For Zemax ...	153
APPENDIX D: Series Expansion Of The Cosine-To-The-Fourth-Power Of The Ray Angle.....	168
APPENDIX E: Derivation Of The Irradiance Function At The Image Plane Of An Optical System With The Aperture Vector At The Exit Pupil	170
APPENDIX F: Derivation Of The Irradiance Function At The Exit Pupil Of An Optical System With The Aperture Vector At The Exit Pupil	172
APPENDIX G: Derivation Of The Irradiance Function At The Image Plane Of An Optical System With The Aperture Vector At The Entrance Pupil.....	177
APPENDIX H: Derivation Of The Irradiance Function At The Exit Pupil Of An Optical System With The Aperture Vector At The Entrance Pupil.....	182
APPENDIX I: Derivation Of The Irradiance Function At The Image Plane Of An Optical System With The Aperture Vector At The Intermediate Position.....	183
APPENDIX J: Aberration Substraction Method For Chromatic Aberration Correction And Athermalization Of Imaging System Zemax Macro.....	186
APPENDIX K: Aspheric/Freeform Surface C Code For User Defined Surfaces In Zemax	190

LIST OF FIGURES

Fig. 2.1 Optical path difference (scale is 1 wave) and field curvature plots for a typical mobile camera lens. Excellent performance is achieved by utilizing highly aspherical surfaces. As a result, the plots show oscillatory behavior that makes the interpretation more difficult. 33

Fig. 2.2 The procedure to calculate intrinsic and extrinsic aberration components for the secondary mirror of a Cassegrain Telescope: (a) – the total aberration of the primary and secondary mirror is determined; (b) – the total aberration of the primary mirror is evaluated at the intermediate image plane; (c) – an optical system that consists of the secondary mirror alone is set up, and the intrinsic aberration component of the secondary mirror is calculated. To establish a set of common coordinates, rays are always aimed at the exit pupil. 41

Fig. 2.3 Optical layout of a mobile camera. (a) - Conventional flat field design based on existing patent (benchmark lens). (b) - Representative lens imaging on a curved sensor (evaluation lens).43

Fig. 2.4 Comparison of the first-order properties for a flat field design (benchmark lens) and a representative lens imaging on a curved sensor (evaluation lens). (a) – Optical power evaluation utilizing parameters w_j and W . (b) – Symmetry evaluation utilizing parameters s_j and S . A curved image surface allows a more symmetrical power distribution around the aperture stop. 45

Fig. 2.5 Ray refraction invariant. (a) - Conventional flat field design based on existing patent (benchmark lens). (b) - Representative lens imaging on a curved sensor (evaluation lens). The angles of refraction of the marginal ray are slightly higher for the faster design with a curved image surface.	46
Fig. 2.6 Surface contributions to the symmetry of spherical aberration and associated Pegel diagram for a flat field design (benchmark lens).	47
Fig. 2.7 Surface contributions to the symmetry of spherical aberration and associated Pegel diagram for a flat field design (benchmark lens) and a representative lens imaging on a curved sensor (evaluation lens).	48
Fig. 2.8 Residual spherical and oblique spherical aberration across the field of view of a representative lens imaging on a curved sensor (evaluation lens).	49
Fig. 2.9 Field curves. (a) - Conventional flat field design based on existing patent (benchmark lens). (b) - Representative lens imaging on a curved sensor (evaluation lens). A curve image surface compensates the field curvature aberration.	50
Fig. 2.10 Surface contributions to the symmetry of field curvature and astigmatism for a flat field design (benchmark lens).	51
Fig. 2.11 MTF a mobile camera: (a) – Conventional flat field design at $f/2.2$ (benchmark lens); (b) – Representative lens imaging on a curved sensor at $f/1.6$ (evaluation lens). The lens imaging on a curved sensor is one f-stop faster and shows more uniform performance over the field.	52

Fig. 2.12 Sensitivity to lens element decenter and tilt. Lens imaging on a curved sensor shows better as-build performance. The horizontal line indicates the nominal criterion value.....	53
Fig. 3.1 Second- and fourth-order apodization terms at the exit pupil given by the coefficients of the pupil irradiance function.	60
Fig. 3.2 Geometry defining the transfer of radiant energy from differential area dA in object space to the conjugate area dA' in image space. The radiant flux through all cross sections of the beam is the same.	61
Fig. 3.3 Geometry defining a pinhole camera. Images are formed as light passes through the aperture.....	63
Fig. 3.4 (a) – Layout of the Landscape lens used to compare irradiance coefficients computed analytically and numerically. (b) – the optical path difference plots for on-axis and full-field positions (plot scale is two waves). In order to minimize fitting errors, the FOV of the lens is limited to 30°	76
Fig. 3.5 Stadium at the University of Arizona simulated through an early wide-angle lens (US 2031792). Sharp decrease in illumination toward the corners of the image limits the field-of-view of the objective.....	78
Fig. 3.6 U.S. Patent 6441971: (a) Layout, (b) field curvature and distortion plot. Distortion correction requirements do not allow one to significantly reduce the chief ray incidence angle on the sensor (CRA). Exit pupil spherical aberration shifts the pupil radially and reduces the CRA. Higher order field curvature and astigmatism are balanced to compensate for the residual Petzval field curvature.	83

Fig. 3.7 Relative illumination plot for U.S. Patent 6441971. Analytical equation and real ray tracing calculation show excellent agreement over the entire FOV.	85
Fig. 3.8 U.S. 2516724: (a) Layout, (b) beam footprints at the exit pupil. The wide-angle objective consists of two nearly symmetrical halves. The form of the greatly curved exterior meniscus lens elements is of great importance. These meniscus lenses contribute substantial pupil coma, which increases the pupil area for oblique beams.	86
Fig. 3.9 Relative illumination plot for U.S. Patent 2516724. Analytical fourth-order approximation and real ray tracing calculation show excellent agreement up to about 20° half field angle.	88
Fig. 3.10 A Double Gauss lens was reoptimized while targeting both second- and fourth-order relative illumination coefficients to zero. The fourth-order approximation is sufficiently accurate for practical purposes and allows efficiently optimizing for a desired relative illumination during the lens design process.	90
Fig. 4.1 Forward and reverse ray tracing flow chart.....	98
Fig. 4.2 Edmund Optics f=20 mm Achromatic Doublet: (a) original lens and (b) on Axis OPD fan, (c) forward and reversed ray tracing model and (d) on Axis OPD fan. The OPD fan full scale is 0.5 waves.	100
Fig. 4.3 Aberration subtraction method flow chart.....	102
Fig. 4.4 Suggested algorithm for application of presented methods.	104
Fig. 4.5 Canon telephoto lens original patent data for infinite (a, c, e) and two meter (b, d, f) object distance configurations. The OPD fan scale is 2 waves.	105

Fig 4.6 OPD fans for infinite object distance at 436nm, 486nm, 588nm and 656nm: (a) - patent data, (b) – iteration 1, (c) – iteration 2, (d) – iteration 3. Plot scale is 2 waves. ..	106
Fig. 4.7 OPD fans for two meter object distance at 436nm, 486nm, 588nm and 656nm: (a) - patent data, (b) – iteration 1, (c) – iteration 2, (d) – iteration 3. Plot scale is 2 waves.	107
Fig. 4.8 Polychromatic MTF from 436nm to 656nm for infinite object distance: (a) - patent data, (b) - final design.	108
Fig. 4.9 Polychromatic MTF for 436nm to 656nm for 2 meters object distance: (a) - patent data, (b) - final design.	108
Fig. 4.10 Zoom lens layout: (a) - original design and (b) - reoptimized lens.....	110
Fig. 4.11 Zoom lens polychromatic MTF: f=24mm: original design (a) and reoptimized lens (b); f=60mm: original design (c) and reoptimized lens (d); f=100mm: original design (e) and reoptimized lens (f); f=140mm: original design (g) and reoptimized lens (h). ..	112
Fig. 4.12 Aberration subtraction method for athermalization flow chart.	115
Fig. 4.13 Optical layout of a single focus athermal lens.	116
Fig. 4.14 OPD fans: (a) – original design @ 20°C , (b) - final design@ 20°C , (c) – original design @ -20°C , (d) - final design @ -20°C , (e) – original design @ 60°C , (f) - final design @ 60°C . Plot scale is 5 waves.	117
Fig. 4.15 Polychromatic MTF for 420nm to 1700nm: (a) – original design @ 20°C , (b) - final design@ 20°C ,(c) – original design @ -20°C , (d) - final design @ -20°C , (e) – original design @ 60°C , (f) - final design @ 60°C	118

Fig. 5.1 (a) - Pedal curve to the ellipse. (b) - Coordinate axes on half the pedal curve. Note the change of curvature from the curve center to the edge.	126
Fig. 5.2 Aspheric lens with profile $z_1(r)$ and ray trace to the target surface.	128
Fig. 5.3 Freeform lens with surface profile $z_2(r)$ and ray trace to the target surface. ..	128
Fig. 5.4 Relative illumination (a) and transverse ray error in mm (b) of the aspheric lens at the target surface $z_2(r)$ for lens of Fig. 5.2. Transverse ray error plot scale is $\pm 3mm$	130
Fig. 5.5 Relative illumination (a) and transverse ray error in mm (b) of the freeform lens at the target surface for lens of Fig. 3. Transverse ray error plot scale is $\pm 0.3mm$	130
Fig. 5.6 The geometry defining global and local coordinates of the off-axis conic segment.	132
Fig. 5.7 Three-mirror system with confocal surfaces. (a) Layout (b) OPD plots over the entire FOV (the plot scale is 10 waves at $3\mu m$).....	135
Fig. 5.8 Layout of the three-mirror system. (a) The design with NURBS freeform surfaces reported by Chrisp; (b) The current design with $z_3(r)$ aspherical surfaces.	136
Fig. 5.9 Reoptimized three-mirror system. 8 th order plane symmetric polynomials are added to the mirror surfaces. (a) OPD plots (the plot scale is 1 wave at $3\mu m$); (b) Spot diagrams.....	136
Fig. 5.10 The RMS Spot Size of the three-mirror system. (a) The design with NURBS freeform surfaces reported by Chrisp; (b) The current design with $z_3(r)$ aspherical surfaces.	137

Fig. 5.11 Geometrical variables used to define the Cartesian Oval surface.	139
Fig. 5.12 Pedal surface profile used for the fit. (a) The major axis $a = 5mm$ and the minor axis $b = 15mm$; (b) The major axis $a = 5mm$ and the minor axis $b = 10mm$	140
Fig. 5.13 Layout of the 1:1 relay imaging lens designed with the extended Cartesian Oval surfaces.	142
Fig. 5.14 OPD plots of the reoptimized 1:1 relay lens (the plot scale is 5 waves at 588nm).	143

LIST OF TABLES

Table 2.1 Common values of W and S for widely-used lens design forms.....	30
Table 2.2 Sixth-order aberrations organized in groups according to aberration symmetry and formulas used to calculate four aberration symmetries of the total aberration.	39
Table 2.3 Design specifications of a typical mobile camera lens.....	44
Table 3.1a. Second-order irradiance coefficients $I_{k,l,m}^{(2)}$ at the image plane of an optical system.	65
Table 3.1b. Fourth-order irradiance coefficients $I_{k,l,m}^{(4)}$ at the image plane of an optical system.	65
Table 3.2a. Second-order irradiance coefficients $\bar{I}_{k,l,m}^{(2)}$ at the exit pupil plane of an optical system.	67
Table 3.2b. Fourth-order irradiance coefficients $\bar{I}_{k,l,m}^{(4)}$ at the exit pupil plane of an optical system.	68
Table 3.3. Relationships between second-order irradiance coefficients $I_{k,l,m}^{(2)}$ and $\bar{I}_{k,l,m}^{(2)}$..	69
Table 3.4a. Second-order irradiance coefficients $I_{k,l,m}^{(2)}$ at the image plane of an optical system with the field vector at the object plane and the aperture vector at the entrance pupil plane.....	71
Table 3.4b. Fourth-order irradiance coefficients $I_{k,l,m}^{(4)}$ at the image plane of an optical system with the field vector at the object plane and the aperture vector at the entrance pupil plane.....	71

Table 3.5a. Second-order irradiance coefficients $\bar{I}_{k,l,m}^{(2)}$ at the exit pupil plane of an optical system with the field vector at the object plane and the aperture vector at the entrance pupil plane.....	72
Table 3.5b. Fourth-order irradiance coefficients $\bar{I}_{k,l,m}^{(4)}$ at the exit pupil plane of an optical system with the field vector at the object plane and the aperture vector at the entrance pupil plane.....	72
Table 3.6 Iterative algorithm that was used to fit irradiance coefficients.	75
Table 3.7 Comparison of irradiance coefficients computed analytically and numerically. The agreement to eighth digits supports the correctness of the formulas.	76
Table 3.8. Second and fourth-order irradiance coefficients describing relative illumination.	80
Table 3.9 Contributions to the relative illumination for US. 6441971.	84
Table 4.1 Optical glasses used in designs on Fig. 4.6 and 4.7.	109
Table 4.2 Optical glasses used in designs on Fig. 4.10 and 4.11.	113
Table 4.3 Optical glasses used in designs in Fig. 4.14 and Fig.4.15.....	119
Table 5.1 Coefficients defining the surface profile $z_1(r)$	129
Table 5.2 Coefficients defining the surface profile $z_2(r)$	129
Table 5.3 Design requirements for the three-mirror telescope.....	134
Table 5.4 The RMS fit error between Pedal surfaces see in Fig. 5.12 and aspherical surface $z_4(r)$ as a function of the number of optimization variables used in the fit.....	141

Table 5.5 Coefficients defining the surface profile $z_4(r)$ used to fit the Pedal surface in Fig. 5.12a.	141
Table 5.6 Coefficients defining the surface profile $z_4(r)$ used to fit the Pedal surface in Fig. 5.12b.	141

ABSTRACT

Many advances have occurred in the field of optical design during the past decade. Some of the newer topics and concepts associated with the design and use of optical systems are complex and require comprehensive understanding of theory, expertise in state-of-the-art technology, and extensive computer simulations.

This dissertation focuses on development of practical methods and tools for successful lens design and evaluation of state-of-the-art imaging and illumination systems. The dissertation addresses several current topics in modern optical engineering and utilizes approaches to provide insights into the inner workings of optical systems. Examples of modern mobile camera lenses are provided to show how specific methods can help to better understand these lens designs and to expand the imaging capabilities of miniature camera systems.

Two simple but effective real ray tracing methods for correcting chromatic aberrations in imaging systems are described. The proposed methods separate monochromatic and chromatic aberration correction into two independent problems. This two-step approach provides effective alternatives in correcting chromatic aberrations.

A number of unique calculations have been performed and some novel and interesting theoretical results, including the fourth-order theory of irradiance changes in axially symmetric optical systems, are reported. The specific relationships between the irradiance distribution and wavefront aberration coefficients to fourth order are derived for the first time. The practical case of relative illumination at the image plane of an optical system is also discussed in some detail.

CHAPTER I: INTRODUCTION

This dissertation consists of a collection of topics that address current problems and applications of state-of-the-art imaging and illumination system lens design. Each topic is presented independently as a self-contained and self-explanatory chapter. Every chapter contains an introduction, references, detailed description of the proposed methods, several practical examples, and concluding remarks.

The goal of this dissertation is to supply useful tools and methods for successful lens design, to discuss the applicability of the proposed methods, and to explain their advantages and disadvantages in comparison to other methods used in optical design.

The chapter-wise organization of the dissertation is as follows.

In Chapter II, a set of novel tools and techniques for analysis and aberration correction of optical imaging systems is presented. These tools are developed and implemented as an aid to lens designers who are concerned about the detailed and comprehensive understanding of complex lenses and trade-offs involved in the design. Examples of state-of-the-art imaging systems are provided to show how these specific methods can help to better understand the design trade-offs and to expand the imaging capabilities of modern camera systems. The author contributions to the material described in different sections of this chapter varied. Some sections provide an overview of previously published methods and offer software implementation of these methods, other sections elaborate and extend ideas and results published elsewhere.

In Chapter III, the concept of the aberration function is extended to define two functions that describe the light irradiance distribution at the exit pupil plane and at the image plane of an axially symmetric optical system. Similarly to the wavefront aberration function, the irradiance function is expanded as a polynomial, where individual terms represent basic irradiance distribution patterns. The conservation of flux in optical imaging systems is used to derive the specific relation between the irradiance coefficients and wavefront aberration coefficients. It is shown that coefficients of the irradiance functions can be expressed in terms of wavefront aberration coefficients and first-order system parameters.

The practical case of relative illumination at the image plane of an optical system is also discussed in some detail. Several examples of lenses that instead of the \cos^4 -law show improved illumination at the focal plane are analyzed. The theory of irradiance changes in axially symmetric optical systems is used to provide insight into the role of individual aberration coefficients in the image illumination fall-off of these lenses.

Chapter IV describes two simple but effective real ray tracing methods for correcting chromatic aberrations in imaging systems. The correction of chromatic aberrations is typically performed using aberration formulas or by using real ray tracing. While the use of aberration formulas might be effective for some simple optical systems, it has limitations for complex and fast lenses. The chromatic aberration correction for such systems may be insufficient due to the presence of higher order aberrations that are not accounted for by aberration coefficients. For this reason, chromatic aberration correction is usually accomplished with real ray tracing, which allow for an accurate view of the actual image.

Real ray tracing optimization is performed by minimization of some error function. However, existing optimization tools in lens design software typically mix the correction of monochromatic and chromatic aberrations by construction of an error function that minimizes both aberrations at the same time. This mixing makes the correction of one aberration type dependent on the correction of the other aberration type.

The proposed real ray tracing methods separate monochromatic and chromatic aberration correction into two independent problems. This two-step approach can be used to correct both chromatic change of focus and chromatic change of magnification, as well as higher-order chromatic aberrations. The chromatic aberration correction is not limited to a specific bandwidth. Several design examples for the use of these methods are presented and discussed. It is shown that the proposed methods provide effective alternatives in correcting chromatic aberrations. Furthermore, a similar algorithm to that used for chromatic aberration correction is applied to athermalize an optical system.

Chapter V proposes several useful base optical surface descriptions in closed form that resemble the ideal profile that is required in a lens and provide effective solutions to imaging and non-imaging problems. Using the concept of base surface, an aspheric polynomial surface can be constructed by power expansion of the base term, and a freeform surface can also be constructed by superposition of several base surfaces having different parameter values. These aspherical/freeform surfaces substantially describe the desired surface for uniform illumination on a target plane or allow stigmatic imaging surface after surface along a particular ray. Examples of the performance of these surfaces that were

constructed using the concept of base surface are provided, and some of their properties are discussed.

The theoretical results and practical methods described in this dissertation serve to provide new insights into the optical design of state-of-the-art imaging and illumination systems, to advance the theory of aberrations, and to enhance the optics community's collective understanding of present topics in modern optical engineering.

Some of the material included in the chapters have been previously published as journal papers or conference proceedings papers [1-6]. The parts of this dissertation that use previously published material make clear references to the corresponding papers.

REFERENCES

1. D. Reshidko, J. Sasian and M. Nakanato, "Ray tracing methods for correcting chromatic aberrations in imaging systems," *Int. Journal of Optics*, Vol. **2014**, Article ID 351584 (2014)
2. D. Reshidko and J. Sasian, "Algorithms and examples for chromatic aberration correction and athermalization of complex imaging system," *SPIE Proceedings*, Vol. **9293** (2014)
3. D. Reshidko and J. Sasian, "Optical analysis of miniature lenses with curved imaging surfaces," *Applied Optics*, Vol. **54**, Issue 28, pp. E216-E223 (2015)
4. D. Reshidko and J. Sasian, "The role of aberrations in the relative illumination of a lens system," *Optical Engineering*, In Press (2016)
5. D. Reshidko and J. Sasian, "Geometrical Irradiance Changes in a Symmetric Optical System," *Optical Engineering*, In Review (2016)
6. J. Sasian, D. Reshidko and Chia-Ling Li, "Aspheric/freeform optical surface description for controlling illumination from point-like light sources," *Optical Engineering*, In Press (2016)

CHAPTER II: ANALYSIS AND EVALUATION TOOLS IN MODERN LENS DESIGN

2.1 INTRODUCTION

Lens design today is realized almost entirely on a personal computer. Lens design programs have evolved over the past decade to meet the increasing performance demands of state-of-the-art optical systems, which are often difficult to design, fabricate, and align. These programs can handle various optical design tasks and configurations including but not limited to rotationally symmetric systems, free-space systems, non-sequential and sequential systems, systems that include highly aspherized, diffractive or other free-form surfaces, and so on. Some programs include additional simulation tools necessary to manufacture a practical optical system, such as consideration of manufacturing methods to be used, mechanical and thermal considerations, manufacturing tolerancing, antireflective coatings, and other complex factors to be included in the total design of an optical system [1].

The evaluation of an optical system in lens design software is performed by tracing real rays. Real ray tracing is a powerful tool that provides strong analysis capabilities and gives the designer the information necessary to understand and explain the behavior of a lens and how it may perform in an actual system. Moreover, different mathematical methods for lens design optimization utilize ray tracing to make changes in the system to improve the image quality.

Although only a very minimal amount of manual calculation is usually carried out today during the lens design process, blind use of a lens design program almost certainly fails to find and deliver an efficient and practical solution to any lens problem. Thus, the design produced by applying a blind process may be over constrained, may be difficult to manufacture or align, or it may have marginal performance.

Understanding of the classic principles and lens design techniques and application of these lens design fundamentals is essential to find a preferable solution and also to guide and select alternative optimization paths for the program to follow. Once the acceptable configuration for the lens design efforts is intelligently determined, the engineer may benefit from the power of the computer programs to improve and polish the system [2].

As optical technology evolves, there is a need to enhance software capabilities to model technological innovations. This chapter contains a discussion of the principles and techniques that are appropriate to design work with a modern lens design program and are helpful in analyzing complex optical systems as well as more basic lens types. Furthermore, this chapter presents a set of novel tools. These tools are developed and implemented as an aid to lens designers who are concerned about the detailed and comprehensive understanding of complex lenses and trade-offs involved in the design. In addition, the proposed tools help to identify the main driver of lens performance, to predict sensitivity to manufacturing tolerances, and to determine redundant degrees of freedom.

Examples of state-of-the-art imaging systems are provided to show how specific methods can help to better understand these lens designs and to expand the imaging capabilities of modern camera systems.

2.2 LENS EVALUATION FROM POWER DISTRIBUTION AND SYMMETRY PRINCIPLES

The way optical elements are combined determines the basic properties of a lens system and, as a consequence, selecting a proper configuration is a key for a successful design effort. Two important types of design schemes are distinguished that allow achieving the desired optical performance.

The first approach is the concept of lens relaxation, which was reviewed by Glatzel and Shafer [3, 4]. In a relaxed lens, the surface Seidel aberrations contributions must each be kept very small to prevent higher-order aberrations from arising. To achieve this, a fairly large number of low-power elements are needed, and the lens inevitably becomes longer. When used successfully, the relaxation design technique results in very broad merit function minimum, smooth wavefronts, loose manufacturing tolerances, and small chromatic variation in aberrations.

A different approach suggests compensating for aberrations by introducing large amounts of opposing aberrations within the design. It should be taken into account that balancing large amounts of aberration brings with it an additional stress which in most cases produces aberrations of a higher order. By contrast to relaxed lenses, in lenses with a lot of stress built into them, even a small parameter change gives a substantial change in the system performance. The advantage, however, is that it allows for a shorter optical system.

Both relaxed and stressed optical configurations have advantages and disadvantages. It is desirable to incorporate relaxation into a design. However, it is very difficult to apply

this approach to any simple lens since more variables are required to successfully utilize the relaxation principle. Some stress is often introduced to make the system less complicated and to meet other design requirements.

There is a need for lens design tools that allow one to quantify relevant lens attributes and provide adequate comparison between different optical configurations. Two applicable parameters for quantifying optical power distribution and lens symmetry are proposed by Sasian and reviewed in this section [5]. These parameters are independent of lens scaling, conjugate, aperture size and field angle, and thus provide consistent comparison between different designs and fair evaluation of the imaging potential of a lens.

The paraxial refraction equation is considered to define the first parameter W ,

$$n'u' = nu - y\Phi. \quad (2.1)$$

In this equation $n'u'$ and nu are products of the index of refraction and first-order marginal ray slope before and after refraction at the surface; y is the marginal ray height at the surface; and Φ is the surface optical power. Therefore, the weighted surface power is given by the difference between the marginal first-order ray slope before and after refracting at the surface. To obtain the normalization factor, the sum of the terms in Eq. 2.1 is taken over all the system surfaces. It follows that

$$w_j = -\frac{1}{1-m} \frac{y_j}{n'u'_N} \Phi \quad (2.2)$$

and

$$\sum_{j=1}^N w_j = 1, \quad (2.3)$$

where w_j is the weighted optical power of surface j . In Eq. 2.3, the surfaces with large w_j contribute more to provide the overall optical power, and therefore, w_j indicates where the optical power originates within the system. Moreover, w_j is independent of scaling, aperture and conjugate since all parameters in Eq. 2.2 are invariant under these transformations and also independent of the field angle by construction. The parameter W is established as the square root of the averaged and squared weighted powers w_j of the lens surfaces as

$$W = \sqrt{\frac{1}{N} \sum_{j=1}^N w_j^2}. \quad (2.4)$$

In a relaxed optical system, all surfaces contribute equally to provide the complete optical power of the system, and W is minimized. Thus, small values of W indicate efficient use of the optical power and relaxed design, while large values of W suggest a stressed optical system.

Symmetry of the lens is another important attribute that has a substantial impact on defining the form of the lens. In a symmetrical optical system, some aberrations tend to cancel out or may be avoided; these outcomes permit a higher level of aberration correction. Thus, symmetrical or nearly symmetrical lens design forms allow for a more balanced and better corrected solution.

The second proposed parameter measures the degree of symmetry to which the surfaces of an optical system are used. Optical surfaces that satisfy aplanatic condition $\Delta(u/n) = u'/n' - u/n = 0$ or are concentric to the stop or to its images $\bar{A} = ni = 0$ refract

the light in a symmetrical way. Here \bar{A} is the product of the index of refraction and the chief ray incidence angle at the surface j . It follows that s_j , based on the aplanatic and concentric conditions,

$$s_j = \frac{1}{1-m} \frac{1}{ni_{stop} (n'u'_N)} \bar{A}_j \Delta(u/n)_j, \quad (2.5)$$

indicates the degree of symmetry of an optical surface. The normalization factor is chosen to make s_j independent of the lens field of view and conjugate. The parameter S is established as the square root of the averaged and squared s_j of the lens surfaces as

$$S = \sqrt{\frac{1}{N} \sum_{j=1}^N s_j^2}. \quad (2.6)$$

In optical systems with a high degree of symmetry, all surfaces satisfy aplanatic or concentric conditions, and S is minimized. Small values of S provide conditions for aberration cancelation and superior image quality over large fields.

Symmetry and power distribution are recognized to have an important impact on the imaging performance of lenses, and parameters W and S help to identify optimum design forms. Lenses with low values of W and S are expected to perform better than lenses with high values. Common values of W and S for several well-known design configurations are presented in Table 2.1. The parameters are calculated in Zemax OpticStudio by evaluating Eq. 2.4 and Eq. 2.6 in a macro program [6]. The macro is given in Appendix A.

Petzval optical configurations follow the relaxation principle: low-power optical elements are located near the aperture where the marginal ray height is large, while stronger optical elements are located next to the image plane where the marginal ray height is small.

In a typical Petzval lens, a small amount of aberration is generated over a small field of view, and a value for W is smaller than S .

On the other hand, Double-Gauss lenses exhibit values of W that are larger than S . Typical Double-Gauss design takes advantage of symmetry properties of surfaces and is distinguished by a small amount of field-dependent aberrations across a large field of view.

A Cook Triplet lens compromise between the attributes of power distribution and symmetry. These lenses typically exhibit almost equal values of W and S . Relatively large values of W and S for a Cook Triplet lens indicate stress inherent in this design form.

Another major design type that can cover a respectable field of view at a relative high speed is a Reversed Telephoto lens. When reversed telephoto form is pushed to extreme, the total field of 180° or more can be covered. To support a large field of view, at least a roughly symmetrical construction is required. A well-corrected Reversed Telephoto Wide-Angle lens demonstrates almost equal and relatively low values of W and S .

Table 2.1 Common values of W and S for widely-used lens design forms.

Lens Form	W	S
Petzval Lens	0.4	0.7
Double-Gauss	1	0.3
Cook Triplet	1	0.9
Reversed Telephoto Wide-Angle	0.4	0.3

2.3 EVALUATION OF THE RAY INVARIANT PRODUCT

It is well known in lens design that minimizing the angles of incidence and refraction of rays at element surfaces is important to reduce aberrations: large or small ray angles are

indicative of stress or relaxation in a lens system. The evaluation of the incidence angles can be used to detect surfaces that contribute to the particular aberration or generate significant high-order aberrations. Therefore, reducing the angles of incidence is usually a change that causes the aberrations to be reduced. Moreover, for high-incidence angles, Snell's Law of refraction becomes very non-linear, producing a highly unstable situation, which often explodes as the lens construction parameters are incremented during the design and optimization or depart slightly from the nominal values due to manufacturing tolerances.

However, the angle of incidence and angle of refraction are different in air versus in the material, and this can lead to confusion in assessing a given lens. As an insightful tool, the ray invariant product $n \cdot \sin(i)$ is instead evaluated surface by surface where i is the ray incidence angle and n is the index of refraction of the media. The rays used are the real marginal and chief rays. A macro program that calculates and plots the ray invariant product for each surface of a lens is given in Appendix B.

The examination of the angles of incidence for a critical ray path is a widely applicable technique that will often indicate the source of a design problem.

2.4 EVALUATION OF LIMITING ABERRATIONS IN A LENS

During the lens design and evaluation, it is important to identify the main driver of lens performance and to determine specific limiting aberrations in the system. Once the designer understands all limiting factors, the lens may be further improved by applying an

appropriate technique to correct these specific limiting aberrations or choosing a different lens configuration to keep them from arising.

The elements comprising the total aberration of an optical system can be computed directly from the specific ray trace data. For example, the slopes of the ray fans at the origins indicate the presence of field curvature, astigmatism or defocus aberrations. The complete aberration content can be further estimated by decomposing the ray or wave aberration curves into individual components.

However, real optical systems are degraded by multiple aberrations and interpretation of such plots may not be quite obvious. Different orders of aberrations are balanced over the aperture and over the field to achieve the best image quality, and limiting aberrations suffered by a particular lens design are not readily evident. Moreover, if higher-order aspheres are used in the design, the transverse aberration plots or field curvature plots often show oscillatory behavior, as in Fig. 2.1. Introducing more orders of aspherical expansion may improve the lens performance, but results in the higher spatial frequency of the errors and makes the interpretation of the plots even more difficult.

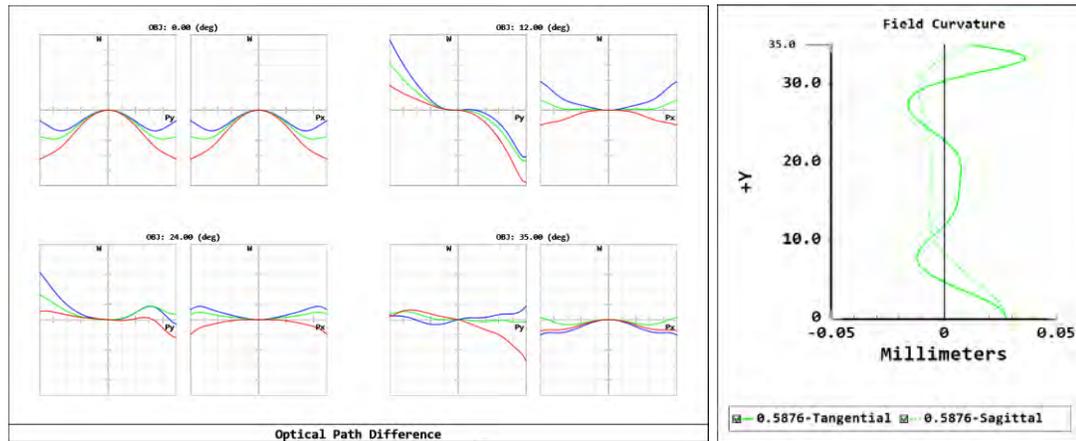


Fig. 2.1 Optical path difference (scale is 1 wave) and field curvature plots for a typical mobile camera lens. Excellent performance is achieved by utilizing highly aspherical surfaces. As a result, the plots show oscillatory behavior that makes the interpretation more difficult.

An alternative approach to determine the limiting aberration in a lens is introducing fictitious degrees of freedom and using optimization routines in the lens design software to effectively target and remove this particular aberration. The magnitude and order of the aberration that is analyzed can be estimated from the coefficients of the optimization variables. The contribution of the aberration to the overall lens performance is derived by analyzing the emerging wave or ray aberration plots of the reoptimized lens.

The particularly onerous aberrations to consider are spherical, oblique spherical and field curvature aberrations. Three techniques to introduce fictitious degrees of freedom to target different orders of the above aberrations are discussed in detail in this section.

The field-independent aberrations comprise defocus and spherical aberration; both aberrations are constant over the entire field of view of the lens and are functions only of the pupil coordinates. The direct evaluation of spherical aberration is a simple matter. First, all active variables in the design are removed and a marginal ray height solve for the surface

thickness that precedes the image plane is assigned. The solve ensures that defocus aberration is completely excluded from the analysis. Next, a zero-thickness aspherical plate is introduced at the stop aperture of the lens. The lens is then reoptimized while targeting an RMS Wavefront Error for the on-axis field point and using aspheric coefficients of the plate as optimization variables. In the reoptimized lens, the aspheric coefficients of the plate are proportional to different orders of spherical aberration. For example, if fourth-order asphere was used, the fourth-order spherical aberration coefficient W_{040} is given by

$$W_{040} = A_4 \cdot (n - 1) \cdot r_{apt}^4, \quad (2.7)$$

where A_4 is the aspheric plate coefficient, n is the index of refraction of the plate, and r_{apt} is the semi-diameter of the stop aperture. Moreover, an overall contribution of spherical aberration to the total performance of the lens can be assessed from the relative change in the merit function value before and after optimization. A significant change in the value of the merit function suggests that spherical aberration is indeed limiting the performance of the lens.

Oblique spherical aberration is a type of high order aberration that, similarly to primary spherical aberration, has at least fourth power dependence in aperture. However, oblique spherical aberration also has at least second power field dependence. Oblique spherical aberration is found by constructing multiple configurations in lens design software. Each configuration is similar to the arrangement constructed for evaluation of primary spherical aberration and represents a single field of the lens. Field curvature aberration is removed by adjusting the focal plane of each configuration to the sagittal focus. The RMS Wavefront

Error in the sagittal direction is minimized to separate oblique spherical aberration from astigmatism. This model does not allow one to calculate the magnitude of the specific order of oblique spherical aberration, nevertheless it is useful in analyzing the field behavior of a lens.

Field curvature is a field-dependent displacement of the image formed by a lens system from the Gaussian image plane. In many modern lens designs, the residual Petzval field curvature is corrected by balancing higher orders of field curvature and astigmatism. The field curves in these designs are wavy with multiple crossing across the field of view, which makes the aberration analysis difficult.

The field curvature contribution to the total aberration of a lens can be estimated to any order by constructing a simple model. First, all active variables in the design are removed and a marginal ray height solve for the surface thickness that precedes the image plane is assigned. The solve ensures that defocus aberration is completely excluded from the analysis. Next, the aperture-dependent components such as spherical and oblique spherical aberrations are removed by stopping down the lens to $f/100$ or so. Finally, the image surface is turned into an aspherical surface, and the lens is then reoptimized while targeting the RMS Wavefront Error in the sagittal direction and using aspheric coefficients of the image surface as optimization variables. In the reoptimized lens, the aspheric coefficients of the image surface are proportional to different orders of field curvature aberration. For example, if a second-order asphere was used, the fourth-order field curvature coefficient W_{220} is given by

$$W_{220} = -\frac{A_2 \cdot r_i^2}{8 \cdot (f / \#)^2}, \quad (2.8)$$

where A_2 is the aspheric coefficient, $f / \#$ is the focal ratio of the lens, and r_i is the maximum image height. Moreover, if the aperture of the lens is restored to the original value, field curvature aberration is still compensated by the aspherical image surface and is removed from the total aberration of a lens.

The methods discussed in this section decouple spherical, oblique spherical and field curvature aberrations from the total aberration that degrades image quality of a lens system, thereby providing insight into the specific role of these aberrations in the optical design process. Examples of analysis utilizing these techniques are provided in subsequent sections.

2.5 EVALUATION OF HIGHER ORDER ABERRATION CONTENT

Aberration theory has proven being useful in problem solving and guiding an intelligent optical design process. It allows calculating the magnitude of image quality degradation, provides insight on the source of aberrations, and explains the nature of the optical image formation.

Current developments in theory and application of aberrations have their advantages and disadvantages. Seidel theory is well understood and provides a fair approximation for some simple lenses. Higher-order aberrations in such lenses are small and the fourth-order theory is sufficient for predicting performance. The sixth-order aberration theory grants new insights, but is too elaborated and complex to use. In addition to five primary

aberrations, the designer needs to consider nine more aberration terms. Moreover, the sixth-order and higher-order aberration coefficients can be further differentiated into intrinsic and extrinsic (or induced) components. When the incoming beam has no aberrations in it, a refracting surface contributes only intrinsic aberrations. The extrinsic aberrations arise when aberrations due to previous surfaces are present in the incoming beam. As the result, it is practically impossible to keep track of all twenty-three aberration terms to six order.

Approximations involving the aberration theory is the reason that aberration correction in complex state-of-the-art lenses is usually accomplished with real ray tracing which enables an accurate view of the actual image. On the other hand, real ray tracing optimization lacks insight into how aberrations are corrected and balanced within the lens.

There is a need for simple and efficient tools for aberration analysis of optical imaging systems that enhance the detailed and comprehensive understanding of complex lenses and explain trade-offs involved in the design. In addition, evaluation of the higher-order aberration content helps to identify the main driver of lens performance, to predict sensitivity to manufacturing tolerances, and to determine the redundant degrees of freedom. This knowledge is essential for designing efficient imaging lenses and for correctly interpreting the results of the real ray tracing optimization.

In an axially symmetric optical system, only certain wavefront deformations are possible. Instead of keeping track of all aberration terms, these terms are combined according to four aberration symmetries: spherical aberration, coma, astigmatism, and field curvature. The magnitude of these four aberration symmetries is calculated by using real ray tracing and evaluating the optical path difference for a given ray. This calculation,

unlike the classical methods of aberration analysis, is made without decomposing the aberration into specific orders. The computation of aberrations using real ray data is not an approximation of the aberrations, but is accurate. As a consequence, the method permits analyzing stressed lens design solutions where fourth-order aberrations are balanced with higher-order aberrations.

Table 2.2 provides the sixth-order aberrations of an axially symmetric system organized in groups according to aberration symmetry [7]. For example, the symmetry of spherical aberration includes all field-independent aberrations of at least the sixth order, while the symmetry of coma consists of all sixth and higher-order aberration terms that have an odd power dependence in aperture. To determine the magnitude of aberration symmetries, real rays are traced through the system. The wave aberration with respect to an ideal wavefront is described by the optical path difference $OPD(H_x, H_y, \rho_x, \rho_y)$ for rays specified by the normalized field (H_x, H_y) and normalized pupil (ρ_x, ρ_y) coordinates referenced initially to the optical path of the chief ray. The formulas used to calculate the aberration symmetries are given in Table 2.2. These definitions assume that there is no focus error. Vignetting also has not been considered for this investigation. The calculation is limited to rotationally symmetric optical systems with spherical, aspheric or any axially symmetric free-form surfaces.

A macro program was written that implements equations in Table 2.2 to calculate four aberration symmetries and to divide the surface contributions to the total aberration of an optical system into intrinsic and extrinsic components (see Appendix C).

Table 2.2 Sixth-order aberrations organized in groups according to aberration symmetry and formulas used to calculate four aberration symmetries of the total aberration.

Wavefront aberration	Symmetry	Formula
$W_{060}(\vec{\rho} \cdot \vec{\rho})$	Spherical aberration symmetry	$OPD(0,0,0,1) - W_{040}$
$W_{331}(\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho})$	Coma aberration symmetry	$\frac{OPD(0,1,0,1) - OPD(0,-1,0,1)}{2}$ $-W_{131}$
$W_{151}(\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho})^2$		
$W_{333}(\vec{H} \cdot \vec{\rho})^3$		
$W_{420}(\vec{H} \cdot \vec{H})^2(\vec{\rho} \cdot \vec{\rho})$	Field curvature aberration symmetry	$OPD(0,1,1,0) - OPD(0,0,0,1)$ $-W_{220}$
$W_{240}(\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho})^2$		
$W_{422}(\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho})^2$	Astigmatism aberration symmetry	$OPD(0,1,0,1) - OPD(0,1,1,0)$ $-COMA - W_{131} - W_{222}$
$W_{242}(\vec{H} \cdot \vec{\rho})^2(\vec{\rho} \cdot \vec{\rho})$		

The images of the aperture stop in object and image spaces are the entrance and exit pupils, respectively. There are also intermediate images of the stop in each space of the system that serve as intermediate entrance and exit pupils for each optical surface. The intermediate exit pupil for a surface is equivalent to the entrance pupil for the following surface. Each surface generally introduces aberrations to the wavefront, as well as errors in mapping coordinates between intermediate pupils. In the following calculation, it is important to use a common set of coordinates that accounts for the distortion of the pupil coordinates. In each step, a uniform grid of rays is defined at the intermediate exit pupil. The rays passing through the uniform exit pupil coordinates are found by using a built-in iterative method in lens design software often referred to as ray aiming.

The procedure to determine intrinsic and extrinsic aberration components of the surface j is straightforward and is similar to one described by Hoffman [8]. First, the rays are traced to the intermediate exit pupil of the surface j and the total aberration for all surfaces up to the surface j is calculated. Next, the total aberration for all surfaces up to the surface $j-1$ is found by tracing rays to the intermediate exit pupil of the surface $j-1$. Since a common set of coordinates on the intermediate exit pupils is used for both calculations, the complete aberration of the surface j is calculated by simply subtracting the total aberration values calculated in steps one and two. Finally, the intrinsic aberration component of the surface j is found by setting up a separate optical system that consists of the surface j alone. The system parameters are defined by the intermediate first-order object and image locations; the intermediate entrance and exit pupils; and the intermediate first-order object coordinate. The extrinsic aberration of the surface j is found by subtracting the intrinsic aberration component calculated in the previous step from the complete aberration of the surface j found earlier. Fig 2.2 shows an example of the calculation of intrinsic and extrinsic aberration components for the secondary mirror of a Cassegrain Telescope.

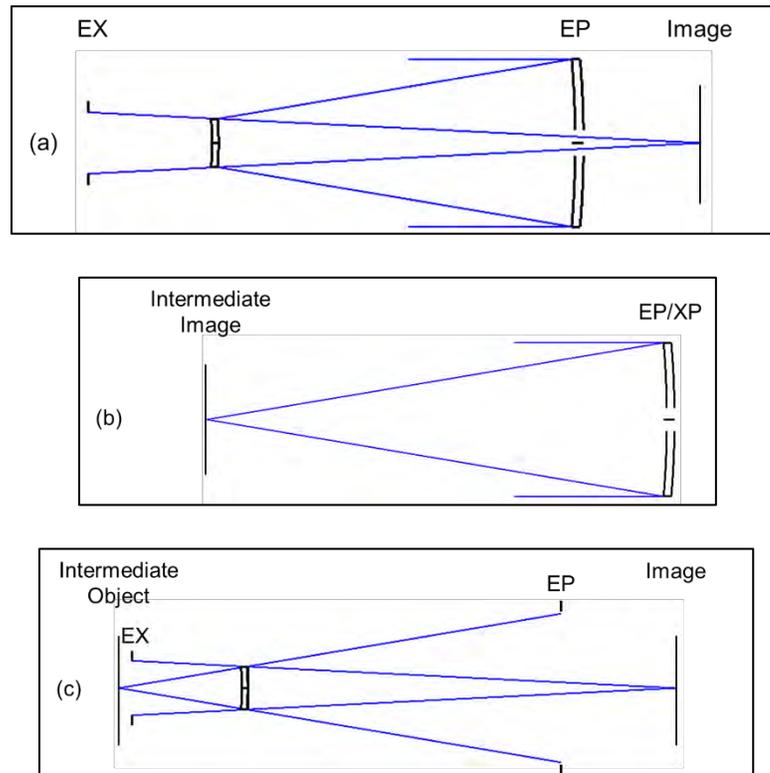


Fig. 2.2 The procedure to calculate intrinsic and extrinsic aberration components for the secondary mirror of a Cassegrain Telescope: (a) – the total aberration of the primary and secondary mirror is determined; (b) – the total aberration of the primary mirror is evaluated at the intermediate image plane; (c) – an optical system that consists of the secondary mirror alone is set up, and the intrinsic aberration component of the secondary mirror is calculated. To establish a set of common coordinates, rays are always aimed at the exit pupil.

Much attention is given to the presentation of aberrations. Siedel aberrations are often presented as individual surface aberration contributions. Ideally, all optical surfaces contribute a small amount of aberration which adds up to close to zero total aberration. Surfaces with maximum contributions are decisive for the total aberration correction. An alternative presentation is referred to in literature as Siedel Pegel diagrams [9]. Siedel Pegel diagrams plot the partial sums of individual surface contributions or, in other words, the

cumulative aberration propagating in the optical system. A large cumulative aberration generated within the system often indicates an additional stress inherent in the design and can be easily read from the associated Pegel diagrams.

The goal of this section is to present the method for calculating aberration symmetries and to gain insight into the higher-order aberration behavior of lens systems. Hofmann and Shafer have previously shown that the extrinsic aberration component is an essential and fundamental element of the higher-order aberration balance in many well-corrected lenses [8, 10-12]. The proposed aberration presentation as four aberration symmetries simplifies their analysis and provides additional insight.

2.6 ANALYSIS EXAMPLE: MOBILE PHONE CAMERA

Mobile camera technology and devices is a very fast growing field in the imaging market and is impacting the industry by deliberately replacing conventional photographic cameras. The design and packaging of a miniature camera lens module imposes unique optical design challenges. In this section, methods and techniques discussed previously are applied to analyze several state-of-the-art mobile camera lenses and to explore how a curved image surface can benefit the lens design of these optical systems [13].

A curved image sensor can potentially improve the chief ray incidence angle (CRA) on the sensor, as well as the aberration balancing, image quality, packaging, and manufacturing tolerance sensitivity.

2.6.1 First-Order Properties

The layouts of two lenses, one with a flat imaging surface called the benchmark lens and another with a curved imaging surface called the evaluation lens, are shown in Fig. 2.3.

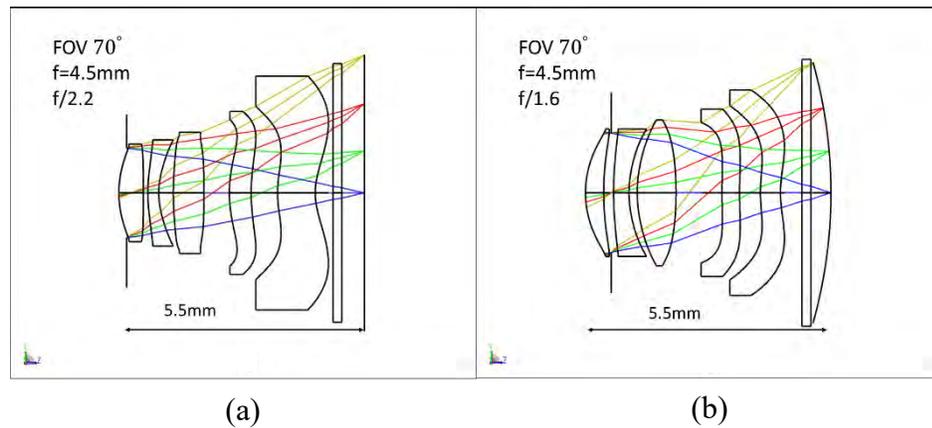


Fig. 2.3 Optical layout of a mobile camera. (a) - Conventional flat field design based on existing patent (benchmark lens). (b) - Representative lens imaging on a curved sensor (evaluation lens).

The design specifications are summarized in Table 2.3. The requirements have been derived by comparing products in the market, from patent data, and from publications in the mobile platform optical design and fabrication sector.

Table 2.3 Design specifications of a typical mobile camera lens.

Requirement	Value
Sensor format	1/3"
f/#	2.2
FOV [deg]	70
f	4.5
Total length [mm]	<5.5
Distortion	<1%
Number of lenses	5
Materials	COC, OKP4
Edge thickness [mm]	>0.1
Center thickness [mm]	>0.3
Air gap [mm]	>0.1
Surface slope [deg]	<55
Element aspect ratio	<1:5
IR cut filter [mm]	0.2
Surface slope [deg]	<55

Digital image sensors become less efficient when the incident light is at higher obliquity. The field of view (FOV) of the mobile camera is large, and the CRA proportional to the FOV. Therefore, for better CRA control, the aperture stop in a conventional mobile lens is placed close to the front, away from the image plane. The CRA is usually limited to no more than 30 degrees. However, if the image sensor is curved, the CRA is significantly reduced. Thus, the stop location has more flexibility and can be moved to make the lens system less unsymmetrical about the stop.

