

UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

DEPARTMENT OF INFORMATION ENGINEERING
MASTER DEGREE IN ICT FOR INTERNET AND MULTIMEDIA

Neural network-based design of freeform off-axis three-mirror telescopes for space applications

MASTER CANDIDATE

Lorenzo Borsoi

Student ID 2056099

SUPERVISOR

Prof. Giampiero Naletto

University of Padova

ACADEMIC YEAR

2022/2023

5-12-2023

*To my Grandfather Gastone,
who has been a source of encouragement and inspiration throughout my life.*

Abstract

This work explores the development of an innovative Neural Network-based framework to automate the design of freeform off-axis three-mirror imaging systems. These optical systems, consisting of three freeform optical components arranged in a non-collinear manner, have enormous potential in fields such as space exploration and astronomy, due to their compactness and superior imaging capabilities.

Starting with a comprehensive overview of freeform optics, this thesis provides an in-depth explanation of the mathematical representations, fabrication, and metrology of freeform surfaces. The challenges of realizing complex optical systems are highlighted, emphasizing the need for efficient designs. Furthermore, we analyze the advantages of freeform off-axis three-mirror imaging systems in space exploration when compared to conventional designs, providing valuable context for the developed framework. In this thesis, we propose a methodology based on Neural Networks to generate effective starting points in the design process. The framework comprises several significant phases. To begin with, we identify the key parameters of the representative system which include the Field of View, F-number, and Entrance Pupil Diameter. Next, we establish the System Parameter Space (SPS) by taking into account the design requirements and the parameters involved in the system. Then, we create a dataset through systematic sampling within the SPS, using a system evolution approach to derive the corresponding surface parameters that can fully describe the location and shape of the surfaces. The Feed-Forward Neural Network (FFNN) is trained rigorously with the given dataset. Once it is validated and proven effective, the trained FFNN can quickly produce the corresponding surface parameters when specific system parameter combinations are provided. As a result, it serves as an optimal starting point for subsequent optimizations, significantly reducing the amount of manual effort required during the design process.

This novel framework represents a step forward in the fusion of advanced machine learning techniques with optical design principles. By automating and streamlining the design process, this framework sets the stage for a new era in the creation of high-performance optical systems, paving the way for future advancements in space exploration, astronomy, and various other domains.

Contents

List of Figures	xi
List of Tables	xiii
List of Acronyms	xix
1 Introduction	1
1.1 Background and context	1
1.1.1 Challenges and limitations of traditional telescopes	1
1.1.2 What is freeform optics?	2
1.1.3 Historical background and space industry trends	3
1.2 Problem definition	7
1.3 Objectives of the study	8
2 Freeform Optics	11
2.1 Mathematical descriptions	11
2.1.1 Analytical functions	12
2.1.2 Numerical orthogonal polynomials	24
2.1.3 Representation techniques for freeform surface with strong slope variation	25
2.2 Manufacturing	25
2.2.1 Ultraprecision machining	25
2.2.2 Polishing	27
2.3 Metrology	28
2.4 Investigating mid-spatial frequency errors	29
3 All-Reflective Unobscured TMA Telescopes	33
3.1 Imaging system characteristics	33
3.2 First-order mirror formula	35

CONTENTS

3.3	Classification of three-mirror configurations	37
3.3.1	Design parameters of a k-mirror telescope	39
3.4	Third-order correction of rotationally symmetric three-mirror telescopes	41
3.5	Aberration theory	44
3.5.1	Spherical aberration	48
3.5.2	Coma	48
3.5.3	Astigmatism and field curvature	49
3.5.4	Distortion	50
3.6	Design methods for freeform imaging systems	51
3.6.1	Aberration-based method	52
3.6.2	Direct design method	52
4	Freeform Off-axis Three-mirror Telescope Design: A Deep Learning-Enabled Approach	57
4.1	Framework	58
4.2	Dataset	61
4.3	Supervised learning	63
4.3.1	Network architecture	63
4.3.2	Network training and validation	65
5	Example Demonstration	69
5.1	Base system design	71
5.1.1	Control of structure constraints: the 3D point-to-line distance	74
5.2	Dataset generation and analysis	76
5.3	Network design and analysis	81
6	Conclusions and Future Works	87
A	Appendix	89
A.1	System evolution algorithm	89
A.2	Neural network training code	106
	References	109
	Ringraziamenti	121

List of Figures

1.1	Freeform surface: (a) 3D plot and (b) simulated interferogram (image source [1]).	2
1.2	Technology focus and market needs for freeform optics (image source [1]).	5
1.3	Comparison between sphere, asphere and freeform in terms of Field-of-View (FoV), f -number and Compactness.	8
2.1	Decomposition of a freeform surface (left) into the basic shape (middle) and higher-order deformations (right) (image source [25]).	12
2.2	The first 21 Zernike polynomials, ordered vertically by radial degree and horizontally by azimuthal degree.	15
2.3	First terms of Q-polynomials (image source [35]).	16
2.4	First terms in x and y for 2D Chebyshev first kind, on the left, and second kind, on the right (image source [35]).	18
2.5	First terms in x and y for the 2D Legendre polynomials (image source [35]).	19
2.6	First terms of A-polynomials first kind (on the left) and second kind (on the right) (image source [35]).	20
2.7	First terms in x and y for the XY-polynomials using the standard single-index sorting (image source [35]).	22
2.8	Power spectral density plot (image source [46]).	31
3.1	Two-surfaces imaging system (image source [51]).	33
3.2	A concave spherical mirror (image source [52]).	36
3.3	General geometry of a three-mirror system (image source [53]).	37
3.4	Zero-Petzval three-mirror telescope configurations. (image source [53]).	39

LIST OF FIGURES

3.5	Three-mirror anastigmat telescope patented by W. Wetherell and D. Womble (image source [53]).	43
3.6	Tangential (on the left) and sagittal (on the right) rays, CR represents the chief ray (image source [56]).	45
3.7	Illustration of ray and wave aberrations (image source [51]).	46
3.8	The image formation in the presence of spherical aberration (image source [61]).	49
3.9	The image formation in the presence of coma (image source [61]).	49
3.10	On the left, the image formation in the presence of astigmatism. On the right, the tangential and sagittal foci (image source [61]).	50
3.11	On the left, image of a squared grid affected by barrel distortion. On the right, image of a square grid affected by pincushion distortion (image source [61]).	51
3.12	A comprehensive selection of type Ia geometries for unobscured three-mirror imager (image source [64]).	53
4.1	Illustration of the design framework.	60
4.2	Illustration of a generic FFNN.	65
5.1	Ziz-zag geometry of freeform off-axis three-mirror telescope and its structure constraints.	70
5.2	On-axis rotationally symmetric Three-Mirror Anastigmat (TMA) layout.	72
5.3	BaseSys ₁ layout.	74
5.4	BaseSys ₁ RMS spots.	74
5.5	BaseSys ₁ Seidel aberrations diagram.	75
5.6	The point-to-line distance in three-dimensional space.	75
5.7	The average RMS spot size of each $F/2$ (on the left) and $F/2.25$ (on the right) imaging system as a function of Full-XFov, Full-YFoV, and ENPD.	77
5.8	The average RMS spot size of each $F/2.5$ (on the left) and $F/2.75$ (on the right) imaging system as a function of Full-XFov, Full-YFoV, and ENPD.	78
5.9	The average RMS spot size of each $F/3$ (on the left) and $F/3.25$ (on the right) imaging system as a function of Full-XFov, Full-YFoV, and ENPD.	78

5.10	The average RMS spot size of each $F/3.5$ (on the left) and $F/3.75$ (on the right) imaging system as a function of Full-XFov, Full-YFoV, and ENPD.	78
5.11	The average RMS spot size of each $F/4$ imaging system as a function of Full-XFov, Full-YFoV, and ENPD.	79
5.12	The maximum relative distortion and the total track length of each imaging system.	79
5.13	The Feed-Forward Neural Network design.	81
5.14	From the top to the bottom and from the left to the right. The network performance, the training state values, the error histogram, and the linear regression.	82
5.15	The average RMS spot size of each $F/2$ (on the left) and $F/2.4$ (on the right) imaging system as a function of Full-XFov, Full-YFoV, and ENPD.	83
5.16	The average RMS spot size of each $F/2.8$ (on the left) and $F/3.2$ (on the right) imaging system as a function of Full-XFov, Full-YFoV, and ENPD.	83
5.17	The average RMS spot size of each $F/3.6$ (on the left) and $F/4$ (on the right) imaging system as a function of Full-XFov, Full-YFoV, and ENPD.	84
5.18	The maximum relative distortion and the total track length of each imaging system.	84

List of Tables

2.1	Relation between conic constant values and surface shapes in non-zero surface curvature cases.	13
2.2	First terms of XY-polynomials with the corresponding interpretation and polynomial order.	21
3.1	The four classes of three-mirror imaging systems, distinguished by Roman numerals.	38
3.2	Parameters of a general k -mirror system.	40
3.3	Seidel aberration coefficients.	41
5.1	Surface data of $F/4$ on-axis rotationally symmetric TMA with ENPD = 50 mm.	72
5.2	BaseSys ₁ surface data.	73
5.3	Typical base systems obtained through system evolution.	80
5.4	Typical predicted systems obtained through our Feed-Forward Neural Network (FFNN).	85

List of Acronyms

- API** Application Programming Interface
- CAD** Computer-aided Design
- CHAPS** Compact Hyperspectral Air Pollution Sensor
- COM** Component Object Model
- DESIS** Deutsche Zentrum für Luft- und Raumfahrt (DLR) Earth Sensing Imaging Spectrometer
- DL** Deep Learning
- DLR** Deutsche Zentrum für Luft- und Raumfahrt
- ENPD** Entrance-Pupil Diameter
- ESA** European Space Agency
- FFNN** Feed-Forward Neural Network
- FoV** Field-of-View
- HSF** High-spatial Frequency
- IRMOS** Infrared Multi-Object Spectrometer
- ISS** International Space Station
- LSF** Low-spatial Frequency
- JWST** James Webb Space Telescope
- MLP** Multilayer Perceptron
- MSE** Mean Squared Error

LIST OF TABLES

MSF Mid-spatial Frequency

MUSES Multi-User-System for Earth Sensing

NSO Netherlands Space Office

NURBS Non-Uniform Rational B-Splines

RBF Radial Basis Function

RMS Root Mean Square

SCUBA-2 Submillimetre Common-User Bolometer Array 2

SL Supervised Learning

TANGO ANthropogenic Greenhouse Gas Observers

TMA Three-Mirror Anastigmat

TOTR Total Track Length

TROPOMI TROPOspheric Monitoring Instrument

ZPL Zemax Programming Language



Introduction

The chapter begins by tracing the challenges and limitations faced by traditional telescopes, such as central obscuration, rotational symmetry constraints, limited design options, size and weight issues, and narrow FoV. It introduces the concept of freeform optics as a revolutionary solution to these challenges. The chapter also explores the historical context of freeform optics and its applications in various fields, particularly in space exploration. It declares the problem statement and outlines the study's objectives: this sets the stage for the subsequent sections, focusing on the study's framework and methodologies.

1.1 BACKGROUND AND CONTEXT

1.1.1 CHALLENGES AND LIMITATIONS OF TRADITIONAL TELESCOPES

Traditional telescope designs have faced several challenges and limitations that have hindered their performance and capabilities. These challenges include:

- **Central Obscuration:** traditional telescopes often suffer from central obscuration, which is the blocking of light by the secondary mirror or other components in the optical path. This can lead to reduced image quality and decreased resolution.
- **Rotational Symmetry:** traditional telescope designs are limited by rotationally symmetric optics. These components restrain the degrees of freedom available for optimizing the system and minimizing aberrations.
- **Limited Design Options:** traditional telescopes have been constrained by the available materials and optical surfaces. Conversely, freeform optical

1.1. BACKGROUND AND CONTEXT

components are attractive because they can be designed to behave in ways traditional optics cannot, offering optical design flexibility not possible with conventional optics.

- **Size and Weight:** traditional telescope designs often result in larger and heavier systems due to the need for multiple optical elements and the constraints of rotational symmetry. This can make them less portable and more challenging to deploy in various applications.

These limitations have led to a necessity for alternative design approaches, such as freeform optics, to overcome these challenges and improve telescope performance.