Fig. 2.4 plots values of w_j and s_j defined in Section 2.2 per optical element for both lenses. Total values correspond to parameters W and S also defined in Section 2.2. The plot provides an indication of where the optical power originates within the system. Both designs have similar configuration: the first three lenses provide most of the optical power while the last two elements are weak correcting lenses. However, the power distribution in the lens imaging on a curved sensor is more symmetrical about the stop. In a symmetrical or nearly symmetrical optical system, all odd aberrations tend to cancel out, permitting a higher level of aberration correction. W and S values of the evaluation lens are around 0.5 and slightly lower comparing to the benchmark lens.

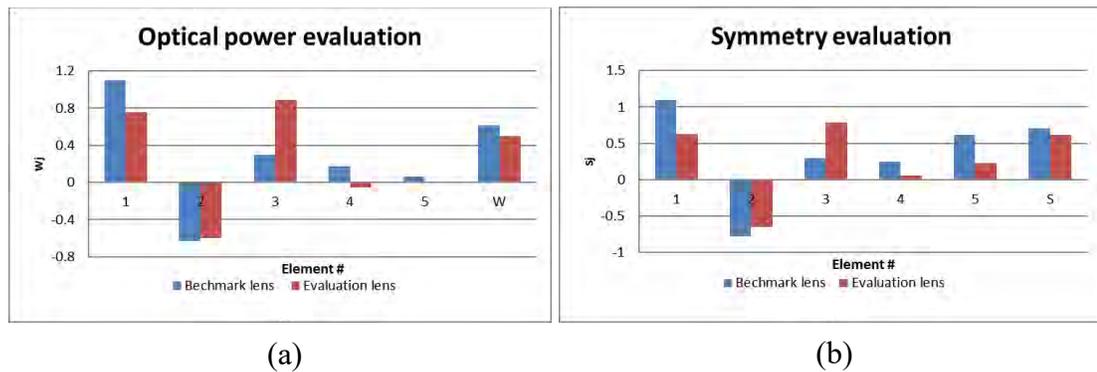


Fig. 2.4 Comparison of the first-order properties for a flat field design (benchmark lens) and a representative lens imaging on a curved sensor (evaluation lens). (a) – Optical power evaluation utilizing parameters w_j and W . (b) – Symmetry evaluation utilizing parameters s_j and S . A curved image surface allows a more symmetrical power distribution around the aperture stop.

2.6.2 Angles of Refraction

In Fig. 2.5, the ray refraction invariant $n \cdot \sin(I)$ discussed in Section 2.3 is evaluated surface by surface for both lenses. The rays used are the real marginal and chief rays. As

shown, the angles of refraction of the marginal ray for the evaluation lens are slightly larger indicating more refraction and lens stress. Though, the evaluation design is faster.

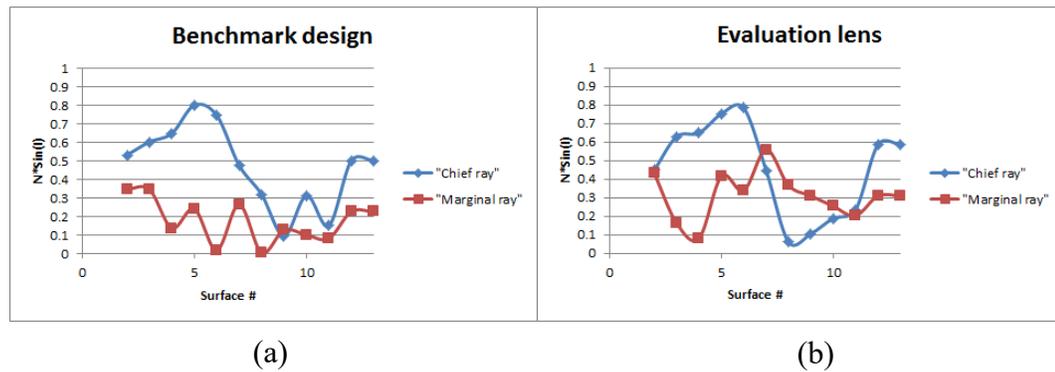


Fig. 2.5 Ray refraction invariant. (a) - Conventional flat field design based on existing patent (benchmark lens). (b) - Representative lens imaging on a curved sensor (evaluation lens). The angles of refraction of the marginal ray are slightly higher for the faster design with a curved image surface.

2.6.3 Spherical Aberration Correction Analysis

Diagrams of surface contributions to the symmetry of spherical aberration for the benchmark lens are shown in Fig. 2.6. The calculation of aberration symmetries was discussed in Section 2.5. The individual surface components are often very large. For example, the front and back surfaces of the second element contribute significant fourth- and high-order spherical aberration of opposite sign that balance one another. The associated Pegel diagram clearly shows tens waves of spherical aberration propagating inside the element, and, consequently, tight surface alignment tolerances are expected for this component. In the benchmark lens, the fourth-order spherical aberration is prevailing over high-order spherical aberration. Notice that large intrinsic aberration components induce extrinsic spherical aberration on the back surface of the second element. However, since the propagation distance inside the element is relatively small, the extrinsic spherical

aberration is also small. From the Pegel diagram, we conclude that correction of spherical aberration is predominantly accomplished with first two elements.



Fig. 2.6 Surface contributions to the symmetry of spherical aberration and associated Pegel diagram for a flat field design (benchmark lens).

The limiting aberrations for the evaluation lens are spherical aberration, oblique spherical aberration and sphero-chromatism. The correction of spherical aberration is accomplished within the front group of elements. However, in contrast to the benchmark lens, the surface contributions to the spherical aberrations generally all have similar

magnitude, but the opposite sign for the fourth- and high-order components, giving large system sums that propagate and must be balanced against one another. Fig. 2.8 shows the variation of spherical aberration and oblique spherical aberration across the field of view. The plot in Fig. 2.8 is generated with the limiting aberration analysis methods discussed in Section 2.4. Spherical and oblique spherical aberrations are balanced at about middle of the FOV, leaving a small amount of residual aberration elsewhere.



Fig. 2.7 Surface contributions to the symmetry of spherical aberration and associated Pegel diagram for a flat field design (benchmark lens) and a representative lens imaging on a curved sensor (evaluation lens).

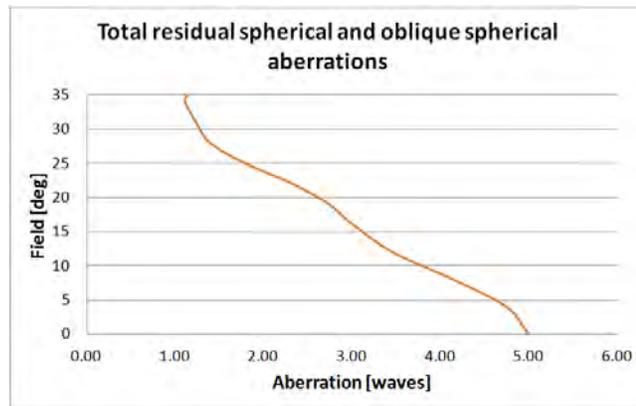


Fig. 2.8 Residual spherical and oblique spherical aberration across the field of view of a representative lens imaging on a curved sensor (evaluation lens).

2.6.4 Field Curvature Correction Analysis

In a flat field lens, field curvature correction is obtained by introducing negative optical power, and this leads to more overall optical power. Field curvature correction optically stresses a lens, and aberration residuals grow larger. For a given image quality, a lens with a curved imaging surface can have faster optical speed due to reduced optical stress. Field curvature aberration is compensated by the curved sensor.

The field curves of the benchmark design, shown in Fig. 2.9, are typical for a flat field mobile lens. The field curves are wavy with multiple crossing across the FOV. The Petzval radius is -19.12mm. Sharp imaging on a flat surface is achieved without satisfying the classical requirement of having a Petzval sum nearly zero. As shown in Fig. 2.10, the aspheric optical elements located close to the image plane contribute higher-order field curvature and astigmatism. Different orders of the field curvature and astigmatism are balanced to compensate for any residual Petzval curvature. The induced aberrations of

elements four and five are very significant in magnitude. The induced aberration component is an inherent part of the aberration balance in a mobile lens and is essential as an effective design variable.

In contrast, the field curves for the evaluation lens are much more smooth. The Petzval radius is -8.74mm (about two times the focal length); this clearly indicates that the field curvature is compensated by the curved image surface. The image surface radius of curvature is -10.86mm .

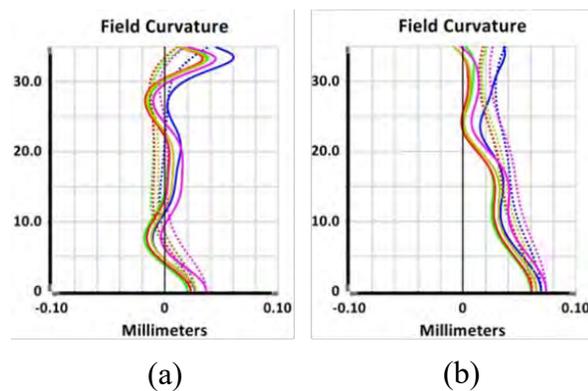


Fig. 2.9 Field curves. (a) - Conventional flat field design based on existing patent (benchmark lens). (b) - Representative lens imaging on a curved sensor (evaluation lens). A curved image surface compensates the field curvature aberration.

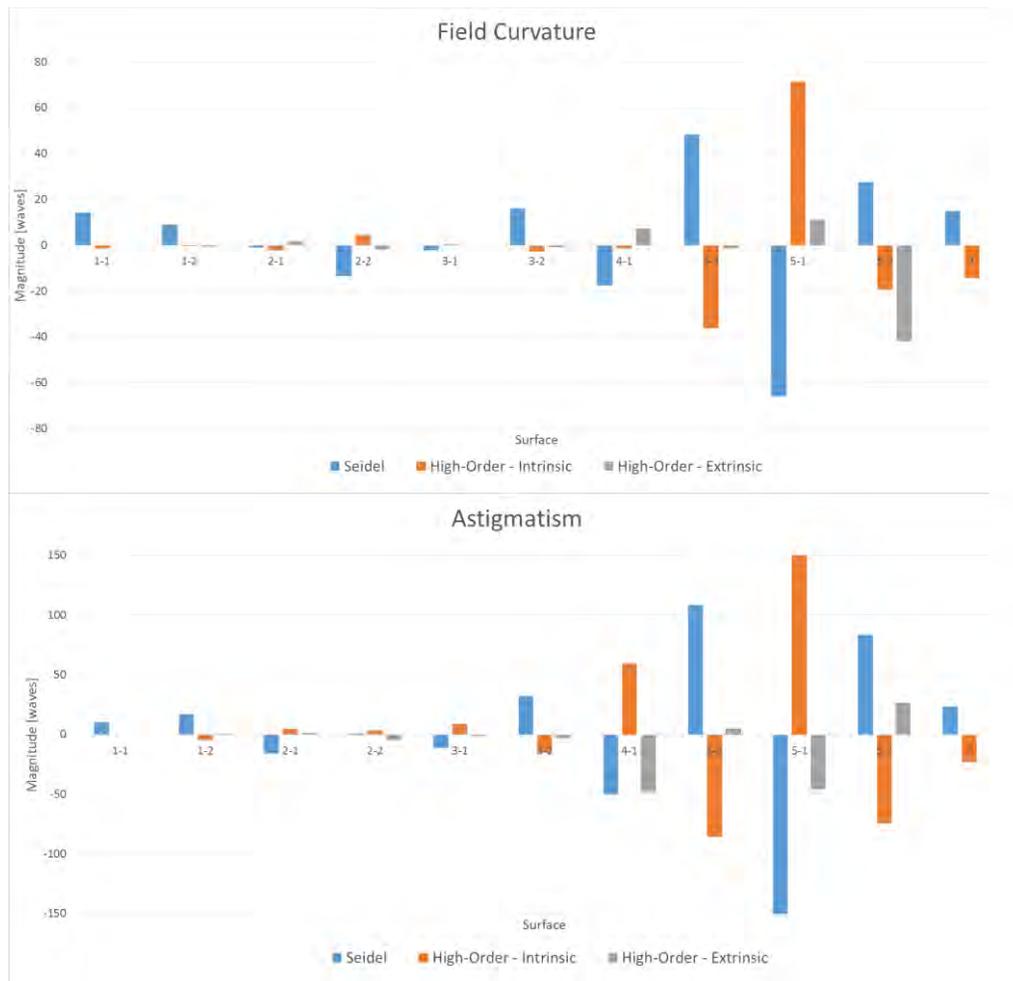


Fig. 2.10 Surface contributions to the symmetry of field curvature and astigmatism for a flat field design (benchmark lens).

2.6.5 Image Quality

Figure 2.11 plots the modulation transfer function (MTF) for both lenses. Both designs show very good performance over the entire FOV with an average MTF of about 70% at 112 lp/mm (grey scale Ny/4 frequency for a 1.1 μ m pixel) and over 45% MTF at 225 lp/mm (grey scale Ny/2 frequency for a 1.1 μ m pixel). However, the lens designed for a curved image sensor is not only one f-number faster compared with the conventional design, but

also shows more uniform MTF over the field. It would be very unlikely to achieve similar aberration correction for this f-number for a flat sensor with five lens elements. In the benchmark lens, the MTF at high frequencies varies significantly with the field angle. The MTF of the evaluation lens is more uniform over the field.

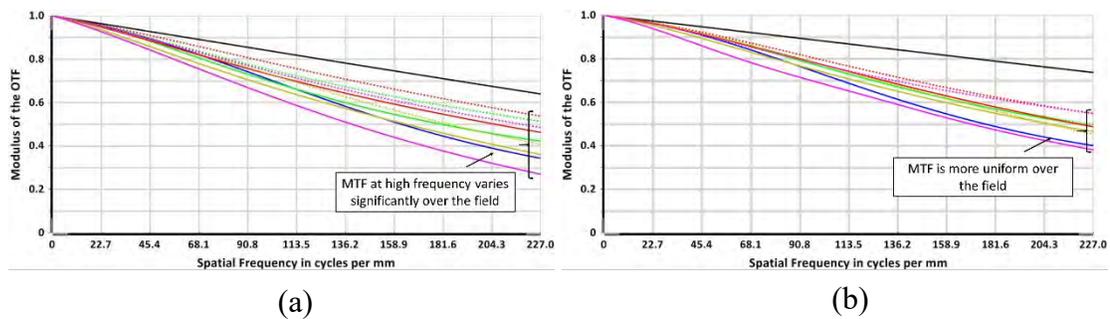


Fig. 2.11 MTF a mobile camera: (a) – Conventional flat field design at $f/2.2$ (benchmark lens); (b) – Representative lens imaging on a curved sensor at $f/1.6$ (evaluation lens). The lens imaging on a curved sensor is one f-stop faster and shows more uniform performance over the field.

2.6.6 Tolerance Analysis

It is important to provide insight into the sensitivity to manufacturing tolerances. Tilts and decenters have the largest effect on the as-built performance of the mobile lens. The effect of misalignments is evaluated for a lens element decenter of $5\mu\text{m}$ and tilt of 0.1° . These values are considered being commercial tolerances for small molded plastic optical elements; however, in practice much tighter tolerances are specified to increase lens manufacturing yield. The RMS wavefront error, root summed squared over the field, is used as a criterion. Twenty-one field points are analyzed: five field points in each $+X$, $-X$, $+Y$, $-Y$ directions and the on-axis field.

Since the tolerance sensitivity strongly depends on the focal ratio of the lens, both flat field and curved sensor designs are compared at $f/2.2$. The tolerance sensitivity of the evaluation lens is also estimated at $f/1.6$. The results are summarized in Fig. 2.12.

As expected, the $f/1.6$ lens is the most tolerance sensitive: the manufacturability may be a limiting factor for this fast design. Fortunately, the fabrication technologies of molded optics are constantly improving, allowing tighter tolerances. Comparison of lenses at $f/2.2$ shows that the evaluation lens is performing better under manufacturing tolerances than the benchmark design.

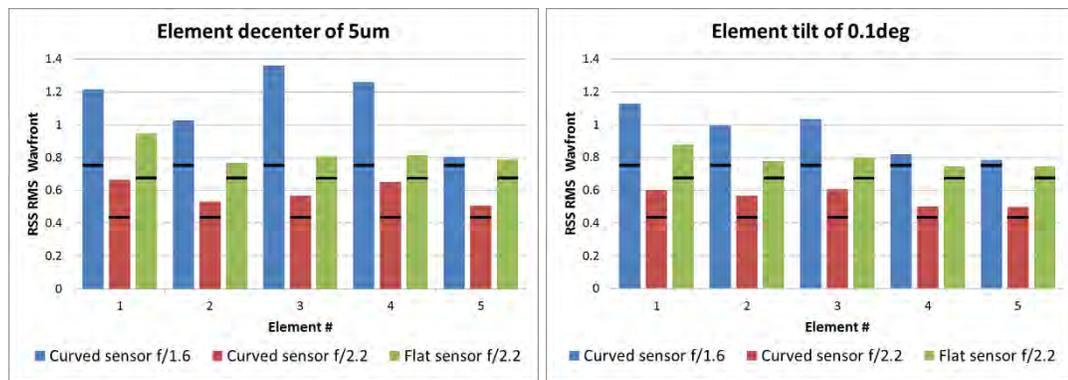


Fig. 2.12 Sensitivity to lens element decenter and tilt. Lens imaging on a curved sensor shows better as-build performance. The horizontal line indicates the nominal criterion value.

2.6.7 Total Length

The total length of a mobile camera is an important design parameter. A shorter lens, such as a telephoto lens, imposes optical stress on the system, and departs more from the symmetry by introducing more optical power in the individual elements. As the aberrations substantially increase with lens stress, it would be nearly impossible to control the remaining aberrations for the required FOV and focal ratio ($f/\#$). Thus, in practice, no substantial reduction in length is obtained using a curved image sensor.

2.7 CONCLUSION

Successful optical design has always required a combination of computational methods and the experience of the lens designer. In this chapter, several appropriate principles and techniques for evaluation of complex optical imaging systems have been presented and discussed. These tools have been developed and implemented in a commercial lens design program as an aid to lens designers who are concerned about the detailed and comprehensive understanding of the inner workings of lenses and trade-offs involved in the design.

The ideas proposed in this chapter are that (1) two applicable parameters for quantify relevant lens attributes and providing adequate comparison between different optical configurations are the optical power distribution and lens symmetry; (2) the examination of the ray invariant product $n \cdot \sin(i)$ for a critical ray path is a widely applicable technique that will often indicate the source of a design problem; (3) understanding the limiting aberrations of a lens is critical in order to further improve the design by applying an appropriate technique to correct these specific limiting aberrations or choosing a different lens configuration to keep them from arising; (4) a better understanding of the intrinsic and extrinsic surface contributions to the overall aberration balance is essential for designing efficient imaging lenses.

Methods and techniques discussed in this chapter were applied to analyze state-of-the-art mobile camera lenses and explore how a curved image surface can benefit the lens design of these optical systems. In summary, the curved image surface allows producing an equivalently performing design with faster focal ratio ($f/\#$) than the conventional design.

It was found that about one f-stop improvement in speed can be achieved while preserving uniform image quality over the entire FOV. Small $f/\#$ number lenses provide a better and very desirable low light imaging and can accommodate for a larger number of sensor pixels leading to better resolution. The aperture stop location between first and second elements is optimal for aberration balancing and controlling the total length of the system. The author believes that in practice a curved image surface will not allow substantial reduction in length of a mobile camera.

The radius of curvature of the sensor in the presented design is about 11mm for a 4.5 mm focal length lens. Although the author is unaware of commercially available curved sensors suitable for mobile applications, potential benefits may force further development of this technology allowing a new generation of faster compact mobile cameras.

REFERENCES

1. W. J. Smith, "Modern Lens Design," McGraw-Hill, Second edition (2004)
2. R. Kingslake, "Lens Design Fundamentals," Academic Press, second edition (2010)
3. D. Shafer, "Optical design and the relaxation response," Proc. SPIE **766** (1987)
4. E. Glatzel, "New lenses in microlithography," SPIE Vol. **237** (1980)
5. J. Sasian and M. R. Descour, "Power distribution and symmetry in lens systems," Opt. Eng. **37**(3) (1998)
6. OpticStudio 15, Product of Zemax LLC., Redmond, WA (2016)
7. J. Sasian, "Introduction to Aberrations in Optical Imaging Systems," Cambridge University Press (2013)

8. J. M. Hoffman, "Induced aberrations in optical systems," PhD Dissertation, the University of Arizona (1993)
9. W. E. Woeltche, "Structure and image forming properties of asymmetrical wide angle lenses for 35-mm photography," *Applied Optics*, Vol. 7 (2) (1968)
10. D. Shafer, "Aberration theory and the meaning of life," *Proc. of SPIE*, Vol. **0554** (1985)
11. D. Shafer. "Some usefull correspondence in aberration theory," *Proc. of SPIE*, Vol. **0655** (1986)
12. D. Shafer, "Induced aberrations complicate lens design," *Laser Focus World* (1991)
13. D. Reshidko and J. Sasian, "Optical analysis of miniature lenses with curved imaging surfaces," *Applied Optics*, Vol **54** (28) (2015)

CHAPTER III: GEOMETRICAL IRRADIANCE CHANGES IN A SYMMETRIC OPTICAL SYSTEM

3.1 INTRODUCTION

In the development of the theory of aberrations in optical imaging systems, emphasis has been given to the study of image aberrations, which are described as wave, angular, transverse, or longitudinal quantities [1]. The light irradiance variation, specifically at the exit pupil plane and at the image plane of an optical system, is a radiometric aspect of the system that is also of interest. An accurate diffraction calculation of the system's point spread function requires not only knowledge of the wavefront phase but also of its amplitude. These, the phase and amplitude are usually calculated geometrically at the exit pupil of the system.

The light irradiance distribution at the exit pupil plane and/or at the image plane of an optical system can be derived from basic radiometric principles, such as conservation of flux [2]. This chapter provides a study of the relationship between irradiance at these two planes and the system's wavefront aberration coefficients. This study is based on geometrical optics.

The concept of the wavefront aberration function is well established. The wavefront aberration function $W(\vec{H}, \vec{\rho})$ of an axially symmetric system gives the geometrical wavefront deformation at the exit pupil as a function of the normalized field \vec{H} and aperture $\vec{\rho}$ vectors. The field vector and the aperture vector may be defined in either object

or image spaces. Two vectors uniquely specify any ray propagating in the lens system. The wavefront aberration function is expanded into polynomial series of the rotational invariants as dot products of the field and aperture vectors, specifically $(\vec{H} \cdot \vec{H})$, $(\vec{H} \cdot \vec{\rho})$ and $(\vec{\rho} \cdot \vec{\rho})$, and to the fourth order of approximation on \vec{H} and $\vec{\rho}$ is written as

$$\begin{aligned}
W(\vec{H}, \vec{\rho}) &= \sum_{j,m,l} W_{k,l,m} (\vec{H} \cdot \vec{H})^j (\vec{H} \cdot \vec{\rho})^m (\vec{\rho} \cdot \vec{\rho})^n \\
&= W_{000} + W_{200} (\vec{H} \cdot \vec{H}) + W_{111} (\vec{H} \cdot \vec{\rho}) + W_{020} (\vec{\rho} \cdot \vec{\rho}) \\
&\quad + W_{040} (\vec{\rho} \cdot \vec{\rho})^2 + W_{131} (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho}) + W_{222} (\vec{H} \cdot \vec{\rho})^2 \\
&\quad + W_{220} (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho}) + W_{311} (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho}) + W_{400} (\vec{H} \cdot \vec{H})^2,
\end{aligned} \tag{3.1}$$

where each aberration coefficient $W_{k,l,m}$ represents the amplitude of basic wavefront deformation forms [3]. Similarly, a pupil aberration function $\bar{W}(\vec{H}, \vec{\rho})$ can be defined to describe aberration between the pupils. This function is constructed by interchanging the role of the field and aperture vectors, and to fourth order it is

$$\begin{aligned}
\bar{W}(\vec{H}, \vec{\rho}) &= \sum_{j,m,l} \bar{W}_{k,l,m} (\vec{H} \cdot \vec{H})^j (\vec{H} \cdot \vec{\rho})^m (\vec{\rho} \cdot \vec{\rho})^n \\
&= \bar{W}_{000} + \bar{W}_{200} (\vec{\rho} \cdot \vec{\rho}) + \bar{W}_{111} (\vec{H} \cdot \vec{\rho}) + \bar{W}_{020} (\vec{H} \cdot \vec{H}) \\
&\quad + \bar{W}_{040} (\vec{H} \cdot \vec{H})^2 + \bar{W}_{131} (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho}) + \bar{W}_{222} (\vec{H} \cdot \vec{\rho})^2 \\
&\quad + \bar{W}_{220} (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho}) + \bar{W}_{311} (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho}) + \bar{W}_{400} (\vec{\rho} \cdot \vec{\rho})^2,
\end{aligned} \tag{3.2}$$

where the pupil aberration coefficients are barred to distinguish them from the image aberration coefficients.

In analogy with the wavefront aberration function, the irradiance function $I(\vec{H}, \vec{\rho})$ is defined that gives the irradiance at the image plane of an optical system. To the fourth order of approximation, this irradiance function is expressed as

$$\begin{aligned}
I(\vec{H}, \vec{\rho}) &= \sum_{j,m,n} I(\vec{H} \cdot \vec{H})^j \cdot (\vec{H} \cdot \vec{\rho})^m \cdot (\vec{\rho} \cdot \vec{\rho})^n \\
&= I_{000} + I_{200} (\vec{H} \cdot \vec{H}) + I_{111} (\vec{H} \cdot \vec{\rho}) + I_{020} (\vec{\rho} \cdot \vec{\rho}) \\
&\quad + I_{040} (\vec{\rho} \cdot \vec{\rho})^2 + I_{131} (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho}) + I_{222} (\vec{H} \cdot \vec{\rho})^2 \\
&\quad + I_{220} (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho}) + I_{311} (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho}) + I_{400} (\vec{H} \cdot \vec{H})^2,
\end{aligned} \tag{3.3}$$

where the irradiance coefficients $I_{k,l,m}$ represent basic illumination distribution patterns at the image plane. In addition, another irradiance function $\bar{I}(\vec{H}, \vec{\rho})$ is also defined that gives the irradiance of the beam at the exit pupil plane of an optical system,

$$\begin{aligned}
\bar{I}(\vec{H}, \vec{\rho}) &= \sum_{j,m,l} \bar{I}_{l,k,m} (\vec{H} \cdot \vec{H})^j (\vec{H} \cdot \vec{\rho})^m (\vec{\rho} \cdot \vec{\rho})^n \\
&= \bar{I}_{000} + \bar{I}_{200} (\vec{\rho} \cdot \vec{\rho}) + \bar{I}_{111} (\vec{H} \cdot \vec{\rho}) + \bar{I}_{020} (\vec{H} \cdot \vec{H}) \\
&\quad + \bar{I}_{040} (\vec{H} \cdot \vec{H})^2 + \bar{I}_{131} (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho}) + \bar{I}_{222} (\vec{H} \cdot \vec{\rho})^2 \\
&\quad + \bar{I}_{220} (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho}) + \bar{I}_{311} (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho}) + \bar{I}_{400} (\vec{\rho} \cdot \vec{\rho})^2,
\end{aligned} \tag{3.4}$$

where each irradiance coefficient $\bar{I}_{l,k,m}$ represents basic apodization distributions at the exit pupil, as shown in Fig. 3.1. In these functions, it is assumed that the irradiance is greater than zero for any point defined by the normalized field \vec{H} and aperture $\vec{\rho}$ vectors.

It is worth noting the symmetry between the image aberration function $W(\vec{H}, \vec{\rho})$ and the pupil aberration function $\bar{W}(\vec{H}, \vec{\rho})$, and likewise the symmetry between the image irradiance function $I(\vec{H}, \vec{\rho})$ and the pupil irradiance function $\bar{I}(\vec{H}, \vec{\rho})$.

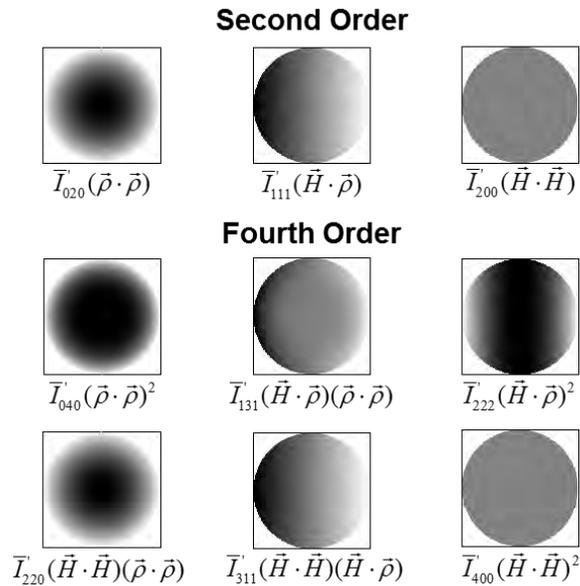


Fig. 3.1 Second- and fourth-order apodization terms at the exit pupil given by the coefficients of the pupil irradiance function.

The question the author pose and answer is: what is the relationship between the wavefront aberration functions $W(\vec{H}, \vec{\rho})$ and $\bar{W}(\vec{H}, \vec{\rho})$, and the irradiance functions $I(\vec{H}, \vec{\rho})$ and $\bar{I}(\vec{H}, \vec{\rho})$? The conservation of flux in an optical system is used to determine the relationship between wavefront and irradiance coefficients. The particular case of relative illumination at the image plane is also discussed in some detail. The theoretical results are in agreement with real ray tracing. Overall, this chapter provides new insights into the irradiance changes in an optical system, and furthers the theory of aberrations.

3.2 RADIATIVE TRANSFER IN AN OPTICAL SYSTEM

This section reviews radiometry and shows how aberrations relate to conservation of optical flux Φ in an optical system. Fig. 3.2 illustrates the basic elements of an axially symmetric optical system and the geometry defining the transfer of radiant energy. Rays from a differential area dA in the object plane pass through the optical system and converge at the conjugate area dA' in the image plane. An arbitrary plane S in object space and a plane S' conjugate to S in image space are defined. The differential cross sections of the beam dS in S and dS' in S' are optically conjugated to first order.

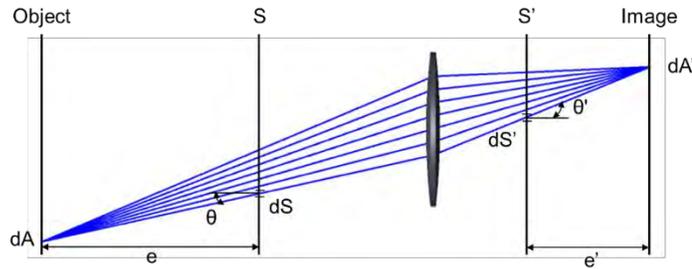


Fig. 3.2 Geometry defining the transfer of radiant energy from differential area dA in object space to the conjugate area dA' in image space. The radiant flux through all cross sections of the beam is the same.

In a lossless and passive optical system, the element of radiant flux $d\Phi$ is conserved through all cross sections of the beam. If we choose planes S and S' at the entrance and exit pupils respectively, the equation for the conservation of flux becomes

$$d\Phi = L_o \frac{dA dS \cos^4(\theta)}{e^2} = L_o \frac{dA' dS' \cos^4(\theta')}{e'^2} = d\Phi', \quad (3.5)$$

where L_o is the source radiance. In Eq. 3.5, a Lambertian source is assumed and, consequently, L_o is constant. The object space angle θ is between the ray connecting dA

and dS , and the optical axis of the lens system. Similarly, θ' is the image space angle between the ray connecting dA' and dS' , and the optical axis. e (e') is the axial distance between the object (image) plane and the entrance (exit) pupil plane respectively.

Eq. 3.5 provides the radiant flux along a particular ray in an optical system. Since the normalized field \vec{H} and aperture $\vec{\rho}$ vectors uniquely specify any ray propagating in the optical system, Eq. 5 gives the radiant flux as a function of normalized field \vec{H} and pupil coordinates $\vec{\rho}$.

The irradiance on the exit pupil plane is obtained by dividing the radiant flux that is incident on a surface by the unit area. It follows that

$$d\bar{I}(\vec{H}, \vec{\rho}) = \frac{d\Phi'}{dS'} = L_o \frac{dA'}{e^2} \cos^4(\theta') = L_o \frac{dA}{e^2} \frac{dS}{dS'} \cos^4(\theta) = \frac{d\Phi}{dS}, \quad (3.6)$$

Where $d\bar{I}(\vec{H}, \vec{\rho})$ is the differential of irradiance on the exit pupil plane. Similarly, the irradiance on the image plane is obtained by dividing the radiant flux by the image surface unit area as in

$$dI(\vec{H}, \vec{\rho}) = \frac{d\Phi'}{dA'} = L_o \frac{dS'}{e^2} \cos^4(\theta') = L_o \frac{dS'}{e^2} \frac{dA}{dA'} \cos^4(\theta) = \frac{d\Phi}{dA}, \quad (3.7)$$

And $dI(\vec{H}, \vec{\rho})$ is now the differential of irradiance on the image plane. The differentials of irradiance $d\bar{I}(\vec{H}, \vec{\rho})$ and $dI(\vec{H}, \vec{\rho})$ are functions of the field and aperture vectors.

Eq. 3.6 and Eq. 3.7 are general and do not involve any approximations. However, care must be taken in evaluating these expressions, since the angles θ and θ' may vary due to aberrations. In addition, the differential areas dA , dA' , dS and dS' may also vary for different points in the aperture and in the field of the lens.

3.3 IRRADIANCE FUNCTION OF A PINHOLE CAMERA

This section calculates coefficients of the irradiance function defined in Eq. 3.6 and Eq. 3.7 for a pinhole camera. The pinhole camera produces images of illuminated objects as light passes through an aperture, as shown in Fig. 3.3.

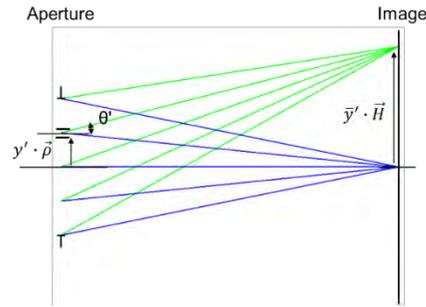


Fig. 3.3 Geometry defining a pinhole camera. Images are formed as light passes through the aperture.

In this model, the field vector \vec{H} is defined on the image plane and the aperture vector $\vec{\rho}$ is set on the plane of the aperture. The irradiance functions $\bar{I}(\vec{H}, \vec{\rho})$ and $I(\vec{H}, \vec{\rho})$ give the relative irradiance along the ray specified by the field and aperture vectors. It follows from Eq. 3.6 and Eq.3.7 that irradiance distribution at a point specified by \vec{H} on the focal plane or at a point specified by $\vec{\rho}$ on the aperture plane are given by $\cos^4(\theta')$ of the particular ray. To calculate the coefficients of the irradiance functions to fourth order, $\cos^4(\theta')$ is expressed in terms of first-order system parameters as (see Appendix D)

$$\begin{aligned}
 I_{pinhole}(\vec{H}, \vec{\rho}) &= \bar{I}_{pinhole}(\vec{H}, \vec{\rho}) = \cos^4(\theta') = \\
 &= 1 - 2 \cdot \bar{u}'^2 (\vec{H} \cdot \vec{H}) - 2 \cdot u'^2 (\vec{\rho} \cdot \vec{\rho}) - 4 \cdot u' \bar{u}' (\vec{H} \cdot \vec{\rho}) \dots \\
 &+ 3 \cdot \bar{u}'^4 (\vec{H} \cdot \vec{H})^2 + 12 \cdot u'^3 \bar{u}' (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho}) + 12 \cdot u'^2 \bar{u}'^2 (\vec{H} \cdot \vec{\rho})^2 \dots \\
 &+ 6 \cdot u'^2 \bar{u}'^2 (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho}) + 12 \cdot u' \bar{u}'^3 (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho}) + 3 \cdot u'^4 (\vec{\rho} \cdot \vec{\rho})^2,
 \end{aligned} \tag{3.8}$$

where u' and \bar{u}' are the first-order marginal and chief ray slopes in image space.

In aberration theory, the reference sphere is used as the reference to measure the wavefront deformation. In analogy, the geometry in Fig. 3.3 can be used to define a model of irradiance $I_{pinhole}(\vec{H}, \vec{\rho})$ which is helpful as a reference and for further calculations. In an actual system, image and pupil aberrations affect the irradiance distribution.

3.4 IRRADIANCE ON THE IMAGE PLANE

The aperture vector $\vec{\rho}$ is selected at the exit pupil plane and, thus, it defines the ray intersection with this plane. Rays at the exit pupil pass through a uniform grid by construction, and differential areas dS' are constant given that we choose to define rays at the exit pupil. To calculate the image plane irradiance, Eq. 3.7 is evaluated in image space as

$$dI = I_0 \cdot \cos^4(\theta') \quad (3.9)$$

where

$$I_0 = L_o \frac{dS'}{e'^2}, \quad (3.10)$$

Where $I_0 = 1$ is a constant corresponding to the irradiance value for the on-axis field point. In contrast to a pinhole camera, in an actual system, the ray angle in Eq. 3.11 is modified by image aberrations. To calculate the irradiance distribution at the image plane of an optical system with the aperture stop at the exit pupil, $\cos^4(\theta')$ is expressed in terms of first-order system parameters and wavefront aberration coefficients (see Appendix E). Table 3.1a provides a summary of the second-order image plane irradiance coefficients. It is interesting to note that the second-order irradiance terms are not affected by image

aberrations and that the image plane irradiance to second order is equivalent to the ideal irradiance of a pinhole camera

$$I(\vec{H}, \vec{\rho})^{(2)} = I_{pinhole}(\vec{H}, \vec{\rho})^{(2)} = 1 - 2u'^2(\vec{\rho} \cdot \vec{\rho}) - 4u'\bar{u}'(\vec{H} \cdot \vec{\rho}) - 2\bar{u}'^2(\vec{H} \cdot \vec{H}). \quad (3.11)$$

Table 3.1b summarizes the fourth-order image plane irradiance coefficients. The fourth-order irradiance coefficients are the sum of two components. The first component is represented by products of the first-order ray slopes u' and \bar{u}' in image space. The second component includes additional terms that are functions of the fourth-order image aberration coefficients and first-order ray slopes.

Table 3.1a. Second-order irradiance coefficients $I_{k,l,m}^{(2)}$ at the image plane of an optical system.

$I_{020}(\vec{\rho} \cdot \vec{\rho}) = -2u'^2(\vec{\rho} \cdot \vec{\rho})$
$I_{111}(\vec{H} \cdot \vec{\rho}) = -4u'\bar{u}'(\vec{H} \cdot \vec{\rho})$
$I_{200}(\vec{H} \cdot \vec{H}) = -2\bar{u}'^2(\vec{H} \cdot \vec{H})$

Table 3.1b. Fourth-order irradiance coefficients $I_{k,l,m}^{(4)}$ at the image plane of an optical system.

$I_{040}(\vec{\rho} \cdot \vec{\rho})^2 = [3u'^4 + \frac{16}{\lambda\mathcal{K}}W_{040}u'\bar{u}'](\vec{\rho} \cdot \vec{\rho})^2$
$I_{131}(\vec{\rho} \cdot \vec{\rho})(\vec{H} \cdot \vec{\rho}) = [12u'^3\bar{u}' + \frac{16}{\lambda\mathcal{K}}W_{040}\bar{u}'^2 + \frac{12}{\lambda\mathcal{K}}W_{131}u'\bar{u}'](\vec{\rho} \cdot \vec{\rho})(\vec{H} \cdot \vec{\rho})$

$$\begin{aligned}
I_{222}(\vec{H} \cdot \vec{\rho})^2 &= [12u'^2 \bar{u}'^2 + \frac{8}{\mathcal{K}} W_{131} \bar{u}'^2 + \frac{8}{\mathcal{K}} W_{222} u' \bar{u}'] (\vec{H} \cdot \vec{\rho})^2 \\
I_{220}(\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho}) &= [6u'^2 \bar{u}'^2 + \frac{4}{\mathcal{K}} W_{131} \bar{u}'^2 + \frac{8}{\mathcal{K}} W_{220} u' \bar{u}'] (\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho}) \\
I_{311}(\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho}) &= [12u' \bar{u}'^3 + \frac{8}{\mathcal{K}} W_{222} \bar{u}'^2 + \frac{8}{\mathcal{K}} W_{220} \bar{u}'^2 + \frac{4}{\mathcal{K}} W_{311} u' \bar{u}'] (\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho}) \\
I_{400}(\vec{H} \cdot \vec{H})^2 &= [3\bar{u}'^4 + \frac{4}{\mathcal{K}} W_{311} \bar{u}'^2] (\vec{H} \cdot \vec{H})^2
\end{aligned}$$

3.4 IRRADIANCE ON THE EXIT PUPIL PLANE

To calculate the irradiance distribution on the exit pupil plane of an optical system, Eq.

3.6 is evaluated in image space as

$$d\bar{I} = \bar{I}_0 \cdot \frac{dA'}{dA} \cdot \cos^4(\theta') \quad (3.12)$$

and

$$\bar{I}_0 \equiv \left(L_o \frac{dA}{e^2} \right), \quad (3.13)$$

where $\bar{I}_0 = 1$ is a constant corresponding to the irradiance value for the on-axis field point.

A comparison of Eq. 3.12 and Eq. 3.9 reveals an additional term $\frac{dA'}{dA}$ that is given by the

ratio between the elements of area at the object and image planes. In an optical system with

aberrations, the ratio $\frac{dA'}{dA}$ may vary over the pupil and over the field. Two differential areas

are related by the determinant of the Jacobian $J(\vec{H}, \vec{\rho})$ of the transformation

$$dA' = J(\vec{H}, \vec{\rho}) dA. \quad (3.14)$$

If the Jacobian determinant is expressed in terms of wavefront aberration coefficients, the differential of irradiance on the exit pupil plane is calculated by evaluating Eq. 3.12

and keeping terms to fourth order (see Appendix F). Table 3.2a and Table 3.2b provide a summary of the exit pupil plane irradiance coefficients to the fourth order of approximation.

The coefficients $\bar{I}_{k,l,m}$ are the sum of several components. The first component is represented by products of the first-order ray slopes u' and \bar{u}' in image space and is equivalent to the first component of $I_{k,l,m}$ coefficients. The second component includes additional terms that are functions of the fourth-order image aberration coefficients and first-order ray slopes. Note that the second-order terms of the exit pupil irradiance function are modified by image aberrations, in contrast to the second-order terms of the image plane irradiance. The third component of the coefficients $\bar{I}_{k,l,m}^{(4)}$ is proportional to the six-order image aberration coefficients. Finally, there are additional terms that involve products of the fourth-order aberrations.

In Table 3.1 and Table 3.2, the field vector \vec{H} is set at the object plane and the aperture vector $\vec{\rho}$ is defined at the exit pupil plane. The derivation assumes that the aperture stop coincides with the location of the aperture vector which is the exit pupil plane. The formulas would change depending on whether the field vector \vec{H} is at the object or image plane, and on whether the aperture vector $\vec{\rho}$ and the stop are at the entrance pupil, the exit pupil, or at an intermediate pupil.

Table 3.2a. Second-order irradiance coefficients $\bar{I}_{k,l,m}^{(2)}$ at the exit pupil plane of an optical system.

$$\bar{I}_{020}(\vec{H} \cdot \vec{H}) = \left[-2\bar{u}'^2 - \frac{4}{\mathcal{K}} W_{311} \right] (\vec{H} \cdot \vec{H})$$

$$\begin{aligned}\bar{I}_{111}(\vec{H} \cdot \vec{\rho}) &= [-4u'u' - \frac{4}{\mathcal{K}}W_{220} - \frac{6}{\mathcal{K}}W_{222}](\vec{H} \cdot \vec{\rho}) \\ \bar{I}_{200}(\vec{\rho} \cdot \vec{\rho}) &= [-2u'^2 - \frac{4}{\mathcal{K}}W_{131}](\vec{\rho} \cdot \vec{\rho})\end{aligned}$$

Table 3.2b. Fourth-order irradiance coefficients $\bar{I}_{k,l,m}^{(4)}$ at the exit pupil plane of an optical system.

$$\begin{aligned}\bar{I}_{040}(\vec{H} \cdot \vec{H})^2 &= [3\bar{u}'^4 - \frac{6}{\mathcal{K}}W_{511} + \frac{3}{\mathcal{K}}W_{311}\bar{u}'^2 + \frac{3}{\mathcal{K}^2}W_{311}W_{311}](\vec{H} \cdot \vec{H})^2 \\ \bar{I}_{131}(\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho}) &= [12u'u'\bar{u}'^3 - \frac{10}{\mathcal{K}}W_{422} - \frac{8}{\mathcal{K}}W_{420} + \frac{5}{\mathcal{K}}W_{222}\bar{u}'^2 + \frac{2}{\mathcal{K}}W_{220}\bar{u}'^2 + \frac{6}{\mathcal{K}}W_{311}u'u' \\ &\quad + \frac{4}{\mathcal{K}^2}W_{220}W_{311} + \frac{10}{\mathcal{K}^2}W_{222}W_{311}](\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho}) \\ \bar{I}_{222}(\vec{H} \cdot \vec{\rho})^2 &= [12u'^2\bar{u}'^2 - \frac{12}{\mathcal{K}}W_{333} - \frac{4}{\mathcal{K}}W_{331} - \frac{2}{\mathcal{K}}W_{131}\bar{u}'^2 + \frac{12}{\mathcal{K}}W_{222}u'u' + \frac{8}{\mathcal{K}}W_{220}u'u' - \frac{2}{\mathcal{K}}W_{311}u'^2 \\ &\quad + \frac{8}{\mathcal{K}^2}W_{222}W_{222} + \frac{16}{\mathcal{K}^2}W_{222}W_{220} - \frac{4}{\mathcal{K}^2}W_{311}W_{131}](\vec{H} \cdot \vec{\rho})^2 \\ \bar{I}_{220}(\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho}) &= [6u'^2\bar{u}'^2 - \frac{6}{\mathcal{K}}W_{331} + \frac{5}{\mathcal{K}}W_{131}\bar{u}'^2 - \frac{2}{\mathcal{K}}W_{222}u'u' - \frac{4}{\mathcal{K}}W_{220}u'u' + \frac{5}{\mathcal{K}}W_{311}u'^2 \\ &\quad + \frac{10}{\mathcal{K}^2}W_{311}W_{131} - \frac{8}{\mathcal{K}^2}W_{222}W_{220}](\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho}) \\ \bar{I}_{311}(\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho}) &= [12u'^3\bar{u}' - \frac{10}{\mathcal{K}}W_{242} - \frac{8}{\mathcal{K}}W_{240} + \frac{5}{\mathcal{K}}W_{222}u'^2 + \frac{2}{\mathcal{K}}W_{220}u'^2 + \frac{6}{\mathcal{K}}W_{131}u'u' \\ &\quad + \frac{4}{\mathcal{K}^2}W_{220}W_{131} + \frac{10}{\mathcal{K}^2}W_{222}W_{131}](\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho}) \\ \bar{I}_{400}(\vec{\rho} \cdot \vec{\rho})^2 &= [3u'^4 - \frac{6}{\mathcal{K}}W_{151} + \frac{3}{\mathcal{K}}W_{131}u'^2 + \frac{3}{\mathcal{K}^2}W_{131}W_{131}](\vec{\rho} \cdot \vec{\rho})^2\end{aligned}$$

3.5 COEFFICIENT RELATIONSHIP

For the case in consideration of having the field vector at the object plane and the aperture vector at the exit pupil plane, the second order relationships in Table 3.3 follow.

Table 3.3. Relationships between second-order irradiance coefficients $I_{k,l,m}^{(2)}$ and $\bar{I}_{k,l,m}^{(2)}$.

$$\begin{array}{l}
 \hline \hline
 I_{020} - \bar{I}_{200} = \frac{4}{\mathcal{K}} W_{131} \\
 \hline
 I_{111} - \bar{I}_{111} = \frac{4}{\mathcal{K}} W_{220} + \frac{6}{\mathcal{K}} W_{222} \\
 \hline
 I_{200} - \bar{I}_{020} = \frac{4}{\mathcal{K}} W_{311} \\
 \hline \hline
 \end{array}$$

Thus it is possible to determine aberration coefficients from measurements of the irradiance at the exit pupil plane and at the image plane of an optical system [4]. However, since the field vector is defined at the object plane, the image plane irradiance refers to image points $\vec{H} + \Delta\vec{H}$ and, as a consequence, measurements should be made at conjugate object-image points.

3.6 COMBINATION OF IRRADIANCE COEFFICIENTS

In the derivation of the irradiance coefficients, it has been assumed that the object has a constant or Lambertian radiance. However, it may be desirable to determine the irradiance coefficients when the object has a different emission profile. In this case, the source radiance is expanded as polynomial series of dot products of the field and aperture vectors and to the second order, for example, can be written as

$$L(\vec{H}, \vec{\rho}) = L_0 \cdot (1 + A(\vec{\rho} \cdot \vec{\rho}) + B(\vec{H} \cdot \vec{\rho}) + C(\vec{H} \cdot \vec{H})), \quad (3.15)$$

where A , B and C are coefficients describing the emission profile. We substitute Eq. 3.15 into Eq. 3.9 and Eq. 3.10 and write the irradiance coefficients to second order as

$$I(\vec{H}, \vec{\rho})^{(2)} = 1 + [-2u^2 + A](\vec{\rho} \cdot \vec{\rho}) + [-4u'u' + B](\vec{H} \cdot \vec{\rho}) + [-2\bar{u}^2 + C](\vec{H} \cdot \vec{H}). \quad (3.16)$$

Thus, second order variations in the source radiance produce second order variations in irradiance at the image plane. In addition, other higher-order irradiance terms, not presented here, result from different combinations of the source radiance terms with irradiance terms in Table 3.1a and Table 3.1b. These higher order terms can be considered as extrinsic irradiance aberrations that result from the interaction of the incoming irradiance variations and aberration in the system.

However, Eq. 3.16 shows that with respect to the irradiance of a pinhole camera $I_{pinhole}(\vec{H}, \vec{\rho})$ second order variations in object space simply add to obtain the irradiance in image space.

3.7 IRRADIANCE COEFFICIENTS AND CHOICE OF COORDINATES

The irradiance coefficients depend on the location of the field and aperture vectors. Tables 3.1 and 3.2 give the coefficients for the case of having the field vector at the object plane and the aperture vector at the exit pupil plane. This case is an important one as for a diffraction calculation knowledge of the amplitude of the field at the exit pupil is necessary; this amplitude is taken to be equal to the square root of the irradiance function $\bar{I}(\vec{H}, \vec{\rho})$. For completeness, Table 3.4 and Table 3.5 present the corresponding formulas for other cases on the location of the field and aperture vectors. The derivation of coefficients in Table 3 and Table 4 is outlined in Appendixes G and H.

Table 3.4a. Second-order irradiance coefficients $I_{k,l,m}^{(2)}$ at the image plane of an optical system with the field vector at the object plane and the aperture vector at the entrance pupil plane.

$$\begin{aligned}
 I_{020}(\vec{\rho} \cdot \vec{\rho}) &= [-2u'^2 + \frac{4}{\mathcal{K}} \bar{W}_{311}] (\vec{\rho} \cdot \vec{\rho}) \\
 I_{111}(\vec{H} \cdot \vec{\rho}) &= [-4u' \bar{u}' + \frac{4}{\mathcal{K}} \bar{W}_{220} + \frac{6}{\mathcal{K}} \bar{W}_{222}] (\vec{H} \cdot \vec{\rho}) \\
 I_{200}(\vec{H} \cdot \vec{H}) &= [-2\bar{u}'^2 + \frac{4}{\mathcal{K}} \bar{W}_{131}] (\vec{H} \cdot \vec{H})
 \end{aligned}$$

Table 3.4b. Fourth-order irradiance coefficients $I_{k,l,m}^{(4)}$ at the image plane of an optical system with the field vector at the object plane and the aperture vector at the entrance pupil plane.

$$\begin{aligned}
 I_{040}(\vec{\rho} \cdot \vec{\rho})^2 &= [3u'^4 + \frac{16}{\mathcal{K}} W_{040} u' \bar{u}' + \frac{6}{\mathcal{K}} \bar{W}_{511} - \frac{3}{\mathcal{K}} \bar{W}_{311} \cdot u'^2 \\
 &\quad - \frac{48}{\mathcal{K}^2} \bar{W}_{222} W_{040} - \frac{48}{\mathcal{K}^2} \bar{W}_{220} W_{040} + \frac{3}{\mathcal{K}^2} \bar{W}_{311} \bar{W}_{311}] (\vec{\rho} \cdot \vec{\rho})^2 \\
 I_{131}(\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho}) &= [12u'^3 \bar{u}' + \frac{16}{\mathcal{K}} W_{040} \bar{u}'^2 + \frac{12}{\mathcal{K}} W_{131} u' \bar{u}' + \frac{10}{\mathcal{K}} \bar{W}_{422} + \frac{8}{\mathcal{K}} \bar{W}_{420} - \frac{5}{\mathcal{K}} \bar{W}_{222} \cdot u'^2 \\
 &\quad - \frac{2}{\mathcal{K}} \bar{W}_{220} \cdot u'^2 - \frac{6}{\mathcal{K}} \bar{W}_{311} \cdot u' \bar{u}' - \frac{112}{\mathcal{K}^2} \bar{W}_{131} W_{040} - \frac{28}{\mathcal{K}^2} \bar{W}_{220} W_{131} \\
 &\quad - \frac{30}{\mathcal{K}^2} \bar{W}_{222} W_{131} + \frac{4}{\mathcal{K}^2} \bar{W}_{220} \bar{W}_{311} + \frac{10}{\mathcal{K}^2} \bar{W}_{222} \bar{W}_{311}] (\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho}) \\
 I_{222}(\vec{H} \cdot \vec{\rho})^2 &= [12u'^2 \bar{u}'^2 + \frac{8}{\mathcal{K}} W_{131} \bar{u}'^2 + \frac{8}{\mathcal{K}} W_{222} u' \bar{u}' + \frac{12}{\mathcal{K}} \bar{W}_{333} + \frac{4}{\mathcal{K}} \bar{W}_{331} \\
 &\quad + \frac{2}{\mathcal{K}} \bar{W}_{131} \cdot u'^2 - \frac{12}{\mathcal{K}} \bar{W}_{222} \cdot u \bar{u}' - \frac{8}{\mathcal{K}} \bar{W}_{220} \cdot u \bar{u}' + \frac{2}{\mathcal{K}} \bar{W}_{311} \cdot \bar{u}'^2 \\
 &\quad - \frac{48}{\mathcal{K}^2} \bar{W}_{131} W_{131} - \frac{16}{\mathcal{K}^2} \bar{W}_{222} W_{222} - \frac{8}{\mathcal{K}^2} \bar{W}_{220} W_{222} - \frac{64}{\mathcal{K}^2} \bar{W}_{040} W_{040} \\
 &\quad + \frac{8}{\mathcal{K}^2} \bar{W}_{222} \bar{W}_{222} + \frac{16}{\mathcal{K}^2} \bar{W}_{222} \bar{W}_{220} - \frac{4}{\mathcal{K}^2} \bar{W}_{311} \bar{W}_{131}] (\vec{H} \cdot \vec{\rho})^2 \\
 I_{220}(\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho}) &= [6u'^2 \bar{u}'^2 + \frac{4}{\mathcal{K}} W_{131} \bar{u}'^2 + \frac{8}{\mathcal{K}} W_{220} u' \bar{u}' + \frac{6}{\mathcal{K}} \bar{W}_{331} \\
 &\quad + \frac{2}{\mathcal{K}} \bar{W}_{222} \cdot u' \bar{u}' + \frac{4}{\mathcal{K}} \bar{W}_{220} \cdot u' \bar{u}' - \frac{5}{\mathcal{K}} \bar{W}_{311} \cdot \bar{u}'^2 - \frac{5}{\mathcal{K}} \bar{W}_{131} \cdot u'^2
 \end{aligned}$$

$$\begin{aligned}
& -\frac{96}{\mathcal{K}^2} \bar{W}_{040} W_{040} - \frac{16}{\mathcal{K}^2} \bar{W}_{131} W_{131} - \frac{16}{\mathcal{K}^2} \bar{W}_{222} W_{222} - \frac{16}{\mathcal{K}^2} \bar{W}_{220} W_{220} - \frac{4}{\mathcal{K}^2} \bar{W}_{220} W_{222} \\
& + \frac{10}{\mathcal{K}^2} \bar{W}_{131} \bar{W}_{311} - \frac{8}{\mathcal{K}^2} \bar{W}_{220} \bar{W}_{222}] (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho}) \\
\hline
I_{311} (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho}) &= [12u' \bar{u}'^3 + \frac{8}{\mathcal{K}} W_{222} \bar{u}'^2 + \frac{8}{\mathcal{K}} W_{220} \bar{u}'^2 + \frac{4}{\mathcal{K}} W_{311} u' \bar{u}' + \frac{10}{\mathcal{K}} \bar{W}_{242} + \frac{8}{\mathcal{K}} \bar{W}_{240} \\
& - \frac{5}{\mathcal{K}} \bar{W}_{222} \cdot \bar{u}'^2 - \frac{2}{\mathcal{K}} \bar{W}_{220} \cdot \bar{u}'^2 - \frac{6}{\mathcal{K}} \bar{W}_{131} \cdot u' \bar{u}' \\
& - \frac{80}{\mathcal{K}^2} \bar{W}_{040} W_{131} - \frac{28}{\mathcal{K}^2} \bar{W}_{131} W_{222} - \frac{32}{\mathcal{K}^2} \bar{W}_{131} W_{220} - \frac{6}{\mathcal{K}^2} \bar{W}_{222} W_{311} - \frac{4}{\mathcal{K}^2} \bar{W}_{220} W_{311} \\
& + \frac{4}{\mathcal{K}^2} \bar{W}_{220} \bar{W}_{131} + \frac{10}{\mathcal{K}^2} \bar{W}_{222} \bar{W}_{131}] (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho}) \\
\hline
I_{400} (\vec{H} \cdot \vec{H})^2 &= [3\bar{u}'^4 + \frac{4}{\mathcal{K}} W_{311} \bar{u}'^2 + \frac{6}{\mathcal{K}} \bar{W}_{151} - \frac{3}{\mathcal{K}} \bar{W}_{131} \cdot \bar{u}'^2 \dots \\
& - \frac{32}{\mathcal{K}^2} \bar{W}_{040} W_{220} - \frac{24}{\mathcal{K}^2} \bar{W}_{040} W_{222} - \frac{8}{\mathcal{K}^2} \bar{W}_{131} W_{311} + \frac{3}{\mathcal{K}^2} \bar{W}_{131} \bar{W}_{131}] (\vec{H} \cdot \vec{H})^2 \\
\hline
\end{aligned}$$

Table 3.5a. Second-order irradiance coefficients $\bar{I}_{k,l,m}^{(2)}$ at the exit pupil plane of an optical system with the field vector at the object plane and the aperture vector at the entrance pupil plane.

$$\begin{aligned}
\bar{I}_{020} (\vec{H} \cdot \vec{H}) &= [-2\bar{u}'^2 - \frac{4}{\mathcal{K}} W_{311}] (\vec{H} \cdot \vec{H}) \\
\bar{I}_{111} (\vec{H} \cdot \vec{\rho}) &= [-4u' \bar{u}' - \frac{4}{\mathcal{K}} W_{220} - \frac{6}{\mathcal{K}} W_{222}] (\vec{H} \cdot \vec{\rho}) \\
\bar{I}_{200} (\vec{\rho} \cdot \vec{\rho}) &= [-2u'^2 - \frac{4}{\mathcal{K}} W_{131}] (\vec{\rho} \cdot \vec{\rho}) \\
\hline
\end{aligned}$$

Table 3.5b. Fourth-order irradiance coefficients $\bar{I}_{k,l,m}^{(4)}$ at the exit pupil plane of an optical system with the field vector at the object plane and the aperture vector at the entrance pupil plane.

$$\bar{I}_{040} (\vec{H} \cdot \vec{H})^2 = [3\bar{u}'^4 - \frac{6}{\mathcal{K}} W_{511} + \frac{3}{\mathcal{K}} W_{311} \bar{u}'^2 - \frac{16}{\mathcal{K}} \bar{W}_{040} u' \bar{u}']$$

$$-\frac{48}{\mathcal{K}^2} \cdot W_{222} \bar{W}_{040} - \frac{48}{\mathcal{K}^2} \cdot W_{220} \bar{W}_{040} + \frac{3}{\mathcal{K}^2} W_{311} W_{311}] (\bar{H} \cdot \bar{H})^2$$

$$\begin{aligned} \bar{I}_{131} (\bar{H} \cdot \bar{H}) (\bar{H} \cdot \bar{\rho}) = & [12u' \bar{u}'^3 - \frac{10}{\mathcal{K}} W_{422} - \frac{8}{\mathcal{K}} W_{420} + \frac{5}{\mathcal{K}} W_{222} \bar{u}'^2 + \frac{2}{\mathcal{K}} W_{220} \bar{u}'^2 + \frac{6}{\mathcal{K}} W_{311} u' \bar{u}' \\ & - \frac{16}{\mathcal{K}} \bar{W}_{040} u'^2 - \frac{12}{\mathcal{K}} \bar{W}_{131} u' \bar{u}' - \frac{112}{\mathcal{K}^2} W_{131} \bar{W}_{040} - \frac{28}{\mathcal{K}^2} W_{220} \bar{W}_{131} - \frac{30}{\mathcal{K}^2} W_{222} \bar{W}_{131} \\ & + \frac{4}{\mathcal{K}^2} W_{220} W_{311} + \frac{10}{\mathcal{K}^2} W_{222} W_{311}] (\bar{H} \cdot \bar{H}) (\bar{H} \cdot \bar{\rho}) \end{aligned}$$

$$\begin{aligned} \bar{I}_{222} (\bar{H} \cdot \bar{\rho})^2 = & [12u'^2 \bar{u}'^2 - \frac{12}{\mathcal{K}} W_{333} - \frac{4}{\mathcal{K}} W_{331} - \frac{2}{\mathcal{K}} W_{131} \bar{u}'^2 + \frac{12}{\mathcal{K}} W_{222} u' \bar{u}' + \frac{8}{\mathcal{K}} W_{220} u' \bar{u}' - \frac{2}{\mathcal{K}} W_{311} u'^2 \\ & - \frac{8}{\mathcal{K}} \bar{W}_{131} u'^2 - \frac{8}{\mathcal{K}} \bar{W}_{222} u' \bar{u}' - \frac{48}{\mathcal{K}^2} W_{131} \bar{W}_{131} - \frac{16}{\mathcal{K}^2} W_{222} \bar{W}_{222} - \frac{8}{\mathcal{K}^2} W_{220} \bar{W}_{222} - \frac{64}{\mathcal{K}^2} W_{040} \bar{W}_{040} \\ & + \frac{8}{\mathcal{K}^2} W_{222} W_{222} + \frac{16}{\mathcal{K}^2} W_{222} W_{220} - \frac{4}{\mathcal{K}^2} W_{311} W_{131}] (\bar{H} \cdot \bar{\rho})^2 \end{aligned}$$

$$\begin{aligned} \bar{I}_{220} (\bar{H} \cdot \bar{H}) (\bar{\rho} \cdot \bar{\rho}) = & [6u'^2 \bar{u}'^2 - \frac{6}{\mathcal{K}} W_{331} + \frac{5}{\mathcal{K}} W_{131} \bar{u}'^2 - \frac{2}{\mathcal{K}} W_{222} u' \bar{u}' - \frac{4}{\mathcal{K}} W_{220} u' \bar{u}' + \frac{5}{\mathcal{K}} W_{311} u'^2 \\ & - \frac{4}{\mathcal{K}} \bar{W}_{131} u'^2 - \frac{8}{\mathcal{K}} \bar{W}_{220} u' \bar{u}' - \frac{96}{\mathcal{K}^2} W_{040} \bar{W}_{040} - \frac{16}{\mathcal{K}^2} W_{131} \bar{W}_{131} - \frac{16}{\mathcal{K}^2} W_{222} \bar{W}_{220} - \frac{16}{\mathcal{K}^2} W_{220} \bar{W}_{220} - \frac{4}{\mathcal{K}^2} W_{220} \bar{W}_{222} \\ & + \frac{10}{\mathcal{K}^2} W_{311} W_{131} - \frac{8}{\mathcal{K}^2} W_{222} W_{220}] (\bar{H} \cdot \bar{H}) (\bar{\rho} \cdot \bar{\rho}) \end{aligned}$$

$$\begin{aligned} \bar{I}_{311} (\bar{H} \cdot \bar{\rho}) (\bar{\rho} \cdot \bar{\rho}) = & [12u'^3 \bar{u}' - \frac{10}{\mathcal{K}} W_{242} - \frac{8}{\mathcal{K}} W_{240} + \frac{5}{\mathcal{K}} W_{222} u'^2 + \frac{2}{\mathcal{K}} W_{220} u'^2 + \frac{6}{\mathcal{K}} W_{131} u' \bar{u}' \\ & - \frac{8}{\mathcal{K}} \bar{W}_{222} u'^2 - \frac{8}{\mathcal{K}} \bar{W}_{220} u'^2 - \frac{4}{\mathcal{K}} \bar{W}_{311} u' \bar{u}' \\ & - \frac{80}{\mathcal{K}^2} W_{040} \bar{W}_{131} - \frac{28}{\mathcal{K}^2} W_{131} \bar{W}_{222} - \frac{32}{\mathcal{K}^2} W_{131} \bar{W}_{220} - \frac{6}{\mathcal{K}^2} W_{222} \bar{W}_{311} - \frac{4}{\mathcal{K}^2} W_{220} \bar{W}_{311} \\ & + \frac{4}{\mathcal{K}^2} W_{220} W_{131} + \frac{10}{\mathcal{K}^2} W_{222} W_{131}] (\bar{H} \cdot \bar{\rho}) (\bar{\rho} \cdot \bar{\rho}) \end{aligned}$$

$$\begin{aligned} \bar{I}_{400} (\bar{\rho} \cdot \bar{\rho})^2 = & [3u'^4 - \frac{6}{\mathcal{K}} W_{151} + \frac{3}{\mathcal{K}} W_{131} u'^2 - \frac{4}{\mathcal{K}} \bar{W}_{311} u'^2 \\ & - \frac{32}{\mathcal{K}^2} W_{040} \bar{W}_{220} - \frac{24}{\mathcal{K}^2} W_{040} \bar{W}_{222} - \frac{8}{\mathcal{K}^2} W_{131} \bar{W}_{311} + \frac{3}{\mathcal{K}^2} W_{131} W_{131}] (\bar{\rho} \cdot \bar{\rho})^2 \end{aligned}$$

3.8 COEFFICIENTS VERIFICATION

In order to support the analytical derivation, the magnitude of the irradiance coefficients was determined both through the formulas derived and numerically. A macro program was written to calculate the irradiance coefficients by making an iterative fit to a selected set of

irradiance values across the aperture and field of an optical system. The iterative algorithm is similar to one used by Sasian to fit aberration coefficients [5].

For example, for an optical system with the stop aperture at the entrance pupil, the routine in Table 3.6 was executed to find the magnitude of the irradiance coefficients $I_{020}(\vec{\rho} \cdot \vec{\rho})$ and $I_{040}(\vec{\rho} \cdot \vec{\rho})^2$. According to Eq. 3.7, the normalized irradiance at the specified field and aperture points is given by $\frac{dS'}{dS} \cdot \cos^4(\theta'(\vec{H}, \vec{\rho}))$. The quantity $\frac{dS'}{dS} \cdot \cos^4(\theta'(\vec{H}, \vec{\rho}))$ was computed in a lens design program by defining a small circle at the entrance pupil and tracing real rays to calculate the area of the corresponding ellipse at the exit pupil.

After a few iterations of the loop in Table 3.6, the coefficients $I_{020}(\vec{\rho} \cdot \vec{\rho})$ and $I_{040}(\vec{\rho} \cdot \vec{\rho})^2$ converged to the theoretical values with insignificant error. A similar approach was applied to validate the remaining irradiance coefficients. The iterative fit methodology was used to test the coefficients values at several conjugate distances and aperture stop positions for both single surface and a system of several surfaces. The obvious agreement of the formulas with the coefficients found with the iterative fit supports the validity of the theory.

As an example, the irradiance coefficients for a Landscape lens [6] shown in Fig. 3.4 were calculated both ways. The lens operates at $f/8$ and the field of view (FOV) is limited to 30° . Table 3.7 presents a comparison of coefficients where the differences in the eighth decimal place are likely due to numerical computation errors.

Table 3.6 Iterative algorithm that was used to fit irradiance coefficients.

FOR $i = 1$ to 100

$\rho = 0.2$

$$I_{real} = \frac{dS}{dS'} \cos^4(\theta) \{0, \rho\}$$

$$I_{200} = (I_{real}^{-1} - I_{400}\rho^4 - I_{600}\rho^6 - I_{800}\rho^8 - I_{1000}\rho^{10}) \cdot \rho^{-2}$$

$\rho = 0.4$

$$Ri = \frac{dS}{dS'} \cos^4(\theta) \{0, \rho\}$$

$$I_{400} = (I_{real}^{-1} - I_{200}\rho^2 - I_{600}\rho^6 - I_{800}\rho^8 - I_{1000}\rho^{10}) \cdot \rho^{-4}$$

$\rho = 0.6$

$$Ri = \frac{dS}{dS'} \cos^4(\theta) \{0, \rho\}$$

$$I_{600} = (I_{real}^{-1} - I_{200}\rho^2 - I_{400}\rho^4 - I_{800}\rho^8 - I_{1000}\rho^{10}) \cdot \rho^{-6}$$

$\rho = 0.8$

$$Ri = \frac{dS}{dS'} \cos^4(\theta) \{0, \rho\}$$

$$I_{800} = (I_{real}^{-1} - I_{200}\rho^2 - I_{400}\rho^4 - I_{600}\rho^6 - I_{1000}\rho^{10}) \cdot \rho^{-8}$$

$\rho = 1$

$$Ri = \frac{dS}{dS'} \cos^4(\theta) \{0, \rho\}$$

$$I_{1000} = (I_{real}^{-1} - I_{200}\rho^2 - I_{400}\rho^4 - I_{600}\rho^6 - I_{800}\rho^8) \cdot \rho^{-10}$$

NEXT

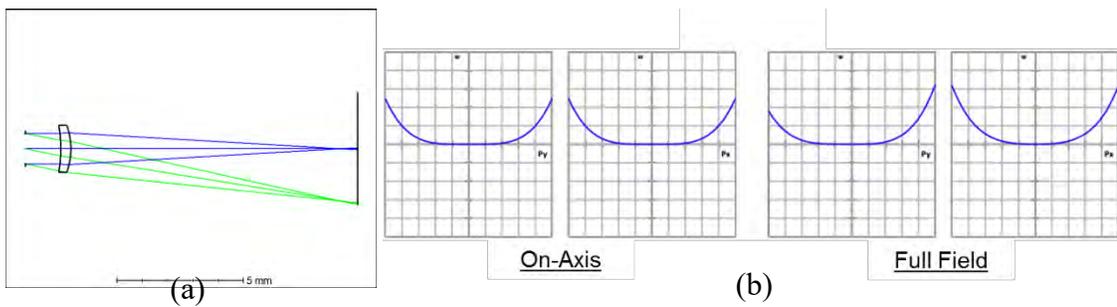


Fig. 3.4 (a) – Layout of the Landscape lens used to compare irradiance coefficients computed analytically and numerically. (b) – the optical path difference plots for on-axis and full-field positions (plot scale is two waves). In order to minimize fitting errors, the FOV of the lens is limited to 30°.

Table 3.7 Comparison of irradiance coefficients computed analytically and numerically. The agreement to eighth digits supports the correctness of the formulas.

Irradiance coefficient $I_{k,l,m}$	Analytical formula	Numerical calculation
$I_{020}(\vec{\rho} \cdot \vec{\rho})$	-0.0051310721	-0.0051310695
$I_{111}(\vec{H} \cdot \vec{\rho})$	0.0051413131	0.0051413122
$I_{200}(\vec{H} \cdot \vec{H})$	-0.0563954333	-0.0563954267
$I_{040}(\vec{\rho} \cdot \vec{\rho})^2$	-0.0002767939	-0.0002768392
$I_{131}(\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho})$	0.0003447530	0.0003447656
$I_{222}(\vec{H} \cdot \vec{\rho})^2$	0.0003958170	0.0003958224
$I_{220}(\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho})$	0.0004320213	
$I_{311}(\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho})$	-0.0001299094	-0.0001299152
$I_{400}(\vec{H} \cdot \vec{H})^2$	0.0028603412	0.0028602385

3.9 RELATIVE ILLUMINATION

A practical and important case is the relative illumination $RI(\vec{H})$ at the image plane of an optical system, which can be written to fourth order as

$$RI(\vec{H}) = 1 + I_{200}(\vec{H} \cdot \vec{H}) + I_{400}(\vec{H} \cdot \vec{H})^2. \quad (3.16)$$

This is given in the limit of a small aperture by the terms $I_{200}(\vec{H} \cdot \vec{H})$ and $I_{400}(\vec{H} \cdot \vec{H})^2$ of the irradiance function $I(\vec{H}, \vec{\rho})$. It has been shown in previous sections that the irradiance coefficients depend on the location of the field and aperture vectors. The theory of irradiance changes in axially symmetric optical systems provides a closed form solution that approximates to fourth order the image illumination fall-off and accounts for the illumination effects of aberrations in an unvignetted optical system. In this section, the role of individual aberration terms in the relative illumination is discussed and the results are validated with real ray tracing.

3.9.1 Historical Background

The distribution of light at the focal plane of an optical system is an important aspect of a lens performance. Joseph Petzval was well aware of the image plane illumination and back in 1858 stated, "A third quality of the new combination of lenses is the equal strength of light from the center to the utmost corners of a surface of the image [7]..." The term relative illumination was used as early as the 1900's, when photographers were concerned with the non-uniform illumination in different regions of the negative [8, 9]. In the 1950s, the increasing demand for wide-angle lenses used for aerial photography and other purposes renewed interest in the subject. As shown in Fig. 3.5, it was found that the decrease in relative illumination of wide-angle lenses limited the angle of view of the early objectives to approximately 95°-105°. The interest in creating accurate topographic maps

led to the development of new lens design solutions capable of producing usable imagery over a greater angular field of view [10].

Relative illumination is defined as the ratio of the irradiance on the focal plane at off-axis field positions to the irradiance at the center of the field. The relationship between the illumination and the field angle is derived from basic radiometric principles. For many practical purposes, the \cos^4 -law has been considered a reasonable approximation to the illumination fall-off produced by a lens at its focal plane; however, assumptions involved in the derivation of the \cos^4 -law are often found to be incorrectly interpreted in the optical literature [2].



Fig. 3.5 Stadium at the University of Arizona simulated through an early wide-angle lens (US 2031792). Sharp decrease in illumination toward the corners of the image limits the field-of-view of the objective.

The \cos^4 -law states that the irradiance of different parts of the image formed by the optical system vary as the fourth power of the cosine of the chief ray angle in object space. As some topographic lenses were found to not follow the \cos^4 -law the difference was referred to as the ‘lens effect’ [10]. Reiss showed that the \cos^4 -law is precisely followed only by an optical system corrected for all image aberrations with an aperture stop that

precedes the lens and the object at infinite distance [11]. Other authors have emphasized the importance of considering the effect of pupil aberrations and stop position on the illumination. It was found that it is possible to design a lens system in such a way that the decrease in illumination towards the edge of the field occurs more gradually than the \cos^4 -law [12-14].

In 1946, Roossinov patented a wide-angle lens that instead of the \cos^4 -law realized a \cos^3 -law. The area of the entrance aperture of this objective increases towards the edge of the field of view. This phenomenon is achieved by maximum divergence from Abbe's conditions of sines for the front half of the objective, with the object located in the plane of the diaphragm [15].

In 1986, Rimmer outlined a method to accurately calculate relative illumination by real ray tracing. His method is based upon the theory developed by Hopkins and requires determining the size of the exit pupil in the direction cosine space [16, 17]. Similar computation methods of relative illumination are used in modern lens design software [18]. Although it is well known that the relative illumination in an optical system is a function of many variables, there is still confusion regarding the specific role of distortion, pupil aberrations and aperture stop position on the distribution of light at the focal plane.

3.9.2 Relative Illumination Coefficient Summary

Table 3.8 presents the second I_{200} and fourth order I_{400} coefficients of the relative illumination function $RI(\vec{H})$ for several cases in the location of the aperture and field vectors. When the stop is between system components, the system is then divided into part

A preceding the stop and part B following the stop, and aberrations of each part are then used to define the irradiance coefficients. Note that both image and pupil aberration coefficients are used (see Appendix I) [19].

Examination of the functional form of the coefficients allows one to make several points of interest. First, to second order pupil coma influences the relative illumination and can nullify the coefficient I_{200} in some cases. This is known as Slyusarev effect. Alternatively, the relationship between image and pupil aberrations

$$\bar{W}_{131} = W_{311} - \frac{\lambda\mathcal{K}}{2} \Delta\{\bar{u}^2\}, \quad (3.17)$$

can be used to determine how image distortion W_{311} influences relative illumination [20].

Second, if the system is telecentric in image space $\bar{u}' = 0$ and is also aplanatic $\bar{W}_{131} = 0$, $\bar{W}_{151} = 0$, and $\bar{W}_{040} = 0$ then the relative illumination is uniform to fourth order. Such a system would have a $f \sin(\theta)$ image height mapping [21]. Other cases for uniform illumination are also possible that require $-2\bar{u}^2 + \frac{4}{\lambda\mathcal{K}} \bar{W}_{131} = 0$.

Table 3.8. Second and fourth-order irradiance coefficients describing relative illumination.

Case	I_{200}	I_{400}
Pinhole camera.	$-2 \cdot \bar{u}^2$	$3 \cdot \bar{u}^4$
Field vector at the image plane.	$-2 \cdot \bar{u}^2$	$3 \cdot \bar{u}^4$

Stop and aperture vector at the exit pupil plane.		
Field vector at the object plane. Stop and aperture vector at the exit pupil plane.	$-2 \cdot \bar{u}^{\prime 2}$	$3 \cdot \bar{u}^{\prime 4} + \frac{4}{\mathcal{K}} W_{311} \cdot \bar{u}^{\prime 2}$
Field vector at the image plane. Stop and aperture vector at the entrance pupil plane.	$-2 \cdot \bar{u}^{\prime 2} + \frac{4}{\mathcal{K}} \bar{W}_{131}$	$3 \cdot \bar{u}^{\prime 4} + \frac{6}{\mathcal{K}} \bar{W}_{151} - \frac{3}{\mathcal{K}} \bar{W}_{131} \bar{u}^{\prime 2} + \frac{3}{\mathcal{K}^2} \bar{W}_{131} \bar{W}_{131}$ $-\frac{1}{\mathcal{K}^2} [32 \bar{W}_{040} W_{220} + 24 \bar{W}_{040} W_{222}]$
Field vector at the object plane. Stop and aperture vector at the entrance pupil plane.	$-2 \cdot \bar{u}^{\prime 2} + \frac{4}{\mathcal{K}} \bar{W}_{131}$	$3 \cdot \bar{u}^{\prime 4} + \frac{4}{\mathcal{K}} W_{311} \bar{u}^{\prime 2} + \frac{6}{\mathcal{K}} \bar{W}_{151} - \frac{3}{\mathcal{K}} \bar{W}_{131} \bar{u}^{\prime 2} + \frac{3}{\mathcal{K}^2} \bar{W}_{131} \bar{W}_{131}$ $-\frac{1}{\mathcal{K}^2} [32 \bar{W}_{040} W_{220} + 24 \bar{W}_{040} W_{222} + 8 \bar{W}_{131} W_{311}]$
Stop and aperture vector at plane between components of the lens. Field vector at the image plane.	$-2 \cdot \bar{u}^{\prime 2} + \frac{4}{\mathcal{K}} \bar{W}_{131B}$	$3 \cdot \bar{u}^{\prime 4} + \frac{6}{\mathcal{K}} \bar{W}_{151B} - \frac{3}{\mathcal{K}} \bar{W}_{131B} \bar{u}^{\prime 2} + \frac{3}{\mathcal{K}^2} \bar{W}_{131B} \bar{W}_{131B}$ $-\frac{1}{\mathcal{K}^2} [32 \bar{W}_{040B} W_{220B} + 24 \bar{W}_{040B} W_{222B}]$
Stop and aperture vector at plane between components of the lens. Field vector at the object plane.	$-2 \cdot \bar{u}^{\prime 2} + \frac{4}{\mathcal{K}} \bar{W}_{131B}$	$3 \cdot \bar{u}^{\prime 4} + \frac{4}{\mathcal{K}} W_{311A} \bar{u}^{\prime 2} + \frac{4}{\mathcal{K}} W_{311B} \bar{u}^{\prime 2} + \frac{6}{\mathcal{K}} \bar{W}_{151B} - \frac{3}{\mathcal{K}} \bar{W}_{131B} \bar{u}^{\prime 2}$ $-\frac{1}{\mathcal{K}^2} [32 \bar{W}_{040B} W_{220B} + 24 \bar{W}_{040B} W_{222B} + 8 \bar{W}_{131B} W_{311B}]$ $+\frac{3}{\mathcal{K}^2} \bar{W}_{131B} \bar{W}_{131B} - \frac{8}{\mathcal{K}^2} \bar{W}_{131B} W_{311A}$

Third, if the system is telecentric in object space $\bar{u} = 0$, and if the Herschel condition of the pupils is satisfied

$$\bar{W}_{131} = \frac{\mathcal{K}}{8} \Delta(\bar{u}^2), \quad (3.18)$$

then the relative illumination would follow to second order a $\cos^3(\theta')$ rule in a system where the aperture vector is defined at the entrance pupil plane or at the intermediate pupil. This follows since the cosine-to-the-third-power of the angle in image space is given by

$$\cos^3(\theta') = 1 - \frac{3}{2}\bar{u}'^2 + \dots \quad (3.19)$$

Fourth, contrary to intuition pupil spherical aberration \bar{W}_{040} may not influence the relative illumination when the aperture vector is at the entrance pupil plane and $W_{220} = W_{222} = 0$, though it may change the chief ray angle of incidence at the image plane.

3.9.3 Example: Mobile Phone Camera Lens

Miniature cameras for consumer electronics and mobile phones are a rapidly growing technology. The system level requirements such as manufacturing cost, packaging, and sensor characteristics impose unique challenges for optical designers. The relative illumination is one of the interesting characteristics of miniature camera lens designs.

The typical distortion requirement in mobile camera lenses is <2% and the FOV is large (common FOV values are 65 - 75°). On the other hand, the chief ray incidence angle (CRA) on the sensor is usually limited to no more than 30°. The CRA impacts the relative illumination, which often is set to 50% at the sensor corners. For better CRA control, the aperture stop in a conventional mobile lens is placed close to the front, away from the image plane. The aperture stop position and the strong aspheric next to the image plane generate exit pupil spherical aberration, which reduces the CRA.

Mobile lenses are well known for the extensive use of aspheric surfaces. The interaction of multiple aspherics within the design enables a high level of control over aberrations. Particularly, sharp imaging on a flat surface can be achieved without satisfying the classical requirement of having a Petzval sum nearly zero. The aspheric optical elements located close to the image plane contribute higher-order field curvature and astigmatism. Different orders of the field curvature and astigmatism are balanced to compensate for any residual Petzval curvature [22].

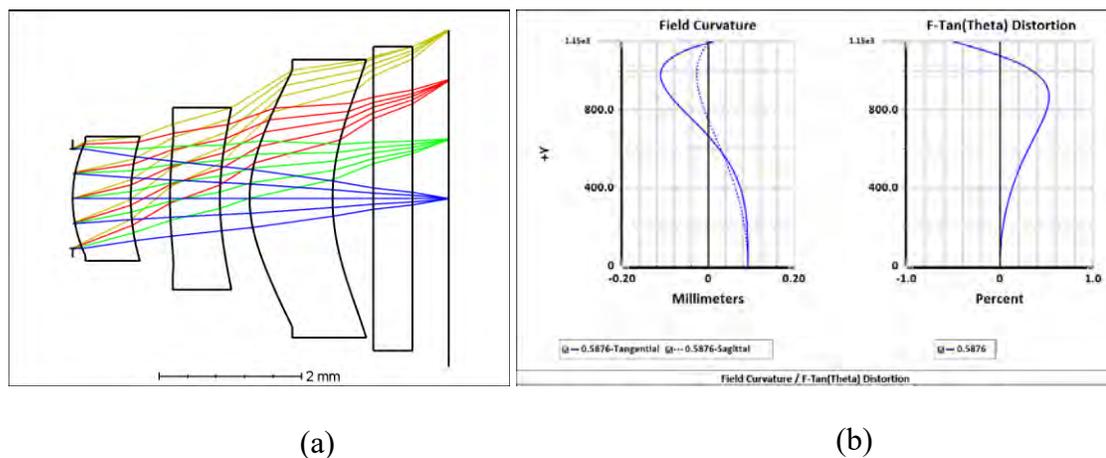


Fig. 3.6 U.S. Patent 6441971: (a) Layout, (b) field curvature and distortion plot. Distortion correction requirements do not allow one to significantly reduce the chief ray incidence angle on the sensor (CRA). Exit pupil spherical aberration shifts the pupil radially and reduces the CRA. Higher order field curvature and astigmatism are balanced to compensate for the residual Petzval field curvature.

To illustrate how aberrations affect the relative illumination of a mobile camera lens, different terms in the close-from solution are evaluated for one of the early miniature digital camera patents [23]. The three-element design shown in Fig. 3.6a covers a FOV of 64° at $f/2.8$. Field curvature and distortion plots are shown in Fig. 3.6b. A small positive fourth-order distortion is compensated with a higher-order negative distortion resulting in $<0.5\%$

of total distortion. The Petzval radius is about two times the focal length; this clearly indicates that the residual field curvature is compensated by balancing higher orders of the field curvature and astigmatism. Aberration contributions to the relative illumination, sorted in descending order, are summarized in Table 3.9.

Table 3.9 Contributions to the relative illumination for US. 6441971.

2nd order coefficient (I_{200})	Numerical value	4th order coefficient (I_{400})	Numerical value
$-2\bar{u}^2$	-0.745977	$3\bar{u}^4$	0.417361
$\frac{4}{\mathcal{K}}\bar{W}_{131}$	0.034043	$\frac{6}{\mathcal{K}}\bar{W}_{151}$	-0.055200
		$-\frac{24}{\mathcal{K}^2}\bar{W}_{040}W_{222}$	0.025298
		$\frac{4}{\mathcal{K}}W_{311}\bar{u}^2$	0.018905
		$-\frac{32}{\mathcal{K}^2}\bar{W}_{040}W_{220}$	-0.016739
		$-\frac{3}{\mathcal{K}}\bar{W}_{131}\bar{u}^2$	-0.009523
		$-\frac{8}{\mathcal{K}^2}\bar{W}_{131}W_{311}$	0.000863
		$\frac{3}{\mathcal{K}^2}\bar{W}_{131}\bar{W}_{131}$	0.000217
Total	-0.711934		0.343373

As expected, the major contribution to the deviation from the \cos^4 -law comes from the pupil coma and distortion. However, since the distortion is small and pupil coma is proportional to the distortion, the ability to manipulate the relative illumination through these terms is limited. The products of residual field curvature, astigmatism, and pupil spherical aberration contribute significantly to the change in relative illumination of a

mobile lens to fourth-order and may provide an effective degree of freedom during the lens design process. The relative illumination curves calculated with real ray tracing and using the analytical approximation are presented in Fig. 3.7. The plot shows excellent agreement between real ray tracing and the analytical solution over the entire FOV, confirming that the assumptions used in the derivation are still valid for relatively fast lenses with relatively large FOVs.

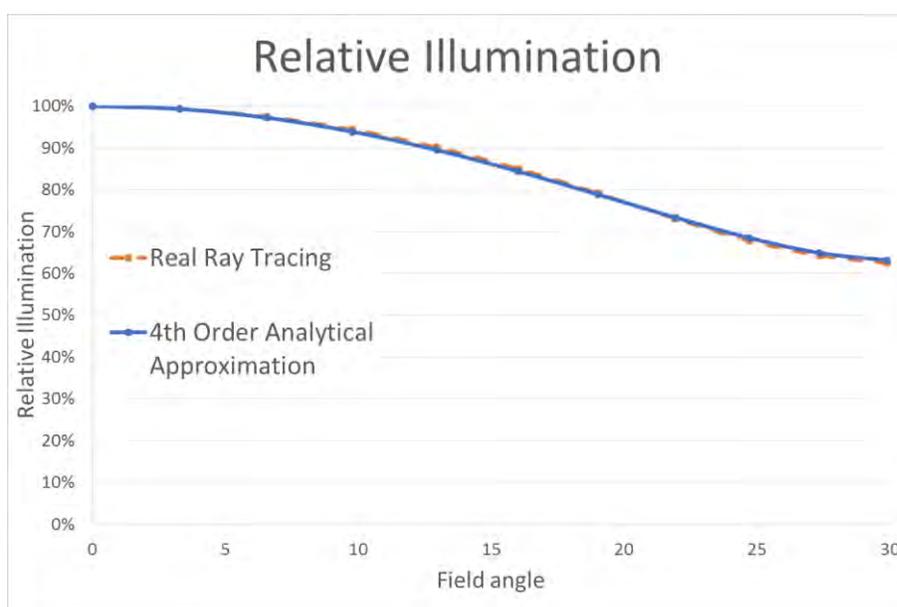


Fig. 3.7 Relative illumination plot for U.S. Patent 6441971. Analytical equation and real ray tracing calculation show excellent agreement over the entire FOV.

3.9.4 Example: Wide-Angle Lens

The field of view of a wide angle objective that follows the \cos^4 -law is limited to approximately 95° - 105° due to a sharp decrease in illumination at the margin of the image. In US. Patent 2516724, shown in Fig. 3.8a, the decrease in the illumination follows the \cos^3 -law and this makes it possible the widening of the angle of view up to 120° or more.

In this design the area of the entrance pupil increases towards the edge of the field of view approximately by the factor of two, and as the result, the area of oblique beams entering the objective differs negligibly from the area of the axial beams. The author of the patent explains this phenomenon by maximum divergence from Abbe's conditions of sines for the front half of the objective with an object located in the plane of the diaphragm, or in other words, by maximizing pupil coma for the front half of the objective as shown in Fig. 3.8b [15].

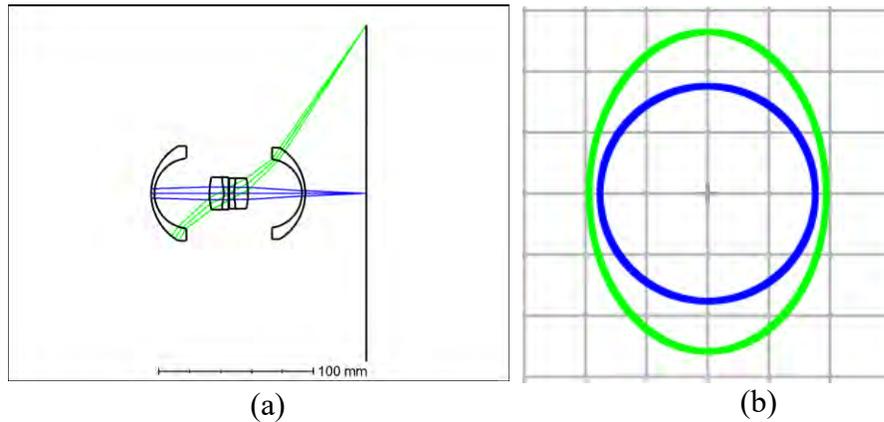


Fig. 3.8 U.S. 2516724: (a) Layout, (b) beam footprints at the exit pupil. The wide-angle objective consists of two nearly symmetrical halves. The form of the greatly curved exterior meniscus lens elements is of great importance. These meniscus lenses contribute substantial pupil coma, which increases the pupil area for oblique beams.

Aberration theory allows quantifying the required amount of pupil coma to achieve the \cos^3 -law decrease in illumination. Similarly to Eq. 3.8, the cosine-to-the-third-power of the angle in the object space is given by

$$\cos^3(\theta) \approx \left[1 - \frac{3}{2} \cdot \bar{u}^2 (\vec{H} \cdot \vec{H}) + \frac{15}{8} \cdot \bar{u}^4 (\vec{H} \cdot \vec{H})^2 \right]. \quad (3.20)$$

US. Patent 2516724 consists of two more or less symmetrical halves. If two exactly symmetrical halves are assumed, the distortion of parts A and B has the same magnitude but opposite sign. Moreover, each part of the symmetrical optical system is separately corrected for field curvature and astigmatism. A simplified expression for the relative illumination of a symmetrical optical system follows as

$$\begin{aligned}
 Ri_{EP}(\vec{H}) \approx & 1 + [-2\bar{u}'^2 + \frac{4}{\lambda\mathcal{K}}\bar{W}_{131B}](\vec{H} \cdot \vec{H}) \dots \\
 & + [3\bar{u}'^4 + \frac{6}{\lambda\mathcal{K}}\bar{W}_{151B} - \frac{3}{\lambda\mathcal{K}}\bar{W}_{131B}\bar{u}'^2 + \frac{3}{\lambda\mathcal{K}^2}\bar{W}_{131B}\bar{W}_{131B}](\vec{H} \cdot \vec{H})^2.
 \end{aligned} \tag{3.21}$$

In a symmetrical lens the paraxial chief ray slope in object and image spaces are equal. The value of fourth- and six-order pupil coma are calculated by setting Eq. 3.21 equal to Eq. 3.20 and solving the following system of equations

$$\begin{aligned}
 -2\bar{u}'^2 + \frac{4}{\lambda\mathcal{K}}\bar{W}_{131B} &= -\frac{3}{2}\bar{u}'^2, \\
 3\bar{u}'^4 + \frac{6}{\lambda\mathcal{K}}\bar{W}_{151B} - \frac{3}{\lambda\mathcal{K}}\bar{W}_{131B}\bar{u}'^2 + \frac{3}{\lambda\mathcal{K}^2}\bar{W}_{131B}\bar{W}_{131B} &= \frac{15}{8}\bar{u}'^4, \\
 \bar{u}' &= \bar{u}.
 \end{aligned} \tag{3.22}$$

The term $\frac{3}{\lambda\mathcal{K}^2}\bar{W}_{131B}\bar{W}_{131B}$ is small and can be neglected. If the image pupil relationship in Eq. 3.20 and the fact that distortion in a symmetrical optical system is zero are considered, the required amount of pupil coma is estimated to be

$$\bar{W}_{131B} = \frac{\lambda\mathcal{K}}{8}\bar{u}'^2, \tag{3.23}$$

$$\bar{W}_{151B} = -\frac{\lambda\mathcal{K}}{8}\bar{u}'^4. \tag{3.24}$$

Fig. 3.9 shows relative illumination curves calculated with real ray tracing alongside the analytical solution. Since the analytical solution provides a fourth-order approximation to relative illumination, there is excellent agreement with real ray tracing calculation only up to about 20° half field angle. For larger field angles, the six- and higher-order relative illumination terms, not presented here, contribute significantly to the illumination of the lens. The fourth-order approximation is not accurate enough to describe the relative illumination beyond 20° half field angle.

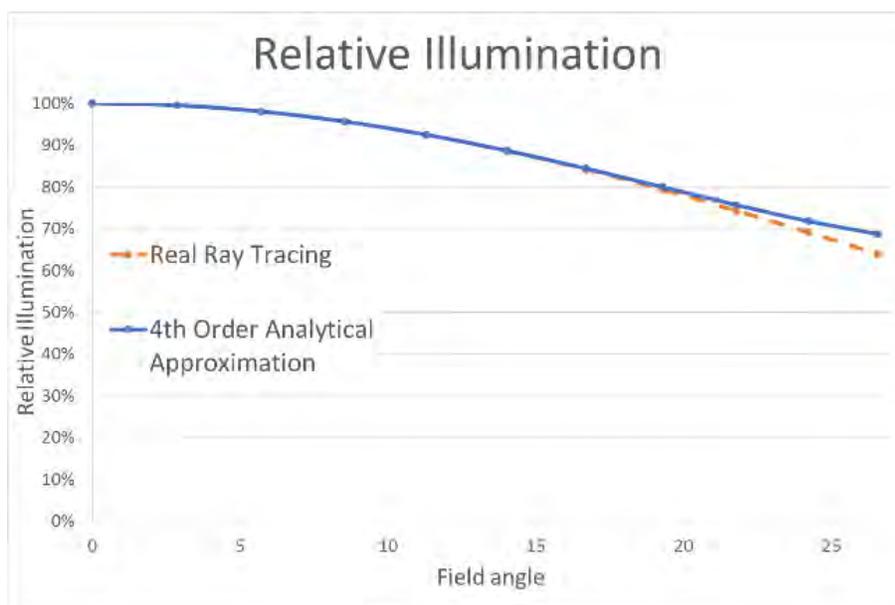


Fig. 3.9 Relative illumination plot for U.S. Patent 2516724. Analytical fourth-order approximation and real ray tracing calculation show excellent agreement up to about 20° half field angle.