1.1.2 WHAT IS FREEFORM OPTICS?

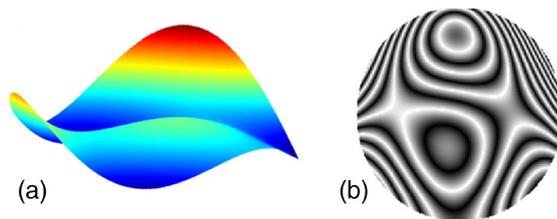


Figure 1.1: Freeform surface: (a) 3D plot and (b) simulated interferogram (image source [1]).

The exploration of freeform optical surfaces marks a transformative leap in the domain of optical systems, challenging the traditional paradigms of rotational symmetry. Historically, optical surfaces were predominantly designed with rotational symmetry, due to the limited manufacturing capabilities and the complexities involved in assembly and integration. However, the unceasing demand for scientific advancement, particularly in fields like astronomy, Earth observation, and planetary exploration, necessitates optical instruments of exceptional performance, compactness, and adaptability.

Freeform optics is a research field that involves the use of non-rotationally symmetric surfaces in optical systems, opening avenues for revolutionary designs in lenses and mirrors. These surfaces lack a fixed axis [2], allowing for diverse and complex shapes that traditional rotational surfaces cannot achieve (a freeform surface is illustrated in Fig. 1.1). The departure from rotational symmetry is not merely a conceptual shift; it represents a radical improvement from the conventional design landscape, ushering in a new era of possibilities.

Freeform optics finds its *raison d'être* in scenarios where conventional rotational symmetry fails to deliver optimal results. In the field of off-axis optical systems - where components are deliberately tilted and positioned away from the central optical axis - conventional symmetric surfaces encounter significant challenges. In such circumstances, off-axis configurations often yield tilt-induced optical aberrations which exhibit complicated field dependencies [3]. Freeform optics, however, breaks away from the constraints of symmetry, presenting innovative solutions to mitigate these tilt-induced aberrations [4]. By embracing asymmetry, freeform optics empowers optical systems to function effectively in highly intricate, folded, and off-axis configurations. This capability holds immense significance in critical fields like space exploration, where precision and efficiency are paramount. Through this paradigm shift, freeform optics reshapes the landscape of optical engineering, offering tailored solutions for complex challenges previously considered impossible.

While freeform optics offer groundbreaking solutions, their integration into existing systems poses challenges. Technological risks and the complexities of disruptive innovation necessitate careful consideration, especially in risk-averse sectors like space exploration. However, ongoing research, industry-university partnerships, and the establishment of specialized centers indicate a promising trajectory. The continuous advancements in freeform optics have been gradually reducing the gap between traditional and freeform optics. As a result, freeform optics is emerging as the new frontier in optical design innovation, promising to revolutionize the field of optics with its cutting-edge capabilities.

1.1.3 HISTORICAL BACKGROUND AND SPACE INDUSTRY TRENDS

Freeform optical systems have significantly transformed various sectors, marking a paradigm shift in optical design and application. The origins of this revolutionary technology can be traced back to pioneering innovations in the early 20th century. The inception of freeform optics is exemplified by the anamorphic lens, introduced during World War I, which utilized toroidal surfaces with non-rotational symmetry [1]. This design, featuring circular profiles but two radii along orthogonal axes, was initially employed in periscopes to widen the FoV outside tanks, showcasing the earliest use of non-symmetric optical surfaces.

A major milestone emerged in 1927 when Henri Chrétien designed the Hy-

1.1. BACKGROUND AND CONTEXT

pergonar lens for photography and motion capture [1]. This development laid the foundation for the cinematic revolution of the 1950s, demonstrating the potential of freeform optics in shaping the future of visual technologies. Concurrently, progressive ophthalmic lenses pioneered the integration of freeform optics into mass-produced consumer products, exemplified by the Alvarez lens in the mid-20th century [5]. This innovative lens design utilized cubic-shaped lenses displaced laterally, enabling variable focus in visual instruments and contributing to the early applications of freeform optics in everyday devices.

The consumer market witnessed another breakthrough in 1972 with the introduction of the SX-70 Polaroid camera, featuring non-rotational aspheres in its viewfinder [6]. Developed by James Baker, William Plummer, and their collaborators, this camera became a commercial success, with millions sold until its interruption in 2005. This widespread adoption highlighted the practical applicability of freeform optics in consumer electronics, setting the stage for further innovations.

In the field of imaging optics, the 1970s and 1980s marked a crucial period of experimentation and development. During this time, notable figures like Offner, Cook, Wetherell, Womble, Tatian, and Shafer explored various designs that incorporated freeform surfaces into unobscured three-mirror imagers. Offner, for example, employed off-axis conic sections to create an afocal TMA for efficient beam reduction [7]. Building upon Offner's work, Cook introduced a focal version of the TMA in 1979 [8]. Simultaneously, Wetherell and Womble contributed to the field by developing a reimaging-free TMA, a configuration later identified as the reflective triplet [9]. Independently, Tatian and Shafer delved into integrating freeform surfaces in unobscured three-mirror imagers. Tatian utilized plane-symmetric X - Y polynomials up to the 10th order, meticulously constructed upon a quadric surface [10]. In parallel, Shafer employed two-axis plate aspheres, reaching up to the 10th order, with one decentered element, thus broadening the scope of freeform optics applications [11].

Figure 1.2 shows a visual representation of the several applications enabled by both imaging and nonimaging optics. Within the realm of lighting and illumination, freeform surfaces play a key role in customizing light emitted from a specific source, achieving predetermined illumination patterns with remarkable efficiency. Similarly, quantum cryptography uses freeform optical transformations for high-efficiency outcomes. In these applications, centered around the common themes of illumination and sorting, the precise contours of optics are

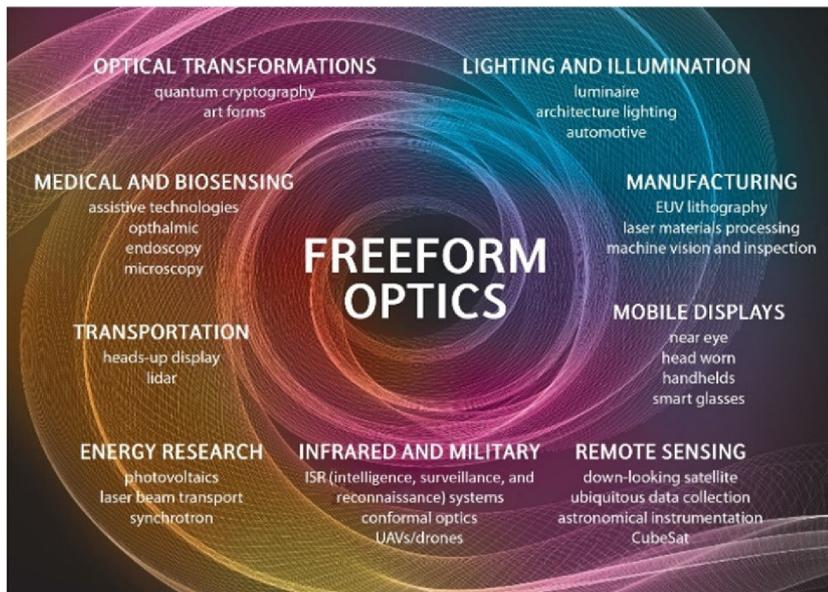


Figure 1.2: Technology focus and market needs for freeform optics (image source [1]).

not as rigorously defined as in traditional imaging, where nanometer-scale precision is typically indispensable for correcting aberrations.

Currently, the integration of freeform optics has permeated various sectors, including remote sensing and military instrumentation, energy research, transportation, manufacturing, as well as medical and biosensing technologies. Freeform optics shows significant potential in both refractive and all-reflective unobscured systems. Both approaches yield compact and high-performance solutions, with the added benefit of lightweight and achromatic capabilities specifically attributed to all-reflective designs.

In the field of space exploration, freeform optics has become increasingly important. Off-axis TMA telescopes, such as those used on the James Webb Space Telescope (JWST), utilize conicoid surfaces to greatly improve both performance and compactness [12]. This evolution is underscored by the confidence in metrology for space-borne systems, paving the way for freeform optics to become commonplace in space optics.

The adoption of freeform optics in space missions has been marked by a cautious yet progressive approach. Ground-based astronomy applications, exemplified by NASA's Infrared Multi-Object Spectrometer (IRMOS) [13] and the Submillimetre Common-User Bolometer Array 2 (SCUBA-2) [14] on the James

1.1. BACKGROUND AND CONTEXT

Clerk Maxwell telescope, have underlined the efficacy of freeform optics.

In the traditionally risk-averse space industry, there is a notable shift towards innovation. Leading organizations like NASA, European Space Agency (ESA), and the Netherlands Space Office (NSO) have invested significantly in freeform optics research. NASA, in particular, has acknowledged the immense potential of freeform surfaces and actively supports groundbreaking initiatives such as the MiniSpec proposal [15], focusing on miniaturized imaging spectrometers for CubeSat platforms. Collaborations between ESA and NSO on the Sentinel-5P satellite further exemplify this trend, with TROPospheric Monitoring Instrument (TROPOMI) freeform primary mirror enabling unprecedented performance with smaller ground pixels and enhanced signal-to-noise ratios [16].

The Sentinel-5 telescope, part of the European Earth Observation program Copernicus, is a prime example of the application of freeform optics in space missions [17]. It is equipped with a UV1 spectrometer based on an Offner-type spectrometer which has been adapted to employ freeform optics and an aspheric off-axis grating. Sentinel-5 will be placed on Metop-SG-A to monitor air quality and investigate atmosphere composition-climate interaction. Under the industrial prime contractor Airbus Defence and Space GmbH Germany, NSO is in charge of the design and manufacturing of the telescope assemblies as well as the UV1 spectrometer for Sentinel-5. Moreover, NSO has been involved in other significant space projects, such as the Compact Hyperspectral Air Pollution Sensor (CHAPS)-D instrument, designed for spaceborne applications, and planned for initial airborne deployment [18]. Additionally, NSO is a key player in the Twin ANthropogenic Greenhouse Gas Observers (TANGO) mission, comprising two agile satellites, TANGO-Carbon and TANGO-Nitro, equipped with compact spectrometers from the NSO Spectrolite family [19]. TANGO will monitor and quantify emissions of methane and carbon dioxide at the level of individual industrial facilities and power plants.

Additionally, the DLR has pioneered the development of DLR Earth Sensing Imaging Spectrometer (DESI), leveraging freeform surfaces to enhance the performance of its Offner-type spectrometer [20]. This technology has been integrated into the Multi-User-System for Earth Sensing (MUSES) on the International Space Station (ISS), representing a significant leap in space-borne observation capabilities.

These innovative initiatives, coupled with dynamic industry-university partnerships such as establishing the Center for Freeform Optics at the University

of Rochester, underscore the trajectory of freeform optics toward maturity and widespread adoption in future space missions. As continuous technological advancements propel the field forward, freeform optics is poised to revolutionize space exploration and various industries, opening doors to unprecedented possibilities in optical design and performance. The future holds the promise of a new era in optics, where freeform technology will continue redefining the boundaries of optically achievable, both on Earth and in the depths of space.

1.2 PROBLEM DEFINITION

Freeform off-axis three-mirror telescopes are needed in space applications due to several reasons. Firstly, the constant drive to develop more compact and lightweight optical instruments with enhanced performance necessitates the use of off-axis configurations. These configurations enable the elimination of central obscuration - which improves image contrast and quality, - but break the rotational symmetry of the system, making it difficult for rotationally symmetric surfaces to compensate for tilt-induced optical aberrations. Secondly, freeform surfaces in off-axis mirror systems enable the correction of aberrations and the achievement of diffraction-limited image quality [21]. By utilizing the correction ability of freeform surfaces, these telescopes can provide high-resolution imaging with a large clear aperture and low f -number [22]. Furthermore, freeform optics have the potential to reduce the size and weight of optical instruments by maintaining the same optical performance [23]. This size reduction is particularly important in space applications where weight and size constraints are critical. Eventually, we must state that traditional telescopes with on-axis designs have a limited FoV due to the obstruction caused by the secondary mirror. Conversely, freeform off-axis three-mirror telescopes can provide a wider and unobstructed FoV [24]. This is particularly advantageous for space missions requiring simultaneous observations of large regions of the sky, such as wide-field surveys. An illustration of the main advantages provided by freeform optics is depicted in Figure 1.3.