3.9.5 Example: Lens Optimization for a Target Relative Illumination

The analytical solutions presented in this paper allow efficient optimization for a desired relative illumination during the lens design process. A straightforward way to calculate the relative illumination in a lens design program is to determine the exit pupil area by tracing

the rays in reverse from the image point in the direction cosines space. The relative illumination is proportional to the area of the pupil or to the number of rays that pass through the system [17]. Although this method is precise, the numerical integration requires tracing a large number of rays and therefore significantly slows down the optimization. On the other hand, it is possible to calculate the wave aberration coefficients by tracing only two first-order rays and to estimate the relative illumination over the entire field of view of a lens.

To demonstrate that the fourth-order approximation to the relative illumination is sufficiently accurate for practical purposes, the close-form solution is used to design a lens that has uniform illumination over the field. The objective being modified is a Double Gauss lens that operates at $f/4$ and has a FOV of 40° . The distortion is limited to be $<5\%$, and both second- and fourth-order relative illumination coefficients are targeted to zero. The resulting system layout and relative illumination plot are shown in Fig. 3.10. The image is equally illuminated over its entire area.

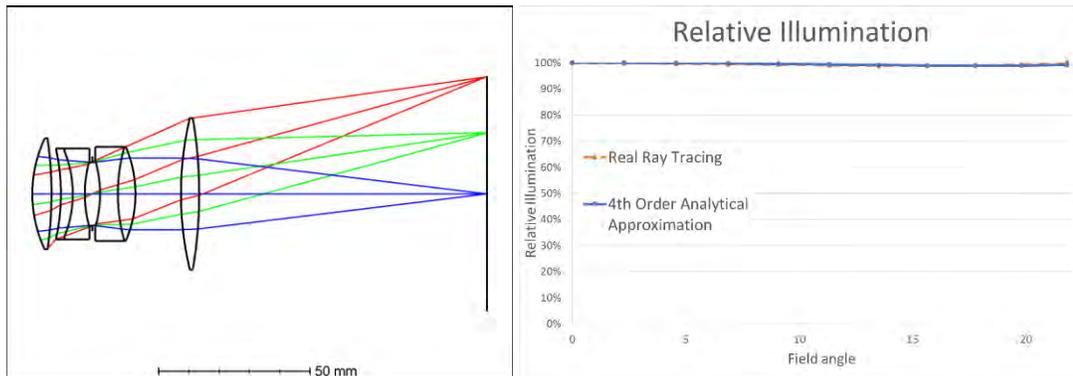


Fig. 3.10 A Double Gauss lens was reoptimized while targeting both second- and fourth-order relative illumination coefficients to zero. The fourth-order approximation is sufficiently accurate for practical purposes and allows efficiently optimizing for a desired relative illumination during the lens design process.

3.9.6 Relative Illumination: Conclusion

Among other image quality requirements, the relative illumination may have a strong impact on the performance of a lens. A number of authors have shown that the standard \cos^4 -law of illumination fall off is not accurate for an objective with image and pupil aberrations. However, a detailed and comprehensive investigation of the relation between relative illumination and individual aberration coefficients to fourth-order has not been previously discussed.

In this section, the problem has been approached from the aberration theory point of view. A closed form solution that approximates the image illumination fall-off and accounts for the illumination effects of aberrations in an unvignetted optical system is derived and presented.

Three different possible aperture stop positions are considered: as the last element, as the first element and between components of the lens system. Only the piston irradiance

term is considered in the derivation, which is equivalent to the specification of a small diaphragm. Table 3.8 gives the relative illumination to the fourth-order of approximation in terms of aberration coefficients for each case.

If the aperture stop of the system follows the lens, only third-order distortion contributes to the fourth-order relative illumination coefficient. On the other hand, if the aperture stop precedes the lens, both image and pupil aberrations affect the relative illumination. In this case, it is possible to have sufficient pupil aberrations such that the illumination will be more uniform or even constant over the entire field. If the aperture stop is located between the components of the lens, image and pupil aberration contributions of the rear half and distortion of the front half of the objective are considered separately. This result reveals a new path to improve image quality and required illumination by balancing aberrations of the two halves. In several special cases, it is also possible to come up with a more compact, reduced form of the equations.

Several examples of lenses that instead of the \cos^4 -law show improved illumination at the focal plane are discussed in detail. Aberration theory is used to provide insight into the role of individual aberration coefficients in the relative illumination of these lenses. Although the specification of a small diaphragm was used in the derivation, this approximation has shown to be sufficiently accurate for practical purposes.

3.9 CONCLUSION

This chapter presents a second- and a fourth-order theory of irradiance changes in axially symmetric optical systems. The concept of the irradiance function is reviewed and

an interpretation of the irradiance aberrations is discussed. The irradiance function terms represent basic distribution patterns in the irradiance of a beam at the exit pupil plane or at the image plane of an imaging system.

The irradiance coefficients are found via basic radiometric principles, such as conservation of flux, and are derived from purely geometric considerations. This approach gives us the specific relationship between the irradiance distribution and wavefront aberration coefficients. The edge diffraction effects are not considered in this study, and unclipped and unfolded beams are assumed.

Table 3.1 and Table 3.2 give the irradiance coefficients in terms of wavefront aberration coefficients and first-order system parameters to the fourth order of approximation. Specific formulas are provided for irradiance at the image and at the exit pupil of an optical system. The irradiance coefficients depend on the selection of coordinates. In this chapter, the field vector is defined at the object plane, while the aperture vector is defined at the exit pupil plane. The formulas for the irradiance coefficients in Table 1 and Table 2 show excellent agreement with the results from real ray tracing. Relative illumination and several cases of interest where the coordinate position changes have also been discussed.

The theory of irradiance aberrations enhances our knowledge about the behavior of light as it propagates in optical systems and provides insights into how individual wavefront aberration terms affect the light irradiance produced by a lens system at its image plane or at the exit pupil plane.

REFERENCES

1. J. Sasian, "Introduction to Aberrations in Optical Imaging Systems," Cambridge University Press (2013)
2. J. M. Palmer and B. G. Grant, "The Art of Radiometry," SPIE Press (2010)
3. K. P. Thompson, "Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry," *J. Opt. Soc. Am. A* **22**, pp. 1389-1401 (2005)
4. C. Roddier and F. Roddier, "Wave-front reconstruction from defocused images and the testing of ground-based optical telescopes," *J. Opt. Soc. Am. A*, Vol. **10**, No. 11 (1993)
5. J. Sasian, "Theory of six-order wave aberrations," *Applied Optics*, Vol. **49**, No. 16 (2010)
6. R. Kingslake, "Illumination in Optical System, Applied Optics and Optical Engineering," Academic Press, Vol. **II** (1965)
7. J. Petzval, "Professor Petzval New Lens," *Photographic Notes*, pp. 118-121 (1858)
8. M. von. Rohr, "Distribution of light on the focal plane of a photographic objective," *Zeitschr. Instrymtenk.* 18, pp. 171-180 and 197-205 (1898)
9. T. R. Dallmeyer, "Technical meeting," *Photographic Journal*, Vol. **19**, pp. 221-232 (1895)
10. F. E. Washer, "Characteristics of wide-angle airplane camera lenses," *Journal of Research of the National Bureau of Standards*, Vol. **29**, pp. 233-246 (1942)
11. M. Reiss, "The Cos^4 Law of Illumination," *Journal of Optical Society of America*, Vol. **35**, N. 4, pp. 283-288 (1945)
12. M. Reiss, "Notes on the Cos^4 Law of Illumination," *Journal of Optical Society of America*, Vol. **38**, N. 11, pp. 980-986 (1948)
13. I. C. Gardner, "Validity of the Cosine-Fourth-Power Law of Illumination," *Journal of Research of the National Bureau of Standards*, Vol. **39**, pp. 213-219 (1947)

14. F. Wachendorf, "The Condition of Equal Irradiance and the Distribution of Light in Images Formed by Optical Systems without Artificial Vignetting," *Journal of Optical Society of America*, Vol. **43**, N. 12, pp. 1205-1208 (1953)
15. M. M. Roossinov, "Wide angle orthoscopic anastigmatic photographic objective," US 2516724 (1950)
16. H. H. Hopkins, "Canonical and Real-Space Coordinates Used in the Theory of Image Formation," *Applied Optics and Optical Engineering*, Vol. **9**, Chap. 8 (1983)
17. M. P. Rimmer, "Relative illumination calculations," *Proc. SPIE*, Vol. **655**, pp. 99-104 (1986)
18. Zemax LLC., "OpticStudio Manual" (2016)
19. D. Reshidko and J. Sasian, "The role of aberrations in the relative illumination of a lens system," *SPIE Proceedings*, Vol. **9948** (2016)
20. J. Sasian, "Interpretation of pupil aberrations in imaging systems," *SPIE Proceedings*, Vol. **6342** (2006)
21. R. Kingslake, "Lens Design Fundamentals," Academic Press, second edition (2010)
22. D. Reshidko and J. Sasian, "Optical analysis of miniature lenses with curved imaging surfaces," *Appl. Opt.* **54** (**28**), pp. E216–E223 (2015)
23. A. Ning, "Compact lens with external aperture stop," US 6441971 B2 (1999)
24. J. Sasian, "Theory of six-order wave aberrations: Errata," *Applied Optics*, Vol. **49**, No. 33 (2010)

CHAPTER IV: RAY TRACING METHODS FOR CORRECTING CHROMATIC ABERRATIONS IN IMAGING SYSTEMS

4.1 INTRODUCTION

Correction of chromatic aberrations in imaging systems has been the subject of research for many generations of scientists and still remains one of the most challenging problems in lens design. The early experiments of Isaac Newton, which later were developed into his color theory, started a detailed understanding of chromatic effects in optical systems. Since Conrady's solution to the powers for an apochromatic triplet that uses the partial dispersion vs. V-number graph, many graphical methods have been derived for glass selection. In particular Hoogland's method, is considered the best according to Rayces [1, 2]. More comprehensive glass selection methods have been also developed for finding proper glass combinations to design systems corrected for two, three, four and even more wavelengths. A notable example is the method of Robb and Mercado [3, 4] that is based on Buchdahl chromatic coordinates. This work has been extended by Albuquerque et. al. [5] to include higher order aberration and to include some minimization metrics. Sigler [6], points out that chromatic aberration correction through the selection of optical glass has been one of the most extensively studied subjects in the field of lens design.

While excellent results for color correction are reported in literature, there is still a gap between theory and practice and many methods can be considered purely academic. Aberration formulas are typically based on the first-order properties of the system or

include only finite number of higher-order aberration coefficients. However, as the complexity of actual systems grows, these simple metrics are not sufficient to estimate the final chromatic performance. Other reported techniques require additional mathematical calculation which cannot be implemented directly in the lens design software and require external tools. The use of additional software interrupts the lens designer work flow and adds a significant amount of time and effort.

The correction of chromatic aberrations still remains challenging in high aperture large field systems with a broad band. Modern optical instruments, such as used in medicine, astronomy, semi-conductor, defense and security markets, take advantage of different spectral bands spanning from short UV to long IR. Depending on the application, these systems are sometimes designed to support several spectral bands. The new generation of sensors which are sensitive to a larger bandwidth, for example from visible to short-wave infrared, are already available commercially. Moreover, some state-of-the-art applications may use several sensors supporting different bandwidth within the same optical channel.

Approximations involving the theoretical methods is the reason that aberration correction in complex state-of-the-art lenses is usually accomplished with real ray tracing which enables an accurate view of the actual image. Real ray tracing optimization is performed by minimization of some error function. However, the standard error function in lens design software typically mixes the correction of monochromatic and chromatic aberrations. In that case correction of chromatic aberration is strongly dependent on the correction of the monochromatic aberration.

In chapter, the chromatic aberrations correction is separated from the correction of monochromatic aberrations of a lens system. Two simple but effective real ray tracing methods are presented [7, 8]. These two methods can be implemented within optical design software and can be used to correct both chromatic change of focus and chromatic change of magnification of imaging systems, as well as higher-order chromatic aberrations. The chromatic aberration correction is not limited to a specific bandwidth. The proposed methods are capable of finding solutions for complex lens systems, as well as simple lens types.

The first method of forward and reverse ray tracing is explained in Section 4.2 and, its limitations are discussed. In Section 4.3, the second method is presented which is an algorithm for writing an error function that separates aberrations. Some examples and guidelines to application of the methods are provided in Section 4.4. Section 4.5 introduces a modification to the second method to aid in the athermalization of imaging systems. Section 4.6 concludes this chapter.

4.2 FORWARD AND REVERSE RAY TRACING

This section presents the method of forward and reversed ray tracing. One of the fundamental axioms of geometrical optics is the principle of reversibility of the optical path of rays [9, 10]. According to this principle any ray traced through an optical system, if reversed in propagation direction, will retrace the same path backwards. Consequently, for any arbitrary optical system, all rays traced forward and then backwards will have the same optical path, i.e. both chromatic and monochromatic aberrations cancel out.

To implement the method in sequential lens design software, a virtual mirror is introduced at the image plane of the original optical system, and a reversed copy of the system is added. Note that in our case, in contrast to regular reflection from the mirror, the sign of thicknesses remains the same. The part of the system before the mirror is called the real part, and the part after the mirror is referenced to as the virtual part. This sequential model allows us to keep the same corresponding angles for the ray propagating in the real and virtual parts of the system. In addition, for any ray, the optical path in the real and virtual part will have the same value, but opposite sign, i.e. both chromatic and monochromatic aberrations cancel out as desired. The schematic representation of the ray tracing flow is shown in Fig. 4.1.

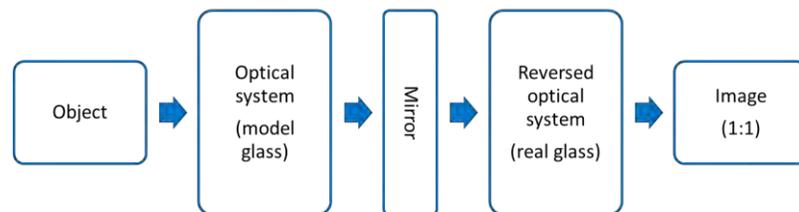


Fig. 4.1 Forward and reverse ray tracing flow chart.

In order to make the chromatic aberrations show, the optical glasses in the real part of the system is changed to fictitious, non-dispersive model glasses. For the model glass the refractive index of the original glass at the primary wavelength of the design is used and zero dispersion is specified, i.e. zero glass V-number, and zero partial dispersion. In the real part of the system, rays that have the same object and pupil coordinates but different wavelength will follow the same path. However, in the virtual part, the optical path will be different due to dispersion of the actual glasses. With this real and virtual lens model, for a given object point, the optical path difference will be zero for the primary wavelength

due to cancellation. For other wavelengths the chromatic aberration will show because there is no cancellation between the real and virtual parts. Thus the chromatic aberration content is separated from the monochromatic aberration content.

Since only chromatic aberrations are present, the full power of an optimizer can be dedicated for color correction. This results in an advantage. Global optimization together with iterative substitution of similar glass types is used to reduce the chromatic aberration through the selection of optical glass in the virtual part of the system. The real part of the system remains unchanged which will force the monochromatic solution to be close to the original design.

The focal length and focal ratio ($f/\#$) of the virtual system is controlled during optimization. In addition, the forward and reverse ray tracing model should have unit magnification. If original lens was designed for infinite conjugate, a paraxial lens with focal length equal to the focal length of the original system is used to focus the beam.

An example of forward and reversed ray tracing model for an Edmund Optics $f=20$ mm achromatic doublet is shown in Fig. 4.2. The primary wavelength is the d-line. Both the chromatic and the monochromatic aberrations are present for the on-axis field in the original design. The forward and reversed ray tracing model is then constructed according to the procedure described above. Since the doublet is used at an infinite conjugate, an $f=20$ mm paraxial lens is added to the modified system to focus the beam. Notice in Fig. 4.2 (d) that for the d-line the Optical Path Difference (OPD) is zero across the aperture. All monochromatic aberrations are canceled and only chromatic aberrations for the F and C

lines are present in the forward and reversed model system. This is the starting point for the optimization process.

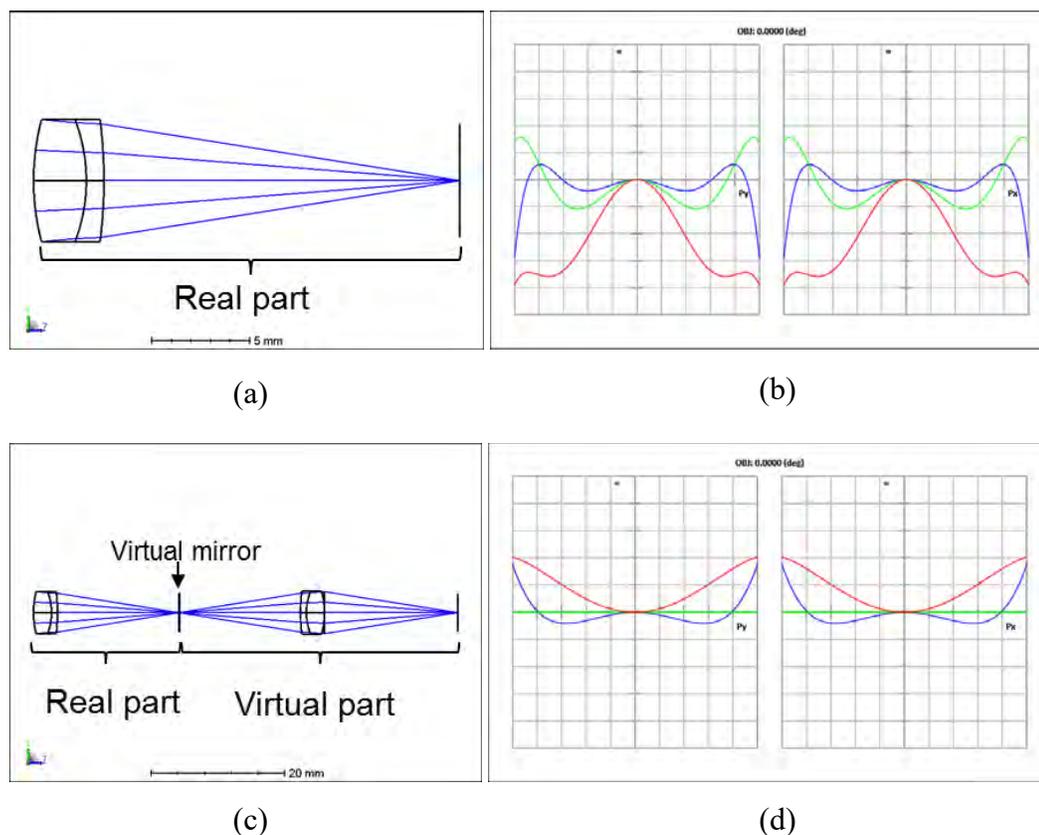


Fig. 4.2 Edmund Optics $f=20$ mm Achromatic Doublet: (a) original lens and (b) on Axis OPD fan, (c) forward and reversed ray tracing model and (d) on Axis OPD fan. The OPD fan full scale is 0.5 waves.

This forward and reversed ray tracing model demonstrates the geometrical method of separation of chromatic and monochromatic aberrations. However, this method has some limitations and drawbacks. The model glass, used in the real part of the system, changes propagation of the rays for the wavelength that differs from the primary wavelength. The deviations of these rays from the original path accumulate as the rays propagate through real and virtual parts of the system. In certain cases it may cause total internal reflection, vignetting of some of the rays, or cause singularities in the optical path for particular object

and pupil coordinates. Moreover, the model has twice as many optical surfaces as the original design, which significantly slows down ray tracing time. Nevertheless, it is a useful method.

4.3 ABERRATION SUBTRACTION METHOD

In this section, the chromatic and monochromatic aberrations are separated directly in the error function. The aberration of the primary wavelength is simply subtracted from the aberration due to other wavelengths.

In software, the lens designer can choose to minimize the wave aberrations or the ray aberrations. An error function is written where the optimization targets are changed as follows. For any wavelength λ and primary wavelength λ_{prime} the wave aberration target is modified to be

$$\left(OPD_{\lambda} \cdot \lambda - OPD_{primary} \cdot \lambda_{primary} \right) / \lambda. \quad (4.1)$$

For ray aberrations, the primary wavelength is subtracted directly without applying any normalization. The optimization target is then modified for every field point and every pupil coordinate in the error function. The chief ray at the primary wavelength is used as a reference for both ray or wave aberrations calculations. This procedure can be easily implemented by writing a macro program (see Appendix J). Fig. 4.3 shows the aberration subtraction procedure.

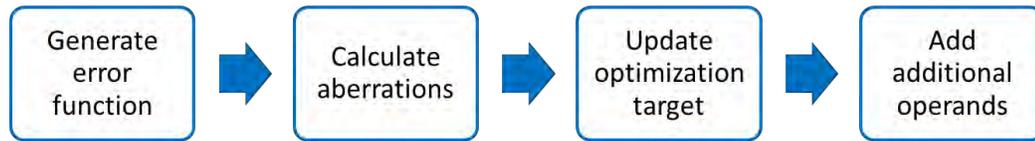


Fig. 4.3 Aberration subtraction method flow chart.

Several comments on the dynamics of minimization of the modified error function are provided below. In the original design the primary wavelength's monochromatic aberration will match the target aberration value, by construction. Initially, and similarity to the forward and reversed ray tracing technique, the primary wavelength will not contribute to the error function. During the optimization, the original lens will change and the primary wavelength's monochromatic aberration will be different. The weight placed on the primary wavelength, in the modified error function, will force the monochromatic solution to be close to the original design. Further minimization of the modified error function together with iterative substitution of similar glass types will reduce, if proper glass combinations exist, all chromatic aberrations to the same monochromatic aberration value.

The direct subtraction of the aberrations in the error function is a similar concept to the previous method of forward and reverse ray tracing. The chromatic aberrations are separated from the monochromatic aberration and minimized to the same value. While the method of Section 4.2 has clear geometrical model, the aberration subtraction method of this section has several noticeable advantages. First, the method is simpler: there is no need to build a complicated sequential model. Ray tracing time will not be significantly increased. Second, rays are traced only forward and there will be no issues with TIR and vignetting as discussed in the previous section. Finally, the actual amount of aberration is controlled in the design. Other optimization operands can be added to the modified error

function to allow better control of various design parameters. For example, if the design requirements are on the Modulation Transfer Function (MTF), we may choose to add MTF operands to the error function.

4.4 GUIDELINES OF APPLICATION OF THE METHODS

The author emphasize that the presented methods are not aim to improve the monochromatic performance of a system, but to reduce the chromatic aberration to the monochromatic value. Both presented methods could be applied at the intermediate design stage, when the monochromatic performance is close to the requirements and chromatic aberrations are corrected to some extent. In order to gradually reduce the overall amount of aberration in the system, the methods could be applied in an iterative process. Each iteration includes the following steps:

- Optimize the lens with a given error function. During this optimization one may choose to put addition weight on the primary wavelength. Putting additional weight on the primary wavelength will improve the monochromatic performance and allow a better target for application of the methods in the next step.
- Apply the forward and reversed ray tracing method or aberration subtraction method. Use global optimization with iterative glass substitution.
- Proceed to the next iteration till no new combinations of glass are found in a reasonable amount of time.

The chart on Fig. 4.4 summaries the algorithm.

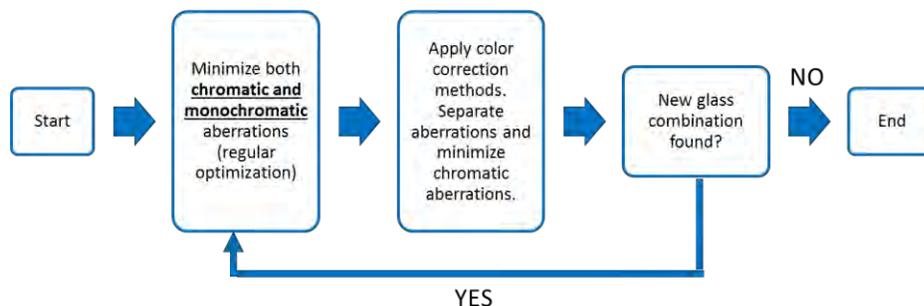


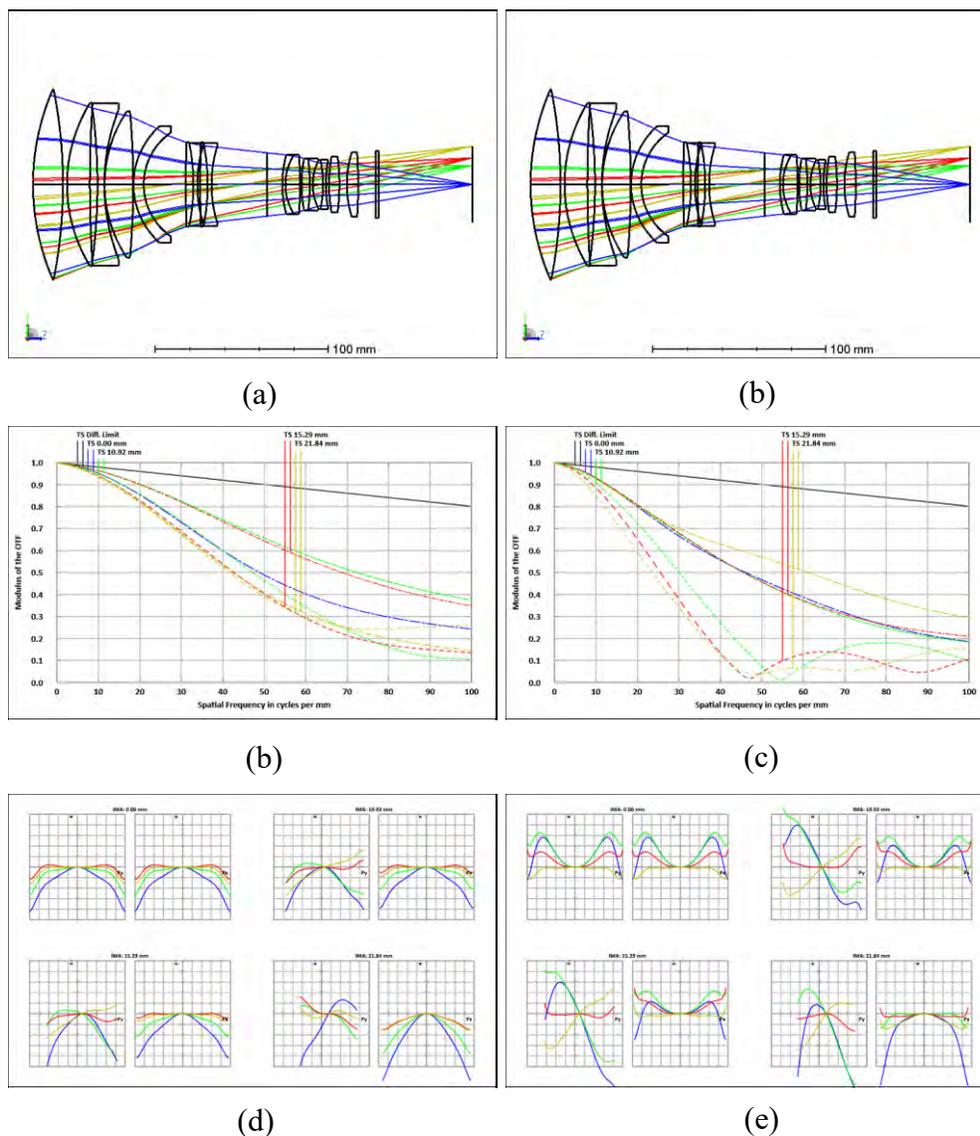
Fig. 4.4 Suggested algorithm for application of presented methods.

4.4 EXAMPLE OF OPTIMIZATION USING THE METHODS

The proposed methods were successfully applied to correct chromatic aberration and improve imaging performance of several complex lenses. This section provides examples of optimization using the aberration subtraction method.

4.4.1 Canon Telephoto Lens

As an example, the presented methods were applied to a Canon telephoto lens [11]. This is a long-focal-length (300mm) optical system that operates at $f/2.9$. Focusing is performed by moving one of the lens units. The lens operates in the visible from 436 nm to 656 nm. The primary wavelength of the design which is also the reference for the monochromatic aberrations is 588 nm. An original patent data for an infinite and a two meter object distance configurations is shown in Fig. 4.5.



iterations 1, 2 and 3 for the infinite and two meter object distance configurations are shown in Figs. 4.6 and 4.7. The glass combinations used in the original design and iterations 1, 2 and 3 are given in Table 4.1.

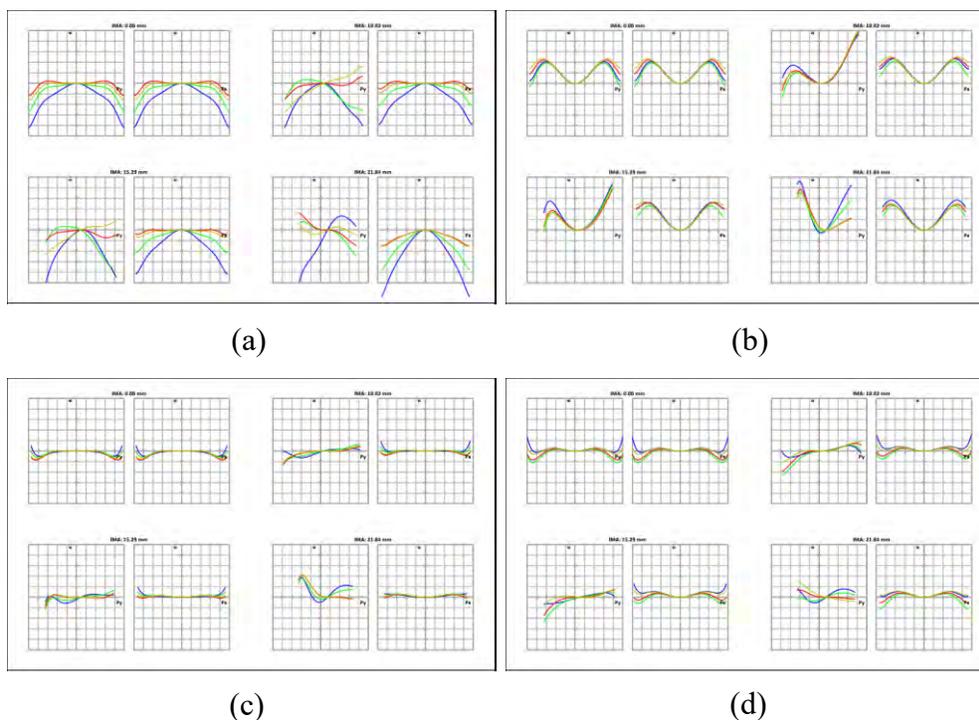


Fig 4.6 OPD fans for infinite object distance at 436nm, 486nm, 588nm and 656nm: (a) - patent data, (b) – iteration 1, (c) – iteration 2, (d) – iteration 3. Plot scale is 2 waves.

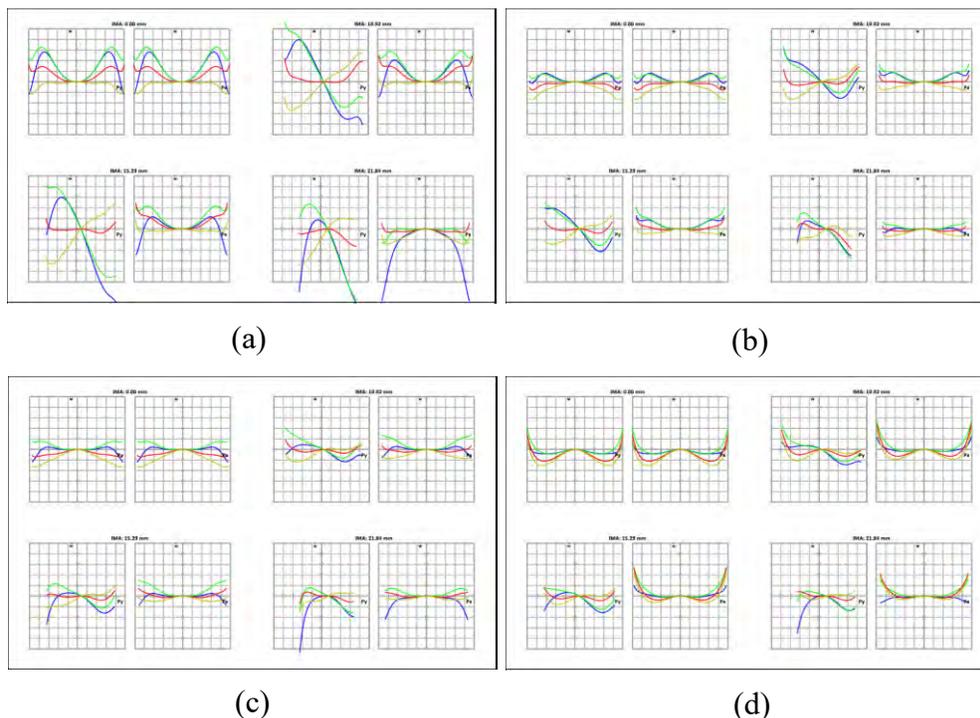
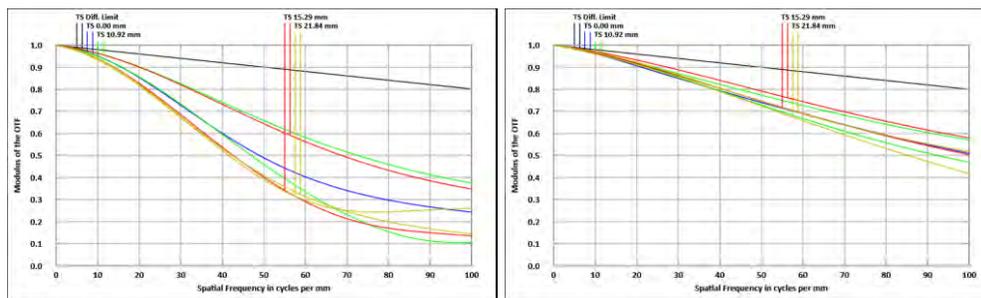


Fig. 4.7 OPD fans for two meter object distance at 436nm, 486nm, 588nm and 656nm: (a) - patent data, (b) - iteration 1, (c) - iteration 2, (d) - iteration 3. Plot scale is 2 waves.

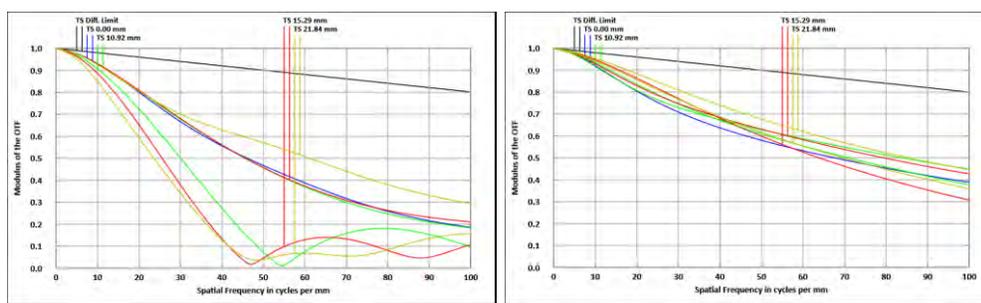
In each iteration the chromatic aberration is made to match better the monochromatic aberration value. Between the iterations the system was reoptimized to minimize the monochromatic aberration. Overall, the amount of aberrations in the system is gradually reduced. The polychromatic MTF improvement is also shown on Figs. 4.8 and 4.9.



(a)

(b)

Fig. 4.8 Polychromatic MTF from 436nm to 656nm for infinite object distance: (a) - patent data, (b) - final design.



(a)

(b)

Fig. 4.9 Polychromatic MTF for 436nm to 656nm for 2 meters object distance: (a) - patent data, (b) - final design.

This significant improvement of MTF over the entire field of view and a good balance was obtained for both the infinite and the two meter object distances.

Table 4.1 Optical glasses used in designs on Fig. 4.6 and 4.7.

Element #	Patent	Iteration 1	Iteration 2	Iteration 3
1	FCD1	N-FK51A	PFK85	LITHOTECH-CAF2
2	FCD1	LITHOTECH-CAF2	LITHOTECH-CAF2	LITHOTECH-CAF2
3	S-LAM54	N-LAF21	LAFN21	TAF4
4	LITHOTECH-CAF2	LITHOTECH-CAF2	LITHOTECH-CAF2	LITHOTECH-CAF2
5	S-TIM1	BAFD16	PBM2	F3
6-1	S-TIH53	S-NPH1	S-NPH1	SFL57
6-2	S-LAH55V	TAF5	YGH52	YGH52
7-1	S-NPH53	S-NPH1	S-NPH1	S-NPH1
7-2	S-LAL14	LAKL12	LACL1	LAKN6
8-1	S-TIH53	PBH23	FDS2	SF10
8-2	S-LAL7	LAK12	LAKN13	LAFL1
9	S-LAH55V	S-LAM2	SF9	SF19
10	S-NSL3	BSL7	BAK2	L-PHL2
11	S-BAL41	P-SK57Q1	PSK2	BSC7
12	S-BSL7	S-BSL7	S-BSL7	S-BSL7

4.4.2 SWIR Zoom Lens

The shortwave infrared (SWIR) range (wavelength from 900 to 1700 nm) allows day-to-night high resolution imaging. Recently, a new generation of detectors was developed allowing sensing in SWIR band. Since conventional optical glass transmits SWIR light, such cameras provide a cost effective solution for a wide variety applications and industries.

Additional example shows the correction of chromatic aberration by utilizing presented methods in a 24-140mm zoom lens [12]. The system operates at f/3.8 in the SWIR spectral

band. Zoom function is realized by moving three groups of elements. In lens design software zoom function is simulated using four lens configurations. The primary wavelength of the design which is also the reference for the monochromatic aberrations is 1300 nm.

As pointed out earlier, a starting system which has a good monochromatic performance and moderate chromatic aberration correction is required. The goal is to improve MTF of the system by reducing the chromatic aberration. In addition, balanced performance over the entire zoom range is desired. Finally, it is desirable to eliminate special glass types to allow a cost effective solution. The iteration algorithm described in Section 4.4 was used, and three iterations with aberration subtraction method for chromatic aberration correction were performed. Original system and final system layout for a single zoom position are shown in Figure 4.10.

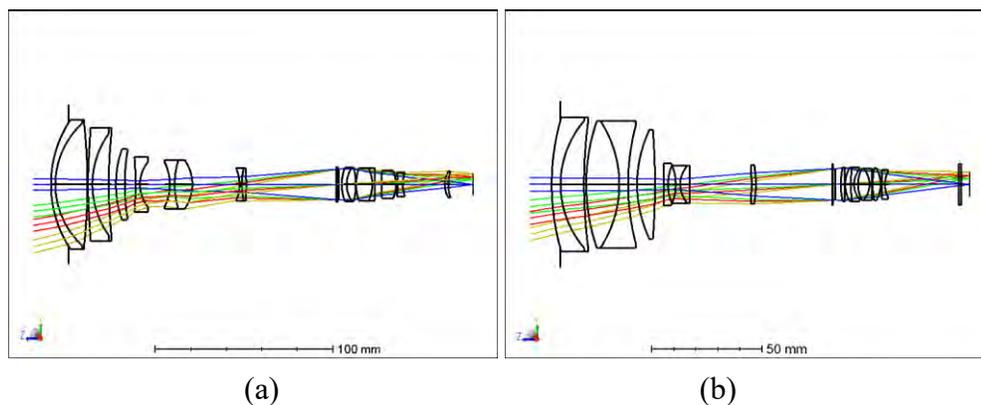


Fig. 4.10 Zoom lens layout: (a) - original design and (b) - reoptimized lens.

In each iteration the chromatic aberration is made to match better the monochromatic aberration value. Between the iterations the system was reoptimized to minimize the monochromatic aberration. Overall, the amount of aberrations in the system is gradually reduced. Special glass types, for example N-SK10, were replaced with standard glass. The

polychromatic MTF of the original system and improved polychromatic MTF of the final system are shown on Fig. 4.11. The glass combinations used in the original and final designs are given in Table 4.2.

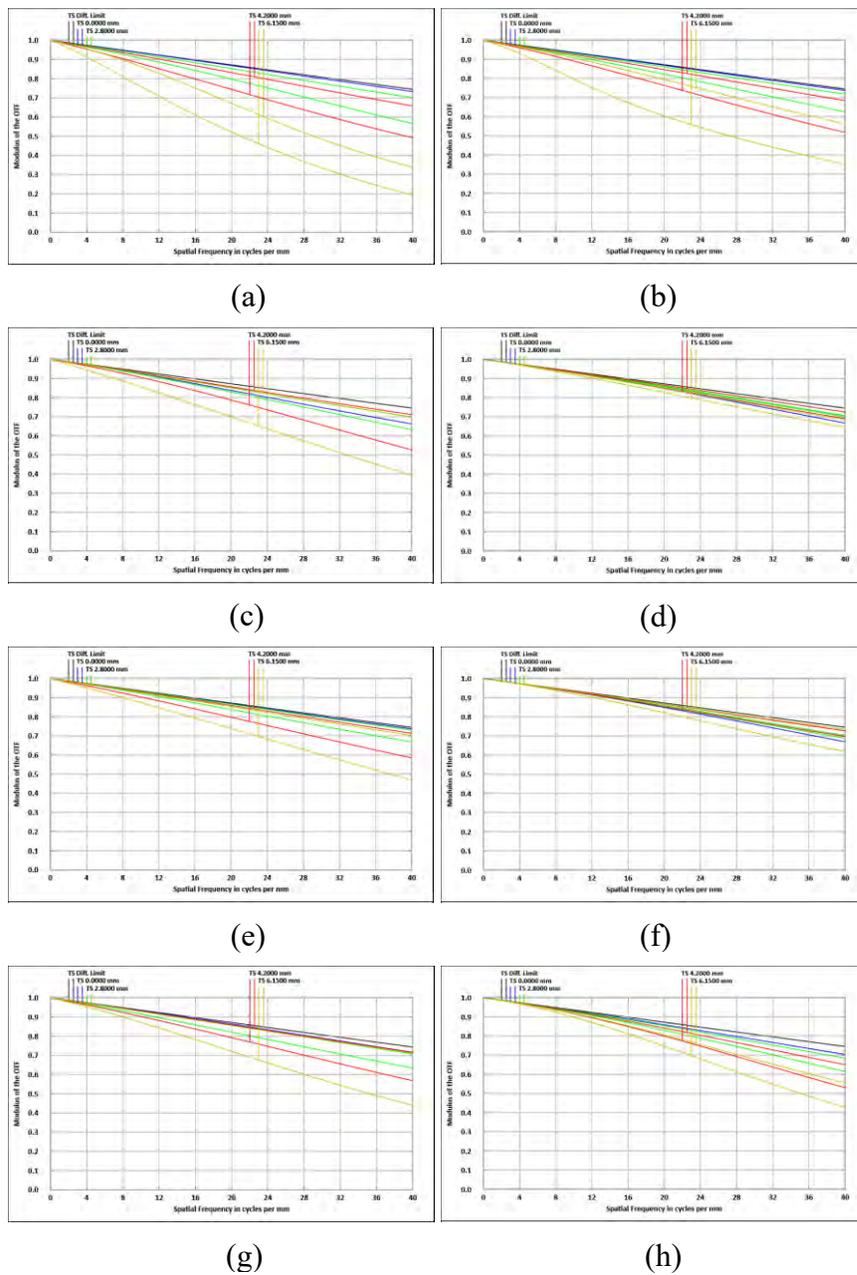


Fig. 4.11 Zoom lens polychromatic MTF: $f=24\text{mm}$: original design (a) and reoptimized lens (b); $f=60\text{mm}$: original design (c) and reoptimized lens (d); $f=100\text{mm}$: original design (e) and reoptimized lens (f); $f=140\text{mm}$: original design (g) and reoptimized lens (h).

This significant improvement of polychromatic MTF over the entire field of view and a good balance over the entire zoom range was obtained.

Table 4.2 Optical glasses used in designs on Fig. 4.10 and 4.11.

<i>Element #</i>	<i>Original design</i>	<i>Final design</i>
1	N-SK2HT	S-BSL7
2	N-LAK33B	S-LAM59
3	S-BAM4	S-PHM52
4-1	N-LASF44	S-LAH52
4-2	S-FLP53	S-FPL53
5-1	S-FPL53	BSL7Y
5-2	N-SK4	S-PHM52
6	S-BSL7	S-LAM59 L-BAL35
7-1	S-LAH65V	S-TIH10
7-2	S-FPM2	
8-1	S-FTM16	S-TIM2
8-2	N-SK10	S-FPM2
9	S-LAM52	L-LAH87
10	N-LASF44	L-LAH87
11-1	S-FPL53	S-TIM2
11-2	N-LASF43	S-FPM2
12-1	S-FPM2	S-FPM2
12-2	S-BAL42	S-LAH65

4.5 ABERRATION SUBTRACTION METHOD FOR ATHERMALIZATION OF IMAGING SYSTEMS

Thermal stability is another key requirement for many lens systems. Optical systems can be exposed to a wide range of temperatures. A temperature change has the major effects of changing the index of refraction of the optical glass and the geometry of the lens elements. First-order imaging properties of the lens are affected and additional aberrations are induced due to perturbation of the system from the nominal design.

Typically a lens is designed for some mean temperature. Once a nominal design is finished, a set of glasses that satisfies both the monochromatic and chromatic aberration correction requirements has been found. Such design may fail to maintain the desirable image quality over the whole range of possible temperatures. In order to insure reasonable performance, the effect of a uniform, homogeneous temperature change, is simulated. Athermalization, the condition of not varying with temperature, is achieved, similarly to chromatic aberration correction, by selection of materials. The goal is to keep the aberration correction invariant over the whole required temperature range. The aberration subtraction method described in Section 4.3 is modified to separate aberrations induced by thermal effects from aberrations in the nominal design.

In lens design software, a uniform temperature change can be modeled by adding lens configurations. Each configuration simulates some particular temperature by adjusting the index of refraction of the optical materials, changing the shape of the optical elements, and adjusting the lens spacings. A separate error function is generated for each configuration. In the aberration subtraction method for athermalization of imaging systems the aberration

of the nominal configuration is simply subtracted from the aberration of other configurations. Then an error function is written where the optimization targets are changed as follows. For any wavelength, field point and pupil coordinate, the wave or ray aberration targets are modified to be the aberration value at this point in the nominal configuration. The chief ray at the primary wavelength is used as a reference for both ray or wave aberrations calculations. Once again this procedure can easily be implemented by writing a macro program (see Appendix J). Fig. 4.12 shows the aberration subtraction procedure for athermalization.

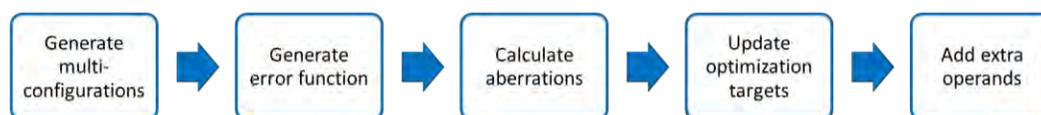


Fig. 4.12 Aberration subtraction method for athermalization flow chart.

Several comments on the dynamics of minimization of the modified error function are provided below. In the original design the nominal configuration's aberration will match the target aberration value, by construction. Initially the nominal configuration will not contribute to the error function. During the optimization, the original lens will change and the nominal configuration's aberration will be different. The weight placed on the nominal configuration, in the modified error function, will force the aberration correction to be close to the original design. Further minimization of the modified error function together with iterative substitution of similar glass types will reduce, if proper glass combinations exist, all other lens configuration's aberrations to the same nominal configuration's aberration value. Therefore a similar monochromatic and chromatic performance is maintained for

the whole temperature range. The aberrations induced by thermal effects are separated from the nominal design aberration and minimized to the same value.

As an example, the presented method of athermalization is applied to a single focus lens [13]. This is a 14 mm focal-length optical system that operates at $f/1.75$. The lens supports a bandwidth from 420nm in the visible to 1700 nm in the near infra-red (NIR). The lens is athermal from -20°C to $+60^{\circ}\text{C}$. An original system layout is shown in Fig. 4.13.

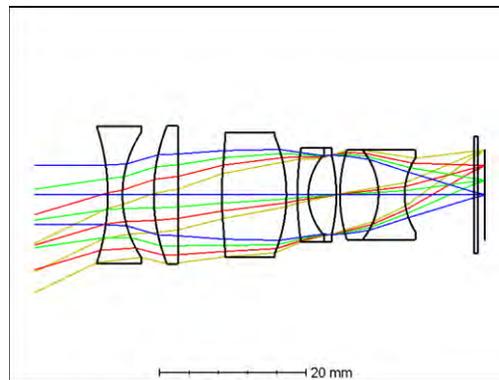


Fig. 4.13 Optical layout of a single focus athermal lens.

The iterative algorithm described in Section 4.4 was used. The OPD fans and MTF plots of the original design and final design are shown in Fig. 4.14 and 4.15. The glass combinations used in the original design and final iteration are given in Table 4.3.