The objective is to design a framework that can generate multiple three-mirror imaging systems with off-axis alignment, using a given set of system parameters. The framework should enable designers to have more options and flexibility in choosing and filtering the output systems. The challenge lies in the large number of input and output parameters, as well as the wide range of

1.3. OBJECTIVES OF THE STUDY

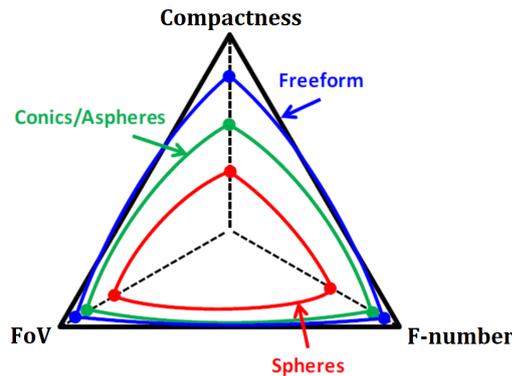


Figure 1.3: Comparison between sphere, asphere and freeform in terms of FoV, f -number and Compactness.

parameter values. To overcome this, this thesis proposes the use of a FFNN and supervised learning to generate high-quality freeform systems.

1.3 OBJECTIVES OF THE STUDY

With our neural network-based design approach, our goal is to achieve optimized and innovative designs for various applications. Here are some objectives we aim to accomplish:

1. **Improved Design Efficiency:** neural networks can analyze large datasets and learn from them to identify patterns and trends. By utilizing this approach, we aim to streamline and accelerate the design process. It allows us to explore a wide range of design possibilities in a shorter time frame, leading to more efficient and effective solutions.
2. **Enhanced Performance:** neural networks can help us to optimize designs by considering multiple factors simultaneously. We can train the network on relevant design parameters, constraints, and performance metrics. Hence, directly finding optimal designs that maximize desired outcomes, such as minimizing aberrations, improving image quality, or maximizing compactness.
3. **Innovative and Non-intuitive Designs:** neural networks can learn complex relationships between design variables and outcomes. This can help to uncover non-intuitive design configurations that traditional approaches may not discover. By leveraging the power of neural networks, we can push the boundaries of design possibilities and explore novel solutions.
4. **Design Automation:** our neural network-based approach enables us to automate the design process to a certain extent. Once the network is trained, it can generate design recommendations based on input parameters and

user requirements. This automation saves time and resources for designers, allowing them to focus on higher-level tasks and decisions.

5. Optimal Trade-offs: design optimization often involves trade-offs between different design objectives and constraints. Neural networks can assist in finding optimal solutions considering these trade-offs. By training the network on various range of design scenarios and performance metrics, we can generate designs that strike the right balance between conflicting requirements.

Hence, by utilizing neural networks, we aim to deliver innovative and optimized solutions tailored to designers' specific requirements.

2

Freeform Optics

2.1 MATHEMATICAL DESCRIPTIONS

Freeform surfaces can be defined as surfaces with no axis of rotational invariance (within or beyond the optical part). A crucial question at the beginning of the concept and design of a freeform system is how to mathematically describe the surface itself. In this scenario, not only the optical design, but also the mechanical design, the manufacturing, and the assembly of the component inside the whole system are meaningful. In this section, we will directly refer to the works of Broemel [25], Gross [26] and Ye [27].

From the viewpoint of practical work and efficiency, there are several criteria for this selection; the selected surface representation should allow for a fast ray trace and calculation of intersection points and local slopes in the ray trace. Moreover, the parametrization of the surface should be flexible with a small number of parameters, allowing for easy manipulation and optimization. Thus, the surface parameters optimization should be robust and converge quickly, providing a good result in the design process. Nevertheless, the surface description should allow for easy access to aberrations, enabling the designer to analyze and correct for any optical imperfections. It also should have a direct relation to tolerancing, allowing for easy analysis of the impact of deviations on system performance. On the other hand, the surface description should allow for a simple extension of the region of interest to a larger area for fixing the mounting, without high gradients at the edges.

2.1. MATHEMATICAL DESCRIPTIONS

Representation techniques of freeform surfaces may be

- Analytical functions: orthogonal polynomials and non-orthogonal functions;
- Discrete points-based methods;
- Hybrid or combination methods.

2.1.1 ANALYTICAL FUNCTIONS

Freeform surfaces can be described using analytical formulae. We commonly recognize two main parts: the basic shape and the higher-order term. The basic shape, which can be a sphere, conic, or biconic, incorporates the quadratic contributions around the axis and determines the paraboloidal behavior of the surface. The higher-order term, responsible for aberration correction, contains the freeform contributions and describes additional deformations. Despite typically having a smaller amplitude, the second term is crucial for achieving precise optical performance. See figure 2.1 for illustration.

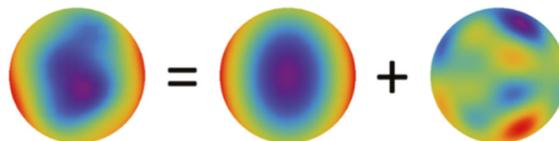


Figure 2.1: Decomposition of a freeform surface (left) into the basic shape (middle) and higher-order deformations (right) (image source [25]).

Basic shapes Understanding the fundamental shapes of optical surfaces forms the cornerstone of optical design. These shapes also serve as a starting point for creating more complex freeform surfaces. By leveraging basic shapes, engineers can confine the sag departure of higher-order freeform components, facilitating easier manufacturing processes and ensuring feasibility for production. This approach strikes a balance between creative design freedom and practicality, enabling the development of innovative optical systems that push the boundaries of what is achievable in optics.

The biconic optical surface is the most general basic shape. The second-order mathematical function represents it and its formula is written as

$$z_{biconic}(x, y) = \frac{c_x x^2 + c_y y^2}{1 + \sqrt{1 - (1 + k_x)c_x^2 x^2 - (1 + k_y)c_y^2 y^2}} \quad (2.1)$$

where x and y are the cartesian coordinates on the surface, z is the surface "sag" or z -coordinate, while the curvatures c_x , c_y and the conic constants k_x , k_y are specified in the two perpendicular cross sections, respectively.

If $c = c_x = c_y$ and $k = k_x = k_y$ are fulfilled at the same time, the surface is rotationally symmetric. This surface, known as conic, can be written as

$$z_{conic}(x, y) = \frac{c(x^2 + y^2)}{1 + \sqrt{1 - (1 + k)(x^2 + y^2)}} \quad (2.2)$$

The relations between the value of the conic constant and the surface shape in the corresponding cross-section are summarized in Table 2.1.

Table 2.1: Relation between conic constant values and surface shapes in non-zero surface curvature cases.

Conic constant	Surface shape
$k < -1$	Hyperboloid
$k = -1$	Paraboloid
$-1 < k < 0$	Prolate ellipsoid
$k = 0$	Sphere
$k > 0$	Oblate ellipsoid

A special case is represented by the planar surface where the curvatures in both cross sections are zero, thus, its surface sag is

$$z_{planar}(x, y) = 0 \quad (2.3)$$

2.1.1.1 ORTHOGONAL POLYNOMIALS

Orthogonal polynomials are often used to analyze optical surface deviations and wavefront aberrations due to their elegant mathematical performance. In this article, we describe mathematical orthogonality and different analytical orthogonal polynomials.

2.1. MATHEMATICAL DESCRIPTIONS

Polynomials Orthogonality Given $\mathcal{P}_m(x, y), \mathcal{P}_n(x, y)$ as Polynomial functions and $\omega(x, y)$ as weighting function over a specified area of support (S), we speak of *orthogonal* polynomials when eq. 2.4 is satisfied, while, we refer to *slope orthogonal* polynomials when eq. 2.5 is verified.

$$\langle \mathcal{P}_m, \mathcal{P}_n \rangle = \iint_S \omega(x, y) \mathcal{P}_m(x, y) \mathcal{P}_n(x, y) dx dy = \delta_{mn} c_n \quad (2.4)$$

$$\langle \vec{\nabla} \mathcal{P}_m, \vec{\nabla} \mathcal{P}_n \rangle = \iint_S \omega(x, y) \vec{\nabla} \mathcal{P}_m(x, y) \vec{\nabla} \mathcal{P}_n(x, y) dx dy = \delta_{mn} c_n \quad (2.5)$$

where δ_{mn} is the Kronecker delta function, c_n is a constant and the subscripts m and n are the non-negative integers. Note that when $m = n$ then $\delta_{mn} = 1$, conversely, when $m \neq n$ then $\delta_{mn} = 0$. In the case of an orthonormal set, $c_n = 1$ for all n .

The main advantages of orthogonal polynomials over non-orthogonal ones are here summarized from Takaki's work [28].

- Orthogonal polynomials uniquely contribute to surface shape and, following eq. 2.4, do not cancel or overlap, giving coefficient values physical insight into surface shape.
- Each surface corresponds to a unique set of coefficients for a given polynomial basis set, since every set of orthogonal polynomials is also linearly independent, avoiding unnecessarily complex descriptions.

Zernike The Zernike polynomials are a spatially orthogonal set, which is well known for describing wavefront errors and aberrations. The terms are defined in polar coordinates and have a constant weighting function [29]. The Standard convention is:

$$z(\rho, \theta) = z_{basic}(\rho, \theta) + \sum_{m,n} a_{mn} Z_n^m(\rho, \theta) \quad \text{with } m, n = 0, 1 \dots N \quad (2.6)$$

The set is defined as a product of a radial part $R_n^m(\rho)$ with an azimuthal part $\Theta^m(\theta)$ and a normalization factor $N(m, n)$.

$$Z_n^m(\rho, \theta) = N(m, n) R_n^m(\rho) \Theta^m(\theta) \quad (2.7)$$

In the standard convention, the Zernike polynomials are normalized to the Root Mean Square (RMS):

$$N(m, n) = \sqrt{\frac{2(n+1)}{1 + \delta_{m0}}} \quad (2.8)$$

The radial part is written as

$$R_n^m(\rho) = \sum_{k=0}^{n-|m|/2} (-1)^k \binom{n-m}{k} \binom{n-2k}{\frac{n-m}{2}-k} \rho^{n-2k} \quad (2.9)$$

Finally, the azimuthal part is described for angles θ measured against the x -axis according to the following formula:

$$\Theta^m(\theta) = \begin{cases} \sin(|m|\theta), & m < 0 \\ \cos(|m|\theta), & m > 0 \\ 1, & m = 0 \end{cases} \quad (2.10)$$

In Fig. 2.2 the Zernike terms are shown graphically.

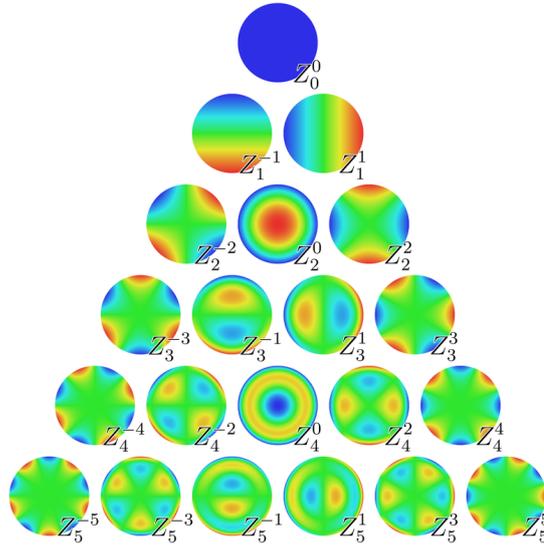


Figure 2.2: The first 21 Zernike polynomials, ordered vertically by radial degree and horizontally by azimuthal degree.

Q-poly (Forbes) Forbes proposed a set of polynomials to represent aspheric surfaces with large deviations and one for constrained slopes [30][31][32]. These

2.1. MATHEMATICAL DESCRIPTIONS

polynomials were based on Jacobi polynomials, and they were found to be useful in improving the manufacturability and testability of aspheres [33]. Building upon his previous work, a new polynomial set (Q_n^m) was developed to characterize optical freeform surfaces [34]. Similar to the Zernike polynomial, it represents the deviation between the freeform surface and the best-fit sphere (*bfs*) along the normal direction. Overall, this technique helps to overcome the numerical deficiencies of even-order polynomials.

If we define $u = \rho/\rho_{max}$ where ρ_{max} is the radius of the tightly enclosing cylinder, then the sag of the freeform surface along the z -axis is

$$\begin{aligned}
 z(\rho, \theta) &= z_{bfs}(\rho, \theta) + z_{mild-asphere}(\rho, \theta) + z_{Q-poly}(\rho, \theta) \\
 &= \frac{c_{bfs}\rho^2}{1 + \sqrt{1 - c_{bfs}^2\rho^2}} + \frac{1}{\sqrt{1 - c_{bfs}^2\rho^2}} \left[u^2(1 - u^2) \sum_{n=0}^N a_n^0 Q_n^0(u^2) \right] + \\
 &\quad \frac{1}{\sqrt{1 - c_{bfs}^2\rho^2}} \left[\sum_{m=1} u^m \sum_{n=0} [a_n^m \cos(m\theta) + b_n^m \sin(m\theta)] Q_n^m(u^2) \right] \\
 &\quad \text{with } m, n = 0, 1 \dots M, N \quad (2.11)
 \end{aligned}$$

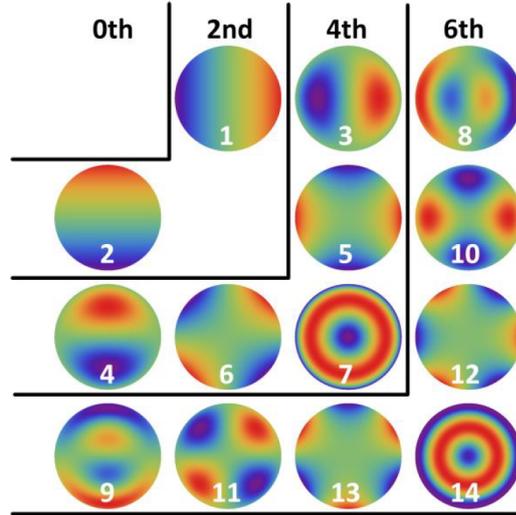


Figure 2.3: First terms of Q-polynomials (image source [35]).

Chebyshev The Chebyshev 2D polynomials are Cartesian products of the 1D Chebyshev polynomials. The resulting set is spatial-orthogonal on a unit square.

An explicit expression for Chebyshev polynomials of the first kind, in the interval $[-1, 1]$ is given by

$$T_m(x) = \cos(m \arccos(x)) \quad (2.12)$$

The one-dimensional term $T_m(x)$ is orthogonal in the range $[-1, 1]$ with a weighting function $\omega(x) = 1/\sqrt{1-x^2}$. Therefore, a 2D Chebyshev polynomial of the first kind can be written via a simple product:

$$z(x, y) = z_{basic}(x, y) + \sum_{m,n}^{M,N} a_{mn} T_m(x) T_n(y) \quad (2.13)$$

It is straightforward to show the orthogonality on the unit square of these 2D polynomials:

$$\langle C_{mn}^{1st}, C_{m'n'}^{1st} \rangle = \int_{-1}^1 \int_{-1}^1 \frac{T_m(x) T_n(y) \cdot T_{m'}(x) T_{n'}(y)}{\sqrt{(1-x^2)(1-y^2)}} dx dy = K \delta_{mm'} \delta_{nn'} \quad (2.14)$$

where K is a constant value used for the normalization of 2-D Chebyshev polynomials

$$K = \begin{cases} \pi^2 & \text{for } n = m = 0 \\ \pi^2/4 & \text{for } n = m \neq 0 \\ \pi^2/2 & \text{otherwise} \end{cases} \quad (2.15)$$

An elegant expression for Chebyshev polynomials of the second kind is given by

$$U_m(x) = \sum_{k=0}^{\lfloor m/2 \rfloor} \binom{m+1}{2k+1} x^{m-2k} (x^2-1)^k \quad (2.16)$$

Chebyshev polynomials of the second kind are orthogonal in the interval $[-1, 1]$ with a weighting function $\omega(x) = \sqrt{1-x^2}$. Therefore, a 2D Chebyshev polynomial of the second kind can be written via a simple product:

$$z(x, y) = z_{basic}(x, y) + \sum_{m,n}^{M,N} a_{mn} U_m(x) U_n(y) \quad (2.17)$$

2.1. MATHEMATICAL DESCRIPTIONS

As before, it is straightforward to show the orthogonality on the unit square of these 2D polynomials:

$$\begin{aligned} \langle C_{mn}^{2nd}, C_{m'n'}^{2nd} \rangle &= \int_{-1}^1 \int_{-1}^1 U_m(x)U_n(y) \cdot U_{m'}(x)U_{n'}(y) \sqrt{(1-x^2)(1-y^2)} dx dy \\ &= \frac{\pi^2}{4} \delta_{mm'} \delta_{nn'} \quad (2.18) \end{aligned}$$

The first terms of the Chebyshev first and second kind can be seen in Fig. 2.4.

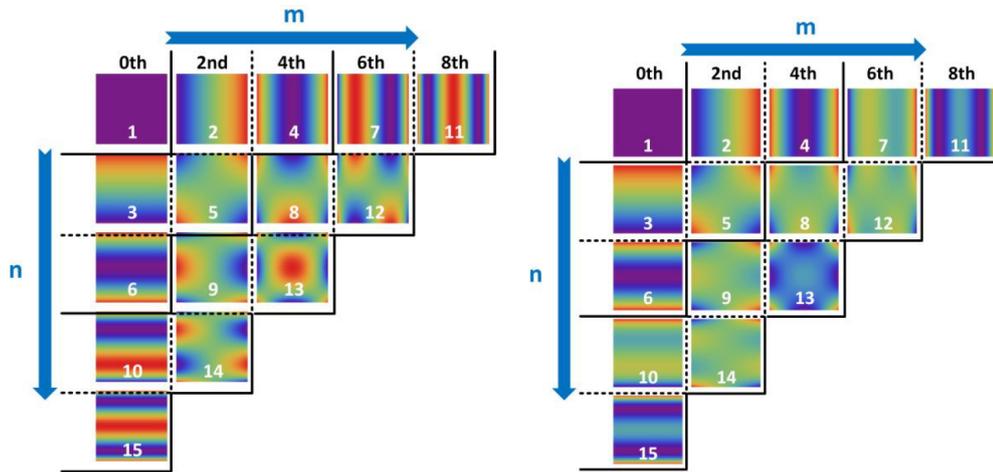


Figure 2.4: First terms in x and y for 2D Chebyshev first kind, on the left, and second kind, on the right (image source [35]).

Legendre The Legendre polynomials are obtained similarly to the Chebyshev polynomials as Cartesian products of one-dimensional functions. An explicit expression for Legendre polynomials is

$$P_m(x) = \frac{1}{2^m} \sum_{k=0}^{\lfloor m/2 \rfloor} (-1)^k \binom{m}{k} \binom{2m-2k}{m} x^{m-2k} \quad (2.19)$$

The Legendre polynomials are orthogonal in the interval $[-1, 1]$ with a unitary weighting function $\omega(x) = 1$. Therefore, a 2D Legendre polynomial of the

first kind can be written via a simple product:

$$z(x, y) = z_{basic}(x, y) + \sum_{m,n}^{M,N} a_{mn} P_m(x) P_n(y) \quad (2.20)$$

It is straightforward to show the orthogonality on the unit square of these 2D polynomials:

$$\langle L_{mn}, L_{m'n'} \rangle = \int_{-1}^1 \int_{-1}^1 L_m(x) L_n(y) \cdot L_{m'}(x) L_{n'}(y) \delta_{nn'} = \frac{4\delta_{mm'}\delta_{nn'}}{(2m+1)(2n+1)} \quad (2.21)$$

The first terms of the Legendre polynomials can be seen in Fig. 2.5.

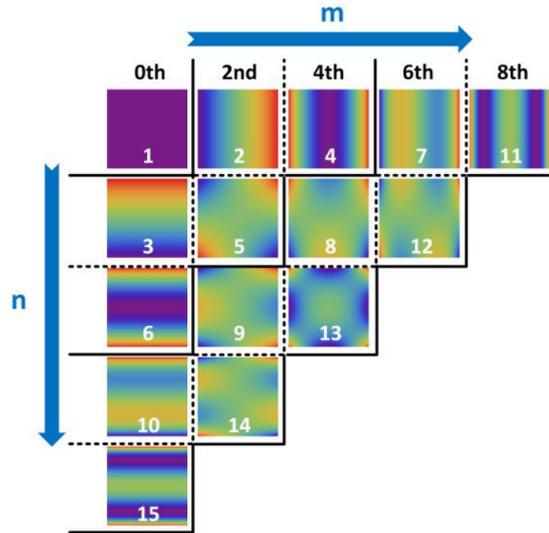


Figure 2.5: First terms in x and y for the 2D Legendre polynomials (image source [35]).

A-poly (Broemel) The A-polynomials are the most effective polynomial set for rectangular domains [25]. They use Forbes' approach to provide superior access to manufacturing and tolerancing, and their biconic shape includes lower-order astigmatism. Thanks to its biconic basic shape, the concept of a "best-fit shape" becomes irrelevant. Previous polynomial sets for rectangular domains have significant drawbacks for design, as detailed in [35]. Therefore, the approach of Bray [36] was employed in this case.

2.1. MATHEMATICAL DESCRIPTIONS

In general, the A-polynomials are described by:

$$z(x, y) = z_{biconic} + \frac{B(x, y)}{P_{biconic}(x, y)} \sum_j^J a_j A_j(x, y) \quad (2.22)$$

Where $B(x, y)$ is the boundary function and $P_{biconic}(x, y)$ is the projection factor of a surface with a biconical basic shape (we refer to [35] for additional informations).

The two sets of A-polynomials were developed using the Gram-Schmidt process [37] and the modified relation to ensure slope orthogonality (see Eq. 2.5).

The first kind A-polynomials (Fig. 2.6 left) are developed with:

$$\langle f_m^{1st}, f_n^{1st} \rangle = \int_{-1}^1 \int_{-1}^1 \vec{\nabla} [f_m^{1st}(x, y)] \cdot \vec{\nabla} [f_n^{1st}(x, y)] dx dy \quad (2.23)$$

The second kind A-polynomials (Fig. 2.6 right) are developed with:

$$\langle f_m^{2nd}, f_n^{2nd} \rangle = \frac{\int_{-1}^1 \int_{-1}^1 \frac{1}{\sqrt{1-x^2}\sqrt{1-y^2}} \vec{\nabla} [(x^2 + y^2)f_m^{2nd}(x, y)] \cdot \vec{\nabla} [(x^2 + y^2)f_n^{2nd}(x, y)] dx dy}{\int_{-1}^1 \int_{-1}^1 \frac{dx dy}{\sqrt{1-x^2}\sqrt{1-y^2}}} \quad (2.24)$$

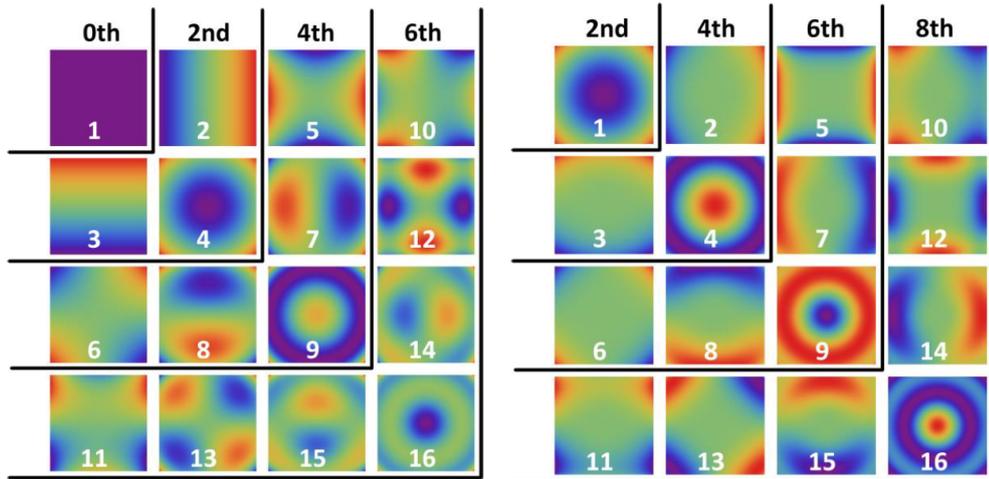


Figure 2.6: First terms of A-polynomials first kind (on the left) and second kind (on the right) (image source [35]).