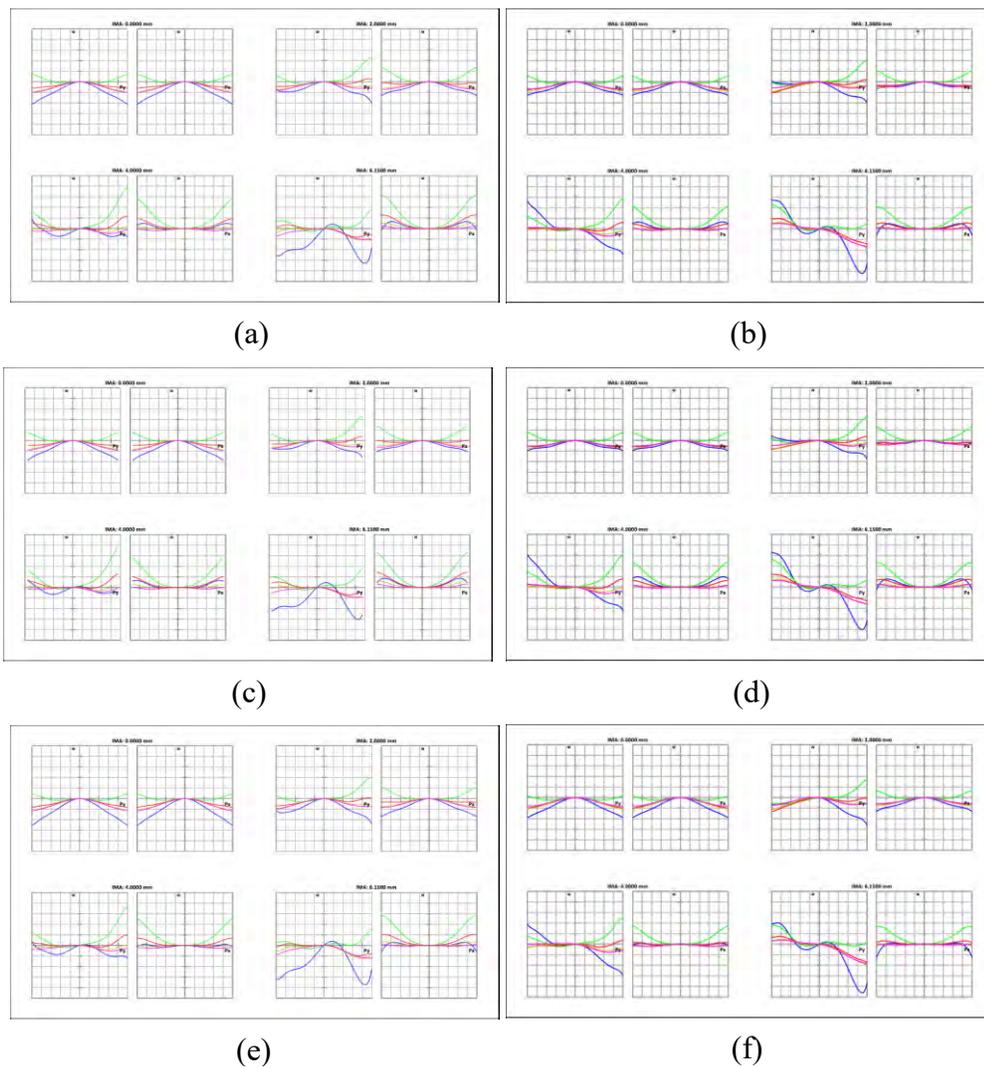


Fig. 4.14 OPD fans: (a) – original design @ 20°C , (b) - final design@ 20°C , (c) – original design @ -20°C , (d) - final design @ -20°C , (e) – original design @ 60°C , (f) - final design @ 60°C . Plot scale is 5 waves.

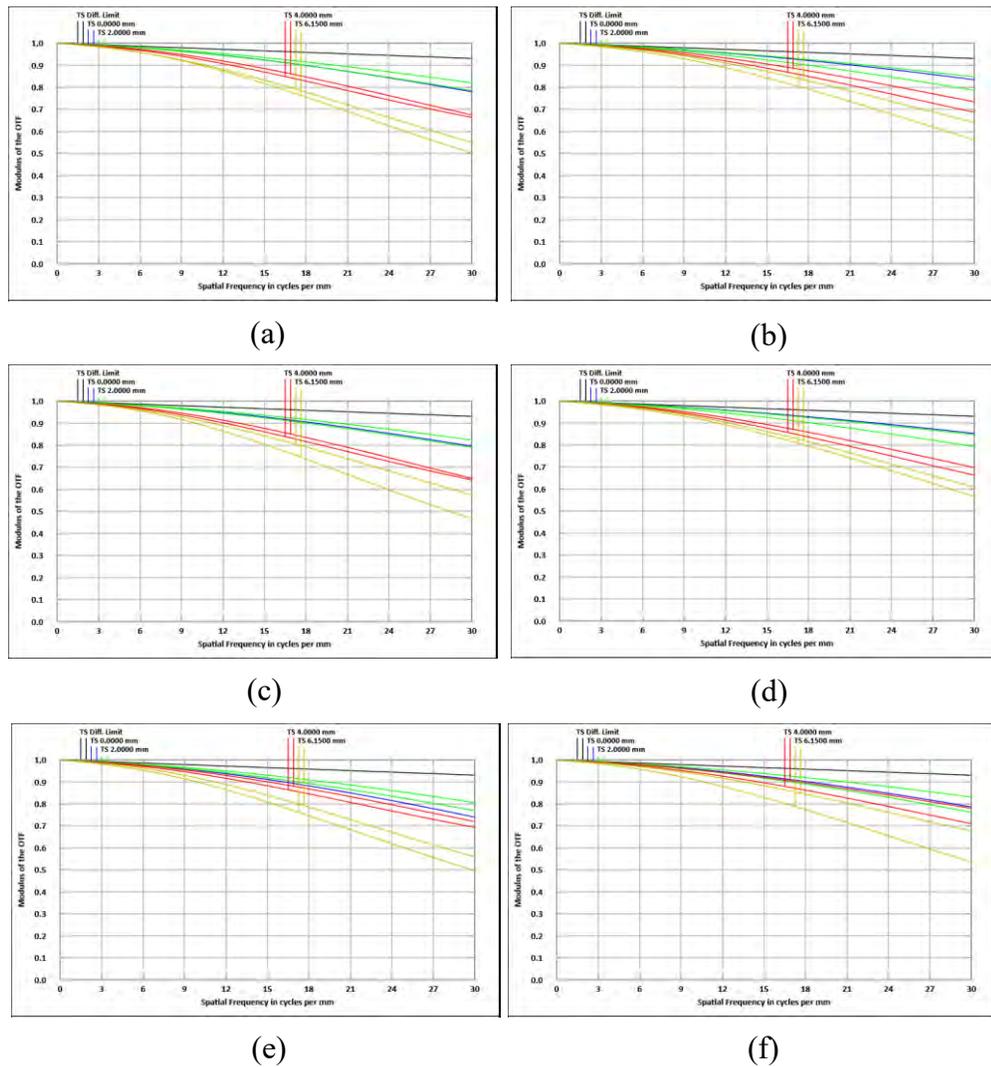


Fig. 4.15 Polychromatic MTF for 420nm to 1700nm: (a) – original design @ 20°C , (b) - final design@ 20°C , (c) – original design @ -20°C , (d) - final design @ -20°C , (e) – original design @ 60°C , (f) - final design @ 60°C .

This slight improvement of MTF over the entire field of view and the good balance was obtained for the whole temperature range.

Table 4.3 Optical glasses used in designs in Fig. 4.14 and Fig.4.15.

<i>Element #</i>	<i>Original design</i>	<i>Final design</i>
1	N-SSK8	K7
2	N-SK57	N-SK57HT
3	N-BAK2	N-BAK2
4-1	N-SF57	N-SK57HT
4-2	N-FK51A	N-PK52A
5-1	N-LASF40	N-LASF40
5-2	N-KZFS2	N-KZFS2

4.5 CONCLUSION

Two methods for chromatic aberration correction have been presented in this chapter. Both methods are based on real ray tracing and can be implemented in commercial lens design software. The idea is to separate the chromatic aberration content from the monochromatic aberration content and use standard lens design software optimization tools, such as global optimization and iterative glass substitution. The separation of aberrations removes the dependence in correction of one aberration type from the correction of other aberration type. In that case the optimization process is focused on reducing chromatic aberration to the monochromatic aberration value. The author finds these methods to be effective in correcting chromatic aberration beyond what a standard error function would do.

The presented methods are not limited to any specific bandwidth, can be used to correct both chromatic change of focus and chromatic change of magnification, support multi-configuration systems, and do not require any additional/external calculations. Unlike other methods found in the literature, the presented methods are not limited to the finite number of aberration coefficients used in an error function, since real ray tracing is used to reduce

the chromatic aberrations. The user can take advantage of any additional lens design software features; for example, additional constraints can be added to the error function or a filter can be used to eliminate unwanted glass types.

The proposed methods work the best when applied in the intermediate design stage, when monochromatic performance is reasonable and chromatic aberrations are corrected to some extent.

As an example of application of the methods, the chromatic aberration correction in a Canon telephoto lens and in an elaborated zoom system was improved. The chromatic correction methods were iterated with regular optimization as described in Section 4.4 to gradually reduce the overall aberration in the system. The final glass combination, which was found by application of the aberration subtraction method, allows reducing both chromatic and monochromatic aberrations while using only standard glass types. Significant improvement in the polychromatic MTF over the entire FOV was obtained. The final solution is not readily found with standard optimization methods as can be tested.

Finally, the methods were extended to a different glass selection problem – athermalization of imaging systems. The aberration subtraction algorithm was modified to separate aberrations induced by temperature effects from the nominal design aberrations. In that case the optimization process is focused on reducing additional aberration due to thermal effects to the nominal aberration value.

Although the algorithms presented here are not limited to any specific lens design software, in this work Zemax OpticStudio lens design program was used. The pupil was sampled using Gaussian quadrature method with 20 rings and 12 arms. Optimization time,

which of course depends on the complexity of particular lens, is typically less than one hour for a simple system.

REFERENCES

1. J. L. Rayces and M. Rosete-Aguilar, "Differential equation for normal glass dispersion and evaluation of the secondary spectrum", *Appl. Opt.* **38**, 2028-2039 (1999)
2. J. L. Rayces and M. Rosete-Aguilar, "Selection of glasses for achromatic doublets with reduced secondary spectrum. I. Tolerance conditions for secondary spectrum, spherochromatism, and fifth-order spherical aberration", *Appl. Opt.* **40**, 5663-5676 (2001)
3. P. N. Robb, "Selection of optical glasses. 1: Two materials", *Appl. Opt.* **24**, 1864-1877 (1985)
4. R. I. Mercado, P. N. Robb, "Color-corrected Optical Systems and Method of Selecting Optical Materials Therefor", U.S Patent **5,210,646** (1993)
5. B. F. C. Albuquerque, J. Sasian, F. L. Sousa and A. S. Montes, "Method of glass selection for color correction in optical system design", *Opt. Express* **20**, 13592-13611 (2012)
6. R. D. Sigler, "Glass selection for airspaced apochromats using the Buchdahl dispersion equation", *Appl. Opt.* **25**, 4311-4320 (1986)
7. D. Reshidko and J. Sasian, "Ray tracing methods for correcting chromatic aberrations in imaging systems", In press, *International Journal of Optics*, Article ID 351584 (2014)
8. D. Reshidko and J. Sasian, "Algorithms and examples for chromatic aberration correction and athermalization of complex imaging system," *Proc. SPIE* **9293** (2014)
9. J. M. Sasián, S. A. Lerner, T. Y. Lin, and L. Laughlin, "Ray and van Cittert–Zernike characterization of spatial coherence", *Appl. Opt.* **40**, 1037-1043 (2001)
10. R. N. Pfisterer, "Clever tricks in optical engineering", *Proc. SPIE* **5524**, 231-239 (2004)

11. M. Misaka, A. Nishio, H. Ogawa, U.S. Patent **6,115,188A** (2000)
12. Camellia lens, SWIR 20 – 140mm F/3.8 Continuous zoom lens, RP Optical Lab Ltd (2014)
13. RP-14-175A, VIS + SWIR 14mm F/1.75 A-thermal mono-focal lens, RP Optical Lab (2014)

CHAPTER V: USEFUL ASPHERIC/FREEFORM SURFACE DESCRIPTIONS FOR IMAGING AND NON-IMAGING PROBLEMS

5.1 INTRODUCTION

Optical design depends on optical surface description and, therefore, the lens designer's choice of the surface type is often a key to provide efficient solutions to imaging and non-imaging problems.

For imaging applications, the well-known axially symmetric conic and polynomial surface of Eq. 5.1 provides solutions for a vast number of problems,

$$z(r) = \frac{cr^2}{1 + \sqrt{1 - (1+k)c^2r^2}} + A_2r^2 + A_4r^4 + A_6r^6 + A_8r^8 \dots, \quad (5.1)$$

where c is the curvature of the surface, k is the conic constant, and A 's are the even aspheric polynomial coefficients.

However, a significant number of terms is required in some cases to provide an adequate solution with aspheric surfaces constructed by superposition of higher-order polynomials. On the other hand, both the optimization convergence and optimization time get worse as the number of higher-order aspheric terms becomes large. In addition, lens systems that contain lens elements that use higher-order aspheric terms are susceptible to creating imaging/non-imaging artifacts when they are slightly misaligned. These artifacts are created because the higher-order terms may create bumps and dips that under perfect

registration cancel out, but that under a slight misalignment add up, thus creating the artifacts.

The case of non-imaging optics also requires surfaces that typically cannot be represented by conic/polynomial-type expansions. Therefore, to solve illumination optics problems, several surface descriptions including splines, implicitly defined surfaces, surfaces based on Bernstein polynomials, and freeform surfaces have been proposed [1-8]. On the other hand, most illumination problems are solved numerically, and the resulting data points are interpolated for ray tracing purposes. Numerical methods provide a solution surface as a set of data points. However, in some applications, it is desirable to describe the solution surface in closed form in order to be specific about the nature of the surface. Thus, there is a need for surface descriptions that effectively solve illumination problems and that can ideally be expressed in closed analytical form.

Some advantages of using closed-form surface descriptions are that the actual surface can be precisely specified, that some tolerancing analyses can be produced, and that parametric studies can be conducted.

This chapter proposes three base surface descriptions in closed form that resemble the ideal profile that is required in a lens. Using the concept of base surface, aspherical/freeform surfaces are constructed and demonstrated in closed form. These surfaces substantially describe the desired surface for uniform illumination on a target plane or allow stigmatic imaging surface after surface along a particular ray. Examples of the performance of these surfaces that were constructed using the concept of base surface are provided, and some of their properties are discussed.

5.2 PEDAL CURVE TO THE ELLIPSE

U.S. Patent 4,734,836 describes a lens for uniform illumination on a planar target using an approximate point source. The first surface of this axially symmetric lens is spherical in shape and is concentric with the point-like source. The second surface is concave near the center and turns smoothly into convex towards the lens edge. Thus, the lens spreads out the light to avoid a bell-shaped illumination profile. The design of the concave-convex surface can be extreme as rays may come up to 75 or more degrees of inclination with respect to the optical axis.

The pedal curve to the ellipse is shown in Fig. 5.1 and is given analytically by the equation

$$a^2x^2 + b^2y^2 = (x^2 + y^2)^2, \quad (5.2)$$

where a is the major axis of the ellipse and b is the minor axis. The sag $S_p(r)$ of the surface can be obtained by rotation of the pedal curve about the z axis and is written as,

$$S_p(r) = b - \sqrt{\frac{b^2 - 2r^2 + \sqrt{b^4 + 4(a^2 - b^2)r^2}}{2}}, \quad (5.3)$$

where $r = \sqrt{x^2 + y^2}$ is the radial distance from the optical axis or z axis. $S_p(r)$ is a base surface for the pedal to the ellipse curve. It is noted that this pedal curve to the ellipse resembles the concave-convex profile that is the desired surface for uniform illumination on a target plane [9].

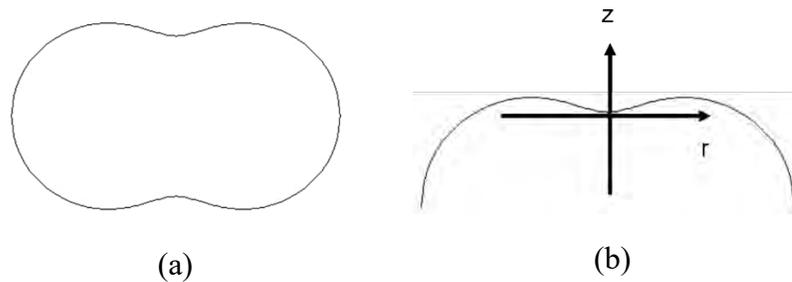


Fig. 5.1 (a) - Pedal curve to the ellipse. (b) - Coordinate axes on half the pedal curve. Note the change of curvature from the curve center to the edge.

Now an aspherical surface $z_1(r)$ is constructed by building a polynomial on the base surface $S_p(r)$ as,

$$z_1(r) = A_1 S_p(r) + A_2 S_p^2(r) + A_3 S_p^3(r) + A_4 S_p^4(r) + A_5 S_p^5(r) + A_6 S_p^6(r) \dots, \quad (5.4)$$

where the A 's represent deformation coefficients. A freeform $z_2(r)$ surface is also constructed by a superposition of base surfaces as,

$$z_2(r) = A_1 S_{p/1}(r) + A_2 S_{p/2}(r) + A_3 S_{p/3}(r) + \dots, \quad (5.5)$$

where each of the $S_{p/n}(r)$ terms have their own independent a , b , and A coefficients.

These two surfaces $z_1(r)$ and $z_2(r)$ were programmed as user-defined surfaces in Zemax OpticStudio optical design software [10] (see Appendix K).

To design a surface, the light source is assumed to be point-like and Lambertian: the intensity decreases as the cosine function of the angle of ray emission θ with respect to the z axis of Fig. 5.1. The source is located on the optical axis, and the target surface is flat and perpendicular to the optical axis.

The optical flux $\Phi(\theta)$ from a Lambertian point source L_0 as a function of angle θ is given by

$$\Phi(\theta) = 2\pi L_0 \int_0^{\theta} \cos(\theta) \sin(\theta) d\theta = \pi L_0 \sin^2(\theta). \quad (5.6)$$

The fractional optical flux from a Lambertian source and from a flat surface that is uniformly illuminated are given by $\sin^2(\theta)/\sin^2(\theta_{\max})$ and Y^2/Y_{\max}^2 , respectively. It follows that for uniform illumination on a flat surface, conservation of optical flux requires that for a given ray emitted at angle θ with respect to the optical axis, the ray intersection Y at the target surface should satisfy

$$Y = \frac{Y_{\max} \sin(\theta)}{\sin(\theta_{\max})}, \quad (5.7)$$

where θ_{\max} is the maximum angle of emission and Y_{\max} is the maximum ray intersection at the target surface.

For the actual surface design, 20 rays from the source were traced equally spaced in angle of emission from $\theta = 0$ to $\theta = \theta_{\max}$. An error function was created as the sum of the squares of the differences of ray intercept and the theoretical ray intercept Y . Then using the damped least square and the orthogonal descent optimization methods, the error function was minimized. It was noted that when ray total internal reflection took place, the ray tracing stopped for that ray and the optimization process stagnated. This indicated that no physical solution was possible due to the index of refraction value or to a target surface in closed proximity to the lens. The surface coefficients were used by the optimizer as

degrees of freedom to reduce the error function. The coefficients were released as variables in sets of two and as the optimizer proceeded, more coefficients were released.

To illustrate the performance of the surfaces, two lenses made out of polycarbonate plastic ($n=1.58546992$ at $\lambda=587.5618$ nm) were designed. The first lens used the aspheric surface profile $z_1(r)$ with a maximum ray angle of $\theta_{\max} = 75^\circ$ degrees. The second lens used the freeform surface profile $z_2(r)$ with also a maximum ray angle of $\theta_{\max} = 75^\circ$.

For both lenses, the marginal ray (at θ_{\max}) angle of incidence on the surface was constrained to zero degrees with respect to the optical axis. In addition, the distance from the Lambertian point source to the aspheric/freeform surface was 5 mm and the distance from the source to the target surface was 20 mm. The lenses and ray trace are shown in Figs. 5.2 and 5.3, respectively, and the surface descriptions are given in Tables 5.1 and 5.2, respectively.

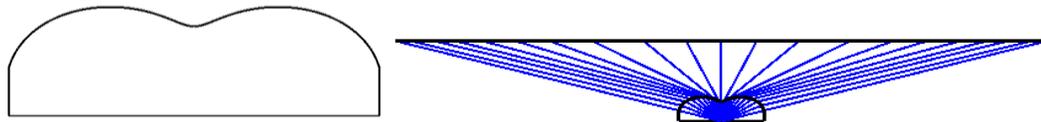


Fig. 5.2 Aspheric lens with profile $z_1(r)$ and ray trace to the target surface.

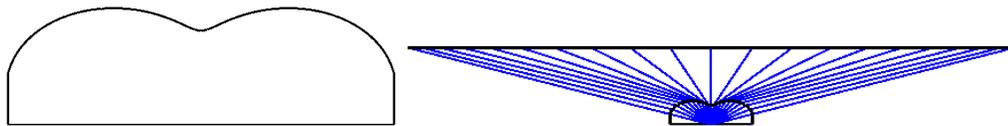


Fig. 5.3 Freeform lens with surface profile $z_2(r)$ and ray trace to the target surface.

Table 5.1 Coefficients defining the surface profile $z_1(r)$.

$z_1(r)$	a [mm]	b [mm]	A1	A2
	10.209629	4.1375154	-0.76745888	-0.0076601472
	A3	A4	A5	A6
	0.0032823004	3.9058491e-005	-0.00036512352	8.911191e-005

Table 5.2 Coefficients defining the surface profile $z_2(r)$.

$z_2(r)$	a [mm]	b [mm]	A
$S_1(r)$	27.089235	5.7955684	-0.34716392
$S_2(r)$	10.330664	11.189056	-0.55695918
$S_3(r)$	117.56108	71.18933	-0.5932761

To evaluate the performance of the aspheric and freeform surfaces, plots of the relative illumination and transverse ray error on the target surface were produced as shown in Fig. 5.4 and Fig. 5.5, respectively.

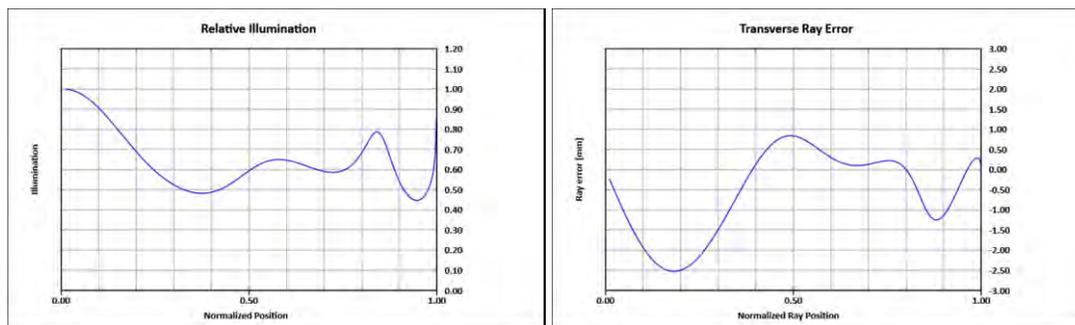


Fig. 5.4 Relative illumination (a) and transverse ray error in mm (b) of the aspheric lens at the target surface $z_2(r)$ for lens of Fig. 5.2. Transverse ray error plot scale is $\pm 3mm$.

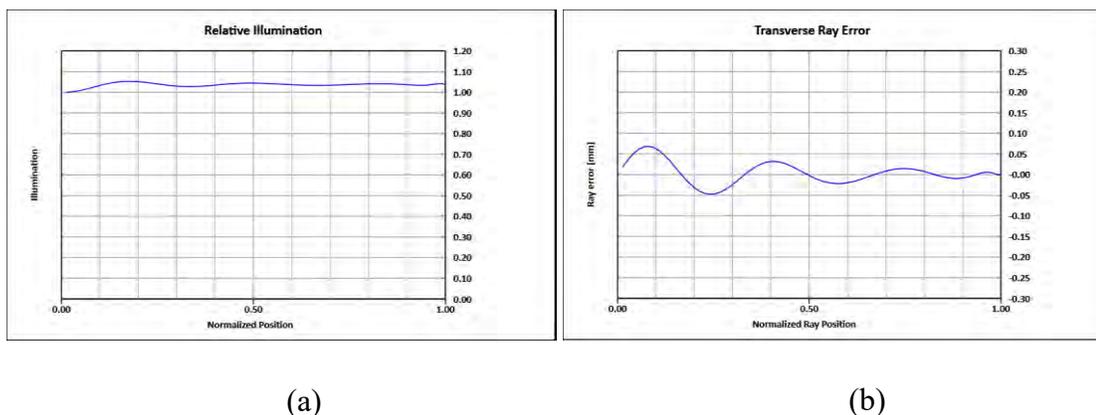


Fig. 5.5 Relative illumination (a) and transverse ray error in mm (b) of the freeform lens at the target surface for lens of Fig. 3. Transverse ray error plot scale is $\pm 0.3mm$.

It is clear from examination of the relative illumination and transverse ray error plots that the freeform surface profile of the lens in Fig. 5.3 is able to best model the ideal surface. In contrast, the relative illumination plot for the aspheric surface varies from about 0.5 to 1 and this is not a good surface match for the ideal surface.

Noteworthy is that the freeform lens plots do not exhibit significant oscillation at the edge of the target. This result can be explained by the absence of higher-order terms in the surface description. This lack of oscillation is in fact a useful outcome as many aspheric

and freeform surfaces constructed by superposition of higher-order polynomials are subject to produce oscillation on the ray behavior.

It is worth mentioning that two factors that contribute to complicating or preventing discovery of solutions are failure of the ray-tracing algorithm to find the ray intersection point on the surface as it becomes steep, and ray total internal reflection. Further, in these type of solutions the angle of incidence can be large near the inflection point of the surface and Fresnel light reflection losses can be significant.

5.3 OFF-AXIS CONIC SURFACE

The conic and polynomial surface of Eq. 5.1 provides solutions for the case of the axially symmetric system. However, when one wishes to design an asymmetric optical system, it is desirable to have a convenient expression for the surface in a coordinate system that is centered on the off-axis surface segment rather than centered on the axis of symmetry. The analytical expression of the conic surface in a coordinate system that is not aligned with the axis of symmetry is more complex and not readily available. In this section, the derivation of the general conic surface in such a coordinate system is reviewed [11]. This off-axis conic surface is used as a base surface to construct an aspherical surface and is successfully applied to design a three-mirror unobscured telescope.

A general expression of an axially symmetric conic surface is rewritten as

$$z(r) = \frac{1}{(1+k)} \left\{ R - \left[R^2 - (1+k) \cdot r^2 \right]^{1/2} \right\} \quad (5.7)$$

where $R=1/c$ is the radius of curvature of the surface. The paraboloid expression is obtained by taking the limit of Eq. 5.7 as $k \rightarrow -1$. Next, the expression of the conic as viewed from a coordinate system that is tangent to the surface at a general point away from the axis of symmetry is derived. The geometry is shown in Fig. 5.6.

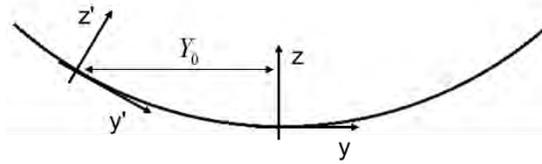


Fig. 5.6 The geometry defining global and local coordinates of the off-axis conic segment.

Without loss of generality, the new coordinate origin is chosen on the Y axis; thus, Y_0 is the distance from the rotation axis to the new coordinate center. From Eq. 5.7 it follows that

$$\tan(\varphi_0) = \frac{\partial z}{\partial x_{x=Y_0}} = \frac{Y_0}{[R^2 - (1+k) \cdot Y_0^2]^{1/2}}, \quad (5.8)$$

$$Z_0 = z(Y_0) = \frac{1}{(1+k)} \left\{ R - [R^2 - (1+k) \cdot Y_0^2]^{1/2} \right\}. \quad (5.9)$$

The rotation of coordinates is now performed according to

$$x = x' \cdot \cos(\varphi_0) - z' \cdot \sin(\varphi_0) + Y_0, \quad (5.10a)$$

$$y = y', \quad (5.10b)$$

$$z = x' \cdot \sin(\varphi_0) + z' \cdot \cos(\varphi_0) + Z_0, \quad (5.10c)$$

$$x' = (x - Y_0) \cdot \cos(\varphi_0) + (z' - Z_0) \cdot \sin(\varphi_0), \quad (5.11a)$$

$$y' = Y, \quad (5.11b)$$

$$z' = -(x - Y_0) \cdot \sin(\varphi_0) + (z' - Z_0) \cdot \cos(\varphi_0). \quad (5.11c)$$

For compactness, the dimensionless variables

$$u = \frac{x'}{R}; v = \frac{y'}{R}; w' = \frac{z'}{R}; \varepsilon = \frac{Y_0}{R}; w = \frac{z}{R} \quad (5.12)$$

and the quantities

$$\begin{aligned} s &\equiv \sin(\varphi_0) = \frac{Y_0}{[R^2 - kY_0^2]^{1/2}}; \\ c &\equiv \cos(\varphi_0) = \left[\frac{R^2 - (k+1)Y_0^2}{R^2 - kY_0^2} \right]^{1/2}; \\ L &\equiv k+1; \\ W_0 &\equiv Z_0/R; \end{aligned} \quad (5.13)$$

are introduced.

Eq. 5.11a, Eq. 5.11b and Eq. 5.7 are now substituted into Eq. 5.10c.

$$\frac{1}{L}(1 - L([uc - ws + \varepsilon]^2 + v^2)^{1/2}) = us + wc + W_0 \quad (5.14)$$

After some algebraic manipulations, Eq. 5.14 is reduced to a quadratic equation as in

$$w^2 + 2w(h + ju) - (fu^2 + gv^2) = 0. \quad (5.15)$$

The solution for $w(u, v)$ is

$$w = -(h + ju) \pm [(h + ju)^2 + fu^2 + gv^2]^{1/2} \quad (5.16)$$

where $f \equiv (s/\varepsilon)^2 g$, $g \equiv -1/(Lc^2 + s^2)$, $h \equiv (\varepsilon/s)g$ and $j \equiv -(L-1)scg$.

An aspherical surface $z_3(r)$ is now constructed by adding a plane-symmetric XY polynomial to the base surface in Eq. 5.16 as

$$z_3(r) = w(r) + A_1x^2 + A_2y^2 + A_3x^2y + A_4y^3 + A_5x^4 + A_6x^2y^2 + A_7y^4 \dots \quad (5.17)$$

The XY polynomial in Eq. 5.17 is centered at the origin of the off-axis conic segment and, thus, provides effective degrees of freedom for the lens design of plane-symmetric mirror systems. The surface $z_3(r)$ was programmed as a user-defined surface in Zemax OpticStudio optical design software (see Appendix K).

The following design shows the performance improvement made by using aspheric surfaces $z_3(r)$ in an $f/2$ three-mirror unobscured design, with the design parameters given in Table 5.3. The design specification is similar to one used by Chrisp [8]. The aperture stop is located at the secondary mirror.

Table 5.3 Design requirements for the three-mirror telescope.

Parameter	Requirement
Field of view (FOV)	10 deg x 9 deg
Focal length	35.7 mm
Focal ratio	$f/2$

For comparison, the design utilizing aspheric surfaces $z_3(r)$ has been created with the mirror sizes, separations and incidence angles that closely resemble the non-uniform rational-basis spline (NURBS) freeform design reported by Chrisp.

The design procedure is the following.

First, the reflective system is constructed with confocal surfaces such that the imaging along the optical axis ray (OAR) is stigmatic surface after surface, as shown in Fig. 5.7. These types of systems have a reduced number of aberrations and potentially can provide better imaging: constant astigmatism, constant coma, anamorphism, and spherical

aberrations exactly cancel out [12]. The mirror surface description $z_3(r)$ simplifies the design of these confocal systems by allowing one to optimize the offset parameter Y_0 for a given geometry. The image plane is constrained to be perpendicular to the OAR.

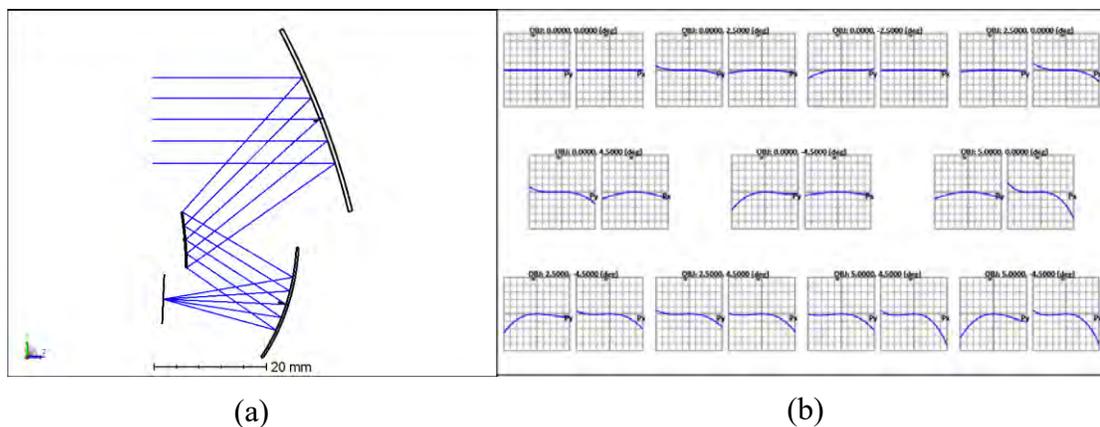


Fig. 5.7 Three-mirror system with confocal surfaces. (a) Layout (b) OPD plots over the entire FOV (the plot scale is 10 waves at $3\mu\text{m}$).

Next, mirror curvatures, conic constants and spacings are removed from the optimization, and polynomial coefficients are released to correct the off-axis aberrations. Although no constraints on distortions were mentioned by Chrisp, in the current design the distortion is limited to $<3\%$. The final system layout is given in Fig. 5.8. The OPD plots and Spot diagrams are shown in Fig. 5.9. The performance is close to being diffraction-limited over the entire FOV.

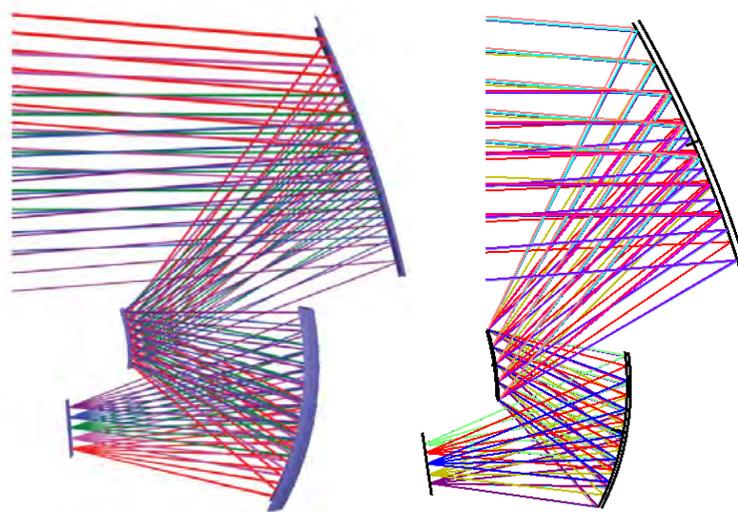


Fig. 5.8 Layout of the three-mirror system. (a) The design with NURBS C^{∞} surfaces reported by Chrisp; (b) The current design with Z_3 aspherical surfaces.

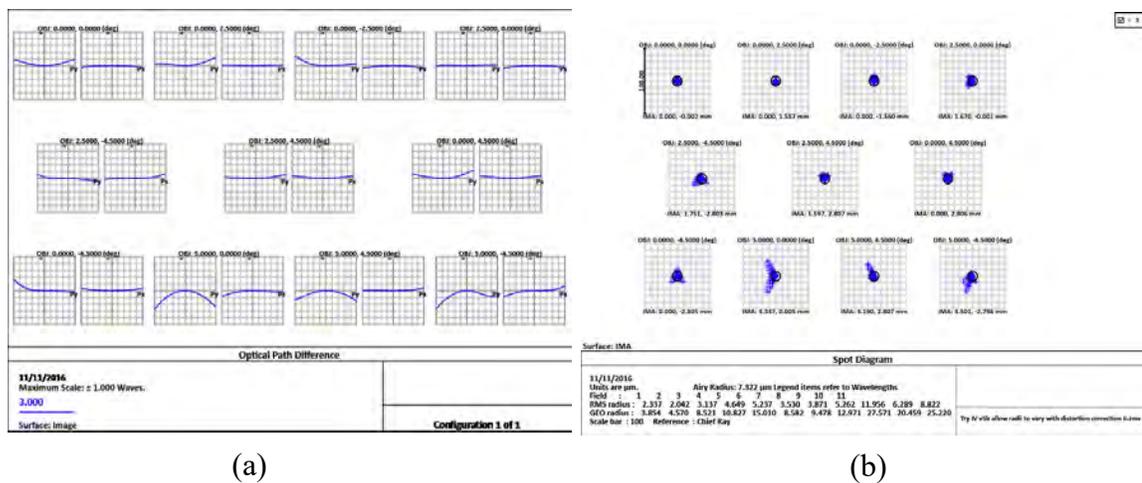


Fig. 5.9 Reoptimized three-mirror system. 8th order plane symmetric polynomials are added to the mirror surfaces. (a) OPD plots (the plot scale is 1 wave at $3\mu\text{m}$); (b) Spot diagrams.

The RMS Spot over the field is shown in Fig. 5.10. In his design, Chrisp reported the average RMS Spot size over the field to be $14\mu\text{m}$. In the current design, the average RMS Spot is about $8.5\mu\text{m}$.

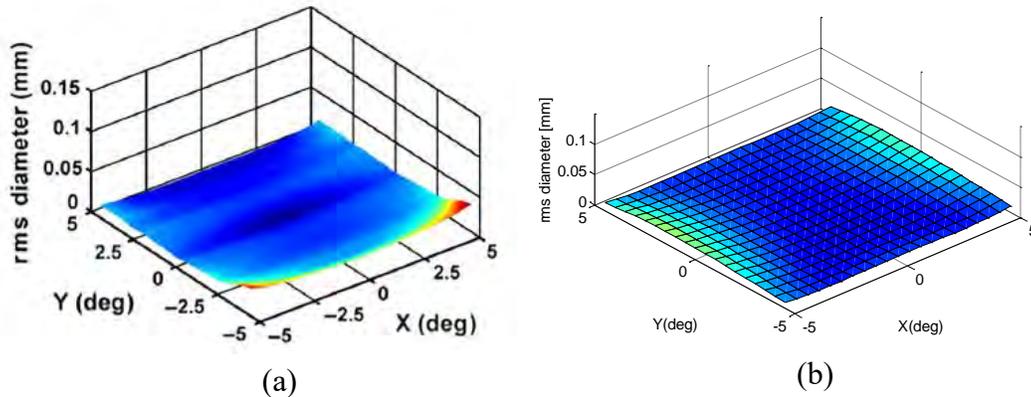


Fig. 5.10 The RMS Spot Size of the three-mirror system. (a) The design with NURBS freeform surfaces reported by Chrisp; (b) The current design with $z_3(r)$ aspherical surfaces.

Although the design presented by Crisp shows excellent performance, the surface representation with NURBS has a number of disadvantages. The major optical design programs are not capable of optimizing NURBS grid-type surfaces in imaging systems. For this reason, the optimization of the design by Chrisp was accomplished with the in-house code. Moreover, the NURBS design represents a “brute force” solution, while the aspheric surface profile of the mirrors $z_3(r)$ is clearly able to best model the ideal surface and to allow one to intelligently approach the optical design.

5.4 EXTENDED CARTESIAN OVAL SURFACE

The Cartesian Oval is an optical surface that separates two homogeneous refracting media and produces a perfect point image of a point object. In the special case of a mirror surface in which the index of refraction of object and image space media have the same magnitude but the opposite sign, the Cartesian Oval solutions are conic surfaces. Other well-known solutions are a sphere for the case of aplanatic and concentric conjugate points

or conic surface with the conic constant equal to the minus square of the index of refraction for the case of having one conjugate point at infinity. However, the general sag equation of the Cartesian Oval is complicated, and the Cartesian Oval is not readily available as a standard surface in commercial lens design software. The solution for the explicit sag of the Cartesian Oval has been previously discussed by other authors. Moreover, an alternate iterative method for the sag of the Cartesian Oval has also been provided [13].

This iterative method solves the defining optical path length equation for the Cartesian Oval for any ray from the object point O to the image point O' , as shown in Fig. 5.11.

$$OPL_p - OPL_{axis} = \left\{ n_1 \cdot \sqrt{(s_1 + S_c(r))^2 + r^2} + n_2 \cdot \sqrt{(s_2 - S_c(r))^2 + r^2} \right\} - \{ n_1 \cdot s_1 + n_2 \cdot s_2 \} = 0, \quad (5.18)$$

where $S_c(r)$ is the sag of the Cartesian Oval; n_1 and n_2 are the indexes of refraction in object and image spaces; and s_1 and s_2 are the object and image distances from the surface vertex.

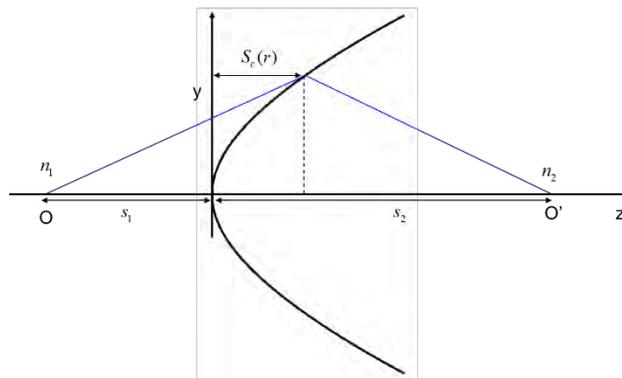


Fig. 5.11 Geometrical variables used to define the Cartesian Oval surface.

The Cartesian Oval has the property of perfectly imaging an object point into an image point with any numerical aperture. However, the imaging performance of the Cartesian Oval degrades rapidly for off-axis field positions. In this section, the Cartesian Oval is used as a base surface to create an aspherical surface. The aspherical surface $z_4(r)$ is constructed by building a polynomial on the base surface $S_c(r)$ as,

$$z_4(r) = A_1 S_c(r) + A_2 S_c^2(r) + A_3 S_c^3(r) + A_4 S_c^4(r) + A_5 S_c^5(r) + A_6 S_c^6(r) \dots, \quad (5.19)$$

where the A 's represent deformation coefficients. A freeform $z_5(r)$ surface is also constructed by a superposition of base surfaces as,

$$z_5(r) = A_1 S_{c/1}(r) + A_2 S_{c/2}(r) + A_3 S_{c/3}(r) + \dots, \quad (5.20)$$

where each of the $S_{s/n}(r)$ terms have their own independent S_1 , S_2 , and A coefficients.

The goal of this polynomial surface is to extend the off-axis imaging capabilities of the Cartesian Oval. Surface $z_4(r)$ and $z_5(r)$ were programmed as the user defined surfaces in Zemax OpticStudio optical design software [10] (see Appendix K).

Noteworthy is that the Cartesian Oval surface can take a variety of shapes according to the conjugate parameters S_1 and S_2 : concave and convex surface shapes, as well as surfaces that are concave near the center and turn smoothly into convex towards the lens edge.

For example, the surface profile $z_4(r)$ is used to fit two Pedal surfaces shown in Fig. 5.12. The Pedal surface in Fig. 5.12a has the major axis $a = 5mm$ and the minor axis $b = 15mm$. This surface is strongly concave near the center and turn smoothly into slightly

convex towards the lens edge. The Pedal surface in Fig. 5.12b has the major axis $a = 5mm$ and the minor axis $b = 10mm$. This surface is slightly concave near the center and turn smoothly into strongly convex towards the lens edge. For this fit, the surface sag at 100 points equally spaced in aperture from $r = 0$ to $r_{\max} = 10mm$ was evaluated for both surfaces. An error function was created as the sum of the squares of the differences of the surface sag $z_4(r)$ and the Pedal surface sag. Then using the damped least square and the orthogonal descent optimization methods, the error function was minimized. The surface $z_4(r)$ coefficients were used by the optimizer as degrees of freedom to reduce the error function. The index of refraction of object and images spaces $n_1 = 1$ and $n_2 = 1.5$ were chosen arbitrary. The conjugate distances s_1 and s_2 , and polynomial coefficients were released as variables and as the optimizer proceeded, more coefficients were released.

Table 5.4 summarizes the RMS fit error as a function of the number of optimization variable used in the fit. The prescription for surfaces in Table 5.4 are provided in Table 5.5 and Table 5.6. Relatively large residual error for both surfaces indicates that the Pedal curve to the ellipse provides a better representation for the concave-convex profile lenses.

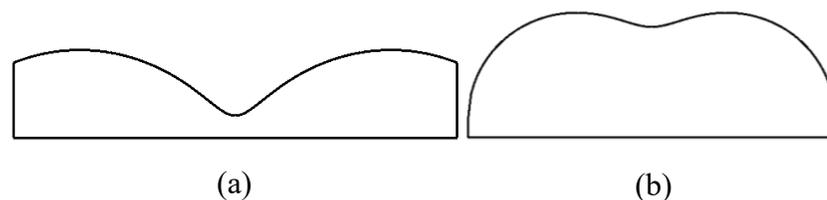


Fig. 5.12 Pedal surface profile used for the fit. (a) The major axis $a = 5mm$ and the minor axis $b = 15mm$; (b) The major axis $a = 5mm$ and the minor axis $b = 10mm$.

Table 5.4 The RMS fit error between Pedal surfaces see in Fig. 5.12 and aspherical surface $z_4(r)$ as a function of the number of optimization variables used in the fit.

Number of Parameters	RMS Fit Error for the surface in Fig. 5.12a [mm]	RMS Fit Error for the surface in Fig. 5.12b [mm]
3	1.92250	0.61958
4	0.60046	0.39904
5	0.22685	0.35787
6	0.10811	0.32312
7	0.09792	0.29325
8	0.08932	0.28577

Table 5.5 Coefficients defining the surface profile $z_4(r)$ used to fit the Pedal surface in Fig. 5.12a.

# of Par.	$s_1[mm]$	$s_2[mm]$	A1	A2	A3	A4	A5	A6
3	4.40735	1.00865	-3.67178					
4	4.41304	1.02403	-5.23497	-2.22104				
5	4.41560	1.03132	-6.33459	-5.99284	-2.82966			
6	4.63623	1.05380	-7.85547	-1.32626	-1.50141	-6.52185		
7	4.63617	1.05365	-8.01806	-1.46306	-1.88774	-1.10109	-1.85167	
8	4.63633	1.05405	-7.93003	-1.42684	-1.87234	-1.13041	-1.31116	7.92262

Table 5.6 Coefficients defining the surface profile $z_4(r)$ used to fit the Pedal surface in Fig. 5.12b.

# of Par.	$s_1[mm]$	$s_2[mm]$	A1	A2	A3	A4	A5	A6
3	4.75887	8.69102	-3.32689					
4	4.29490	8.11311	-2.83046	-6.05352				
5	4.29357	8.09574	-2.70601	-7.89477	-8.82643			
6	4.29356	8.09564	-2.88590	-6.91806	1.31532	-2.58629		
7	4.29264	8.08377	-2.91148	8.41478	2.10056	-1.08043	7.67221	
8	4.29221	8.07830	-2.83740	1.42260	1.90860	-1.58230	2.61426	-1.31240

As an example of the performance of the proposed surface, a 1:1 relay imaging lens composing from two identical singlet elements was designed, as shown in Fig. 5.14. The

system is symmetrical about the aperture stop. Both elements are plano-convex lenses with the aspheric convex surface profile defined by $z_4(r)$. The lens operates at $f/2$ and has a focal length of 20mm.

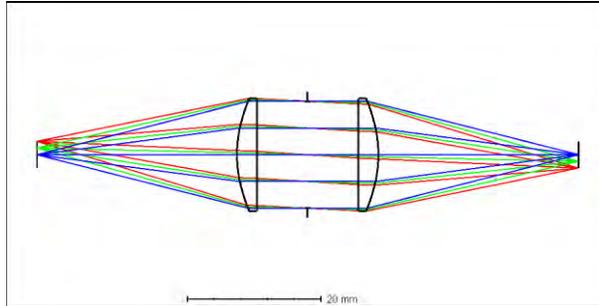


Fig. 5.13 Layout of the 1:1 relay imaging lens designed with the extended Cartesian Oval surfaces.