2.1.1.2 NON-ORTHOGONAL FUNCTIONS

XY-poly XY-polynomials, Monomials, or *Extended Polynomials* (in OpticStudio), are a simple Taylor expansion in x and y with no orthogonality at all. The description can be expressed in Cartesian coordinates as follows:

$$z(x, y) = z_{basic}(x, y) + \sum_{m,n}^{M,N} a_{mn} x^m y^n \quad \text{with } m, n = 0, 1 \dots M, N \quad (2.25)$$

XY-poly, despite being non-orthogonal, allows for a fast, convenient representation. Thus, it was our top choice for designing the off-axis freeform TMA base system in section 4.1. Tab.2.2 and Fig.2.7 displays the first terms of polynomials sorted by aberrational order through a simple transformation of double-indices (m, n) into a single-index j . Furthermore, we explicit the OpticStudio single-indexing (j_{OS}) for *Extended Polynomials*:

$$\begin{cases} j = \frac{(m+n)^2 + m + 3n}{2} + 1 & \text{Standard single-indexing} \\ j_{OS} = \frac{(m+n)^2 + m + 3n}{2} & \text{OpticStudio single-indexing} \end{cases} \quad (2.26)$$

Table 2.2: First terms of XY-polynomials with the corresponding interpretation and polynomial order.

j	j _{OS}	m	n	Term	Interpretation	Polynomial order
1		0	0	1		0 th
2	1	1	0	x	decenter	2 nd
3	2	0	1	y		
4	3	2	0	x^2	with $j = 6$: astigmatism 0° or defocus astigmatism 45°	4 th
5	4	1	1	xy		
6	5	0	2	y^2		
7	6	3	0	x^3	each term corresponds to trefoil (either 0° or 30°) and coma (x or y)	6 th
8	7	2	1	x^2y		
9	8	1	2	xy^2		
10	9	0	3	y^3		
11	10	4	0	x^4	combination of terms correspond to spherical aberration, four sheets (0° and 22.5°) and secondary astigmatism	8 th
12	11	3	1	x^3y		
13	12	2	2	x^2y^2		
14	13	1	3	xy^3		
15	14	0	4	y^4		

2.1. MATHEMATICAL DESCRIPTIONS

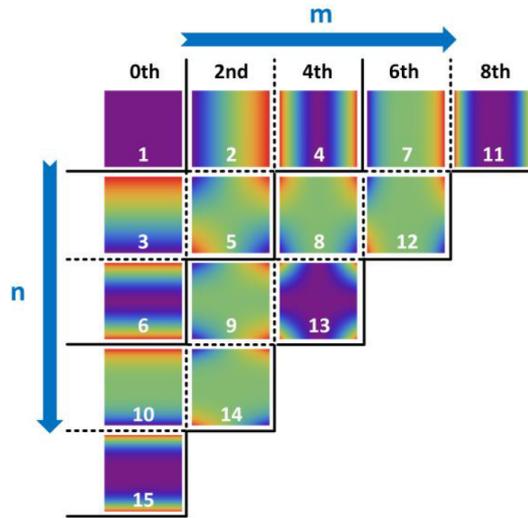


Figure 2.7: First terms in x and y for the XY -polynomials using the standard single-index sorting (image source [35]).

Spline surface (NURBS) Non-Uniform Rational B-Splines (NURBS) are locally described, point-based mathematical functions [38]. They are expressed in parametric form using control points and weights. NURBS surfaces are commonly used in Computer-aided Design (CAD) systems and computer graphics software for modeling and rendering realistic, complex, 3D surfaces.

NURBS surfaces offer several advantages over other surface representations. They provide greater flexibility and control in shaping complex surfaces, allowing for smooth curves and transitions. NURBS surfaces also support local surface descriptions, meaning that changes to control points only affect a small surface neighborhood, making it easier to modify specific areas without affecting the entire surface. The applications of NURBS surfaces are wide-ranging; they are commonly used in automotive design, aerospace engineering, industrial design, and animation. Chrisp [39] designed an imaging freeform optical system using NURBS surfaces.

Its mathematical expression is

$$z(u, v) = \frac{\sum_{i=0}^n \sum_{j=0}^m \omega_{i,j} B_{i,k}(u) B_{j,l}(v) P_{i,j}}{\sum_{i=0}^n \sum_{j=0}^m \omega_{i,j} B_{i,k}(u) B_{j,l}(v)} \quad (2.27)$$

where $P_{i,j}$ is the control function, $\omega_{i,j}$ is the weighted factor. $B_{i,k}(u)$ is the k^{th} -order B-spline function along the u -direction (Eq. 2.28); similarly, $B_{j,l}(v)$ is the

l^{th} -order B-spline function along the v -direction. The subscripts i, j, k , and l are non-negative integers.

$$\begin{cases} k = 0, & B_{i,0}(u) = \begin{cases} 1, & \text{if } u_i \leq u < u_{i+1} \text{ and } u_i < u_{i+1} \\ 0, & \text{otherwise} \end{cases} \\ k \geq 1, & B_{i,k}(u) = \frac{u-u_i}{u_{i+k}-u_i} B_{i,k-1}(u) + \frac{u_{i+k+1}-u}{u_{i+k+1}-u_{i+1}} B_{i+1,k-1}(u) \end{cases} \quad (2.28)$$

Radial Basis Function (RBF) Local basis functions are a mathematical representation used to provide a local representation of shape for optical surfaces. They are a sum of basis functions, where each basis function represents a specific shape or feature of the surface. They offer several advantages in optical design. First, they can accurately represent non-rotationally symmetric surfaces as well as rotationally symmetric surfaces. Second, they facilitate the optimization of freeform systems by allowing for the optimization of slopes on the surface. Third, they provide a flexible and efficient way to describe complex surface shapes. Its mathematical description is a linear combination of radial basis functions with the conic surface:

$$z(x, y) = z_{conic}(x, y) + \sum_{n=1}^N \omega_n \Phi_n \quad (2.29)$$

where Φ_n is the radial basis function and ω_n is the corresponding coefficient. The radial basis function is expressed as Eq. 2.30 over a specified support.

$$\Phi_n(\mathbf{p}) = \phi(\|\mathbf{p} - \mathbf{p}_n\|_2) \quad (2.30)$$

where \mathbf{p} is a point (x, y) in the supported region, \mathbf{p}_n is the supported center (x_n, y_n) , $\|\mathbf{p} - \mathbf{p}_n\|_2$ is the Euclidean norm and ϕ is the basis function for constructing the radial basis function.

The function ϕ is often expressed in the form of a Gaussian. This is advantageous because the Gaussian function is smooth, has approximate local characteristics (meaning that the function decreases rapidly as the distance from the center increases), and has good analyticity of solution [40].

2.1.2 NUMERICAL ORTHOGONAL POLYNOMIALS

It is important to note that analytical orthogonal polynomials, generated using the Gram-Schmidt orthogonalization method with differently shaped apertures, are only theoretically orthogonal over their continuous domain of definition. In practice, using discrete data points or dealing with complex apertures can significantly impact the orthogonality of these polynomials. Therefore, it is crucial to use a set of orthogonal polynomials that can adapt to discrete data points.

Numerical orthogonal polynomials are a set of polynomials that can be used to fit a discrete data point cloud of a freeform surface. These polynomials can be obtained using a matrix transformation method [41]. The method involves deriving the derivatives of numerical orthogonal polynomials in the x and y directions by partial differential. These derivatives are then used as basis functions to obtain the numerical orthogonal gradient polynomials, which can be employed to fit the measured slope or gradient data and transform it to the reconstructed wavefront directly over the general shaped aperture. The numerical orthogonal polynomials are chosen to be orthogonal over the discrete data points and can be expressed as a linear combination of Zernike polynomials owing to its orthogonality and completeness, as expressed in

$$F_l(x_n, y_n) = \sum_{j=1}^J M_{lj} Z_j(x_n, y_n) \quad (2.31)$$

where $F_l(x_n, y_n)$ is the numerical orthogonal polynomial, the subscript l is the ordering number, and (x_n, y_n) is the n^{th} point coordinate out of N in the effective region. M_{lj} is the transformation coefficient, and J is the number of terms of the Zernike polynomials $Z_j(x_n, y_n)$.

Eq. 2.31 can be also written in a matrix form:

$$\mathbf{F} = \mathbf{Z}\mathbf{M}^T \quad (2.32)$$

where the numerical matrix \mathbf{F} and \mathbf{Z} are two $N \times J$ matrices and \mathbf{M}^T is the transpose of the transform matrix \mathbf{M} .

2.1.3 REPRESENTATION TECHNIQUES FOR FREEFORM SURFACE WITH STRONG SLOPE VARIATION

Ye et al. [27] explain that accurately representing freeform surfaces with strong slope variations requires a precise characterization of fine local features. One approach is to use a linear combination of base polynomials for nanometer-accurate surface fitting. However, this method can be computationally inefficient due to the need for hundreds or thousands of terms. Alternatively, a hybrid or combination method can provide high accuracy with fewer polynomial terms. Although local fitting coefficients may not be meaningful for the entire surface, this method effectively characterizes local surface features for freeform surface estimation.

2.2 MANUFACTURING

Manufacturing freeform optics presents challenges due to the absence of rotational symmetry, requiring more than two degrees of freedom for fabrication. Various methods like ultraprecision machining, loose/bound abrasive finishing, molding/replication, and novel techniques are employed to overcome these challenges. In this section, we will directly refer to the works of S. Kumar [42], F. Z. Fang [43] and J. P Rolland [1].

2.2.1 ULTRAPRECISION MACHINING

In recent developments, two important approaches have emerged: Slow Tool Servo (STS) and Fast Tool Servo (FTS). These methods have revolutionized the machining of freeform surfaces, making ultra-precision cutting highly effective.

Slow Tool Servo (STS) STS enhances standard two-axis diamond turning machines by incorporating a high-precision C-axis encoder. This upgrade allows it to operate with C-axis controls and a T-shaped configuration. In STS, the Z-axis oscillates while X and C maintain constant speed. STS is known for its slower operation due to the massive Z-axis and limited speed and acceleration of the driving motor. Despite its longer machining time, STS offers superior surface finish with respect to FTS and is ideal for specific freeform optics applications, such as off-axis aspheric mirrors and structured optics like aspheric lens arrays.

2.2. MANUFACTURING

Fast Tool Servo (FTS) FTS utilizes a linear actuator to drive ultra-precision lathes. It operates in a T-type configuration similar to SSS and enables high-frequency movements in the Z-axis. FTS is commonly used for diamond turning micro prisms, lens arrays, torics, and off-axis aspherics with small sags.

Ultra-Precision Diamond Turning (UPDT) UPDT, also known as Single Point Diamond Turning (SPDT) or Diamond Turning Machining (DTM), originated in the 1960s. The key advantage of SPDT lies in its ability to achieve precision at the nanoscale level, owing to the sharpening of diamond tool edges. By delicately removing thin layers of material, SPDT ensures high form accuracy and smooth surfaces. This technique can handle both ductile and brittle materials, including aluminum, gold, copper alloys, and acrylic plastics. UPDT's effectiveness relies on precise control of process parameters, tooling, material selection, processing algorithms and machine cleanliness.

Ultra-Precision Milling (UPM) UPM is a high-speed machining process utilizing diamond tools fixed on rotating spindles. This technique is ideal for creating microstructures and channels for biomedical, imaging, and lighting applications. Surface quality in UPM is influenced by factors such as spindle rotational motion errors, vibrations, feed rate, spindle speed, tool geometry, and material properties. Achieving high precision with UPM necessitates careful control of the interaction between the cutting tool and workpiece.

Ultra-Precision Fly-Cutting (UPFC) UPFC is an intermittent cutting method employing a diamond tool mounted on a spindle. It can be performed in end-fly-cutting (EFC) or radial-fly cutting (RFC) modes. EFC is suitable for generating large freeform smooth surfaces, while RFC excels in creating special features. Combining these techniques with FTS and STS approaches enables the fabrication of diverse features, including torics, microlens arrays, prisms, and Fresnel microstructures. UPFC, although time-consuming, is highly effective for intricate shapes such as rectangular biconics and components for head-up displays.

Ultra-Precision Grinding (UPG) UPG utilizes grinding wheels with ultrafine grains to remove material. Achieving exceptional accuracy in UPG demands

precise grinding tool maintenance and various grinding kinematics like cross-grinding, parallel grinding, and fixed spot grinding. Parallel grinding, where the grinding tool is tilted at 45° , results in better form accuracy compared to cross-grinding. UPG is particularly suited for hard and brittle materials like ceramics and fused silica. The challenge lies in transitioning from brittle to ductile behavior during material removal.

2.2.2 POLISHING

Finishing processes are crucial to transforming optical ground workpieces into final optics products. Several key processes, including grinding, machining, and polishing, are utilized. However, these processes often leave tool footprints and unexpected patterns on complex freeform surfaces, leading to surface defects and sub-surface damages. To achieve optimum optical performance, these defects need to be removed through superfinishing processes. Various corrective techniques are employed to control topography errors such as low-spatial frequency, Mid-spatial Frequency (MSF), and high-spatial frequency, as defined in ISO 10110-8:2019.

Bonnet Polishing Bonnet polishing is a sub-aperture deterministic polishing technique where a spherical bag made of knitted cloth and rubber lamination serves as the polishing tool. This method efficiently removes defects from optical surfaces, making it suitable for large aperture optics and smaller components.

Magnetorheological (MR) Finishing MR finishing employs a magnetorheological polishing fluid that stiffens under a magnetic field, allowing precise control over surface roughness. This technique is capable of nano-level finishing and is adaptable for various freeform surfaces, using different setups and toolpaths.

Laser Polishing Laser polishing, a non-contact thermodynamic process, utilizes laser energy to irradiate workpiece surfaces, reducing roughness and eliminating defects. This technique offers two variations: cold polishing, involving laser ablation, and thermal polishing, which uses heat accumulation to achieve surface smoothness. Laser polishing is effective for polishing complex surfaces that traditional methods cannot handle.

2.3. METROLOGY

Ion Beam Polishing (IBP) IBP utilizes plasma to remove material at the atomic level, providing nanoscale figure correction. Location-specific processing (LSP) allows controlled material removal, enabling precise corrections. IBP offers stable and controlled removal with minimal impact on surrounding areas, making it suitable for optical components. However, its limitations include environmental dependence and unsuitability for reflective metals.

2.3 METROLOGY

The metrology of freeform optics, characterized by complex shapes and deviations from best-fit spheres, poses unique challenges. Manufacturers require precise measurements, including height resolution better than 1 nm, lateral resolution allowing several hundred measurements across the aperture, traceability to international standards, and verified uncertainty below 1/10th of the applied wavelength.

Various metrology techniques have been developed to address these challenges:

1. **Full-field Interferometry:** this method provides form and MSF measurements with high repeatability, requiring a stable environment, a reference surface, and consideration of surface roughness. Commercial Fizeau interferometers are used, although extreme freeform departures may necessitate optical nulls such as computer-generated holograms. Dynamic adjustable optical nulls and sub-aperture stitching extend interferometer capabilities. White Light Interferometry offers null-free accurate measurements in the MSF range, though retrace errors need correction.
2. **Coherent Measurement Techniques:** techniques like Temporal White Light Interferometry and Phase Retrieval provide measurements without optical nulls. Tilted Wave Interferometer uses multiple point sources and a height map conversion process. Phase Measuring Deflectometry is a null-free full-frame metrology solution, measuring MSF errors quickly. Numerical reconstruction algorithms and techniques like Zonal, Modal, or Hybrid integration methods enhance precision.
3. **Slope Measuring Systems:** lateral shearing interferometry and Shack–Hartman wavefront sensors are used for slope measurements. Spatially dithered distributions of binary pixels and differential Shack–Hartman techniques improve spatial resolution and dynamic range.
4. **Coordinate Measuring Machines (CMMs):** CMMs, including contact and optical probes, provide point-wise measurements over large volumes.

Challenges include large measurement times and normal incidence conditions. Compensation methods, such as using separate CMMs, are employed to mitigate errors.

5. Atomic Force Microscopy: atomic force microscopy computes surface properties from 3D structures, measuring surface roughness and auto-correlation. Limitations include size constraints for measured parts.
6. Surface Scattering Techniques: They analyze scattered light to characterize surface properties, offering unique solutions for surface characterization. Bidirectional scatter distribution function and Total Integrated Scatter measurements provide valuable insights into surface roughness and slope.

Each technique has specific strengths and limitations, and the choice depends on factors such as surface complexity, measurement requirements, and cost considerations. Continuous advancements in these techniques enhance the precision and efficiency of freeform optics metrology.

2.4 INVESTIGATING MID-SPATIAL FREQUENCY ERRORS

After our previous investigation, we found that freeform designs can be accurately described by a few parameters and polynomial representations of moderate or low orders (see Section 2.1). However, in reality, manufactured surfaces exhibit irregular perturbations, local deviations, and higher spatial frequency perturbations due to the material removal process.

Obtaining accurate measurement data for real freeform surfaces is essential to ensure that they meet the required performance specifications. Typically, the surface topology is scanned - by, for example, interferometry - to capture discrete values at various points, creating a point cloud representation. These data points are then used to describe and analyze the surface, allowing for evaluating performance specifications and optimizing the optical design. Any deviations or perturbations from the desired specifications can be identified and addressed by comparing the real surface $z_{real}(x, y)$ with the ideal surface $z_{ideal}(x, y)$ [44]:

$$z_{dev}(x, y) = z_{real}(x, y) - z_{ideal}(x, y) \quad (2.33)$$

Power spectral density (PSD) Optical surface errors in freeform surfaces can be classified based on their spatial frequency by analyzing the power spectral density. PSD is the Fourier transform of the surface topography for a range of

2.4. INVESTIGATING MID-SPATIAL FREQUENCY ERRORS

spatial frequencies, which contains the power - and not the phase - information. In order to properly analyze these imperfections on the surface along with their orientations and distribution, a two-dimensional (2D) evaluation is necessary. The two-dimensional discrete Fourier Transform (2D-DFT) of the surface profile z_{dev} is expressed as [45]

$$U(v_x, v_y) = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} z_{dev}(x, y) \exp\left\{-j2\pi\left(\frac{v_x x}{N_x} + \frac{v_y y}{N_y}\right)\right\} \Delta x \Delta y \quad (2.34)$$

In these expressions, v_x and v_y are the frequency variables, N_x and N_y are the total number of sample points in each domain, respectively. While Δx and Δy are the spacing between the pixels, where it is assumed that $\Delta x = d_x/N_x$ and $\Delta y = d_y/N_y$, with d_x and d_y the widths of the surface in the x - and y -direction. The two-dimensional Power spectral density (2D-PSD) can be derived as

$$PSD_{2D}(v_x, v_y) = \frac{1}{A} |U(v_x, v_y)|^2 = \frac{|U(v_x, v_y)|^2}{d_x d_y} \quad (2.35)$$

If one wants to evaluate the 2D-PSD along one dimension, it just needs to be integrated over any line. Here, we provide the 1D-PSD along the y -coordinate

$$PSD(v) = \sum_{x=0}^{N_x} PSD_{2D}(v_x, v_y) \Delta v_x \quad (2.36)$$

where $\Delta v_x = 1/d_x$ is the data point spacing in the frequency domain along v_x -direction. The 1D-PSD can be used to compute the RMS surface error, which is the area underneath the 1D-PSD curve and can be written as

$$\sigma = \sqrt{\sum PSD(v) \Delta v} \quad (2.37)$$

Fig. 2.8 displays a typical 1D-PSD curve. The whole spectral range is divided into low-, mid- and high-spatial frequency ranges. The curve is mainly linear but with nonlinear deviation on the left end. The slope of the line depends on the fabrication machine. The frequency scale starts at $1/D$ and ends at $1/\lambda$, where D is the surface width and λ is the used wavelength.

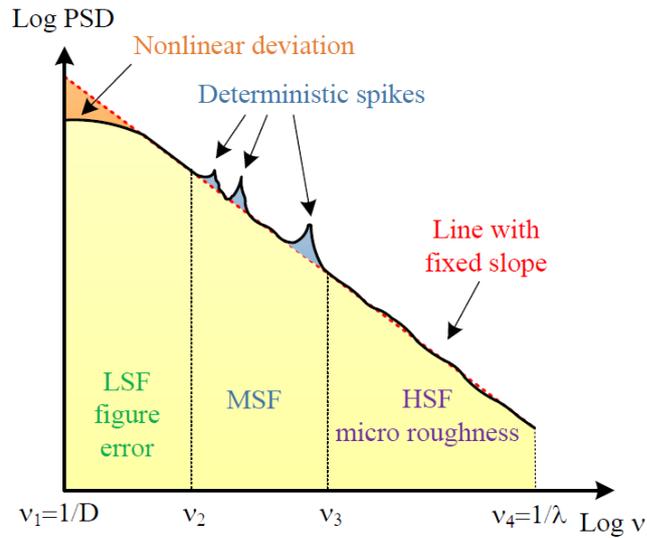


Figure 2.8: Power spectral density plot (image source [46]).

Low-spatial Frequency (LSF) errors are commonly referred to as *figure errors* and occur at a large scale or over a long distance and can cause surface irregularities or aberrations such as defocus, astigmatism, and coma. The LSF errors are fully deterministic and have a lower frequency of occurrence, which can result in blurred or distorted images produced by optical systems. Low-frequency shape deviations can be minimized by various post-processing techniques, like magnetorheological finishing. Usually, they are described by various kinds of polynomials [26] or RBF [44]. The relationship between the number of approximating polynomials and the spatial spectrum of the approximation is that as the order of polynomials increases, the spatial spectrum of the approximation also increases. Therefore, the number of polynomials used in the approximation directly affects the range of spatial frequencies that can be accurately represented. Moreover, a higher number of polynomials allows for a more accurate representation of the surface. However, it is important to note that there is a trade-off between the number of polynomials used and the computational complexity of the approximation process. Therefore, it is necessary to find a balance between accuracy and computational efficiency when determining the appropriate number of polynomials for a given approximation task.

Mid-spatial frequency (MSF) errors are typically separated by a few micrometers and have an intermediate frequency range. These errors consist of both random error, which is the reason for the linear decay of the PSD, and deterministic

2.4. INVESTIGATING MID-SPATIAL FREQUENCY ERRORS

spikes. The deterministic spikes result from computer-controlled sub-aperture techniques, as a combination of machine oscillations, thermal changes and tool size effects. The irregularities in diamond-turned surfaces create significant challenges due to their complex and quasi-periodic nature. In addition to the typical concentric undulations found on rotationally symmetric objects, there are also radially distributed structures that extend from the center of rotation. This means that the surfaces of these objects are not uniform in all directions, and waves of different orientations exist. MSF errors can have a significant impact on the quality of an imaging system and can lead to reduced image sharpness or contrast. In summary, the influence of MSF errors on the system performance comes from the mixture of both statistical and deterministic errors. One mitigation approach could be to avoid the formation of MSF waviness in the first place, for example by adopting pseudo-random tool paths in addition to post-polishing process [47].

A computationally huge number of Zernike [48] or Q-polynomial [49] terms is necessary to represent these effects adequately; therefore, different surface descriptions, such as superpositions of radially symmetric sine and cosine functions [44], are more appropriate for describing the impact of the deviations on the system performance. There are also combinations of surface representations found in literature, which combine smooth polynomials with localized descriptions [50].

High-spatial Frequency (HSF) errors occur at a small scale. These errors include surface roughness or fine polishing marks and can cause scattering or diffraction of light, resulting in reduced image resolution or increased image noise. The HSF errors are the micro-roughness on the surface and are purely statistical. Theoretically, the maximum frequency equals $1/\lambda$. However, the maximum frequency that can be resolved is limited by the spatial resolution. HSF errors are often addressed through polishing processes to improve the surface finish of an optical component. High spatial frequency errors are commonly referred to as *finish* and are quantified using RMS roughness.