The design procedure is the following. First, the system is constructed with Cartesian Oval surfaces such that the imaging on axis is stigmatic surface after surface. Parameters s_1 and s_2 for the Cartesian Oval surfaces are selected to achieve a perfectly collimated beam at the stop. Next, a 3mm field of view is added, and polynomial coefficients are released to balance the off-axis aberrations. In the final design, a balanced performance over the entire field is achieved.

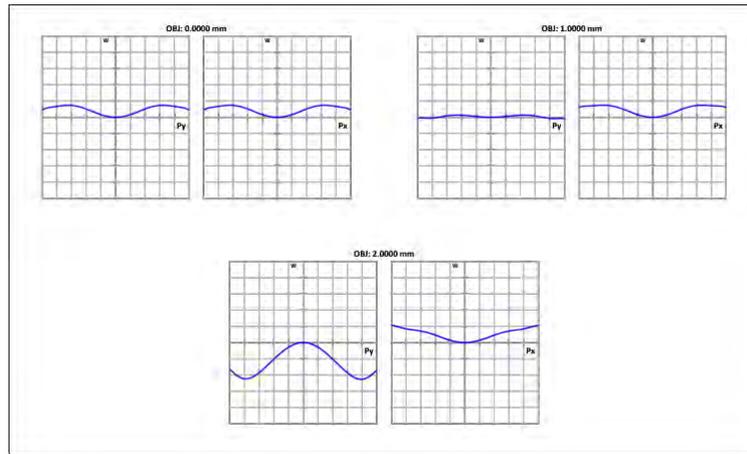


Fig. 5.14 OPD plots of the reoptimized 1:1 relay lens (the plot scale is 5 waves at 588nm).

5.5 CONCLUSION

In this chapter, the concept of a base surface from which an aspheric polynomial surface can be constructed by power expansion of the base term, and from which a freeform surface can also be constructed by superposition of several base surfaces having different parameter values was introduced.

The surfaces $z_1(r)$ and $z_2(r)$ that are presented in this chapter are useful for providing specific illumination distributions. Notably, the freeform surface $z_2(r)$ exhibits little oscillation in the relative illumination or transverse ray error.

The surfaces $z_3(r)$, $z_4(r)$ and $z_5(r)$ are useful for designing systems where the imaging is stigmatic surface after surface along a particular ray. Moreover, these surfaces provide additional degrees of freedom to balance aberration for off-axis field positions.

REFERENCES

1. G. Shultz, *Aspheric surfaces*, Progress in optics Vol. **XXV**, 349-415, (1988).
2. B. Braunecker, R. Hentschel, and H. J. Tiziani, *Advanced Optics Using Aspherical Elements*, SPIE Press, Bellingham Washington (2008).
3. J. E. Stacy, "Asymmetric spline surfaces: characteristics and applications," Appl. Opt. **23**, 2710-2714, (1984).
4. P. Jester, C. Menke, and K. Urban, "B-spline representation of optical surfaces and its accuracy in a ray trace algorithm," Appl. Opt. **50**, 822-828 (2011).
5. H. Gross, et. al., "Overview on surface representations for freeform surfaces," Proc. SPIE **9626**, 96260U (2015).
6. S. Vives, et. al., "Modeling highly Aspheric optical surfaces using a new polynomial formalism into Zemax," Proc. SPIE **8167**, 81670B, (2011).
7. S. Pascal, et. al., "New modeling of freeform surfaces for optical design of astronomical instruments," Proc. SPIE **8450**, 845053-1, (2012).
8. M. Chrisp, "New freeform NURBS imaging design code," Proc. SPIE **9293**, 02930N (2014).
9. J. Sasian, D. Reshidko and C. Li, "An aspheric/freeform optical surface description for controlling illumination from point-like light sources," Optical Engineering (2016)
10. OpticStudio 16 SP2, Zemax LLC.
11. J. Nelson and M. Temple-Raston, "The off-axis expansion of conic surfaces," UC TMT Report **91** (1982)
12. J. Sasian, "How to approach the design of a bilateral symmetric optical system," Optical Engineering, Vol. **33 (6)** (1994)
13. C. Hsueh, T. Elazhary, M. Nakano and J. Sasian, "Closed-form sag solutions for Cartesian Oval surfaces," J. Opt., Vol. **40 (4)** (2011)

CHAPTER VI: CONCLUSION

This dissertation presented of a collection of topics that address current problems and applications of state-of-the-art imaging and illumination system lens design. Each topic was introduced independently as a self-contained and self-explanatory chapter.

In Chapter II, several useful principles and techniques for evaluation of complex optical imaging systems have been presented and discussed. These methods and techniques were applied to analyze state-of-the-art mobile camera lenses and to explore how a curved image surface can benefit the lens design of these optical systems.

The ideas proposed in this chapter are that (1) two applicable parameters for quantify relevant lens attributes and providing adequate comparison between different optical configurations are the optical power distribution and lens symmetry; (2) the examination of the ray invariant product $n \cdot \sin(i)$ for a critical ray path is a widely applicable technique that will often indicate the source of a design problem; (3) understanding the limiting aberrations of a lens is critical in order to further improve the design by applying an appropriate technique to correct these specific limiting aberrations or choosing a different lens configuration to keep them from arising; (4) a better understanding of the intrinsic and extrinsic surface contributions to the overall aberration balance is essential for designing efficient imaging lenses.

Chapter III presented a second- and a fourth-order theory of irradiance changes in axially symmetric optical systems. The concept of the irradiance function was reviewed and an interpretation of the irradiance aberrations was discussed.

The irradiance function terms represent basic distribution patterns in the irradiance of a beam at the exit pupil plane or at the image plane of an imaging system. The irradiance coefficients were found via basic radiometric principles, such as conservation of flux, and were derived from purely geometric considerations. Specific formulas and relationships between the irradiance distribution and wavefront aberration coefficients to fourth order were provided for irradiance at the image and at the exit pupil of an optical system. The practical case of relative illumination at the image plane of an optical system was also discussed in some detail.

The theory of irradiance aberrations enhances our knowledge about the behavior of light as it propagates in optical systems and provides insights into how individual wavefront aberration terms affect the light irradiance produced by a lens system at its image plane or at the exit pupil plane.

Two methods for chromatic aberration correction have been presented in Chapter IV. Both methods are based on real ray tracing and can be implemented in commercial lens design software. The idea is to separate the chromatic aberration content from the monochromatic aberration content and use standard lens design software optimization tools, such as global optimization and iterative glass substitution. The separation of aberrations removes the dependence in correction of one aberration type from the

correction of other aberration type. In that case the optimization process is focused on reducing chromatic aberration to the monochromatic aberration value.

The presented methods are not limited to any specific bandwidth, can be used to correct both chromatic change of focus and chromatic change of magnification, support multi-configuration systems, and do not require any additional/external calculations. Unlike other methods found in the literature, the presented methods are not limited to the finite number of aberration coefficients used in an error function, since real ray tracing is used to reduce the chromatic aberrations. As an example of application of the methods, the chromatic aberration correction in a Canon telephoto lens and in an elaborated zoom system was improved.

The ray tracing methods were further extended to a different glass selection problem – athermalization of imaging systems. The algorithm was modified to separate aberrations induced by temperature effects from the nominal design aberrations. In that case the optimization process is focused on reducing additional aberration due to thermal effects to the nominal aberration value.

In Chapter V, the concept of a base surface from which an aspheric polynomial surface can be constructed by power expansion of the base term, and from which a freeform surface can also be constructed by superposition of several base surfaces having different parameter values was introduced.

Surfaces that were presented in this chapter are useful for providing specific illumination distributions or for designing systems where the imaging is stigmatic surface after surface along a particular ray. Examples of the performance of these surfaces that

were constructed using the concept of base surface were provided, and some of their properties were discussed.

The theoretical results and practical methods described in this dissertation serve to provide new insights into the optical design of state-of-the-art imaging and illumination systems, to advance the theory of aberrations, and to enhance the optics community's collective understanding of present topics in modern optical engineering.


```

! Chief ray
  RAYTRACE 0,1,0,0

  yc=rayy(k)
  uc=RAYM(k)/RAYN(k)
  ucb=RAYM(k-1)/RAYN(k-1)
  B=nkb*(yc/radk+ucb)
  delc=uc/nk-ucb/nkb

! Skip a dummy surface.
  IF (nk==nkb)
    t=t+1
    GOTO 1
  ENDIF

  LG=-nk*(um*yc-uc*ym)

  Wk = -1/(1-pmag)*ym/(ni*pumi)*dnk/radk
  W = W + Wk*Wk
  Wss = Wss + Wk

  Sk = 1/(1-pmag)*B*delm/(Bstop*ni*pumi)
  S = S + Sk*Sk
  Sss = Sss + Sk

  FORMAT 10.2
  IF (k==1)
    Print "   Surface   W       S"
  ENDIF

  Print k," ", wk," ",Sk
  LABEL 1

NEXT
Print
Print "   SUM", " ", wss," ",Sss

W=SQRT(W/(N-t))
S=SQRT(S/(N-t))

PARAXIAL OFF

PRINT " "
PRINT "Power Distribution   W: ", W
PRINT "Degree of Symmetry   S: ", S

Optreturn 1, w
Optreturn 2, s
Optreturn 3, wss
Optreturn 4, sss

```

APPENDIX B:
EVALUATION OF THE RAY INVARIANT PRODUCT
MACRO FOR ZEMAX

```

! Written by Chia-Ling Li, modified by Dmitry Reshidko

DECLARE surf, double, 1, 60
DECLARE chief, double, 1, 60
DECLARE marg, double, 1, 60
DECLARE comb, double, 1, 60

!*****
FORMAT 10.5

PRINT
PRINT "  Surface# NsinI_chief(1,0)  NsinI_marginal(0,1) NsinI_specific(1,1)"
PRINT
!*****

!Starts the surface by surface calculation with a loop

j=0

FOR i=1, NSUR(), 1

  dummy=indx(i-1)-indx(i)
  IF (dummy == 0)
    GOTO 1
  else

    j=j+1
    C=OCOD("RAED")

    !Trace chief ray after surface
    Ic=OPEV(C,i,pwav(),0,1,0,0)
    chief(j)=indx(i)*sine(Ic/180*3.14159265359)

    !Trace marginal ray after surface
    Im=OPEV(C,i,pwav(),0,0,0,1)
    marg(j)=indx(i)*sine(Im/180*3.14159265359)

    !Trace specific ray after surface
    Is=OPEV(C,i,pwav(),0,1,0,1)
    comb(j)=indx(i)*sine(Is/180*3.14159265359)

  surf(j)=i

PRINT " ",surf(j)," ",chief(j), "      ",marg(j), "      ",comb(j)

```

```
    ENDIF
    Label 1
NEXT

!*****
!Make a plot

PLOT NEW
PLOT TITLE, "surface# vs n*sin(I)"
PLOT TITLEx, "surface#"
PLOT TITLEY, "n*sin(I)"
PLOT BANNER, "RESULTS GENERATED USING NsinI.ZPL"
PLOT FORMATX, "%2.0f"
PLOT FORMATY, "%1.2f"
PLOT DATA, surf, chief, j, 0, 0, 0
PLOT DATA, surf, marg, j, 1, 0, 0
PLOT DATA, surf, comb, j, 5, 0, 0
PLOT COMM1, "black curve : chief ray (1,0)"
PLOT COMM2, "blue curve : marginal ray (0,1)"
PLOT COMM3, "pink curve : specific ray (1,1)"
PLOT GO
```

APPENDIX C:

EVALUATION OF HIGHER ORDER ABERRATION

CONTENT MACRO FOR ZEMAX

```

! Written by Dmitry Reshidko
! Float by stop size aperture
! Remove pickups and solves

!Compute higher order contribution
DECLARE DISP_arr,double,1,5

start_surf = 1
nominal_object_units = 0 # 0 - Angle; 1 - Object height
vis = 1
plot_scale = 0
plot_scale_s = 0
DISP_arr (1) = 1 # SA
DISP_arr (2) = 1 # COMA
DISP_arr (3) = 1 # ASTI
DISP_arr (4) = 1 # FC
DISP_arr (5) = 0 # DIST

p_wave = PWAV()
p_wave_um = WAVL(p_wave)
C=OCOD("OPDC")
C1=OCOD("SPHA")
C2=OCOD("COMA")
C3=OCOD("ASTI")
C4=OCOD("FCUR")
C5=OCOD("DIST")
CF=OCOD("FCGS")
CFN=OCOD("SFNO")
CA=OCOD("RAIN")

print
n_surf = NSUR()

! Set ray aiming to real
SYSP 70, 2

DECLARE ABER_A_arr,double,2,5,n_surf
DECLARE ABER_B_arr,double,2,5,n_surf
DECLARE ABER_TOTAL_arr,double,2,5,n_surf
DECLARE ABER_EX_arr,double,2,5,n_surf
DECLARE ABER_INT_arr,double,2,5,n_surf
DECLARE ABER_SEIDEL_arr,double,2,5,n_surf
DECLARE ABER_TOTAL_SUR_arr,double,1,5
DECLARE RADII_arr,double,1,n_surf+2
DECLARE COLOR_arr,double,1,5

```

```

DECLARE ASPH4_arr,double,1,n_surf+2
DECLARE ASPH6_arr,double,1,n_surf+2
DECLARE ASPH8_arr,double,1,n_surf+2
DECLARE ASPH10_arr,double,1,n_surf+2
DECLARE ASPH12_arr,double,1,n_surf+2
DECLARE ASPH14_arr,double,1,n_surf+2
DECLARE ASPH16_arr,double,1,n_surf+2

COLOR_arr(1) = 3
COLOR_arr(2) = 2
COLOR_arr(3) = 7
COLOR_arr(4) = 6
COLOR_arr(5) = 24

! save nominal conjugates
nom_obj_dist = THIC(0)
obj_dist = nom_obj_dist

! Calculate pupil distance and diameter
GETSYSTEMDATA 1
restore_pupil = 0
f_num_w = VEC1(10)
IF (VEC1(23)≠1)

nom_pupil_r = VEC1(11)
nom_pupil_d = VEC1(12)
nom_stop_surf = VEC1(23)

! Move stop being the first surface
INSERT 1
SURF 1,THIC,nom_pupil_d*-1
SURF 1, SDIA, nom_pupil_r/2
SURF 0,THIC,THIC(0)+nom_pupil_d
STOPSURF 1
obj_dist = THIC(0)
n_surf = n_surf + 1
restore_pupil = 1

IF (nom_pupil_d>1e5)
print "System telecentric object space"
ENDIF

ENDIF

start_surf = start_surf + 1

UPDATE
! Determine max field
PARAXIAL ON
RAYTRACE 0, 1, 0, 0, p_wave
IF RAYE() THEN print "ERROR!"
nom_max_field = RAYY(n_surf)

```

```

PARAXIAL OFF
num_fields = NFLD()
nom_max_field_deg = FLDY(num_fields)
SYSP 101, 2

! set solve
INSERT n_surf
n_surf = n_surf + 1
SOLVETYPE n_surf-1, TM, 0, 0, 0, 0
UPDATE

! Consider vignetting
FLAG_B = 0

FOR i_surf, start_surf, n_surf-1, 1

temp$ = $GLASS(i_surf)
SURP i_surf, COMM, temp$
temp$ = ""
SURP i_surf, GLAS, temp$

NEXT

DECLARE ABER_arr, double, 2, 5, n_surf-1
DECLARE SEIDEL_arr, double, 2, 5, n_surf-1

UPDATE

print "W040W131W222W220W311"
print "W0n0WnmlWnmmWnm0Wn11"
print

SA_siedel_total = 0
COMA_siedel_total = 0
ASTI_siedel_total = 0
FC_siedel_total = 0
DIST_siedel_total = 0

SA_all_total = 0
COMA_all_total = 0
ASTI_all_total = 0
FC_all_total = 0
!DIST_all_total = 0

max_scale = 0
max_scale_s = 0

FOR i_surf, start_surf, n_surf-2, 1

!!!! Calculate aberrations for A !!!!!
STOPSURF 1

```

```

! Restore glass in A
FOR j_surf, 0, i_surf-1, 1
dummy = SPRO(j_surf, 1)
temp$ = $buffer()
SURP j_surf, GLAS, temp$
NEXT

! insert XP
INSERT i_surf

! Remove glass in B
dummy = SPRO(i_surf-1, 1)
temp$ = $buffer()
test$ = "MIRROR"
IF (temp$ $== test$)
temp$ = ""
ENDIF
FOR j_surf, i_surf, n_surf+1, 1
SURP j_surf, GLAS, temp$
NEXT

! Remove fields
SYSP 100, nominal_object_units
SYSP 103, 2, 0

UPDATE

!! Here we go to XP !!

! Calculate XP distance distance

nominal_i_thick = THIC(i_surf-1)
total_i_thick = 0
FOR j_surf, i_surf-1, n_surf, 1
total_i_thick = total_i_thick + THIC(j_surf)
NEXT

GETSYSTEMDATA 1
xp_pupil_diam = VEC1(13)
xp_pupil_dist = VEC1(14)

!!IF (i_surf==3) THEN GOTO DEBUG

! Move stop to the XP

SURP i_surf-1, THIC, total_i_thick+xp_pupil_dist
SURP i_surf, SDIA, xp_pupil_diam/2
STOPSURF i_surf

!! End go to XP !!

IF ((total_i_thick+xp_pupil_dist)>1e5)

```

```

print i_surf,"-A telecentric image space"
ENDIF

IF (THIC(n_surf-1)>1e5)
print i_surf,"-A afocal"
ENDIF

! Change fields
SYSP 100, nominal_object_units
SYSP 103, 2, nom_max_field_deg

! Evaluate aberration for part A

GOSUB EVALUATE_ABERRATIONS

ABER_A_arr(1,i_surf) = SA_all
ABER_A_arr(2,i_surf) = COMA_all
ABER_A_arr(3,i_surf) = ASTI_all
ABER_A_arr(4,i_surf) = FC_all
!ABER_A_arr(5,i_surf) = DIST_all

!IF (i_surf==3) THEN GOTO DEBUG

! Move stop back and restore
STOPSURF 1
SURP i_surf-1,THIC,nominal_i_thick
DELETE i_surf

!!!! Calculate TOTAL aberrations !!!!!

! Restore glass in A
FOR j_surf, 0, i_surf, 1
dummy = SPRO(j_surf, 1)
temp$ = $buffer()
SURP j_surf,GLAS, temp$
NEXT

! insert XP
INSERT i_surf+1

! Remove glass in B
dummy = SPRO(i_surf, 1)
temp$ = $buffer()
test$ = "MIRROR"
IF (temp$ $== test$)
temp$ = ""
ENDIF
FOR j_surf, i_surf+1, n_surf+1, 1
SURP j_surf,GLAS, temp$
NEXT

```

```

! Remove fields
SYSP 100, nominal_object_units
SYSP 103, 2, 0

UPDATE

!! Here we go to XP !!

! Calculate XP distance distance

nominal_i_thick = THIC(i_surf)
total_i_thick = 0
FOR j_surf, i_surf, n_surf, 1
total_i_thick = total_i_thick + THIC(j_surf)
NEXT

GETSYSTEMDATA 1
xp_pupil_diam = VEC1(13)
xp_pupil_dist = VEC1(14)

! Move stop to the XP

SURP i_surf, THIC, total_i_thick + xp_pupil_dist
SURP i_surf + 1, SDIA, xp_pupil_diam / 2
STOPSURF i_surf + 1

!! End go to XP !!

IF ((total_i_thick + xp_pupil_dist) > 1e5)
print i_surf, "-TOTAL telecentric image space"
ENDIF

IF (THIC(n_surf - 1) > 1e5)
print i_surf, "-TOTAL afocal"
ENDIF

! Change fields
SYSP 100, nominal_object_units
SYSP 103, 2, nom_max_field_deg

! Evaluate Total aberrations !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

GOSUB EVALUATE_ABERRATIONS

ABER_TOTAL_arr(1, i_surf) = SA_all
ABER_TOTAL_arr(2, i_surf) = COMA_all
ABER_TOTAL_arr(3, i_surf) = ASTI_all
ABER_TOTAL_arr(4, i_surf) = FC_all
!ABER_TOTAL_arr(5, i_surf) = DIST_all

ABER_SEIDEL_arr(1, i_surf) = OPEV(C1, i_surf, p_wave, 0, 0, 0, 0)
ABER_SEIDEL_arr(2, i_surf) = OPEV(C2, i_surf, p_wave, 0, 0, 0, 0)

```

```

ABER_SEIDEL_arr(3,i_surf) = OPEV(C3,i_surf,p_wave,0,0,0,0)
ABER_SEIDEL_arr(4,i_surf)      =      OPEV(C4,i_surf,p_wave,0,0,0,0)      +
0.5*OPEV(C3,i_surf,p_wave,0,0,0,0)
ABER_SEIDEL_arr(5,i_surf) = OPEV(C5,i_surf,p_wave,0,0,0,0)

! Determine max field
PARAXIAL ON
RAYTRACE 0, 1, 0, 0, p_wave
IF RAYE() THEN print "ERROR!"
nom_max_field = RAYY(n_surf+1)
PARAXIAL OFF

!IF (i_surf==3) THEN GOTO DEBUG

! Move stop back and restore
STOPSURF 1
SURP i_surf,THIC,nominal_i_thick
DELETE i_surf+1

!!!! Calculate aberrations for B !!!!!

! Restore glass in A
FOR j_surf, 0, i_surf-1, 1
dummy = SPRO(j_surf, 1)
temp$ = $buffer()
SURP j_surf,GLAS, temp$
NEXT

! Remove glass in B
dummy = SPRO(i_surf-1, 1)
temp$ = $buffer()
test$ = "MIRROR"
IF (temp$ $== test$)
temp$ = ""
ENDIF
FOR j_surf, i_surf, n_surf+1, 1
SURP j_surf,GLAS, temp$
NEXT

UPDATE

! Create coupling stop

thick_pupil = 0
FOR j_surf, i_surf, n_surf, 1
thick_pupil = thick_pupil + THIC(j_surf)
NEXT

GETSYSTEMDATA 1
xp_pupil_diam = VEC1(13)
xp_pupil_dist = VEC1(14)

```

```

INSERT i_surf

IF (ABSO(xp_pupil_dist)>1e9)
SURP i_surf,THIC,THIC(i_surf-1)
ELSE
SURP i_surf,THIC,(xp_pupil_dist + thick_pupil)*-1
IF ((xp_pupil_dist + thick_pupil)>1e5)
print i_surf,"-B telecentric object space"
ENDIF
ENDIF

SURP i_surf,GLAS, temp$
SURP i_surf, SDIA, xp_pupil_diam/2
STOPSURF i_surf

! Adjust the conjugates
thick_object = 0
FOR j_surf, i_surf-1, 1, -1
thick_object = thick_object + THIC(j_surf)
NEXT

SURP 0,THIC,(-thick_object + xp_pupil_dist)

! insert XP
INSERT i_surf+2

! Remove glass in A
FOR j_surf, 0, i_surf-1, 1
SURP j_surf,GLAS, temp$
NEXT

! Set glass in B
dummy = SPRO(i_surf+1, 1)
temp$ = $buffer()
test$ = "MIRROR"
IF (temp$ $== test$)
SURP i_surf+1,GLAS, temp$
temp$ = ""
FOR j_surf, i_surf+2, n_surf+2, 1
SURP j_surf,GLAS, temp$
NEXT
ELSE
FOR j_surf, i_surf+1, n_surf+2, 1
SURP j_surf,GLAS, temp$
NEXT
ENDIF

! Remove fields
SYSP 100, 2
SYSP 103, 2, 0

UPDATE

```

```

!! Here we go to XP !!

! Calculate XP distance distance

nominal_i_thick = THIC(i_surf+1)
total_i_thick = 0
FOR j_surf, i_surf+1, n_surf+1, 1
total_i_thick = total_i_thick + THIC(j_surf)
NEXT

GETSYSTEMDATA 1
xp_pupil_diam = VEC1(13)
xp_pupil_dist = VEC1(14)

! Move stop to the XP

SURF i_surf+1, THIC, total_i_thick+xp_pupil_dist
SURF i_surf+2, SDIA, xp_pupil_diam/2
STOPSURF i_surf+2

!! End go to XP !!

IF (ABS(total_i_thick+xp_pupil_dist)>1e5)
print i_surf, "-B telecentric image space"
ENDIF

IF (ABS(THIC(n_surf-1))>1e5)
print i_surf, "-B afocal"
ENDIF

! Change fields
SYSP 100, 2
SYSP 103, 2, nom_max_field

!IF (i_surf==6) THEN GOTO DEBUG

! Evaluate aberration for part B !!!!!!!!!!!!!!!
FLAG_B = 1
GOSUB EVALUATE_ABERRATIONS
ABER_B_arr(1, i_surf) = SA_all

IF (nom_max_field>0)
ABER_B_arr(2, i_surf) = COMA_all
print "+", GLCZ(n_surf+2)-GLCZ(i_surf+1)
ELSE
ABER_B_arr(2, i_surf) = -1*COMA_all
print i_surf, "  -", GLCZ(n_surf+2)-GLCZ(i_surf+1)
ENDIF

IF (nominal_object_units)
IF (nom_max_field>0)

```

```

ABER_B_arr(2,i_surf) = -1*COMA_all
print "+", GLCZ(n_surf+2)-GLCZ(i_surf+1)
ELSE
ABER_B_arr(2,i_surf) = COMA_all
print i_surf,"  -", GLCZ(n_surf+2)-GLCZ(i_surf+1)
ENDIF
ENDIF

ABER_B_arr(3,i_surf) = ASTI_all
ABER_B_arr(4,i_surf) = FC_all
!ABER_B_arr(5,i_surf) = DIST_all
FLAG_B=0

! Move stop back and restore
STOPSURF 1
SURP i_surf+1,THIC,nominal_i_thick
DELETE i_surf+2
DELETE i_surf
SURP 0,THIC,obj_dist

! Restore fields
SYSP 100, nominal_object_units
SYSP 103, 2, nom_max_field_deg

UPDATE

NEXT

FOR i_surf, 0, n_surf-1, 1
dummy = SPRO(i_surf, 1)
temp$ = $buffer()
SURP i_surf,GLAS, temp$
NEXT

! Calculating extrinsic aberrations
FOR i_aber=1,5,1
FOR i_surf, start_surf, n_surf-2, 1
ABER_EX_arr(i_aber,i_surf) = ABER_TOTAL_arr(i_aber,i_surf) - ABER_A_arr(i_aber,i_surf) -
ABER_B_arr(i_aber,i_surf)
ABER_INT_arr(i_aber,i_surf) = ABER_B_arr(i_aber,i_surf)-ABER_SEIDEL_arr(i_aber,i_surf)
IF ((ABSO(ABER_EX_arr(i_aber,i_surf))> max_scale) & DISP_arr(i_aber)) THEN
max_scale=ABSO(ABER_EX_arr(i_aber,i_surf))
ABER_TOTAL_SUR_arr(i_aber) = ABER_TOTAL_SUR_arr(i_aber) + ABER_EX_arr(i_aber,i_surf)
+ ABER_INT_arr(i_aber,i_surf)
NEXT
NEXT

IF (plot_scale>0) THEN max_scale=plot_scale

! Restore system
DELETE n_surf-1

```

```

SURP 0,THIC,nom_obj_dist

! restore EP
IF (restore_pupil==1)
STOPSURF nom_stop_surf+1
DELETE 1
ENDIF

UPDATE

! printing
FORMAT 0.4
print
print "W0n0  Wnml  Wnmm  Wnm0  Wn11"
print
FOR i_surf, start_surf, n_surf-2, 1
IF restore_pupil==1
print i_surf-1
ELSE
print i_surf
ENDIF
print "S:",ABER_SEIDEL_arr(1,i_surf), "", ABER_SEIDEL_arr(2,i_surf), "",
ABER_SEIDEL_arr(3,i_surf), "", ABER_SEIDEL_arr(4,i_surf), "", ABER_SEIDEL_arr(5,i_surf)
print "I:",ABER_INT_arr(1,i_surf), "", ABER_INT_arr(2,i_surf), "", ABER_INT_arr(3,i_surf), "",
ABER_INT_arr(4,i_surf)
print "E:",ABER_EX_arr(1,i_surf), "", ABER_EX_arr(2,i_surf), "", ABER_EX_arr(3,i_surf), "",
ABER_EX_arr(4,i_surf)
print "TS:",ABER_INT_arr(1,i_surf)+ABER_EX_arr(1,i_surf), "",
ABER_INT_arr(2,i_surf)+ABER_EX_arr(2,i_surf), "",
ABER_INT_arr(3,i_surf)+ABER_EX_arr(3,i_surf), "",
ABER_INT_arr(4,i_surf)+ABER_EX_arr(4,i_surf)
print
print

NEXT

print
print "TOTAL:"
print OPEV(C1,0,p_wave,0,0,0), "", OPEV(C2,0,p_wave,0,0,0), "", OPEV(C3,0,p_wave,0,0,0), "
", OPEV(C4,0,p_wave,0,0,0) + 0.5*OPEV(C3,0,p_wave,0,0,0)
print ABER_TOTAL_arr(1,n_surf-2)-OPEV(C1,0,p_wave,0,0,0), "", ABER_TOTAL_arr(2,n_surf-2)-
OPEV(C2,0,p_wave,0,0,0), "", ABER_TOTAL_arr(3,n_surf-2)-OPEV(C3,0,p_wave,0,0,0), "",
ABER_TOTAL_arr(4,n_surf-2)-(OPEV(C4,0,p_wave,0,0,0) + 0.5*OPEV(C3,0,p_wave,0,0,0))
print
print "CHECK:"
print ABER_TOTAL_SUR_arr(1), "", ABER_TOTAL_SUR_arr(2), "", ABER_TOTAL_SUR_arr(3), "
", ABER_TOTAL_SUR_arr(4)
print

print "T", "", OPEV(C1,0,p_wave,0,0,0), "",
ABER_TOTAL_arr(1,n_surf-2)-
OPEV(C1,0,p_wave,0,0,0), "", , "", OPEV(C2,0,p_wave,0,0,0), "",

```

```

print ABER_TOTAL_arr(2,n_surf-2)-OPEV(C2,0,p_wave,0,0,0,0), "", "", OPEV(C3,0,p_wave,0,0,0,0) ,
""
print ABER_TOTAL_arr(3,n_surf-2)-OPEV(C3,0,p_wave,0,0,0,0), "", "", OPEV(C4,0,p_wave,0,0,0,0)
+ 0.5*OPEV(C3,0,p_wave,0,0,0,0) , "", ABER_TOTAL_arr(4,n_surf-2)-(OPEV(C4,0,p_wave,0,0,0,0) +
0.5*OPEV(C3,0,p_wave,0,0,0,0))

IF VIS

declare x_arr, double,1,7
declare y_arr, double,1,7

PLOT NEW

FOR i_surf, start_surf, n_surf-2, 1

ind=0.0
FOR i_aber=1,5,1
IF DISP_arr (i_aber)
height = ABER_EX_arr(i_aber,i_surf)
!IF (abso(height)>max_scale) THEN height = max_scale
x_arr(1) = i_surf+ind+0.01
x_arr(2) = i_surf+ind+0.01
x_arr(3) = i_surf+ind+0.19
x_arr(4) = i_surf+ind+0.19
x_arr(5) = i_surf+ind+0.01
x_arr(6) = i_surf+ind+0.19
x_arr(7) = i_surf+ind+0.01
y_arr(1) = 0
y_arr(2) = height
y_arr(3) = height
y_arr(4) = 0
y_arr(5) = height
y_arr(6) = height
y_arr(7) = 0
PLOT DATA, x_arr, y_arr, 7,COLOR_arr(i_aber)
ENDIF
ind = ind+0.2
NEXT
NEXT

PLOT TITLE, "High order surface contribution"
PLOT TITLEX, "Surface number"
PLOT TITLEY, "Aberration [waves]"
PLOT RANGEY, -max_scale,max_scale
PLOT RANGEX, start_surf, n_surf-1
PLOT FORMATX, "%0.0f"
PLOT FORMATY, "%0.2f"
PLOT TICK, 1, max_scale/4
PLOT GO

ENDIF

```

```

LABEL DEBUG

END

SUB EVALUATE_ABERRATIONS

IF !FLAG_B
RADI_arr(i_surf+1) = RADI(i_surf+1)
SURP i_surf+1, CURV, 0
ASPH4_arr (i_surf+1) = PARM(2,i_surf+1)
SURP i_surf+1, PARM, 0, 2
ASPH6_arr (i_surf+1) = PARM(3,i_surf+1)
SURP i_surf+1, PARM, 0, 3
ASPH8_arr (i_surf+1) = PARM(4,i_surf+1)
SURP i_surf+1, PARM, 0, 4
ASPH10_arr (i_surf+1) = PARM(5,i_surf+1)
SURP i_surf+1, PARM, 0, 5
ASPH12_arr (i_surf+1) = PARM(6,i_surf+1)
SURP i_surf+1, PARM, 0, 6
ASPH14_arr (i_surf+1) = PARM(7,i_surf+1)
SURP i_surf+1, PARM, 0, 7
ASPH16_arr (i_surf+1) = PARM(8,i_surf+1)
SURP i_surf+1, PARM, 0, 8
ENDIF

FOR j_surf,i_surf+2,n_surf,1
RADI_arr(j_surf) = RADI(j_surf)
SURP j_surf, CURV, 0
ASPH4_arr (j_surf) = PARM(2,j_surf)
SURP j_surf, PARM, 0, 2
ASPH6_arr (j_surf) = PARM(3,j_surf)
SURP j_surf, PARM, 0, 3
ASPH8_arr (j_surf) = PARM(4,j_surf)
SURP j_surf, PARM, 0, 4
ASPH10_arr (j_surf) = PARM(5,j_surf)
SURP j_surf, PARM, 0, 5
ASPH12_arr (j_surf) = PARM(6,j_surf)
SURP j_surf, PARM, 0, 6
ASPH14_arr (j_surf) = PARM(7,j_surf)
SURP j_surf, PARM, 0, 7
ASPH16_arr (j_surf) = PARM(8,j_surf)
SURP j_surf, PARM, 0, 8
NEXT

IF FLAG_B
FOR j_surf,1,i_surf,1
RADI_arr(j_surf) = RADI(j_surf)
SURP j_surf, CURV, 0
ASPH4_arr (j_surf) = PARM(2,j_surf)
SURP j_surf, PARM, 0, 2
ASPH6_arr (j_surf) = PARM(3,j_surf)
SURP j_surf, PARM, 0, 3

```

```

ASPH8_arr(j_surf) = PARM(4,j_surf)
SURP j_surf, PARM, 0, 4
ASPH10_arr(j_surf) = PARM(5,j_surf)
SURP j_surf, PARM, 0, 5
ASPH12_arr(j_surf) = PARM(6,j_surf)
SURP j_surf, PARM, 0, 6
ASPH14_arr(j_surf) = PARM(7,j_surf)
SURP j_surf, PARM, 0, 7
ASPH16_arr(j_surf) = PARM(8,j_surf)
SURP j_surf, PARM, 0, 8
NEXT
ENDIF

!IF ((i_surf==8) & FLAG_B) THEN GOTO DEBUG

UPDATE

H0 = OPEV(C,0,p_wave,0,0,0,1)
Hy = OPEV(C,0,p_wave,0,1,0,1)
Hym = OPEV(C,0,p_wave,0,-1,0,1)
Hx = OPEV(C,0,p_wave,0,1,1,0)
delta_s = OPEV(CF,0,p_wave,0,1,0,0)

IF (!FLAG_B & (RADII_arr(i_surf+1) != 0))
SURP i_surf+1, CURV, 1/RADII_arr(i_surf+1)
SURP i_surf+1, PARM, ASPH4_arr(i_surf+1),2
SURP i_surf+1, PARM, ASPH6_arr(i_surf+1),3
SURP i_surf+1, PARM, ASPH8_arr(i_surf+1),4
SURP i_surf+1, PARM, ASPH10_arr(i_surf+1),5
SURP i_surf+1, PARM, ASPH12_arr(i_surf+1),6
SURP i_surf+1, PARM, ASPH14_arr(i_surf+1),7
SURP i_surf+1, PARM, ASPH16_arr(i_surf+1),8
ENDIF

FOR j_surf,i_surf+2,n_surf,1
IF (RADII_arr(j_surf) != 0)
SURP j_surf, CURV, 1/RADII_arr(j_surf)
SURP j_surf, PARM, ASPH4_arr(j_surf),2
SURP j_surf, PARM, ASPH6_arr(j_surf),3
SURP j_surf, PARM, ASPH8_arr(j_surf),4
SURP j_surf, PARM, ASPH10_arr(j_surf),5
SURP j_surf, PARM, ASPH12_arr(j_surf),6
SURP j_surf, PARM, ASPH14_arr(j_surf),7
SURP j_surf, PARM, ASPH16_arr(j_surf),8
ENDIF
NEXT

IF FLAG_B
FOR j_surf,1,i_surf,1
IF (RADII_arr(j_surf) != 0)

```

```
SURP j_surf, CURV, 1/RADII_arr(j_surf)
SURP j_surf, PARM, ASPH4_arr(j_surf),2
SURP j_surf, PARM, ASPH6_arr(j_surf),3
SURP j_surf, PARM, ASPH8_arr(j_surf),4
SURP j_surf, PARM, ASPH10_arr(j_surf),5
SURP j_surf, PARM, ASPH12_arr(j_surf),6
SURP j_surf, PARM, ASPH14_arr(j_surf),7
SURP j_surf, PARM, ASPH16_arr(j_surf),8
ENDIF
NEXT
ENDIF

! SA
SA_all = H0

! COMA
COMA_all = (Hy-Hym)/2

! ASTI
ASTI_all = (Hy-Hx) - COMA_all

! FC
FC_all = Hx - SA_all

RETURN
```

APPENDIX D:
SERIES EXPANSION OF THE COSINE-TO-THE-FOURTH-
POWER OF THE RAY ANGLE

By definition, the cosine-to-the-fourth-power of the angle between the ray connecting the point (x'_s, y'_s) on the aperture plane and point (x'_i, y'_i) on the image plane, and the optical axis is

$$\begin{aligned} \cos^4(\theta') &= \left\{ \frac{e'}{[(x'_s - x'_i)^2 + (y'_s - y'_i)^2 + e'^2]^{1/2}} \right\}^4 \\ &= \left\{ \frac{1}{\left[1 + \frac{(x'_s - x'_i)^2}{e'^2} + \frac{(y'_s - y'_i)^2}{e'^2} \right]^{1/2}} \right\}^4 \\ &= \frac{1}{\left[1 + \frac{(x'^2_s + y'^2_s)}{e'^2} + \frac{(x'^2_i - y'^2_i)}{e'^2} - \frac{2 \cdot (x'_s x'_i - y'_s y'_i)}{e'^2} \right]^2}, \end{aligned} \quad (\text{D.1})$$

where e' is the axial distance between the image plane and the aperture plane.

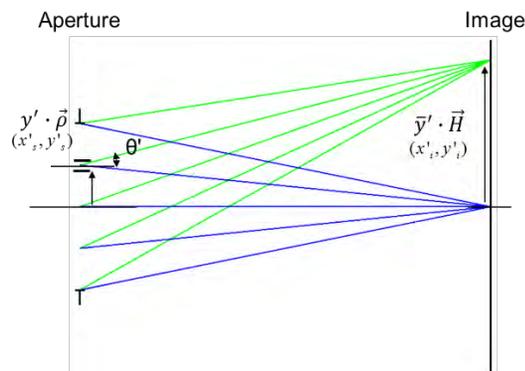


Fig. D.1 Geometrical variables involved in computing irradiance of a pinhole camera.

Points (x'_s, y'_s) and (x'_i, y'_i) are specified by the field \vec{H} aperture $\vec{\rho}$ vectors, as in

Fig. D.1, and are expressed in terms of the first-order system parameters as

$$\begin{aligned}(x_i^2 + y_i^2)^{1/2} &= \bar{y}' \cdot \vec{H} \\ (x_s^2 + y_s^2)^{1/2} &= y' \cdot \vec{\rho}\end{aligned}\tag{D.2}$$

and

$$\begin{aligned}\bar{u}' &= -\frac{\bar{y}'}{e'} \\ u' &= \frac{y'}{e'}.\end{aligned}\tag{D.3}$$

Here \bar{y}' and \bar{u}' are the chief ray height and slope at the image plane, y' and u' are the marginal ray height and slope at the aperture plane.

Eq. D.2 and Eq. D.3 are substituted into Eq. D.1, and the fourth order approximation to the cosine-to-the-fourth-power of the ray angle is obtained from the first two terms of a Taylor series expression

$$\begin{aligned}\cos^4(\theta'(\vec{H}, \vec{\rho})) &= \frac{1}{[1 + \bar{u}'^2(\vec{H} \cdot \vec{H}) + u'^2(\vec{\rho} \cdot \vec{\rho}) + 2 \cdot u' \bar{u}'(\vec{H} \cdot \vec{\rho})]^2} \\ &= 1 - 2 \cdot [\bar{u}'^2(\vec{H} \cdot \vec{H}) + u'^2(\vec{\rho} \cdot \vec{\rho}) + 2 \cdot u' \bar{u}'(\vec{H} \cdot \vec{\rho})] \dots \\ &\quad + 3 \cdot [\bar{u}'^2(\vec{H} \cdot \vec{H}) + u'^2(\vec{\rho} \cdot \vec{\rho}) + 2 \cdot u' \bar{u}'(\vec{H} \cdot \vec{\rho})]^2 \\ &= 1 - 2 \cdot \bar{u}'^2(\vec{H} \cdot \vec{H}) - 2 \cdot u'^2(\vec{\rho} \cdot \vec{\rho}) - 4 \cdot u' \bar{u}'(\vec{H} \cdot \vec{\rho}) \dots \\ &\quad + 3 \cdot \bar{u}'^4(\vec{H} \cdot \vec{H})^2 + 12 \cdot u'^3 \bar{u}'(\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho}) + 12 \cdot u'^2 \bar{u}'^2(\vec{H} \cdot \vec{\rho})^2 \dots \\ &\quad + 6 \cdot u'^2 \bar{u}'^2(\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho}) + 12 \cdot u' \bar{u}'^3(\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho}) + 3 \cdot u'^4(\vec{\rho} \cdot \vec{\rho})^2.\end{aligned}\tag{D.4}$$

$\theta'(\vec{H}, \vec{\rho})$ in Eq. D.4 indicate that the function is evaluated for a ray connecting a point defined by $\vec{\rho}$ on the aperture plane and a point defined by \vec{H} on the image plane.

APPENDIX E:
DERIVATION OF THE IRRADIANCE FUNCTION AT THE
IMAGE PLANE OF AN OPTICAL SYSTEM WITH THE
APERTURE VECTOR AT THE EXIT PUPIL

In an actual system, rays may not pass through the ideal image point due to aberrations, as shown in Fig. E.1. The ray intersection with the image plane is determined by considering the transverse ray errors $\Delta\vec{H}$ and it is defined by the vector $\vec{H} + \Delta\vec{H}$.

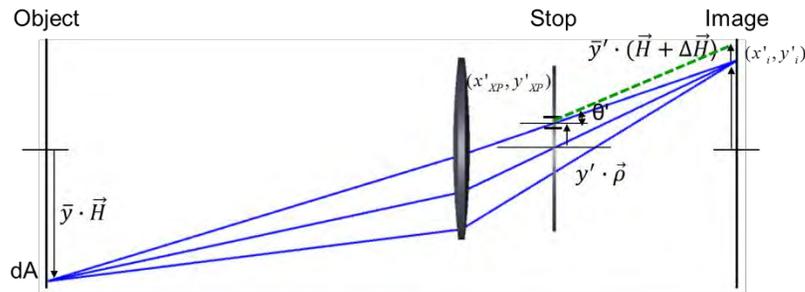


Fig. E.1 Geometrical variables involved in computing irradiance on the focal plane of an optical system with the aperture stop at the exit pupil. Real rays (solid lines) and first-order rays (dashed lines) coincide at the exit pupil. Real rays may not pass through the ideal image point due to aberrations.

In the presence of image aberrations, the cosine-to-the-fourth-power of the ray angle $\cos^4(\theta'(\vec{H} + \Delta\vec{H}, \vec{\rho}))$ is evaluated by writing the differential as

$$\cos^4(\theta'(\vec{H} + \Delta\vec{H}, \vec{\rho})) \approx \cos^4(\theta'(\vec{H}, \vec{\rho})) + \nabla_H \cos^4(\theta'(\vec{H}, \vec{\rho})) \cdot \Delta\vec{H}, \quad (\text{E.1})$$

where $\nabla_H \cos^4(\theta'(\vec{H}, \vec{\rho}))$ is the gradient of the function in Eq. D.4 with the respect to the field vector \vec{H} . With no second-order terms in the aberrations function, the terms $\nabla_H \cos^4(\theta'(\vec{H}, \vec{\rho})) \cdot \Delta\vec{H}$ result in irradiance terms that are at least of fourth order.

Thompson has shown that the gradient operator is given by the derivative of the function with respect to the designated vector [3.3]. It follows that to first order

$$\nabla_H \cos^4(\theta'(\vec{H}, \vec{\rho})) = [-4 \cdot \bar{u}'^2 \cdot \vec{H} - 4 \cdot u' \bar{u}' \cdot \vec{\rho}]. \quad (\text{E.2})$$

Finally, the terms $\nabla_H \cos^4(\theta'(\vec{H}, \vec{\rho})) \cdot \Delta \vec{H}$ are summarized in Table E.1.

Table E.1 Summary of contributions to irradiance from the transverse ray errors $\Delta \vec{H}$.

Ray aberration $\Delta \vec{H}$	Correction term $-4\bar{u}'^2 (\vec{H} \cdot \Delta \vec{H})$	Correction term $-4u' \bar{u}' (\Delta \vec{H} \cdot \vec{\rho})$
$-\frac{4}{\mathcal{K}} W_{040} (\vec{\rho} \cdot \vec{\rho}) \vec{\rho}$	$\frac{16}{\mathcal{K}} W_{040} \bar{u}'^2 (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho})$	$\frac{16}{\mathcal{K}} W_{040} u' \bar{u}' (\vec{\rho} \cdot \vec{\rho})^2$
$-\frac{2}{\mathcal{K}} W_{131} (\vec{H} \cdot \vec{\rho}) \vec{\rho}$	$\frac{8}{\mathcal{K}} W_{131} \bar{u}'^2 (\vec{H} \cdot \vec{\rho})^2$	$\frac{8}{\mathcal{K}} W_{131} u' \bar{u}' (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho})$
$-\frac{1}{\mathcal{K}} W_{131} (\vec{\rho} \cdot \vec{\rho}) \vec{H}$	$\frac{4}{\mathcal{K}} W_{131} \bar{u}'^2 (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho})$	$\frac{4}{\mathcal{K}} W_{131} u' \bar{u}' (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho})$
$-\frac{2}{\mathcal{K}} W_{222} (\vec{H} \cdot \vec{\rho}) \vec{H}$	$\frac{8}{\mathcal{K}} W_{222} \bar{u}'^2 (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho})$	$\frac{8}{\mathcal{K}} W_{222} u' \bar{u}' (\vec{H} \cdot \vec{\rho})^2$
$-\frac{2}{\mathcal{K}} W_{220} (\vec{H} \cdot \vec{H}) \vec{\rho}$	$\frac{8}{\mathcal{K}} W_{220} \bar{u}'^2 (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho})$	$\frac{8}{\mathcal{K}} W_{220} u' \bar{u}' (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho})$
$-\frac{1}{\mathcal{K}} W_{311} (\vec{H} \cdot \vec{H}) \vec{H}$	$\frac{4}{\mathcal{K}} W_{311} \bar{u}'^2 (\vec{H} \cdot \vec{H})^2$	$\frac{4}{\mathcal{K}} W_{311} u' \bar{u}' (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho})$

APPENDIX F:
DERIVATION OF THE IRRADIANCE FUNCTION AT THE
EXIT PUPIL OF AN OPTICAL SYSTEM WITH THE
APERTURE VECTOR AT THE EXIT PUPIL

To obtain the Jacobian determinant, the transverse ray aberration vector $\Delta\vec{H}$ is expressed in orthogonal components along unit vectors $\hat{\rho}$ and \hat{k} as

$$\Delta\vec{H} = \Delta H_{\rho}\hat{\rho} + \Delta H_k\hat{k}, \quad (\text{F.1})$$

where $\hat{\rho}$ is a unit vector along the aperture vector $\vec{\rho}$ and \hat{k} is a unit vector perpendicular to the aperture vector $\vec{\rho}$, as shown in Table F.1.

Table F.1 Third-order transverse ray aberrations in the orthogonal coordinates.