3

All-Reflective Unobscured TMA Telescopes

3.1 IMAGING SYSTEM CHARACTERISTICS

In this section, we will introduce the concepts of stops, pupils and windows which are of fundamental importance in designing first-order (i.e., paraxial) optical instruments.

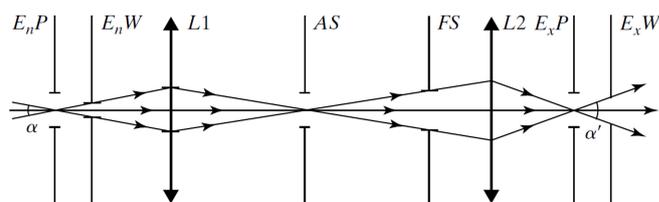


Figure 3.1: Two-surfaces imaging system (image source [51]).

Focal lengths The distance over which collimated (i.e., parallel) rays coming from infinity converge to a point on the optical axis is known as the back or image focal length (f_i). Similarly, the front or object focal length (f_o) is the distance from a focused point for an image at infinity.

Aperture and field stops by their inherently finite nature, capture only a fraction of the wavefront emitted from a point source. The surface clear diameter,

3.1. IMAGING SYSTEM CHARACTERISTICS

acting as an aperture stop (AS in Fig. 3.1), determines the light throughput, with elements like a diaphragm controlling it.

The field stop (FS in Fig. 3.1) limits the object size or solid angle the system can capture, defining the instrument's FoV. In a camera, the film or CCD sensor edge serves as the field stop. While the aperture stop governs the quantity of light reaching the image point, the field stop either permits or obstructs these rays.

Ultimately, adjusting the aperture stop influences the energy cone, affecting irradiance at each image point. Conversely, manipulating the field stop allows previously blocked regions beyond the object to be imaged.

Entrance- and exit-windows The entrance window (E_nW in Fig. 3.1) is the image of the field stop formed by all optical elements preceding it. The entrance window traces the lateral dimensions of the object to be viewed and its angular diameter determines the angular FoV. When the field stop is located in an image plane, the entrance window is positioned in the conjugate object plane, where it directly outlines the lateral dimensions of the object field that is being imaged by the optical system.

The exit window (E_xW in Fig. 3.1) is the image of the field stop formed by all optical elements following it.

Entrance- and exit-pupils The concept of the pupil is crucial in determining whether a ray will traverse the optical system. The entrance pupil (E_nP in Fig. 3.1) is the limiting aperture that the light rays "see" looking into the optical system from any object point, or more formally, it's the image of the controlling aperture stop formed by the imaging elements preceding it. Besides, the exit pupil (E_xP in Fig. 3.1) is the image of the controlling aperture stop formed by the imaging elements following it. Both are crucial for understanding the path of light in an optical system. Be aware that different objects along the axis may correspond to different aperture stops and pupils.

Chief and marginal rays In an optical system, the chief ray is a ray that starts from an off-axis point on an object and goes through the center of the entrance pupil. It is representative of a cone-shaped bundle of rays that originates from a specific point on the object. Correcting aberrations is particularly important in surface design, and the chief ray plays a significant role in this process. In fact,

any ray passing an aberration-free system through the center of the entrance pupil should also pass through the center of the stop and the exit pupil.

On the other hand, the marginal ray is a ray that starts from the central point of an object and travels towards the edge of the entrance pupil in an optical system. Its purpose is to provide information about the outermost parts of the optical system. The marginal ray is often traced alongside the chief ray, and it helps to analyze and understand the behavior of rays at the periphery of the optical system.

Relative aperture and f-number When we form an image of a point source using an optical system, the irradiance is directly proportional to the entrance pupil's area and inversely proportional to the image area. Since the image area will vary as the square of its lateral dimension, which is proportional to the square of the focal length f^2 , we can state the flux density at the image plane varies as $(D/f)^2$.

The ratio D/f is known as the relative aperture, and its inverse is the focal ratio, or f -number, often written $F/\#$, that is $F/\# = f/D$. Thus, for the same focal length, a smaller f -number allows more light to reach the image plane. Accordingly, the exposure time is proportional to the square of the f -number, for this reason, the latter is sometimes referred to as surface speed.

3.2 FIRST-ORDER MIRROR FORMULA

First-order, also named paraxial or Gaussian optics, is the study of image formation in the paraxial region, which is the region near the optical axis. The paraxial equation that relates conjugate object and image points to the physical parameters of a spherical mirror can be derived with the help of Fig. 3.2. By looking at the figure, we observe that the segment CA , which is the normal to the sphere in A , is the bisector of the triangle SAP , because the angle of reflection (θ_r) is equal to the angle of incidence (θ_i). Thus, it must hold the proportionality between segments:

$$\frac{SC}{SA} = \frac{CP}{PA} \quad (3.1)$$

Furthermore, we can write $SC = s_o + R$ and $CP = -R - s_i$. Also, in the paraxial region, it holds $SA \approx s_o$, $PA \approx s_i$. Where s_o , s_i and R are the object distance,

3.2. FIRST-ORDER MIRROR FORMULA

image distance and radius of curvature, respectively. Eventually, Eq. 3.1 becomes

$$\frac{1}{s_o} + \frac{1}{s_i} = -\frac{2}{R} \quad (3.2)$$

which is the first-order mirror formula and can be applied to concave ($R < 0$) or convex ($R > 0$) spherical mirrors.

We can evaluate the object focal length f_o through

$$\lim_{s_i \rightarrow \infty} s_o = f_o \quad (3.3)$$

Similarly, the back image length f_i is defined as

$$\lim_{s_o \rightarrow \infty} s_i = f_i \quad (3.4)$$

Consequently, from Eq. 3.2

$$\frac{1}{f_o} + \frac{1}{\infty} = \frac{1}{\infty} + \frac{1}{f_i} = -\frac{2}{R} \rightarrow f = f_o = f_i = -\frac{R}{2} \quad (3.5)$$

The final mirror formula can be written as

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f} \quad (3.6)$$

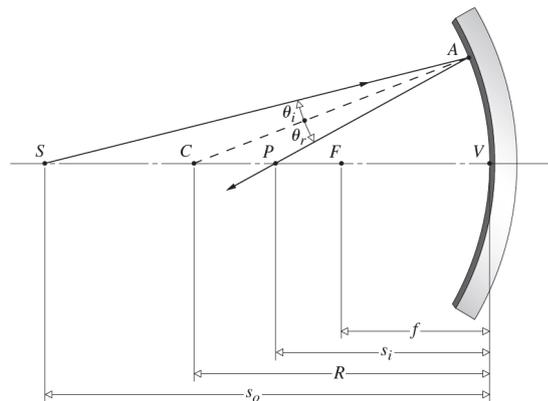


Figure 3.2: A concave spherical mirror (image source [52]).

3.3 CLASSIFICATION OF THREE-MIRROR CONFIGURATIONS

The basic three-mirror system is a coaxial system with three conic mirrors.

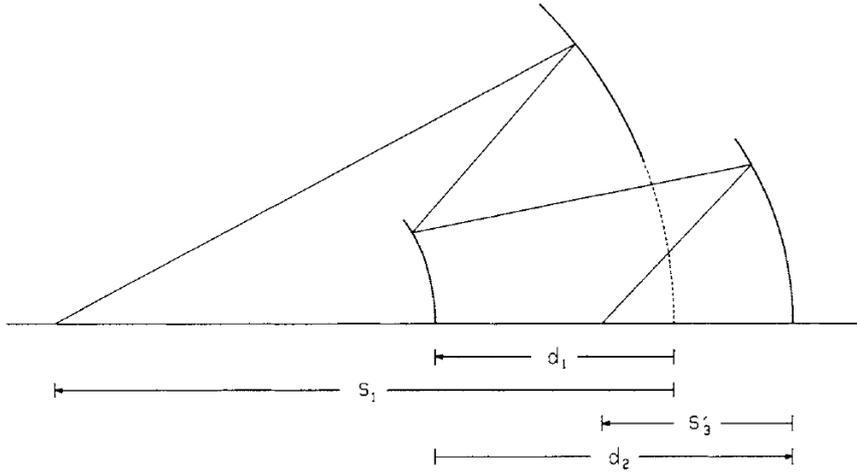


Figure 3.3: General geometry of a three-mirror system (image source [53]).

Referring to Fig. 3.3, it can be demonstrated that real objects exist only if

$$s_1 d_1 > 0 \quad (3.7)$$

and real images must respect the relation

$$s'_3 d_2 < 0 \quad (3.8)$$

moreover, it can also be shown that if s_1 and s'_3 have the same sign, thus

$$s'_3 / s_1 > 0 \quad (3.9)$$

then the initial object and the final image are either both real or virtual. Since it holds that

$$s'_3 / s_1 = -m_s \Omega_s \quad (3.10)$$

we can state that m_s and Ω_s must be opposite in sign. Here m_s is the lateral system magnification, - which is the paraxial image to object coordinate ratio - and Ω_s is the system ray-height ratio, - in which the ray heights are determined on the outermost surface planes by a paraxial ray through the origins of the object and image planes (see Tab. 3.2).

3.3. CLASSIFICATION OF THREE-MIRROR CONFIGURATIONS

According to D. Korsch [53], three-mirror imaging systems can be classified based on the signs of ray-height ratios:

- $\Omega_i > 0$: no image between surfaces i and $i + 1$;
- $\Omega_i < 0$: image between surfaces i and $i + 1$;
- $\Omega_s > 0$: final image upside-down;
- $\Omega_s < 0$: final image right-side-up.

On the basis of these properties, we can recognize four different classes:

	Ω_1	Ω_2	Ω_s	System description
I	+	+	+	no intermediate image and final image upside-down
II	-	+	-	intermediate image after primary and final image right-side-up
III	+	-	-	intermediate image after secondary and final image right-side-up
IV	-	-	+	intermediate images after primary and secondary, final image upside-down

Table 3.1: The four classes of three-mirror imaging systems, distinguished by Roman numerals.

Nonetheless, depending on the orientation of the curved surface (i.e. concave or convex), the geometric configurations may still vary considerably within each class.

Three-mirror telescopes are usually aplanatic, anastigmatic, flat-field systems, i.e., they are corrected for spherical aberration, coma, astigmatism, and field curvature. In order to achieve a flat image surface, these TMA telescopes need to satisfy the Petzval condition, which involves ensuring that the sum of the curvatures of the mirrors is zero. However, this condition cannot be met if all surfaces are either concave or convex, and it is also not possible to have two intermediate images as in the case of class-IV configurations. To help illustrate this, diagrams of unobscured zero-Petzval telescope configurations for every possible combination of concave and convex surfaces for the classes I, II, and III are shown in Fig. 3.4.