Ray aberration $\Delta\vec{H}$	$\hat{\rho}$	\hat{k}
$\Delta\vec{H}_{040}$	$4\rho^3$	
$\Delta\vec{H}_{131}$	$3\rho^2 H_{\rho}$	$\rho^2 H_k$
$\Delta\vec{H}_{222}$	$2\rho H_{\rho}^2$	$2\rho H_{\rho} H_k$
$\Delta\vec{H}_{220}$	$2\rho(H_{\rho}^2 + H_k^2)$	
$\Delta\vec{H}_{311}$	$(H_{\rho}^2 + H_k^2)H_{\rho}$	$(H_{\rho}^2 + H_k^2)H_k$
$\Delta\vec{H}_{400}$		

Then, the transformations $H'_\rho = H_\rho + \Delta H_\rho$ and $H'_k = H_k + \Delta H_k$ is considered, which gives the position of the given ray at the image pupil, and so the Jacobian determinant is obtained as

$$J(\vec{H}, \vec{\rho}) = \frac{\bar{y}'^2}{\bar{y}^2} \left(1 + \nabla_H \Delta \vec{H} + \frac{\partial \Delta \vec{H}_\rho}{\partial \vec{H}_\rho} \frac{\partial \Delta \vec{H}_k}{\partial \vec{H}_k} - \frac{\partial \Delta \vec{H}_\rho}{\partial \vec{H}_k} \frac{\partial \Delta \vec{H}_k}{\partial \vec{H}_\rho} \right), \quad (\text{F.2})$$

where $\nabla_H \Delta \vec{H}$ is the divergence of the transverse ray error vector, and $\frac{\partial \Delta \vec{H}_\rho}{\partial \vec{H}_\rho}$, $\frac{\partial \Delta \vec{H}_k}{\partial \vec{H}_k}$,

$\frac{\partial \Delta \vec{H}_\rho}{\partial \vec{H}_k}$ and $\frac{\partial \Delta \vec{H}_k}{\partial \vec{H}_\rho}$ are derivative of the transverse ray error vector in the orthogonal

coordinates [3.4, 3.5].

With no second order terms in the aberration function, the terms $\frac{\partial \Delta \vec{H}_\rho}{\partial \vec{H}_\rho} \frac{\partial \Delta \vec{H}_k}{\partial \vec{H}_k}$ and

$-\frac{\partial \Delta \vec{H}_\rho}{\partial \vec{H}_k} \frac{\partial \Delta \vec{H}_k}{\partial \vec{H}_\rho}$ result in irradiance terms that are at least of fourth order. The derivatives

of the third-order transverse ray aberrations in orthogonal coordinates are summarized in Table F.2.

Table F.2 Third-order transverse ray aberrations derivatives in orthogonal coordinates.

Ray aberration $\Delta \vec{H}$	$\partial_{H_\rho} \Delta \vec{H}_\rho$	$\partial_{H_k} \Delta \vec{H}_k$	$\partial_{H_k} \Delta \vec{H}_\rho$	$\partial_{H_\rho} \Delta \vec{H}_k$
$\Delta \vec{H}_{040}$				
$\Delta \vec{H}_{131}$	$3\rho^2$	ρ^2		
$\Delta \vec{H}_{222}$	$4\rho H_\rho$	$2\rho H_\rho$		$2\rho H_k$
$\Delta \vec{H}_{220}$	$4\rho H_\rho$		$4\rho H_k$	

$\Delta\vec{H}_{311}$	$(3H_\rho^2 + H_k^2)$	$(H_\rho^2 + 3H_k^2)$	$2H_\rho H_k$	$2H_\rho H_k$
$\Delta\vec{H}_{400}$				

The relationships between wavefront deformations and transverse ray aberrations to fifth order are given in Table F.3. The fifth-order transverse ray errors are related to the wavefront deformation by the gradient of the aberration function and include additional terms that are products of the fourth-order pupil aberration coefficients and paraxial ray slopes \bar{u}' and u' in image space. The subscripts H and ρ refer to the components along the field and aperture vectors respectively [3.1].

Table F.3 Fifth-order transverse ray aberrations for an optical system with the stop aperture at the exit pupil.

Ray aberration $\Delta\vec{H}$	$-\frac{1}{\mathcal{K}}\bar{\nabla}_\rho W^6(\vec{H}, \vec{\rho})$	Terms $O^{(5)}$
$\Delta\vec{H}_{060\rho}(\vec{\rho}\cdot\vec{\rho})^2\vec{\rho}$	$-\frac{6}{\mathcal{K}}W_{060}$	$-\frac{6}{\mathcal{K}}W_{040}u'^2$
$\Delta\vec{H}_{151H}(\vec{\rho}\cdot\vec{\rho})^2\vec{H}$	$-\frac{1}{\mathcal{K}}W_{151}$	$-\frac{1}{\mathcal{K}}[\frac{1}{2}\cdot W_{131}u'^2 + 4\cdot W_{040}u'\bar{u}']$
$\Delta\vec{H}_{151\rho}(\vec{H}\cdot\vec{\rho})(\vec{\rho}\cdot\vec{\rho})\vec{\rho}$	$-\frac{4}{\mathcal{K}}W_{151}$	$-\frac{1}{\mathcal{K}}[4\cdot W_{131}u'^2 + 8\cdot W_{040}u'\bar{u}']$
$\Delta\vec{H}_{333H}(\vec{H}\cdot\vec{\rho})^2\vec{H}$	$-\frac{3}{\mathcal{K}}W_{333}$	$-\frac{1}{\mathcal{K}}[2\cdot W_{131}\bar{u}'^2 + 4\cdot W_{222}u'\bar{u}']$
$\Delta\vec{H}_{331H}(\vec{H}\cdot\vec{H})(\vec{\rho}\cdot\vec{\rho})\vec{H}$	$-\frac{1}{\mathcal{K}}W_{331}$	$-\frac{1}{\mathcal{K}}[2\cdot W_{220}u'\bar{u}' + \frac{1}{2}\cdot W_{311}u'^2 + \frac{3}{2}\cdot W_{131}\bar{u}'^2]$
$\Delta\vec{H}_{331\rho}(\vec{H}\cdot\vec{H})(\vec{H}\cdot\vec{\rho})\vec{\rho}$	$-\frac{2}{\mathcal{K}}W_{331}$	$-\frac{1}{\mathcal{K}}[2\cdot W_{222}u'\bar{u}' + W_{131}\bar{u}'^2 + 4\cdot W_{220}u'\bar{u}' + W_{311}u'^2]$
$\Delta\vec{H}_{242\rho}(\vec{H}\cdot\vec{\rho})^2\vec{\rho}$	$-\frac{2}{\mathcal{K}}W_{242}$	$-\frac{1}{\mathcal{K}}[2\cdot W_{222}u'^2 + 4\cdot W_{131}u'\bar{u}']$
$\Delta\vec{H}_{242H}(\vec{H}\cdot\vec{\rho})(\vec{\rho}\cdot\vec{\rho})\vec{H}$	$-\frac{2}{\mathcal{K}}W_{242}$	$-\frac{1}{\mathcal{K}}[4\cdot W_{040}\bar{u}'^2 + 4\cdot W_{131}u'\bar{u}'^2 + W_{222}u'^2]$
$\Delta\vec{H}_{240H}(\vec{H}\cdot\vec{H})(\vec{\rho}\cdot\vec{\rho})\vec{\rho}$	$-\frac{4}{\mathcal{K}}W_{240}$	$-\frac{1}{\mathcal{K}}[3\cdot W_{220}u'^2 + W_{131}u'\bar{u}' + 2\cdot W_{040}\bar{u}'^2]$
$\Delta\vec{H}_{422H}(\vec{H}\cdot\vec{H})(\vec{H}\cdot\vec{\rho})\vec{H}$	$-\frac{2}{\mathcal{K}}W_{422}$	$-\frac{1}{\mathcal{K}}[3\cdot W_{222}\bar{u}'^2 + 2\cdot W_{220}\bar{u}'^2 + 2\cdot W_{311}u'\bar{u}']$
$\Delta\vec{H}_{420\rho}(\vec{H}\cdot\vec{H})^2\vec{\rho}$	$-\frac{2}{\mathcal{K}}W_{420}$	$-\frac{1}{\mathcal{K}}[W_{220}\bar{u}'^2 + W_{311}u'\bar{u}']$

$\Delta\vec{H}_{511H}(\vec{H}\cdot\vec{H})^2\vec{H}$	$-\frac{1}{\mathcal{K}}W_{511}$	$-\frac{1}{\mathcal{K}}\frac{3}{2}\cdot W_{311}\bar{u}^{\prime 2}$
--	---------------------------------	--

The divergence of the transverse ray errors $\nabla_H\Delta\vec{H}$ results in terms that are at least of second order. Sasian has shown the procedure to calculate the divergence operator of the function with respect to the designated vector [3.5, 3.24]. Following this procedure and taking the required derivatives, the terms in Table F.4 are obtained.

The Jacobian coefficients in Table F.4 are the sum of all components. Thus, for example, the term $J_{400}(\vec{H}\cdot\vec{H})^2$ is given by

$$J_{400} = \left[-\frac{6}{\mathcal{K}}W_{511} - \frac{9}{\mathcal{K}}W_{311}\bar{u}^{\prime 2} + \frac{3}{\mathcal{K}^2}W_{311}W_{311} \right] (\vec{H}\cdot\vec{H})^2. \quad (\text{F.3})$$

Finally, the coefficients in Table 3.2 are obtained by combining equations in Table 3.1, Table F.2 and Table F.4, and keeping only second- and fourth-order terms.

Table F.4 Terms corresponding to the determinant of the Jacobian transformation between the object and image planes for an optical system with the stop aperture at the exit pupil.

Irradiance coefficient $J_{k,l,m}$	Fourth-order aberration contributions	Six-order aberration contributions	Orthogonal derivative contributions
$J_{020}(\vec{\rho}\cdot\vec{\rho})$	$-\frac{4}{\mathcal{K}}W_{131}$		
$J_{111}(\vec{H}\cdot\vec{\rho})$	$-\frac{4}{\mathcal{K}}W_{220}$ $-\frac{6}{\mathcal{K}}W_{222}$		
$J_{200}(\vec{H}\cdot\vec{H})$	$-\frac{4}{\mathcal{K}}W_{311}$		
$J_{040}(\vec{\rho}\cdot\vec{\rho})^2$	$-\frac{1}{\mathcal{K}}[5\cdot W_{131}u^{\prime 2} + 16\cdot W_{040}u^{\prime}u^{\prime}]$	$-\frac{6}{\mathcal{K}}W_{151}$	$\frac{3}{\mathcal{K}^2}W_{131}W_{131}$

$J_{131}(\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho})$	$-\frac{1}{\mathcal{K}}[7 \cdot W_{222}u'^2 + 6 \cdot W_{220}u'^2]$ $-\frac{1}{\mathcal{K}}[22 \cdot W_{131}u'\bar{u}' - 16 \cdot W_{040}\bar{u}'^2]$	$-\frac{10}{\mathcal{K}}W_{242}$ $-\frac{8}{\mathcal{K}}W_{240}$	$\frac{4}{\mathcal{K}^2}W_{220}W_{131}$ $\frac{10}{\mathcal{K}^2}W_{222}W_{131}$
$J_{222}(\vec{H} \cdot \vec{\rho})^2$	$-\frac{1}{\mathcal{K}}[10 \cdot W_{131}\bar{u}'^2 + 20 \cdot W_{222}u'\bar{u}']$ $-\frac{1}{\mathcal{K}}[8 \cdot W_{220}u'\bar{u}' + 2 \cdot W_{311}u'^2]$	$-\frac{12}{\mathcal{K}}W_{333}$ $-\frac{4}{\mathcal{K}}W_{331}$	$\frac{8}{\mathcal{K}^2}W_{222}W_{222}$ $\frac{16}{\mathcal{K}^2}W_{222}W_{220}$ $-\frac{4}{\mathcal{K}^2}W_{311}W_{131}$
$J_{220}(\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho})$	$-\frac{1}{\mathcal{K}}[7 \cdot W_{131}\bar{u}'^2 + 2 \cdot W_{222}u'\bar{u}']$ $-\frac{1}{\mathcal{K}}[12 \cdot W_{220}u'\bar{u}' + 3 \cdot W_{311}u'^2]$	$-\frac{6}{\mathcal{K}}W_{331}$	$\frac{10}{\mathcal{K}^2}W_{311}W_{131}$ $-\frac{8}{\mathcal{K}^2}W_{222}W_{220}$
$J_{311}(\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho})$	$-\frac{1}{\mathcal{K}}[15 \cdot W_{222}\bar{u}'^2 + 14 \cdot W_{220}\bar{u}'^2]$ $-\frac{14}{\mathcal{K}}W_{311}u'\bar{u}'$	$-\frac{10}{\mathcal{K}}W_{422}$ $-\frac{8}{\mathcal{K}}W_{420}$	$\frac{4}{\mathcal{K}^2}W_{220}W_{311}$ $\frac{10}{\mathcal{K}^2}W_{222}W_{311}$
$J_{400}(\vec{H} \cdot \vec{H})^2$	$-\frac{9}{\mathcal{K}}W_{311}\bar{u}'^2$	$-\frac{6}{\mathcal{K}}W_{511}$	$\frac{3}{\mathcal{K}^2}W_{311}W_{311}$

APPENDIX G:
DERIVATION OF THE IRRADIANCE FUNCTION AT THE
IMAGE PLANE OF AN OPTICAL SYSTEM WITH THE
APERTURE VECTOR AT THE ENTRANCE PUPIL

In an optical system with the stop aperture at the entrance pupil, the ray angle in image space is determined by considering the transverse ray errors at both image and exit pupil planes, as shown in Fig. G.1.

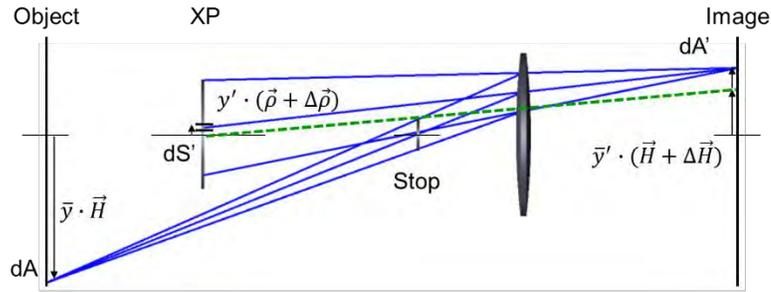


Fig. G.1 Geometrical variables involved in computing irradiance at the focal plane of an optical system with the aperture stop at the entrance pupil. Real rays (solid lines) and first-order rays (dashed lines) coincide at the entrance pupil. Real rays may not pass through the ideal points at the exit pupil and image planes due to aberrations.

In the presence of image and pupil aberrations, the cosine-to-the-fourth-power of the ray angle $\cos^4(\theta')\{\vec{H} + \Delta\vec{H}, \vec{\rho} + \Delta\vec{\rho}\}$ is evaluated by writing the differential as

$$\begin{aligned} \cos^4(\theta')\{\vec{H} + \Delta\vec{H}, \vec{\rho} + \Delta\vec{\rho}\} &\approx \cos^4(\theta')\{\vec{H}, \vec{\rho}\} + \nabla_H \cos^4(\theta')\{\vec{H}, \vec{\rho}\} \cdot \Delta\vec{H} \dots \\ &+ \nabla_\rho \cos^4(\theta')\{\vec{H}, \vec{\rho}\} \cdot \Delta\vec{\rho}, \end{aligned} \quad (\text{G.1})$$

where $\nabla_H \cos^4(\theta')\{\vec{H}, \vec{\rho}\}$ is the gradient of the function in Eq. D.4 with the respect to the field vector \vec{H} and $\nabla_\rho \cos^4(\theta')\{\vec{H}, \vec{\rho}\}$ is the gradient of the function with the respect to

the aperture vector $\vec{\rho}$. The terms $\nabla_H \cos^4(\theta') \{\vec{H}, \vec{\rho}\} \cdot \Delta \vec{H}$ are discussed in Appendix E and are given in Table E.1. The terms $\nabla_\rho \cos^4(\theta') \{\vec{H}, \vec{\rho}\} \cdot \Delta \vec{\rho}$ are calculated in a similar manner and are summarized in Table G.1.

Table G.1 Summary of contributions to irradiance from the transverse ray errors $\Delta \vec{\rho}$.

Ray aberration $\Delta \vec{\rho}$	Correction term $-4u'^2 (\vec{\rho} \cdot \Delta \vec{\rho})$	Correction term $-4u' \bar{u}' (\vec{H} \cdot \Delta \vec{\rho})$
$\frac{4}{\mathcal{K}} \bar{W}_{040} (\vec{H} \cdot \vec{H}) \vec{H}$	$-\frac{16}{\mathcal{K}} \bar{W}_{040} u'^2 (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho})$	$-\frac{16}{\mathcal{K}} \bar{W}_{040} u' \bar{u}' (\vec{H} \cdot \vec{H})^2$
$\frac{1}{\mathcal{K}} \bar{W}_{131} (\vec{H} \cdot \vec{H}) \vec{\rho}$	$-\frac{4}{\mathcal{K}} \bar{W}_{131} u'^2 (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho})$	$-\frac{4}{\mathcal{K}} \bar{W}_{131} u' \bar{u}' (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho})$
$\frac{2}{\mathcal{K}} \bar{W}_{131} (\vec{H} \cdot \vec{\rho}) \vec{H}$	$-\frac{8}{\mathcal{K}} \bar{W}_{131} u'^2 (\vec{H} \cdot \vec{\rho})^2$	$-\frac{8}{\mathcal{K}} \bar{W}_{131} u' \bar{u}' (\vec{H} \cdot \vec{H}) (\vec{H} \cdot \vec{\rho})$
$\frac{2}{\mathcal{K}} \bar{W}_{222} (\vec{H} \cdot \vec{\rho}) \vec{\rho}$	$-\frac{8}{\mathcal{K}} \bar{W}_{222} u'^2 (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho})$	$-\frac{8}{\mathcal{K}} \bar{W}_{222} u' \bar{u}' (\vec{H} \cdot \vec{\rho})^2$
$\frac{2}{\mathcal{K}} \bar{W}_{220} (\vec{\rho} \cdot \vec{\rho}) \vec{H}$	$-\frac{8}{\mathcal{K}} \bar{W}_{220} u'^2 (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho})$	$-\frac{8}{\mathcal{K}} \bar{W}_{220} u' \bar{u}' (\vec{H} \cdot \vec{H}) (\vec{\rho} \cdot \vec{\rho})$
$\frac{1}{\mathcal{K}} \bar{W}_{311} (\vec{\rho} \cdot \vec{\rho}) \vec{\rho}$	$-\frac{4}{\mathcal{K}} \bar{W}_{311} u'^2 (\vec{\rho} \cdot \vec{\rho})^2$	$-\frac{4}{\mathcal{K}} \bar{W}_{311} u' \bar{u}' (\vec{H} \cdot \vec{\rho}) (\vec{\rho} \cdot \vec{\rho})$

To obtain the Jacobian determinant, the transverse pupil ray aberration vector $\Delta \vec{\rho}$ is expressed in orthogonal components along unit vectors \hat{h} and \hat{k} as

$$\Delta \vec{\rho} = \Delta \rho_h \hat{h} + \Delta \rho_k \hat{k}, \quad (\text{G.2})$$

where \hat{h} is a unit vector along the aperture vector \vec{H} and \hat{k} is a unit vector perpendicular to the aperture vector \vec{H} . Then, the transformations $\rho'_h = \rho_h + \Delta \rho_h$ and $\rho'_k = \rho_k + \Delta \rho_k$ is considered, which gives the position of the given ray at the image pupil, and so the Jacobian determinant is found

$$\bar{J}_p(\vec{H}, \vec{\rho}) = \frac{dS'}{dS} = \frac{y'^2}{y^2} (1 + \nabla_\rho \Delta \vec{\rho} + \frac{\partial \Delta \vec{\rho}_h}{\partial \vec{\rho}_h} \frac{\partial \Delta \vec{\rho}_k}{\partial \vec{\rho}_k} - \frac{\partial \Delta \vec{\rho}_h}{\partial \vec{\rho}_k} \frac{\partial \Delta \vec{\rho}_k}{\partial \vec{\rho}_h}), \quad (\text{G.2})$$

where $\nabla_{\rho} \Delta \vec{\rho}$ is the divergence of the transverse pupil ray error vector, and $\frac{\partial \Delta \vec{\rho}_{\rho}}{\partial \vec{\rho}_{\rho}}$, $\frac{\partial \Delta \vec{\rho}_k}{\partial \vec{\rho}_k}$,

$\frac{\partial \Delta \vec{\rho}_{\rho}}{\partial \vec{\rho}_k}$ and $\frac{\partial \Delta \vec{\rho}_k}{\partial \vec{\rho}_{\rho}}$ are derivative of the transverse ray error vector in the orthogonal

coordinates [3.4, 3.5].

Notice that the Jacobian determinant in Eq. G.2 resembles the Jacobian determinant defined in Eq. F.2. It follows that the terms in Eq. G.2 are obtained from Table F.4 by interchanging the image and pupil aberration coefficients, exchanging the field and aperture vectors, interchanging the chief and marginal ray slopes, and changing sign of the Lagrange invariant. Moreover, since the aperture vector is now located at the entrance pupil, we also need to consider the change of the fifth-order transverse pupil aberration coefficients with the aperture vector location. Table G.2 summarizes the Jacobian determinant terms in Eq. G.2. Finally, the results in Table 3.4 are obtained by combining equations in Table E.1, Table G.1 and Table G.2, and keeping terms to fourth order.

Table G.2a Terms corresponding to the determinant of the Jacobian transformation between entrance and exit pupil planes for an optical system with the stop aperture at the entrance pupil.

Irradiance coefficient $\bar{J}_{p/k,l,m}$	Fourth-order pupil aberration contributions	Six-order pupil aberration contributions	Coordinate distortion contributions	Orthogonal derivative contributions
$\bar{J}_{p/020}(\vec{H} \cdot \vec{H})$	$\frac{4}{\mathcal{K}} \bar{W}_{311}$			
$\bar{J}_{p/111}(\vec{H} \cdot \vec{\rho})$	$\frac{4}{\mathcal{K}} \bar{W}_{220}$ $\frac{6}{\mathcal{K}} \bar{W}_{222}$			
$\bar{J}_{p/200}(\vec{\rho} \cdot \vec{\rho})$	$\frac{4}{\mathcal{K}} \bar{W}_{131}$			
$\bar{J}_{p/040}(\vec{H} \cdot \vec{H})^2$	$\frac{5}{\mathcal{K}} \bar{W}_{131} \bar{u}'^2$ $\frac{16}{\mathcal{K}} \bar{W}_{040} u' \bar{u}'$	$\frac{6}{\mathcal{K}} \bar{W}_{151}$	$-\frac{32}{\mathcal{K}^2} \bar{W}_{040} W_{220}$ $-\frac{24}{\mathcal{K}^2} \bar{W}_{040} W_{222}$ $-\frac{8}{\mathcal{K}^2} \bar{W}_{131} W_{311}$	$\frac{3}{\mathcal{K}^2} \bar{W}_{131} \bar{W}_{131}$
$\bar{J}_{p/131}(\vec{H} \cdot \vec{H})(\vec{H} \cdot \vec{\rho})$	$\frac{7}{\mathcal{K}} \bar{W}_{222} \bar{u}'^2$ $\frac{6}{\mathcal{K}} \bar{W}_{220} \bar{u}'^2$ $\frac{22}{\mathcal{K}} W_{131} u' \bar{u}'$ $\frac{16}{\mathcal{K}} W_{040} \bar{u}'^2$	$\frac{10}{\mathcal{K}} \bar{W}_{242}$ $\frac{8}{\mathcal{K}} \bar{W}_{240}$	$-\frac{80}{\mathcal{K}^2} \bar{W}_{040} W_{131}$ $-\frac{28}{\mathcal{K}^2} \bar{W}_{131} W_{222}$ $-\frac{32}{\mathcal{K}^2} \bar{W}_{131} W_{220}$ $-\frac{6}{\mathcal{K}^2} \bar{W}_{222} W_{311}$ $-\frac{4}{\mathcal{K}^2} \bar{W}_{220} W_{311}$	$\frac{4}{\mathcal{K}^2} \bar{W}_{220} \bar{W}_{131}$ $\frac{10}{\mathcal{K}^2} \bar{W}_{222} \bar{W}_{131}$
$\bar{J}_{p/222}(\vec{H} \cdot \vec{\rho})^2$	$\frac{10}{\mathcal{K}} \bar{W}_{131} u'^2$ $\frac{20}{\mathcal{K}} \bar{W}_{222} u' \bar{u}'$ $\frac{8}{\mathcal{K}} \bar{W}_{220} u' \bar{u}'$ $\frac{2}{\mathcal{K}} \bar{W}_{311} \bar{u}'^2$	$\frac{12}{\mathcal{K}} \bar{W}_{333}$ $\frac{4}{\mathcal{K}} \bar{W}_{331}$	$-\frac{48}{\mathcal{K}^2} \bar{W}_{131} W_{131}$ $-\frac{16}{\mathcal{K}^2} \bar{W}_{222} W_{222}$ $-\frac{8}{\mathcal{K}^2} \bar{W}_{220} W_{222}$ $-\frac{64}{\mathcal{K}^2} \bar{W}_{040} W_{040}$	$\frac{8}{\mathcal{K}^2} \bar{W}_{222} \bar{W}_{222}$ $\frac{16}{\mathcal{K}^2} \bar{W}_{222} \bar{W}_{220}$ $-\frac{4}{\mathcal{K}^2} \bar{W}_{311} \bar{W}_{131}$

$\bar{J}_{p/220}$ $(\vec{H} \cdot \vec{H})(\vec{\rho} \cdot \vec{\rho})$	$\frac{7}{\mathcal{K}} \bar{W}_{131} u'^2$ $\frac{2}{\mathcal{K}} \bar{W}_{222} u' \bar{u}'$ $\frac{12}{\mathcal{K}} \bar{W}_{220} u' \bar{u}'$ $\frac{3}{\mathcal{K}} \bar{W}_{311} \bar{u}'^2$	$\frac{6}{\mathcal{K}} \bar{W}_{331}$	$-\frac{96}{\mathcal{K}^2} \bar{W}_{040} W_{040}$ $-\frac{16}{\mathcal{K}^2} \bar{W}_{131} W_{131}$ $-\frac{16}{\mathcal{K}^2} \bar{W}_{222} W_{220}$ $-\frac{16}{\mathcal{K}^2} \bar{W}_{220} W_{220}$ $-\frac{4}{\mathcal{K}^2} \bar{W}_{220} W_{222}$	$\frac{10}{\mathcal{K}^2} \bar{W}_{131} \bar{W}_{311}$ $-\frac{8}{\mathcal{K}^2} \bar{W}_{220} \bar{W}_{222}$
$\bar{J}_{p/311}$ $(\vec{H} \cdot \vec{\rho})(\vec{\rho} \cdot \vec{\rho})$	$\frac{15}{\mathcal{K}} \bar{W}_{222} u'^2$ $\frac{1}{\mathcal{K}} \bar{W}_{220} u'^2$ $\frac{14}{\mathcal{K}} \bar{W}_{311} u' \bar{u}'$	$\frac{10}{\mathcal{K}} \bar{W}_{422}$ $\frac{8}{\mathcal{K}} \bar{W}_{420}$	$-\frac{112}{\mathcal{K}^2} \bar{W}_{131} W_{040}$ $-\frac{28}{\mathcal{K}^2} \bar{W}_{220} W_{131}$ $-\frac{30}{\mathcal{K}^2} \bar{W}_{222} W_{131}$	$\frac{4}{\mathcal{K}^2} \bar{W}_{220} \bar{W}_{311}$ $\frac{10}{\mathcal{K}^2} \bar{W}_{222} \bar{W}_{311}$
$\bar{J}_{p/400} (\vec{\rho} \cdot \vec{\rho})^2$	$\frac{9}{\mathcal{K}} \bar{W}_{311} u'^2$	$\frac{6}{\mathcal{K}} \bar{W}_{511}$	$-\frac{48}{\mathcal{K}^2} \bar{W}_{222} W_{040}$ $-\frac{48}{\mathcal{K}^2} \bar{W}_{220} W_{040}$	$\frac{3}{\mathcal{K}^2} \bar{W}_{311} \bar{W}_{311}$

APPENDIX H:
DERIVATION OF THE IRRADIANCE FUNCTION AT THE
EXIT PUPIL OF AN OPTICAL SYSTEM WITH THE
APERTURE VECTOR AT THE ENTRANCE PUPIL

Rather than following the comprehensive procedure used in previous sections, the irradiance coefficients of an optical system with the stop aperture at the entrance pupil are determined by using a symmetry principle. Since both the field vector \vec{H} and the aperture vector $\vec{\rho}$ are defined in object space, the roles of the object and entrance pupils can be exchanged without loss of generality and then the relationships in Table 3.5a and Table 3.5b follow. These relationships are obtained from the image plane irradiance coefficients in Table 3.4a and Table 3.4b by interchanging the image and pupil aberration coefficients, exchanging the field and aperture vectors, interchanging the chief and marginal ray slopes, and changing sign of the Lagrange invariant.

APPENDIX I:
 DERIVATION OF THE IRRADIANCE FUNCTION AT THE
 IMAGE PLANE OF AN OPTICAL SYSTEM WITH THE
 APERTURE VECTOR AT THE INTERMEDIATE POSITION

In order to derive an expression for the image plane illumination of a lens system with an internal diaphragm, the lens is divided into two parts and the contribution of each part is evaluated separately. As shown in Fig. I.1, Part A consists of all optical elements between the object plane and the diaphragm and thus has the aperture stop at the exit pupil. Part B includes all elements between the diaphragm and the image plane and has the aperture stop at the entrance pupil. By construction, the exit pupil of part A is the entrance pupil of part B. Furthermore, the image formed by A is the object of B.

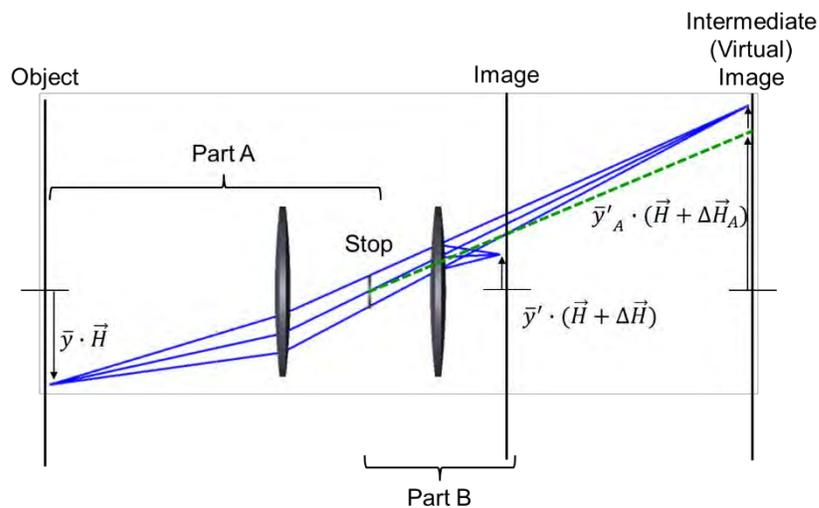


Fig. I.1 Geometrical variables involved in computing irradiance at the focal plane of the optical system with the aperture stop between components the lens. The lens is divided into two parts A and B, and contributions of each part are evaluated individually.

If part A is corrected for all image aberrations, the beams after A converge to the image points defined by the field vector \vec{H} . The incoming beams incident on part B have no aberrations and in the limit of small aperture relative illumination of the system is given by the expression similar to the case of having the aperture stop at the entrance pupil

$$\begin{aligned}
 RI_{EP}(\vec{H}) \approx & 1 + [-2\bar{u}^2 + \frac{4}{\mathcal{K}} \bar{W}_{131B}] (\vec{H} \cdot \vec{H}) \dots \\
 & + [3\bar{u}^4 + \frac{4}{\mathcal{K}} W_{311B} \bar{u}^2 + \frac{6}{\mathcal{K}} \bar{W}_{151B} - \frac{3}{\mathcal{K}} \bar{W}_{131B} \bar{u}^2] (\vec{H} \cdot \vec{H})^2 \dots \\
 & - \frac{1}{\mathcal{K}^2} [32\bar{W}_{040B} W_{220B} + 24\bar{W}_{040B} W_{222B} + 8\bar{W}_{131B} W_{311B}] (\vec{H} \cdot \vec{H})^2 \dots \\
 & + \frac{3}{\mathcal{K}^2} \bar{W}_{131B} \bar{W}_{131B} (\vec{H} \cdot \vec{H})^2,
 \end{aligned} \tag{I.1}$$

where the subscript B refers to aberration coefficients of part B alone. Most objectives with the internal diaphragm are designed to provide the best image quality only at the system focal plane. Parts A and B of the objective are not corrected individually. Part A forms a highly aberrated image, which is reimaged and compensated by the rest of the system. In the limit of small aperture, the distortion of A contributes transverse ray errors $\Delta\vec{H}_A$ to the third order given by

$$\Delta\vec{H}_A = -\frac{1}{\mathcal{K}} W_{311A} (\vec{H} \cdot \vec{H}) \vec{H}. \tag{I.2}$$

The image points of A are now defined by the vector $\vec{H} + \Delta\vec{H}_A$ and the relative illumination of the system is calculated by evaluating Eq. I.1 at $RI_{EP}(\vec{H} + \Delta\vec{H}_A)$. The second-order terms contribute extra fourth-order terms and the expression for the relative illumination of an optical system with internal diaphragm to fourth order is

$$\begin{aligned}
RI_{EP}(\vec{H}) &\approx 1 + [-2\bar{u}^2 + \frac{4}{\mathcal{K}}\bar{W}_{131B}] (\vec{H} \cdot \vec{H}) \dots \\
&+ [3\bar{u}^4 + \frac{4}{\mathcal{K}}W_{311A}\bar{u}^2 + \frac{4}{\mathcal{K}}W_{311B}\bar{u}^2 + \frac{6}{\mathcal{K}}\bar{W}_{151B} - \frac{3}{\mathcal{K}}\bar{W}_{131B}\bar{u}^2] (\vec{H} \cdot \vec{H})^2 \dots \\
&- \frac{1}{\mathcal{K}^2} [32\bar{W}_{040B}W_{220B} + 24\bar{W}_{040B}W_{222B} + 8\bar{W}_{131B}W_{311B}] (\vec{H} \cdot \vec{H})^2 \dots \\
&+ [\frac{3}{\mathcal{K}^2}\bar{W}_{131B}\bar{W}_{131B} - \frac{8}{\mathcal{K}^2}\bar{W}_{131B}W_{311A}] (\vec{H} \cdot \vec{H})^2.
\end{aligned} \tag{I.3}$$

A comparison of Eq. I.1 and Eq. I.3 reveals two additional terms that are caused by the aberrations of part A. The term $\frac{4}{\mathcal{K}}W_{311A}\bar{u}^2$, combined with the term $\frac{4}{\mathcal{K}}W_{311B}\bar{u}^2$, results in the total distortion of the system: $\frac{4}{\mathcal{K}}W_{311}\bar{u}^2$. The term $-\frac{8}{\mathcal{K}^2}\bar{W}_{131B}W_{311A}$ is a product of the pupil aberration of part B and the image aberration of part A. In the special case of an optical system with symmetry around the stop, the distortion of parts A and B has the same magnitude but opposite sign. Moreover, each part of the symmetrical optical system is separately corrected for field curvature and astigmatism. If $W_{311A} = -W_{311B}$ is substituted into Eq. I.3, a simplified expression for the relative illumination of a symmetrical optical system can be written as

$$\begin{aligned}
Ri_{EP}(\vec{H}) &\approx 1 + [-2\bar{u}^2 + \frac{4}{\mathcal{K}}\bar{W}_{131B}] (\vec{H} \cdot \vec{H}) \dots \\
&+ [3\bar{u}^4 + \frac{6}{\mathcal{K}}\bar{W}_{151B} - \frac{3}{\mathcal{K}}\bar{W}_{131B}\bar{u}^2 + \frac{3}{\mathcal{K}^2}\bar{W}_{131B}\bar{W}_{131B}] (\vec{H} \cdot \vec{H})^2.
\end{aligned} \tag{I.4}$$

The relative illumination of a symmetrical optical system does not depend on the aberration contributions of elements that precede the diaphragm. More uniform illumination at the focal plane can be achieved by controlling the pupil coma of the rear part of the system.

APPENDIX J:
 ABERRATION SUBTRACTION METHOD FOR
 CHROMATIC ABERRATION CORRECTION AND
 ATHERMALIZATION OF IMAGING SYSTEM ZEMAX
 MACRO

The following macro updates the Default Merit Function based on the optical path difference according to the aberration subtraction method for chromatic aberration correction.

```

! Written by Dmitry Reshidko
! Inputs
n_waves = NWAV()
p_wave = PWAV()
n_config = NCON()

DECLARE wave_arr, DOUBLE, 1, n_waves

FOR i=1,n_waves,1
wave_arr(i) = WAVL(i)
NEXT

IF DEBUG
print "Wavelength:"
FOR i=1,n_waves,1
print wave_arr(i)
NEXT
ENDIF

! Input # of Rings and Arms as in Gaussian Quadrature
INPUT "Number of rings:",rings
INPUT "Number of arms:",arms

n_fields = NFLD()

! First line in the Default Merit Function
INPUT "Start at:", start_line

```

```

start_line = start_line + 3

LABEL CONFIGS

ind=1
on_axis = 1
FOR i_field=1,n_fields,1
  IF on_axis
    cur_line = start_line
    FOR i_popd=1,rings,1
      popd_val = OPER(cur_line + (p_wave-1)*rings + i_popd, 10)
      ind = 1
      FOR i_opd=i_popd,rings*n_waves,rings
        SETOPERAND      cur_line      +      i_opd      ,      8,
popd_val*wave_arr(p_wave)/wave_arr(ind)
      ind = ind + 1
      NEXT
    NEXT
    on_axis = 0
    cur_line = cur_line + rings*n_waves + 1
  ELSE
    FOR i_popd=1,rings*arms/2,1
      popd_val = OPER(cur_line + (p_wave-1)*rings*arms/2 + i_popd, 10)
      ind = 1
      FOR i_opd=i_popd,rings*arms/2*n_waves,rings*arms/2
        SETOPERAND      cur_line      +      i_opd      ,      8,
popd_val*wave_arr(p_wave)/wave_arr(ind)
      ind = ind + 1
      NEXT
    NEXT
    cur_line = cur_line + arms*rings/2*n_waves + 1
  ENDIF

NEXT

IF n_config>1
  n_config = n_config - 1
  start_line = cur_line + 3
  GOTO CONFIGS
ENDIF

```

The following macro updates the Default Merit Function based on the ray errors according to the aberration subtraction method for athermalization of imaging systems.

```

! Inputs
DEBUG = 1

n_waves = NWAV()
p_wave = PWAV()
n_config = NCON()

DECLARE wave_arr, DOUBLE, 1, n_waves
FOR i=1,n_waves,1
wave_arr(i) = WAVL(i)
NEXT

! Input # of Rings and Arms as in Gaussian Quadrature
INPUT "Number of rings:",rings
INPUT "Number of arms:",arms

n_fields = NFLD();

! First line in the Default Merit Function
INPUT "Start at:", start_line

start_line = start_line + 3

ind=1;

on_axis = 1;

FOR i_field=1,n_fields,1

  IF on_axis

    cur_line = start_line

```

```

FOR i_popd=1,rings*n_waves,1
    popd_val = OPER(cur_line + i_popd, 10)
    ind = 1
    FOR i_opd=1,n_config,1
        SETOPERAND cur_line + i_popd + (rings*n_waves + (n_fields-
1)*rings*arms*n_waves/2 + n_fields + 3)*(i_opd-1), 8, popd_val
        ind = ind + 1
    NEXT
NEXT
on_axis = 0
cur_line = cur_line + rings*n_waves + 1
ELSE
    FOR i_popd=1,rings*arms/2*n_waves,1
        popd_val = OPER(cur_line + i_popd, 10)
        ind = 1
        FOR i_opd=1,n_config,1
            SETOPERAND cur_line + i_popd + (rings*n_waves + (n_fields-
1)*rings*arms*n_waves/2 + n_fields + 3)*(i_opd-1), 8, popd_val
            ind = ind + 1
        NEXT
    NEXT
    cur_line = cur_line + arms*rings/2*n_waves + 1
ENDIF
NEXT

```

APPENDIX K:

ASPHERIC/FREEFORM SURFACE C CODE FOR USER DEFINED SURFACES IN ZEMAX

Zemax user-defined surface C code for the Pedal Curve to the Ellipse.

```

#include <windows.h>
#include <math.h>
#include <string.h>
#include "usersurf.h"

#pragma warning ( disable : 4996 ) // functions like strcpy are now deprecated
for security reasons

/*
Written by Dmitry Reshidko (Sep 10, 2015)

The sag equation for this surface is given by :
base1 = (a1*a - a1*[a ^ 2 - 2 * r ^ 2 + (a ^ 4 + 4 * (a ^ 2 - b ^ 2)*r ^ 2) ^
1 / 2] / 2 } ^ 1 / 2)
Z = StandardConic + a1*base1 + a2*base1^2 + ... + b1*base2 + b2*base2^2 + ...
*/

int __declspec(dllexport) APIENTRY UserDefinedSurface3(USER_DATA *UD,
FIXED_DATA3 *FD);

/* a generic Snells law refraction routine */
int Refract(double thisn, double nextn, double *l, double *m, double *n, double
ln, double mn, double nn);

BOOL WINAPI DllMain(HANDLE hInst, ULONG ul_reason_for_call, LPVOID lpReserved)
{
return TRUE;
}

int __declspec(dllexport) APIENTRY UserDefinedSurface3(USER_DATA *UD,
FIXED_DATA3 *FD)
{
int i, sag_error;
double alpha, power, t;
double x, y, z, Z1, Z2, Zi, Z_tot_extra;
int loop;
double a1_axis, b1_axis, A1, A2, A3, A4, A5, A6;
int num_pedal;
double aa_axis, bb_axis, Ai_1, Ai_2, Ai_3, Ai_4, Ai_5, Ai_6;
double betta, gamma, cvt, bettad, gammad;
double r2, tp, dz, sag, mm, mx, my;

```

```

switch (FD->type)
{
case 0:
/* ZEMAX is requesting general information about the surface */
switch (FD->numb)
{
case 0:
/* ZEMAX wants to know the name of the surface */
/* do not exceed 12 characters */
strcpy(UD->string, "PedalSurface");
break;
case 1:
/* ZEMAX wants to know if this surface is rotationally symmetric */
/* it is not, so return a 1 */
strcpy(UD->string, "1");
break;
case 2:
/* ZEMAX wants to know if this surface is a gradient index media */
/* it is not, so return a null string */
UD->string[0] = '\0';
break;
}
break;
case 1:
/* ZEMAX is requesting the names of the parameter columns */
/* the value FD->numb will indicate which value ZEMAX wants. */
/* they are all "Unused" for this surface type */
/* returning a null string indicates that the parameter is unused. */
switch (FD->numb)
{
case 1:
strcpy(UD->string, "a1: along Z");
break;
case 2:
strcpy(UD->string, "b1: along Y");
break;
case 3:
strcpy(UD->string, "A1");
break;
case 4:
strcpy(UD->string, "A2");
break;
case 5:
strcpy(UD->string, "A3");
break;
case 6:
strcpy(UD->string, "A4");
break;
case 7:
strcpy(UD->string, "A5");
break;
case 8:
strcpy(UD->string, "A6");

```

```

break;
case 9:
strcpy(UD->string, "Number Extra Pedal");
break;
default:
UD->string[0] = '\0';
break;
}
break;
case 2:
/* ZEMAX is requesting the names of the extra data columns */
/* the value FD->numb will indicate which value ZEMAX wants. */
/* they are all "Unused" for this surface type */

if (FD->numb <= FD->param[9] * 8)
{
if (((FD->numb - 1) % 8) + 1) == 1)
{
sprintf(UD->string, "a%i: along Z", FD->numb/8 + 2);
//strcpy(UD->string, "aai");
}
else if (((FD->numb - 1) % 8) + 1) == 2)
{
sprintf(UD->string, "b%i: along Y", FD->numb / 8 + 2);
//strcpy(UD->string, "bii");
}
else
{
sprintf(UD->string, "A%i%i", (FD->numb - 1) / 8 + 2, ((FD->numb - 1) % 8) -
1));
//strcpy(UD->string, "AAA");
}
}
else
{
UD->string[0] = '\0';
}

break;
case 3:
/* ZEMAX wants to know the sag of the surface */
/* if there is an alternate sag, return it as well */
/* otherwise, set the alternate sag identical to the sag */
/* The sag is sag1, alternate is sag2. */

/* Pedal terms */
a1_axis = FD->param[1];
b1_axis = FD->param[2];
A1 = FD->param[3];
A2 = FD->param[4];
A3 = FD->param[5];
A4 = FD->param[6];
A5 = FD->param[7];
A6 = FD->param[8];

```

```

num_pedal = FD->param[9];

UD->sag1 = 0.0;
UD->sag2 = 0.0;

/* invalid input - consider absolute value */
if (a1_axis < 0) return(-1);
if (b1_axis < 0) return(-1);

x = UD->x;
y = UD->y;
r2 = x*x + y*y;

// Standard Sag
alpha = 1 - (1 + FD->k)*FD->cv*FD->cv*r2;
//if (alpha < 0) return(-1);

if (alpha < 0)
{
return -1;
}
else
{
Z1 = (FD->cv*r2) / (1 + sqrt(alpha));
}

Z2 = 0;
/* Pedal terms I */
if (sqrt(r2) > b1_axis)
{
Z2 = 0;
}
else
{
beta = r2*(b1_axis*b1_axis - a1_axis*a1_axis) + a1_axis*a1_axis*a1_axis*a1_axis
/ 4.0;
if (beta < 0)
{
Z2 = 0;
}
else
{
gamma = a1_axis*a1_axis / 2.0 - r2 + sqrt(beta);
if (gamma < 0)
{
Z2 = 0;
}
else
{
Z2 = a1_axis - sqrt(gamma);
}
}
}
}

```

```

Z_tot_extra = 0;
for (i = 1; i <= num_pedal; i++)
{
aa_axis = FD->xdata[(i - 1) * 8 + 1];
bb_axis = FD->xdata[(i - 1) * 8 + 2];
Ai_1 = FD->xdata[(i - 1) * 8 + 3];
Ai_2 = FD->xdata[(i - 1) * 8 + 4];
Ai_3 = FD->xdata[(i - 1) * 8 + 5];
Ai_4 = FD->xdata[(i - 1) * 8 + 6];
Ai_5 = FD->xdata[(i - 1) * 8 + 7];
Ai_6 = FD->xdata[(i - 1) * 8 + 8];

if (aa_axis < 0) return(-1);
if (bb_axis < 0) return(-1);

/* Extra Pedal Terms */
Zi = 0;
if (sqrt(r2) > bb_axis)
{
Zi = 0;
}
else
{
beta = r2*(bb_axis*bb_axis - aa_axis*aa_axis) + aa_axis*aa_axis*aa_axis*aa_axis
/ 4.0;
if (beta < 0)
{
Zi = 0;
}
else
{
gamma = aa_axis*aa_axis / 2.0 - r2 + sqrt(beta);
if (gamma < 0)
{
Zi = 0;
}
else
{
Zi = aa_axis - sqrt(gamma);
}
}
}

Z_tot_extra += Ai_1*Zi + Ai_2*Zi*Zi + Ai_3*Zi*Zi*Zi + Ai_4*Zi*Zi*Zi*Zi +
Ai_5*Zi*Zi*Zi*Zi*Zi + Ai_6*Zi*Zi*Zi*Zi*Zi*Zi;
}

UD->sag1 = A1*(Z1 + Z2) + A2*pow((Z1 + Z2), 2) + A3*pow((Z1 + Z2), 3) +
A4*pow((Z1 + Z2), 4) + A5*pow((Z1 + Z2), 5) + A6*pow((Z1 + Z2),6) + Z_tot_extra;

```

```

// no hyperhemispheric support
UD->sag2 = UD->sag1;

break;
case 4:
/* ZEMAX wants a paraxial ray trace to this surface */
/* x, y, z, and the optical path are unaffected, at least for this surface type
*/
/* for paraxial ray tracing, the return z coordinate should always be zero. */
/* paraxial surfaces are always planes with the following normals */
UD->ln = 0.0;
UD->mn = 0.0;
UD->nn = -1.0;

/* Pedal terms */
a1_axis = FD->param[1];
b1_axis = FD->param[2];
A1 = FD->param[3];
num_pedal = FD->param[9];

// Standard Optical Power
power = A1*(FD->n2 - FD->n1)*FD->cv;

// Pedal contribution
if ((a1_axis != 0) && (b1_axis != 0))
{
cvt = A1*(2.0*a1_axis*a1_axis - b1_axis*b1_axis) / a1_axis / a1_axis / a1_axis;
power += (FD->n2 - FD->n1)*cvt;
}

/* Extra Pedal Terms */
for (i = 1; i <= num_pedal; i++)
{
aa_axis = FD->xdata[(i - 1) * 8 + 1];
bb_axis = FD->xdata[(i - 1) * 8 + 2];
Ai_1 = FD->xdata[(i - 1) * 8 + 3];

if ((aa_axis != 0) && (bb_axis != 0))
{
cvt = Ai_1*(2.0*aa_axis*aa_axis - bb_axis*bb_axis) / aa_axis / aa_axis /
aa_axis;
power += (FD->n2 - FD->n1)*cvt;
}
}

if ((UD->n) != 0.0)
{
(UD->l) = (UD->l) / (UD->n);
(UD->m) = (UD->m) / (UD->n);

(UD->l) = (FD->n1*(UD->l) - (UD->x)*power) / (FD->n2);

```

```

(UD->m) = (FD->n1*(UD->m) - (UD->y)*power) / (FD->n2);

/* normalize */
(UD->n) = sqrt(1 / (1 + (UD->l)*(UD->l) + (UD->m)*(UD->m)));
/* de-paraxialize */
(UD->l) = (UD->l)*(UD->n);
(UD->m) = (UD->m)*(UD->n);
}
break;

case 5:
/* ZEMAX wants a real ray trace to this surface */

/* Pedal terms */
a1_axis = FD->param[1];
b1_axis = FD->param[2];
A1 = FD->param[3];
A2 = FD->param[4];
A3 = FD->param[5];
A4 = FD->param[6];
A5 = FD->param[7];
A6 = FD->param[8];
num_pedal = FD->param[9];

/* invalid input - consider absolute value */
if (a1_axis < 0) return(-1);
if (b1_axis < 0) return(-1);

x = UD->x;
y = UD->y;
z = UD->z;

/* make sure we do at least 1 loop */
t = 100.0;
tp = 0.0;
loop = 0;

// calculate r2 for the first time
r2 = x * x + y * y;

while (fabs(t) > 1e-10)
{
/*
First, compute the sag using whatever the surface sag expression is.
This is given the x and y starting points. The following block of code
will change depending upon the surface shape, the rest of this iteration
is typically common to all surface shapes.
*/

sag_error = 0;
// standard sag
alpha = 1 - (1 + FD->k)*FD->cv*FD->cv*r2;
// if (alpha < 0) return(-1);

```

```

if (alpha < 0)
{
//return (-1);
alpha = 1;
Z1 = 1 / FD->cv;
}
else
{
Z1 = (FD->cv*r2) / (1 + sqrt(alpha));
}

/* Pedal terms I */
Z2 = 0;
if (sqrt(r2) > b1_axis)
{
sag_error = 1;
}
else
{
beta = r2*(b1_axis*b1_axis - a1_axis*a1_axis) + a1_axis*a1_axis*a1_axis*a1_axis
/ 4.0;
if (beta < 0)
{
sag_error = 1;
}
else
{
gamma = a1_axis*a1_axis / 2.0 - r2 + sqrt(beta);
if (gamma < 0)
{
sag_error = 1;
}
else
{
Z2 = a1_axis - sqrt(gamma);
}
}
}

if (sag_error == 1)
{
beta = 0.99*0.99*b1_axis*b1_axis*(b1_axis*b1_axis - a1_axis*a1_axis) +
a1_axis*a1_axis*a1_axis*a1_axis / 4.0;
gamma = a1_axis*a1_axis / 2.0 - 0.99*0.99*b1_axis*b1_axis + sqrt(beta);
Z2 = a1_axis - sqrt(gamma);
sag_error = 0;
}

Z_tot_extra = 0;
sag_error = 0;
for (i = 1; i <= num_pedal; i++)

```

```

{
aa_axis = FD->xdata[(i - 1) * 8 + 1];
bb_axis = FD->xdata[(i - 1) * 8 + 2];
Ai_1 = FD->xdata[(i - 1) * 8 + 3];
Ai_2 = FD->xdata[(i - 1) * 8 + 4];
Ai_3 = FD->xdata[(i - 1) * 8 + 5];
Ai_4 = FD->xdata[(i - 1) * 8 + 6];
Ai_5 = FD->xdata[(i - 1) * 8 + 7];
Ai_6 = FD->xdata[(i - 1) * 8 + 8];

if (aa_axis < 0) return(-1);
if (bb_axis < 0) return(-1);

/* Extra Pedal Terms */
Zi = 0;
if (sqrt(r2) > bb_axis)
{
sag_error = 1;
}
else
{
beta = r2*(bb_axis*bb_axis - aa_axis*aa_axis) + aa_axis*aa_axis*aa_axis*aa_axis
/ 4.0;
if (beta < 0)
{
sag_error = 1;
}
else
{
gamma = aa_axis*aa_axis / 2.0 - r2 + sqrt(beta);
if (gamma < 0)
{
sag_error = 1;
}
else
{
Zi = aa_axis - sqrt(gamma);
}
}
}

if (sag_error == 1)
{
beta = 0.99*0.99*bb_axis*bb_axis*(bb_axis*bb_axis - aa_axis*aa_axis) +
aa_axis*aa_axis*aa_axis*aa_axis / 4.0;
gamma = aa_axis*aa_axis / 2.0 - 0.99*0.99*bb_axis*bb_axis + sqrt(beta);
Zi = aa_axis - sqrt(gamma);
sag_error = 0;
}
}

```

```

}

Z_tot_extra += Ai_1*Zi + Ai_2*Zi*Zi + Ai_3*Zi*Zi*Zi + Ai_4*Zi*Zi*Zi*Zi +
Ai_5*Zi*Zi*Zi*Zi*Zi + Ai_6*Zi*Zi*Zi*Zi*Zi*Zi;

}

sag = A1*(Z1 + Z2) + A2*pow((Z1 + Z2), 2) + A3*pow((Z1 + Z2), 3) + A4*pow((Z1 +
Z2), 4) + A5*pow((Z1 + Z2), 5) + A6*pow((Z1 + Z2), 6) + Z_tot_extra;

/* okay, now with sag in hand, how far are we away in z? */
dz = sag - z;

/* now compute how far along the z axis this is */
/* slow it down by a factor of 0.05 to support steep angles */
t = dz*0.05;

/* propagate the additional "t" distance */
x += UD->l*t;
y += UD->m*t;
z += UD->n*t;

/* add in the optical path */
tp += t;

// update r2
r2 = x * x + y * y;

/* prevent infinite loop if no convergence */
loop++;
if (loop > 100000) return(-1);
}

UD->path = tp;

/* now do the normals */
if (r2 == 0)
{
UD->ln = 0;
UD->mn = 0;
UD->nn = -1;
}
else
{
mm = 0;
alpha = 1.0 - (1.0 + FD->k)*FD->cv*FD->cv*r2;
if (alpha < 0) return(-1); /* ray misses */
Z1 = (FD->cv*r2) / (1 + sqrt(alpha));
alpha = sqrt(alpha);

// Pedal terms I
Z2 = 0;

```

```

if (sqrt(r2) > b1_axis)
{
Z2 = 0;
}
else
{
bettad = 4.0*r2*(b1_axis*b1_axis - a1_axis*a1_axis) +
a1_axis*a1_axis*a1_axis*a1_axis;
betta = r2*(b1_axis*b1_axis - a1_axis*a1_axis) + a1_axis*a1_axis*a1_axis*a1_axis
/ 4.0;

if (betta < 0)
{
Z2 = 0;
}
else
{
gammad = a1_axis*a1_axis - 2.0*r2 + sqrt(bettad);
gamma = a1_axis*a1_axis / 2.0 - r2 + sqrt(betta);
if (gamma < 0)
{
Z2 = 0;
}
else
{
Z2 = a1_axis - sqrt(gamma);
mm -= (A1 + 2.0*A2*(Z1 + Z2) + 3.0*A3*pow((Z1 + Z2), 2) + 4.0*A4*pow((Z1 + Z2),
3) + 5.0*A5*pow((Z1 + Z2), 4) + 6.0*A6*pow((Z1 + Z2),5))*sqrt(2.0) *
((b1_axis*b1_axis - a1_axis*a1_axis) / sqrt(bettad) - 1) / sqrt(gammad);
}
}
}

// Conic part after computing Z2
mm += (A1 + 2.0*A2*(Z1 + Z2) + 3.0*A3*pow((Z1 + Z2), 2) + 4.0*A4*pow((Z1 + Z2),
3) + 5.0*A5*pow((Z1 + Z2), 4) + 6.0*A6*pow((Z1 + Z2),5))*(FD->cv / (1.0 +
alpha))*(2.0 + (FD->cv*FD->cv*r2*(1.0 + FD->k)) / (alpha*(1.0 + alpha)));

sag_error = 0;
for (i = 1; i <= num_pedal; i++)
{
aa_axis = FD->xdata[(i - 1) * 8 + 1];
bb_axis = FD->xdata[(i - 1) * 8 + 2];
Ai_1 = FD->xdata[(i - 1) * 8 + 3];
Ai_2 = FD->xdata[(i - 1) * 8 + 4];
Ai_3 = FD->xdata[(i - 1) * 8 + 5];
Ai_4 = FD->xdata[(i - 1) * 8 + 6];
Ai_5 = FD->xdata[(i - 1) * 8 + 7];
Ai_6 = FD->xdata[(i - 1) * 8 + 8];

if (aa_axis < 0) return(-1);

```

```

if (bb_axis < 0) return(-1);

/* Extra Pedal Terms */
Zi = 0;
if (sqrt(r2) > bb_axis)
{
Zi = 0;
}
else
{
beta = r2*(bb_axis*bb_axis - aa_axis*aa_axis) + aa_axis*aa_axis*aa_axis*aa_axis
/ 4.0;
bettad = 4.0*r2*(bb_axis*bb_axis - aa_axis*aa_axis) +
aa_axis*aa_axis*aa_axis*aa_axis;
if (beta < 0)
{
Zi = 0;
}
else
{
gamma = aa_axis*aa_axis / 2.0 - r2 + sqrt(beta);
gammad = aa_axis*aa_axis - 2.0*r2 + sqrt(bettad);
if (gamma < 0)
{
Zi = 0;
}
else
{
Zi = aa_axis - sqrt(gamma);
mm -= (Ai_1 + 2.0*Ai_2*Zi + 3.0*Ai_3*Zi*Zi + 4.0*Ai_4*Zi*Zi*Zi +
5.0*Ai_5*Zi*Zi*Zi*Zi + 6.0*Ai_6*Zi*Zi*Zi*Zi*Zi)*sqrt(2.0) * ((bb_axis*bb_axis -
aa_axis*aa_axis) / sqrt(bettad) - 1) / sqrt(gammad);
}
}
}

}

// mm now holds (1/r)*(dz/dr)

mx = x * mm;
my = y * mm;

UD->nn = -sqrt(1 / (1 + (mx*mx) + (my*my)));
UD->ln = -mx*UD->nn;
UD->mn = -my*UD->nn;
}

/* restore coordinates */
UD->x = x;
UD->y = y;
UD->z = z;

```

```

if (Refract(FD->n1, FD->n2, &UD->l, &UD->m, &UD->n, UD->ln, UD->mn, UD->nn))
return(-FD->surf);
break;
case 6:
/* ZEMAX wants the index, dn/dx, dn/dy, and dn/dz at the given x, y, z. */

/* This is only required for gradient index surfaces, so return dummy values */
UD->index = FD->n2;
UD->dndx = 0.0;
UD->dndy = 0.0;
UD->dndz = 0.0;
break;
case 7:
/* ZEMAX wants the "safe" data. */
/* this is used by ZEMAX to set the initial values for all parameters and extra
data */
/* when the user first changes to this surface type. */
/* this is the only time the DLL should modify the data in the FIXED_DATA FD
structure */
for (i = 1; i <= 11; i++) FD->param[i] = 0.0;
for (i = 1; i <= 200; i++) FD->xdata[i] = 0.0;
break;
}
return 0;
}

int Refract(double thisn, double nextn, double *l, double *m, double *n, double
ln, double mn, double nn)
{
double nr, cosi, cosi2, rad, cosr, gamma;
if (thisn != nextn)
{
nr = thisn / nextn;
cosi = fabs((*l) * ln + (*m) * mn + (*n) * nn);
cosi2 = cosi * cosi;
if (cosi2 > 1) cosi2 = 1;
rad = 1 - ((1 - cosi2) * (nr * nr));
if (rad < 0) return(-1);
cosr = sqrt(rad);
gamma = nr * cosi - cosr;
(*l) = (nr * (*l)) + (gamma * ln);
(*m) = (nr * (*m)) + (gamma * mn);
(*n) = (nr * (*n)) + (gamma * nn);
}
return 0;
}

```

Zemax user-defined surface C code for the Decentered Conic surface.

```

#include <windows.h>
#include <math.h>
#include <string.h>
#include "usersurf.h"

/*
Written by DR 09042016

base = Decentered Conic
sag = Decentered Conic + Plane Symmetric Polynomial

*/

#define BIG_RAY 1.0E+010

int __declspec(dllexport) APIENTRY UserDefinedSurface(USER_DATA *UD, FIXED_DATA
*FD);

/* a generic Snells law refraction routine */
int Refract(double thisn, double nextn, double *l, double *m, double *n, double
ln, double mn, double nn);

BOOL WINAPI DllMain(HANDLE hInst, ULONG ul_reason_for_call, LPVOID lpReserved)
{
return TRUE;
}

int __declspec(dllexport) APIENTRY UserDefinedSurface(USER_DATA *UD, FIXED_DATA
*FD)
{
int i, loop;
double power, t, tp, x, y, z, dz, sag, slowfactor, alpha;
double R, k, K, u, v, eps, L, s, c, g, f, h, j, w, poly;
double mx, my, poly_mx, poly_my;

switch (FD->type)
{
case 0:
/* ZEMAX is requesting general information about the surface */
switch (FD->numb)
{
case 0:
/* ZEMAX wants to know the name of the surface */
/* do not exceed 12 characters */
strcpy(UD->string, "OffAxisConic");
break;
case 1:
/* ZEMAX wants to know if this surface is rotationally symmetric */
/* it is, so return any character in the string; otherwise, return a null string
*/
UD->string[0] = '\\0';
break;

```

```

case 2:
/* ZEMAX wants to know if this surface is a gradient index media */
/* it is not, so return a null string */
UD->string[0] = '\0';
break;
}
break;
case 1:
/* ZEMAX is requesting the names of the parameter columns */
/* the value FD->numb will indicate which value ZEMAX wants. */
/* they are all "Unused" for this surface type except parameter 1 */
/* returning a null string indicates that the parameter is unused. */
switch (FD->numb)
{
case 1:
strcpy(UD->string, "y0");
break;
case 2:
strcpy(UD->string, "slowfactor");
break;
default:
UD->string[0] = '\0';
break;
}
break;
case 2:
/* ZEMAX is requesting the names of the extra data columns */
/* the value FD->numb will indicate which value ZEMAX wants. */
/* they are all "Unused" for this surface type */
/* returning a null string indicates that the extradata value is unused. */
switch (FD->numb)
{
case 1:
strcpy(UD->string, "Y");
break;
case 2:
strcpy(UD->string, "X2");
break;
case 3:
strcpy(UD->string, "Y2");
break;
case 4:
strcpy(UD->string, "X2Y");
break;
case 5:
strcpy(UD->string, "Y3");
break;
case 6:
strcpy(UD->string, "X4");
break;
case 7:
strcpy(UD->string, "X2Y2");
break;
case 8:

```

```
strcpy(UD->string, "Y4");
break;
case 9:
strcpy(UD->string, "X4Y");
break;
case 10:
strcpy(UD->string, "X2Y3");
break;
case 11:
strcpy(UD->string, "Y5");
break;
case 12:
strcpy(UD->string, "X6");
break;
case 13:
strcpy(UD->string, "X4Y2");
break;
case 14:
strcpy(UD->string, "X2Y4");
break;
case 15:
strcpy(UD->string, "Y6");
break;
case 16:
strcpy(UD->string, "X6Y");
break;
case 17:
strcpy(UD->string, "X4Y3");
break;
case 18:
strcpy(UD->string, "X2Y5");
break;
case 19:
strcpy(UD->string, "Y7");
break;
case 20:
strcpy(UD->string, "X8");
break;
case 21:
strcpy(UD->string, "X6Y2");
break;
case 22:
strcpy(UD->string, "X4Y4");
break;
case 23:
strcpy(UD->string, "X2Y6");
break;
case 24:
strcpy(UD->string, "Y8");
break;
case 25:
strcpy(UD->string, "X8Y");
break;
case 26:
```

```

strcpy(UD->string, "X6Y3");
break;
case 27:
strcpy(UD->string, "X4Y5");
break;
case 28:
strcpy(UD->string, "X2Y7");
break;
case 29:
strcpy(UD->string, "Y9");
break;
case 30:
strcpy(UD->string, "X10");
break;
case 31:
strcpy(UD->string, "X8Y2");
break;
case 32:
strcpy(UD->string, "X6Y4");
break;
case 33:
strcpy(UD->string, "X4Y6");
break;
case 34:
strcpy(UD->string, "X2Y8");
break;
case 35:
strcpy(UD->string, "Y10");
break;
default:
UD->string[0] = '\0';
break;
}
break;
case 3:
/* ZEMAX wants to know the sag of the surface */
/* The sag is sag1, alternate is sag2. */

UD->sag1 = 0.0;
UD->sag2 = 0.0;

/* if a plane, just return */
if (FD->cv == 0) return(0);

R = FD->param[1];

if (FD->cv == 0)
{
k = 0;
}
else
{
k = 1/FD->cv;
}

```

```

K = FD->k;
x = UD->x;
y = UD->y;

u = y / k;
v = x / k;
eps = R / k;
L = K + 1;
s = eps / sqrt(1 - K*eps*eps);
c = sqrt((1 - L*eps*eps) / (1 - K*eps*eps));

if ((R == 0) && (K == -1))
{
z = (x*x + y*y) / 2 / k;
}
else
{
if (R == 0)
{
alpha = k*k - (K + 1)*(x*x + y*y);
if (alpha < 0)
{
return -1;
}
else
{
if (k / (K + 1) <= 0)
{
z = 1 / (K + 1)*(k + sqrt(k*k - (K + 1)*(x*x + y*y)));
}
else
{
z = 1 / (K + 1)*(k - sqrt(k*k - (K + 1)*(x*x + y*y)));
}
}
}
else
{
g = -1 / (L*c*c + s*s);
f = (s / eps)*(s / eps) * g;
h = eps / s*g;
j = -(L - 1)*s*c*g;
alpha = (h + j*u)*(h + j*u) + f*u*u + g*v*v;

if (alpha < 0)
{
return -1;
}
else
{
if (-(h + j*u) >= 0)
{
w = -(h + j*u) - sqrt((h + j*u)*(h + j*u) + f*u*u + g*v*v);
}
}
}
}

```

```

}
else
{
w = -(h + j*u) + sqrt((h + j*u)*(h + j*u) + f*u*u + g*v*v);
}
z = w*k;
}
}

poly = FD->xdata[1] * y + FD->xdata[2] * pow(x, 2) + FD->xdata[3] * pow(y, 2) +
FD->xdata[4] * pow(x, 2) * pow(y, 1) + FD->xdata[5] * pow(y, 3) + FD->xdata[6] *
pow(x, 4);
poly += FD->xdata[7] * pow(x, 2) * pow(y, 2) + FD->xdata[8] * pow(y, 4) + FD-
>xdata[9] * pow(x, 4) * pow(y, 1) + FD->xdata[10] * pow(x, 2) * pow(y, 3);
poly += FD->xdata[11] * pow(y, 5) + FD->xdata[12] * pow(x, 6) + FD->xdata[13] *
pow(x, 4) * pow(y, 2) + FD->xdata[14] * pow(x, 2) * pow(y, 4) + FD->xdata[15] *
pow(y, 6);
poly += FD->xdata[16] * pow(x, 6) * pow(y, 1) + FD->xdata[17] * pow(x, 4) *
pow(y, 3) + FD->xdata[18] * pow(x, 2) * pow(y, 5) + FD->xdata[19] * pow(y, 7);
poly += FD->xdata[20] * pow(x, 8) + FD->xdata[21] * pow(x, 6) * pow(y, 2) + FD-
>xdata[22] * pow(x, 4) * pow(y, 4) + FD->xdata[23] * pow(x, 2) * pow(y, 6) + FD-
>xdata[24] * pow(y, 8);
poly += FD->xdata[25] * pow(x, 8) * pow(y, 1) + FD->xdata[26] * pow(x, 6) *
pow(y, 3) + FD->xdata[27] * pow(x, 4) * pow(y, 5) + FD->xdata[28] * pow(x, 2) *
pow(y, 7) + FD->xdata[29] * pow(y, 9);
poly += FD->xdata[30] * pow(x, 10) + FD->xdata[31] * pow(x, 8) * pow(y, 2) + FD-
>xdata[32] * pow(x, 6) * pow(y, 4) + FD->xdata[33] * pow(x, 4) * pow(y, 6) + FD-
>xdata[34] * pow(x, 2) * pow(y, 8) + FD->xdata[35] * pow(y, 10);

UD->sag1 = z + poly;
UD->sag2 = UD->sag1;
break;

case 4:
/* ZEMAX wants a paraxial ray trace to this surface */
/* x, y, z, and the optical path are unaffected, at least for this surface type
*/
/* for paraxial ray tracing, the return z coordinate should always be zero. */
/* paraxial surfaces are always planes with the following normals */

/* for the ogive surface, we will assume ro = 0 for this computation */

UD->ln = 0.0;
UD->mn = 0.0;
UD->nn = -1.0;
power = (FD->n2 - FD->n1)*FD->cv;
if ((UD->n) != 0.0)
{
(UD->l) = (UD->l) / (UD->n);
(UD->m) = (UD->m) / (UD->n);

(UD->l) = (FD->n1*(UD->l) - (UD->x)*power) / (FD->n2);

```

```

(UD->m) = (FD->n1*(UD->m) - (UD->y)*power) / (FD->n2);

/* normalize */
(UD->n) = sqrt(1 / (1 + (UD->l)*(UD->l) + (UD->m)*(UD->m)));
/* de-paraxialize */
(UD->l) = (UD->l)*(UD->n);
(UD->m) = (UD->m)*(UD->n);
}
break;
case 5:
/* ZEMAX wants a real ray trace to this surface */
/*
/* slow the iteration down if the ray is steep */
slowfactor = FD->param[2];
if (slowfactor <= 0 || slowfactor > 1.0) slowfactor = 1.0;

/* make sure we do at least 1 loop */

R = FD->param[1];
if (FD->cv == 0)
{
k = 0;
}
else
{
k = 1 / FD->cv;
}
K = FD->k;

tp = 0.0;
t = 100.0;
x = UD->x;
y = UD->y;
z = UD->z;
loop = 0;

while (fabs(t) > 1e-10)
{
u = y / k;
v = x / k;
eps = R / k;
L = K + 1;
s = eps / sqrt(1 - K*eps*eps);
c = sqrt((1 - L*eps*eps) / (1 - K*eps*eps));

if ((R == 0) && (K == -1))
{
sag = (x*x + y*y) / 2 / k;
}
else

```

```

{
  if (R == 0)
  {
    alpha = k*k - (K + 1)*(x*x + y*y);
    if (alpha < 0)
    {
      return -1;
    }
    else
    {
      if (k / (K + 1) <= 0)
      {
        sag = 1 / (K + 1)*(k + sqrt(k*k - (K + 1)*(x*x + y*y)));
      }
      else
      {
        sag = 1 / (K + 1)*(k - sqrt(k*k - (K + 1)*(x*x + y*y)));
      }
    }
  }
  else
  {
    g = -1 / (L*c*c + s*s);
    f = (s / eps)*(s / eps) * g;
    h = eps / s*g;
    j = -(L - 1)*s*c*g;
    alpha = (h + j*u)*(h + j*u) + f*u*u + g*v*v;

    if (alpha < 0)
    {
      return -1;
    }
    else
    {
      if (-(h + j*u) >= 0)
      {
        w = -(h + j*u) - sqrt((h + j*u)*(h + j*u) + f*u*u + g*v*v);
      }
      else
      {
        w = -(h + j*u) + sqrt((h + j*u)*(h + j*u) + f*u*u + g*v*v);
      }
      sag = w*k;
    }
  }
}

poly = FD->xdata[1] * y + FD->xdata[2] * pow(x, 2) + FD->xdata[3] * pow(y, 2) +
FD->xdata[4] * pow(x, 2) * pow(y, 1) + FD->xdata[5] * pow(y, 3) + FD->xdata[6] *
pow(x, 4);
poly += FD->xdata[7] * pow(x, 2) * pow(y, 2) + FD->xdata[8] * pow(y, 4) + FD-
->xdata[9] * pow(x, 4) * pow(y, 1) + FD->xdata[10] * pow(x, 2) * pow(y, 3);

```

```

poly += FD->xdata[11] * pow(y, 5) + FD->xdata[12] * pow(x, 6) + FD->xdata[13] *
pow(x, 4) * pow(y, 2) + FD->xdata[14] * pow(x, 2) * pow(y, 4) + FD->xdata[15] *
pow(y, 6);
poly += FD->xdata[16] * pow(x, 6) * pow(y, 1) + FD->xdata[17] * pow(x, 4) *
pow(y, 3) + FD->xdata[18] * pow(x, 2) * pow(y, 5) + FD->xdata[19] * pow(y, 7);
poly += FD->xdata[20] * pow(x, 8) + FD->xdata[21] * pow(x, 6) * pow(y, 2) + FD-
>xdata[22] * pow(x, 4) * pow(y, 4) + FD->xdata[23] * pow(x, 2) * pow(y, 6) + FD-
>xdata[24] * pow(y, 8);
poly += FD->xdata[25] * pow(x, 8) * pow(y, 1) + FD->xdata[26] * pow(x, 6) *
pow(y, 3) + FD->xdata[27] * pow(x, 4) * pow(y, 5) + FD->xdata[28] * pow(x, 2) *
pow(y, 7) + FD->xdata[29] * pow(y, 9);
poly += FD->xdata[30] * pow(x, 10) + FD->xdata[31] * pow(x, 8) * pow(y, 2) + FD-
>xdata[32] * pow(x, 6) * pow(y, 4) + FD->xdata[33] * pow(x, 4) * pow(y, 6) + FD-
>xdata[34] * pow(x, 2) * pow(y, 8) + FD->xdata[35] * pow(y, 10);

sag += poly;

/* okay, now with sag in hand, how far are we away in z? */
dz = sag - z;

/* now compute how far along the z axis this is */
/* note this will crash if n == 0!! */
//t = dz / (UD->n);
t = dz * slowfactor;

/* for some aspheres, it is safer to use dz directly, as it is a smaller number
*/
/* the convergence will be slower if the ray angle is steep, fast if near
parallel to the axis */

/* propagate the additional "t" distance */
x += UD->l*t;
y += UD->m*t;
z += UD->n*t;

/* add in the optical path */
tp += t;

/* prevent infinite loop if no convergence */
loop++;
if (loop > 5000) return(-1);
}

/* okay, we should be at the intercept coordinates now */
UD->x = x;
UD->y = y;
UD->z = z;

/* don't forget the path! */
UD->path = tp;

/* now do the normals, this is the tricky part */

```

```

u = y / k;
v = x / k;
eps = R / k;
L = K + 1;
s = eps / sqrt(1 - K*eps*eps);
c = sqrt((1 - L*eps*eps) / (1 - K*eps*eps));

if ((R == 0) && (K == -1))
{
mx = x / k;
my = y / k;
}
else
{
if (R == 0)
{
alpha = k*k - (K + 1)*(x*x + y*y);
if (alpha < 0)
{
return -1;
}
else
{
if (k / (K + 1) <= 0)
{
mx = -1 * x / sqrt(k*k - (K + 1)*(x*x + y*y));
my = -1 * y / sqrt(k*k - (K + 1)*(x*x + y*y));
}
else
{
mx = x / sqrt(k*k - (K + 1)*(x*x + y*y));
my = y / sqrt(k*k - (K + 1)*(x*x + y*y));
}
}
}
else
{
g = -1 / (L*c*c + s*s);
f = (s / eps)*(s / eps) * g;
h = eps / s*g;
j = -(L - 1)*s*c*g;

alpha = (h + j*u)*(h + j*u) + f*u*u + g*v*v;
if (alpha < 0)
{
return -1;
}
else
{
if (-(h + j*u) >= 0)
{

```

```

mx = -1 * g*v / sqrt((h + j*u)*(h + j*u) + f*u*u + g*v*v);
my = -j - ((h + j*u)*j + f*u) / sqrt((h + j*u)*(h + j*u) + f*u*u + g*v*v);
}
else
{
mx = g*v / sqrt((h + j*u)*(h + j*u) + f*u*u + g*v*v);
my = -j + ((h + j*u)*j + f*u) / sqrt((h + j*u)*(h + j*u) + f*u*u + g*v*v);
}

}
}

poly_mx = 2 * FD->xdata[2] * pow(x, 1) + 2 * FD->xdata[4] * pow(x, 1) * pow(y,
1) + 4 * FD->xdata[6] * pow(x, 3);
poly_mx += 2 * FD->xdata[7] * pow(x, 1) * pow(y, 2) + 4 * FD->xdata[9] * pow(x,
3) * pow(y, 1) + 2 * FD->xdata[10] * pow(x, 1) * pow(y, 3);
poly_mx += 6 * FD->xdata[12] * pow(x, 5) + 4 * FD->xdata[13] * pow(x, 3) *
pow(y, 2) + 2 * FD->xdata[14] * pow(x, 1) * pow(y, 4);
poly_mx += 6 * FD->xdata[16] * pow(x, 5) * pow(y, 1) + 4 * FD->xdata[17] *
pow(x, 3) * pow(y, 3) + 2 * FD->xdata[18] * pow(x, 1) * pow(y, 5);
poly_mx += 8 * FD->xdata[20] * pow(x, 7) + 6 * FD->xdata[21] * pow(x, 5) *
pow(y, 2) + 4 * FD->xdata[22] * pow(x, 3) * pow(y, 4) + 2 * FD->xdata[23] *
pow(x, 1) * pow(y, 6);
poly_mx += 8 * FD->xdata[25] * pow(x, 7) * pow(y, 1) + 6 * FD->xdata[26] *
pow(x, 5) * pow(y, 3) + 4 * FD->xdata[27] * pow(x, 3) * pow(y, 5) + 2 * FD-
>xdata[28] * pow(x, 1) * pow(y, 7);
poly_mx += 10 * FD->xdata[30] * pow(x, 9) + 8 * FD->xdata[31] * pow(x, 7) *
pow(y, 2) + 6 * FD->xdata[32] * pow(x, 5) * pow(y, 4) + 4 * FD->xdata[33] *
pow(x, 3) * pow(y, 6) + 2 * FD->xdata[34] * pow(x, 1) * pow(y, 8);

mx += poly_mx;

poly_my = FD->xdata[1] + 2 * FD->xdata[3] * pow(y, 1) + FD->xdata[4] * pow(x, 2)
+ 3 * FD->xdata[5] * pow(y, 2);
poly_my += 2 * FD->xdata[7] * pow(x, 2) * pow(y, 1) + 4 * FD->xdata[8] * pow(y,
3) + FD->xdata[9] * pow(x, 4) + 3 * FD->xdata[10] * pow(x, 2) * pow(y, 2);
poly_my += 5 * FD->xdata[11] * pow(y, 4) + 2 * FD->xdata[13] * pow(x, 4) *
pow(y, 1) + 4 * FD->xdata[14] * pow(x, 2) * pow(y, 3) + 6 * FD->xdata[15] *
pow(y, 5);
poly_my += FD->xdata[16] * pow(x, 6) + 3 * FD->xdata[17] * pow(x, 4) * pow(y, 2)
+ 5 * FD->xdata[18] * pow(x, 2) * pow(y, 4) + 7 * FD->xdata[19] * pow(y, 6);
poly_my += 2 * FD->xdata[21] * pow(x, 6) * pow(y, 1) + 4 * FD->xdata[22] *
pow(x, 4) * pow(y, 3) + 6 * FD->xdata[23] * pow(x, 2) * pow(y, 5) + 8 * FD-
>xdata[24] * pow(y, 7);
poly_my += FD->xdata[25] * pow(x, 8) + 3 * FD->xdata[26] * pow(x, 6) * pow(y, 2)
+ 5 * FD->xdata[27] * pow(x, 4) * pow(y, 4) + 7 * FD->xdata[28] * pow(x, 2) *
pow(y, 6) + 9 * FD->xdata[29] * pow(y, 8);
poly_my += 2 * FD->xdata[31] * pow(x, 8) * pow(y, 1) + 4 * FD->xdata[32] *
pow(x, 6) * pow(y, 3) + 6 * FD->xdata[33] * pow(x, 4) * pow(y, 5) + 8 * FD-
>xdata[34] * pow(x, 2) * pow(y, 7) + 10 * FD->xdata[35] * pow(y, 9);

```

```

my += poly_my;

UD->nn = -sqrt(1.0 / (1.0 + (mx*mx) + (my*my)));
UD->ln = -mx*UD->nn;
UD->mn = -my*UD->nn;

//}

if (Refract(FD->n1, FD->n2, &UD->l, &UD->m, &UD->n, UD->ln, UD->mn, UD->nn))
return(-FD->surf);
break;
case 6:
/* ZEMAX wants the index, dn/dx, dn/dy, and dn/dz at the given x, y, z. */

/* This is only required for gradient index surfaces, so return dummy values */
UD->index = FD->n2;
UD->dndx = 0.0;
UD->dndy = 0.0;
UD->dndz = 0.0;
break;
case 7:
/* ZEMAX wants the "safe" data. */
/* this is used by ZEMAX to set the initial values for all parameters and extra
data */
/* when the user first changes to this surface type. */
/* this is the only time the DLL should modify the data in the FIXED_DATA FD
structure */
for (i = 1; i <= 8; i++) FD->param[i] = 0.0;
for (i = 1; i <= 200; i++) FD->xdata[i] = 0.0;
break;
}
return 0;
}
int Refract(double thisn, double nextn, double *l, double *m, double *n, double
ln, double mn, double nn)
{
double nr, cosi, cosi2, rad, cosr, gamma;
if (thisn != nextn)
{
nr = thisn / nextn;
cosi = fabs((*l) * ln + (*m) * mn + (*n) * nn);
cosi2 = cosi * cosi;
if (cosi2 > 1) cosi2 = 1;
rad = 1 - ((1 - cosi2) * (nr * nr));
if (rad < 0) return(-1);
cosr = sqrt(rad);
gamma = nr * cosi - cosr;
(*l) = (nr * (*l)) + (gamma * ln);
(*m) = (nr * (*m)) + (gamma * mn);
(*n) = (nr * (*n)) + (gamma * nn);
}
return 0;
}

```

Zemax user-defined surface C code for the Extended Cartesian Oval surface.

```

#include <windows.h>
#include <math.h>
#include <string.h>
#include <stdio.h>
#include "usersurf.h"

/*
Written by Dmitry Reshidko

base = Cartesian Oval
sag = A1*base + A2*base^2 + A3*base^3 + A4*base^4

The routine for computing the sag of Cartesian Oval was adopted from Jose
Sasian:

Four inputs are required 1) the distance from the object
to the surface SM, 2) the distance from the surface to the image SP, 3) The
index of refraction before
the surface n1, and 4) the index of refraction after the surface n2. These
parameters are parameters 1-4 in the
surface editor. The sag of the surface is found by iteration and minimizing the
opd error; a closed form
formula is used to determine the derivatives. For object or image at a large
distance an approximation on the
opd is made as to avoid errors. A good way to use this surface is to pick up the
distances SM and SP.
The distance SP is determined by a solve using the current surface radius of
curvature. The C code was written
by ZEMAX and modified by Jose Sasian 2010 */

int __declspec(dllexport) APIENTRY UserDefinedSurface3(USER_DATA *UD,
FIXED_DATA3 *FD);

/* a generic Snells law refraction routine */
int Refract(double thisn, double nextn, double *l, double *m, double *n, double
ln, double mn, double nn);

BOOL WINAPI DllMain(HANDLE hInst, ULONG ul_reason_for_call, LPVOID lpReserved)
{
return TRUE;
}

int __declspec(dllexport) APIENTRY UserDefinedSurface3(USER_DATA *UD,
FIXED_DATA3 *FD)
{
int i, loop, loop2, sag_error;
double x2, y2, tp, x, y, z, zi, dz, mx, my, cvt, njm, njp, op1, op2, pp, qq,
kk, ww, sag, hh;
double sm, sp, r2, r, dnn, aa4, deno, nume, mag, n1, n2;

```

```

double A1, A2, A3, A4, A5, A6;

double power, t, t2;

switch (FD->type)
{
case 0:
/* ZEMAX is requesting general information about the surface */
switch (FD->numb)
{
case 0:
/* ZEMAX wants to know the name of the surface */
/* do not exceed 12 characters */
strcpy(UD->string, "CARTESIAN");
break;
case 1:
/* ZEMAX wants to know if this surface is rotationally symmetric it is, so
return any character */
UD->string[0] = '1';
break;
case 2:
/* ZEMAX wants to know if this surface is a gradient index media it is not, so
return a null string */
UD->string[0] = '\0';
break;
}
break;
case 1:
/* ZEMAX is requesting the names of the parameter columns the value FD->numb
will indicate which value ZEMAX wants. returning a null string indicates that
the parameter is unused. */
switch (FD->numb)
{
case 1:
strcpy(UD->string, "sm");
break;
case 2:
strcpy(UD->string, "sp");
break;

case 3:
strcpy(UD->string, "n1");
break;
case 4:
strcpy(UD->string, "n2");
break;
case 5:
strcpy(UD->string, "A1");
break;
case 6:
strcpy(UD->string, "A2");
break;
case 7:

```

```

strcpy(UD->string, "A3");
break;
case 8:
strcpy(UD->string, "A4");
break;
case 9:
strcpy(UD->string, "A5");
break;
case 10:
strcpy(UD->string, "A6");
break;
default:
UD->string[0] = '\0';
break;
}
break;
case 2:
/* ZEMAX is requesting the names of the extra data columns the value FD->numb
will indicate which value ZEMAX wants. returning a null string indicates that
the extradata value is unused. */
switch (FD->numb)
{

default:
UD->string[0] = '\0';
break;
}
break;
case 3:
/* ZEMAX wants to know the sag of the surface */

UD->sag1 = 0.0;
UD->sag2 = 0.0;

cvt = FD->cv;

n1 = FD->param[3];
n2 = FD->param[4];

A1 = FD->param[5];
A2 = FD->param[6];
A3 = FD->param[7];
A4 = FD->param[8];
A5 = FD->param[9];
A6 = FD->param[10];

dnn = n2 - n1;
sm = FD->param[1];
sp = FD->param[2];

/* An ad-hoc iterative method to find the sag. The kk and ww parameters allow to
handle the cases when the object is to the right of left of the surface and for
positive and negative radius */

```

```

kk = 1;

if (sm > 0)
{
kk = -1;
}

ww = 1;

if (sp < 0)
{
ww = -1;
}

r2 = UD->x * UD->x + UD->y * UD->y;
r = sqrt(r2);

t = 100.0;
tp = 0.0;
x = UD->x;
y = UD->y;
z = UD->z;

/* Makes an estimate on the sag using gaussian optics and fourth order theory
and starts a loop to caculate the sag*/

cvt = (n2 / sp - n1 / sm) / dnn;
aa4 = -(1 / sm - cvt)*(1 / sm - cvt)*(1 / sp / n2 - 1 / sm / n1)*n1*n1 / 8 /
dnn;
z = r2*cvt / 2 + r2*r2*aa4;

loop = 0;
sag_error = 0;
while (fabs(t) > 1e-10)
{
if ((fabs(sm) <= 100000000))
{
op1 = n1*sqrt((-sm + z)*(-sm + z) + r2) + kk*n1*sm;
}

if ((fabs(sm) > 100000000))
{
op1 = n1*(z - r2 / 2 / sm - z*z / 2 / sm);
}

if ((fabs(sp) <= 100000000))

```

```

{
op2 = n2*sqrt((sp - z)*(sp - z) + r2) - ww*n2*sp;
}

if ((fabs(sp) > 100000000))
{
op2 = n2*(-z + r2 / 2 / sp + z*z / 2 / sp);
}

t = (kk*op1 + ww*op2) / dnn;
z = z + t;

loop++;
if (loop > 1000)
{
sag_error = 1;
break;
}
}

// if can't compute the sag - give an estimate based on the fourth-order theory
if (sag_error)
{
cvt = (n2 / sp - n1 / sm) / dnn;
aa4 = -(1 / sm - cvt)*(1 / sm - cvt)*(1 / sp / n2 - 1 / sm / n1)*n1*n1 / 8 /
dnn;
z = r2*cvt / 2 + r2*r2*aa4;
sag_error = 0;
}

sag = A1*z + A2*z*z + A3*z*z*z + A4*z*z*z*z + A5*z*z*z*z*z + A6*z*z*z*z*z*z;

UD->sag1 = sag;

UD->sag2 = UD->sag1;

break;

case 4:
/* ZEMAX wants a paraxial ray trace to this surface */
/* x, y, z, and the optical path are unaffected, at least for this surface type
*/
/* for paraxial ray tracing, the return z coordinate should always be zero. */
/* paraxial surfaces are always planes with the following normals */

UD->ln = 0.0;
UD->mn = 0.0;
UD->nn = -1.0;
njm = FD->n1;
njp = FD->n2;

sm = FD->param[1];
sp = FD->param[2];

```

```

n1 = FD->param[3];
n2 = FD->param[4];
dnn = n2 - n1;

A1 = FD->param[5];

cvt = (n2 / sp - n1 / sm) / dnn;

power = A1*dnn*cvt;

if ((UD->n) != 0.0)
{
(UD->l) = (UD->l) / (UD->n);
(UD->m) = (UD->m) / (UD->n);

(UD->l) = (FD->n1*(UD->l) - (UD->x)*power) / (FD->n2);
(UD->m) = (FD->n1*(UD->m) - (UD->y)*power) / (FD->n2);

/* normalize */
(UD->n) = sqrt(1 / (1 + (UD->l)*(UD->l) + (UD->m)*(UD->m)));
/* de-paraxialize */
(UD->l) = (UD->l)*(UD->n);
(UD->m) = (UD->m)*(UD->n);
}
break;
case 5:
/* ZEMAX wants a real ray trace to this surface */
/* do not allow n == 0 */
if (UD->n == 0.0) return -1;

/* An ad-hoc iterative method to find the sag. The kk and ww parameters allow to
handle the cases when the
object is to the right of left of the surface and for positive and negative
radius */

cvt = FD->cv;
n1 = FD->param[3];
n2 = FD->param[4];
dnn = n2 - n1;

sm = FD->param[1];
sp = FD->param[2];

A1 = FD->param[5];
A2 = FD->param[6];
A3 = FD->param[7];
A4 = FD->param[8];
A5 = FD->param[9];
A6 = FD->param[10];

kk = 1;

if (sm > 0)

```

```

{
kk = -1;
}

ww = 1;

if (sp < 0)
{
ww = -1;
}

tp = 0.0;
t2 = 100.0;
x = UD->x;
y = UD->y;
z = UD->z;

cvt = (n2 / sp - n1 / sm) / dnn;
aa4 = -(1 / sm - cvt)*(1 / sm - cvt)*(1 / sp / n2 - 1 / sm / n1)*n1*n1 / 8 /
dnn;

loop2 = 0;
while (fabs(t2) > 1e-12)
{

r2 = x * x + y * y;

/* Makes an estimate on the sag using gaussian optics and fourth order theory
and starts a loop to caculate the sag*/

zi = r2*cvt / 2 + r2*r2*aa4;
loop = 0;
t = 100.0;
sag_error = 0;
while (fabs(t) > 1e-12)
{

if ((fabs(sm) <= 100000000))
{
op1 = n1*sqrt((-sm + zi)*(-sm + zi) + r2) + kk*n1*sm;
}

if ((fabs(sm) > 100000000))
{
op1 = n1*(zi - r2 / 2 / sm - zi*zi / 2 / sm);
}
}
}

```

```

if ((fabs(sp) <= 10000000))
{
op2 = n2*sqrt((sp - zi)*(sp - zi) + r2) - ww*n2*sp;
}

if ((fabs(sp) > 10000000))
{
op2 = n2*(-zi + r2 / 2 / sp + zi*zi / 2 / sp);
}

t = (kk*op1 + ww*op2) / dnn;
zi = zi + t;

loop++;
if (loop > 1000) return (1);
}

sag = A1*zi + A2*zi*zi + A3*zi*zi*zi + A4*zi*zi*zi*zi + A5*zi*zi*zi*zi*zi +
A6*zi*zi*zi*zi*zi*zi;

/* okay, now with sag in hand, how far are we away in z? */
dz = sag - z;

/* now compute how far along the z axis this is */
/* slow it down by a factor of 0.01 to support steep angles */
t2 = dz*0.01;

/* for some aspheres, it is safer to use dz directly, as it is a smaller number
*/
/* the convergence will be slower if the ray angle is steep, fast if near
parallel to the axis */

/* propagate the additional "t" distance */
x += UD->l*t2;
y += UD->m*t2;
z += UD->n*t2;

/* add in the optical path */
tp += t2;

/* prevent infinite loop if no convergence */
loop2++;
if (loop2 > 10000) return(-1);
}

/* okay, we should be at the intercept coordinates now */
UD->x = x;
UD->y = y;

```

```

UD->z = z;
x2 = x*x;
y2 = y*y;
r2 = x2 + y2;
r = sqrt(r2);

/* don't forget the path! */
UD->path = tp;          // add fd->n1

// Find Cartesian Oval
zi = r2*cvt / 2 + r2*r2*aa4;
loop = 0;
t = 100.0;
while (fabs(t) > 1e-10)
{
    if ((fabs(sm) <= 100000000))
    {
        op1 = n1*sqrt((-sm + zi)*(-sm + zi) + r2) + kk*n1*sm;
    }

    if ((fabs(sm) > 100000000))
    {
        op1 = n1*(zi - r2 / 2 / sm - zi*zi / 2 / sm);
    }

    if ((fabs(sp) <= 100000000))
    {
        op2 = n2*sqrt((sp - zi)*(sp - zi) + r2) - ww*n2*sp;
    }

    if ((fabs(sp) > 100000000))
    {
        op2 = n2*(-zi + r2 / 2 / sp + zi*zi / 2 / sp);
    }

    t = (kk*op1 + ww*op2) / dnn;
    zi = zi + t;

    loop++;
    if (loop > 1000) return (-1);
}

/* now do the normals */
/*Knowing the sag z then we can have the derivatives in closed form as follows:*/
mag = -(n1*sm)*(1 / sp / n2);
if (((-1 + zi / sm)*(-1 + zi / sm) + r2 / sm / sm) < 0) return (-1);
pp = sqrt((-1 + zi / sm)*(-1 + zi / sm) + r2 / sm / sm);
if (((1 - zi / sp)*(1 - zi / sp) + r2 / sp / sp) < 0) return (-1);

```

```

qq = sqrt((1 - zi / sp)*(1 - zi / sp) + r2 / sp / sp);
deno = mag*(-1 + zi / sm) / sm / pp - (1 - zi / sp) / sp / qq;
nume = mag / sm / sm / pp + 1 / sp / sp / qq;

my = -y*nume / deno*(A1 + 2 * A2*zi + 3 * A3*zi*zi + 4 * A4*zi*zi*zi + 5 *
A5*zi*zi*zi*zi + 6 * A6*zi*zi*zi*zi*zi);
mx = -x*nume / deno*(A1 + 2 * A2*zi + 3 * A3*zi*zi + 4 * A4*zi*zi*zi + 5 *
A5*zi*zi*zi*zi + 6 * A6*zi*zi*zi*zi*zi);

UD->nn = -sqrt(1 / (1 + (mx*mx) + (my*my)));
UD->ln = -mx*UD->nn;
UD->mn = -my*UD->nn;

if (Refract(FD->n1, FD->n2, &UD->l, &UD->m, &UD->n, UD->ln, UD->mn, UD->nn))
return(-FD->surf);

break;

case 6:
/* ZEMAX wants the index, dn/dx, dn/dy, and dn/dz at the given x, y, z. */

/* This is only required for gradient index surfaces, so return dummy values */
UD->index = FD->n2;
UD->dndx = 0.0;
UD->dndy = 0.0;
UD->dndz = 0.0;
break;
case 7:
/* ZEMAX wants the "safe" data. */
/* this is used by ZEMAX to set the initial values for all parameters and extra
data when the user first changes to this surface type. This is the only time the
DLL should modify the data in the FIXED_DATA FD structure */
for (i = 1; i <= 100; i++) FD->param[i] = 0.0;
for (i = 1; i <= 100; i++) FD->xdata[i] = 0.0;
break;
}
return 0;
}

int Refract(double thisn, double nextn, double *l, double *m, double *n, double
ln, double mn, double nn)
{
double nr, cosi, cosi2, rad, cosr, gamma;
if (thisn != nextn)
{
nr = thisn / nextn;
cosi = fabs((*l) * ln + (*m) * mn + (*n) * nn);
cosi2 = cosi * cosi;
if (cosi2 > 1) cosi2 = 1;
rad = 1 - ((1 - cosi2) * (nr * nr));
if (rad < 0) return(-1);
cosr = sqrt(rad);
gamma = nr * cosi - cosr;
}
}

```

```
(*l) = (nr * (*l)) + (gamma * ln);  
(*m) = (nr * (*m)) + (gamma * mn);  
(*n) = (nr * (*n)) + (gamma * nn);  
}  
return 0;  
}
```