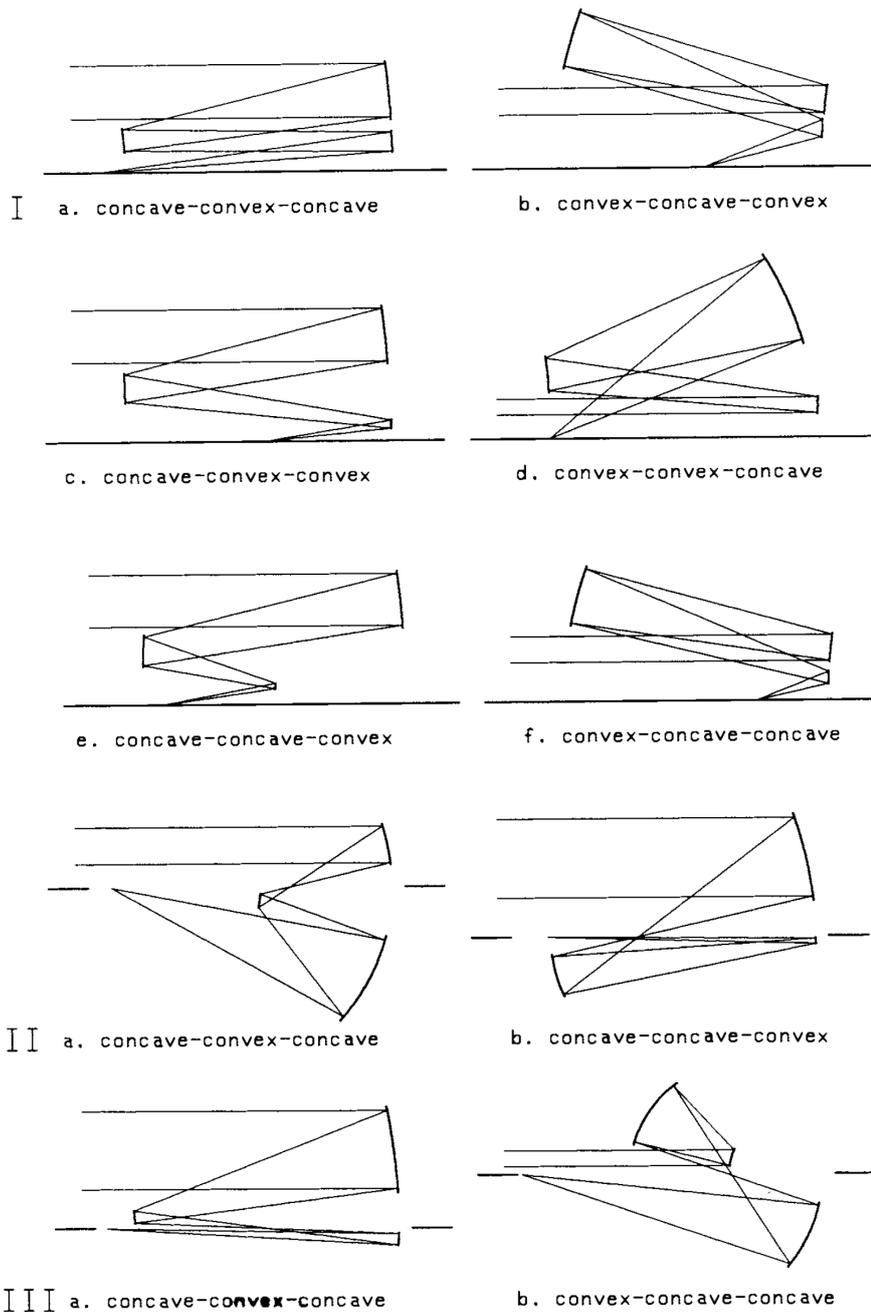


Figure 3.4: Zero-Petzval three-mirror telescope configurations. (image source [53]).

3.3.1 DESIGN PARAMETERS OF A k -MIRROR TELESCOPE

In designing a k -mirror system, that is an optical system realized with k mirrors, it is convenient to define a set of input parameters. The number of parameters necessary and sufficient to completely define the first-order proper-

3.3. CLASSIFICATION OF THREE-MIRROR CONFIGURATIONS

ties of a k -mirror system is $2k + 1$. Usually, it consists of object s_1 and image s'_k distances, mirror separations d_i (with $i = 1$ to $k - 1$), paraxial ray-height ratios Ω_i (with $i = 1$ to $k - 1$) and entrance-pupil distance t_1 ; but it may be different by convenience.

The most useful parameters of a general k -mirror system are listed in Tab. 3.2. Note that in the case of a telescope $v_1 = 1/s_1 = 0$.

Table 3.2: Parameters of a general k -mirror system.

Parameter	Formula
Inverse object distance	$v_i = 1/s_i$
Inverse image distance	$v'_i = 1/s'_i$
Curvature	$c_i = \frac{1}{2}(v_i + v'_i)$
Radius of curvature	$r_i = 1/c_i$
Mirror separation	$d_i = s'_i - s_{i+1} = \frac{v_{i+1} - v'_i}{v'_i v_{i+1}} \quad (i = 1 \text{ to } k - 1)$
Paraxial ray-height ratio	$\Omega_i = \frac{s_{i+1}}{s'_i} = \frac{v'_i}{v_{i+1}} \quad (i = 1 \text{ to } k - 1)$
Sub-system ray-height ratio	$\Omega_{si-1} = \prod_{j=1}^{i-1} \Omega_j \quad \text{with } \Omega_{sk-1} = \Omega_s$
Lateral magnification	$m_i = -\frac{s'_i}{s_i} = -\frac{v_i}{v'_i}$
Lateral sub-system magnification	$m_{si} = \prod_{j=1}^i m_j \quad \text{with } m_{sk} = m_s$
Axial magnification	$m'_i = -m_i^2$
Numerical eccentricity	$\epsilon_i = \frac{v'_i - v_i}{v'_i + v_i}$
Exit-pupil distance	$t'_i = \frac{t_i}{2c_i t_{i-1}}$
Entrance-pupil distance	$t_{i+1} = t'_i - d_i$
Image-scale coefficient	$\tau_i = 1 - v_i t_i$

For completeness, we report the image focal length of a three-mirror system:

$$f_s = -\{2[c_1 - c_2 + c_3 + 2c_1(c_2 - c_3)d_1 + 2c_3(c_1 - c_2)d_2 + 4c_1c_2c_3d_1d_2]\}^{-1} \quad (3.11)$$

3.4 THIRD-ORDER CORRECTION OF ROTATIONALLY SYMMETRIC THREE-MIRROR TELESCOPES

The study of third-order optics builds upon the foundation laid by first-order optics, expanding our understanding of optical systems by exploring the intricacies of aberrations. While first-order optics focuses on the basic geometry of image formation (i.e., axial positions and the radii of the surfaces, the pupil locations and the ideal image location for any given object point), it fails to predict image quality. To overcome this limitation, third-order optics becomes indispensable, especially when dealing with complex optical systems made entirely of reflective surfaces. Finding the mathematical solution for mirror systems that have multiple surfaces is often very challenging, and sometimes it is impossible to do it perfectly. Unlike single conic reflectors, which have straightforward derivations, the math behind multiple surface mirror systems can get complex. Therefore, it is necessary to develop acceptable approximations in situations where it is impossible to find closed-form solutions.

In 1856, Ludwig Seidel laid the foundation for third-order aberration theory [54]. This theory is very useful in understanding and reducing aberrations in optical systems that have more than one surface. The method involves expanding the equation for reflected rays into a power series that depends on the lateral coordinates of the object and the surface (see Section 3.5). The significance of this approach lies not only in its ability to analyze the image-formation characteristics of a given system but also in its capacity to design systems that are free of third-order aberrations.

As we will see later, the Seidel or primary aberrations express the lateral aberration components in the Gaussian (i.e., paraxial) image plane and are represented by the aberration coefficients cataloged in Tab. 3.3. Section 3.5 provides a thorough explanation of the wave nature of these errors.

Table 3.3: Seidel aberration coefficients.

A-term	spherical aberration
B-term	coma
C-term	astigmatism
C- and D-terms	field curvature
E-term	distortion

3.4. THIRD-ORDER CORRECTION OF ROTATIONALLY SYMMETRIC THREE-MIRROR TELESCOPES

The third-order equations for the aberration coefficients of an imaging system with k mirrors are described in Eq. 3.12 [55]. By resolving for $k = 3$ we can get the aberrations for a three-mirror system.

$$\left\{ \begin{array}{ll} A = -\tau_1^{-4} \sum_{i=1}^k (-1)^i \Omega_{s_{i-1}}^4 A_i & \text{with } A_i = c_i^3 \Delta \delta_i \\ B = \tau_1^{-2} \sum_{i=1}^k \Omega_{s_{i-1}}^2 B_i & \text{with } B_i = t_i \tau_i^{-1} A_i - \epsilon_i c_i^2 \\ C = -\sum_{i=1}^k (-1)^i C_i & \text{with } C_i = -t_i^2 \tau_i^{-2} A_i + 2B_i t_i \tau_i + c_i \\ D = -\sum_{i=1}^k (-1)^i D_i & \text{with } D_i = 0.5 C_i - c_i \\ E = \tau_1^2 \sum_{i=1}^k \Omega_{s_{i-1}}^{-2} E_i & \text{with } E_i = t_i \tau_i^{-1} (c_i - t_i \tau_i^{-1} B_i) \end{array} \right. \quad (3.12)$$

here $\Delta \delta_i$ quantifies the aspheric deviation from the Cartesian (i.e., spherical) surface:

$$\Delta \delta_i = \delta_i - \delta_0 = \delta_i + \epsilon_i^2 \quad (3.13)$$

and ϵ_i is the numerical eccentricity of the surface i (refer to Tab.3.2). It is important to note that in Eq. 2.1, we have already defined $\Delta \delta$ by naming it as conic constant k .

The field curvature is given by:

$$c_0 = (-1)^k (C + 2D) = 2(-1)^k C + c_p|_{k=3} = -2C + c_p \quad (3.14)$$

with

$$c_p = 2(-1)^k \sum_{i=1}^k (-1)^i c_i|_{k=3} = 2(c_1 - c_2 + c_3) \quad (3.15)$$

being the Petzval curvature for a three-mirror system.

Allowing any aspheric deviation ($\Delta \delta \neq 0$) permits us to correct three Seidel aberrations without restricting the system geometry. In particular, in the absence of spherical aberration ($A = 0$), coma becomes independent of the primary and secondary aspheric deviation; in the case of aplanatism ($A = 0$ and $B = 0$), astigmatism becomes independent of all three aspheric deviations, in addition, both astigmatism and field curvature are independent of the entrance-pupil location. Finally, in the case of aplanatism and anastigmatism ($A = 0$, $B = 0$ and $C = 0$) distortion is independent of all three aspheric deviations.

Now, we want to properly shape the three surfaces to eliminate spherical aberration, coma, and astigmatism from the system. Thus, we set A , B and C in

Eq. 3.12 together equal to zero and obtain [53]

$$\Delta\delta_3 = \frac{2\Omega_1(v_1 + v'_3) - v_1^2 d_1 - \Omega_1 \Omega_2 v_3'^2 (\Omega_1 \Omega_2 d_1 - 2d_2)}{4\Omega_1 \Omega_2^2 d_2 (\Omega_1 \Omega_2 d_1 - d_2) c_3^3} \quad (3.16)$$

$$\Delta\delta_2 = \frac{v_1^2 - \Omega_1^2 \Omega_2^2 v_3'^2 + 4\Omega_1^2 \Omega_2^3 (\Omega_1 \Omega_2 d_1 - d_2) c_3^3 \Delta\delta_3}{4\Omega_1^3 d_1 c_2^3} \quad (3.17)$$

$$\Delta\delta_1 = \frac{\Omega_1^4}{c_1^3} (c_2^3 \Delta\delta_2 - \Omega_2^4 c_3^3 \Delta\delta_3) \quad (3.18)$$

Since the system is anastigmatic ($C = 0$) then the field curvature, c_0 , coincides with the Petzval curvature, c_p , which is given by Eq. 3.15. To correct for field curvature, one can set c_p to zero, but this comes at the cost of geometrical constraints. Thus, the condition of flat field, when solved, reduces the number of free input parameters by one. Consequently, the set of initial input parameters increases to eight.

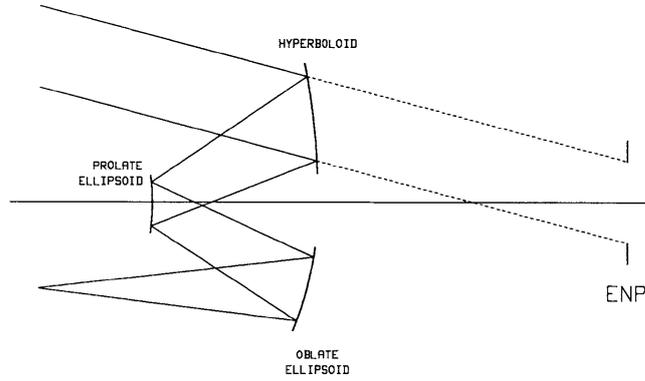


Figure 3.5: Three-mirror anastigmat telescope patented by W. Wetherell and D. Womble (image source [53]).

A renowned example of three-mirror telescope corrected for spherical aberration, coma, astigmatism and field curvature is shown in Fig. 3.5. This was first presented by W. Wetherell and D. Womble in 1980 [9]. The system is of the type Ia (according to the classification in Fig. 3.4), it presents centered pupils and a biased field (i.e., not parallel to the optical axis of the system) to gain access to the final image. The entrance pupil is virtual and located such that the secondary mirror becomes the system stop. This results advantageous in two ways: first, the secondary mirror size is restrained and second, the secondary mirror

3.5. ABERRATION THEORY

is centered on the optical axis. Furthermore, because of the infinite initial object distance, $v_1 = 0$, and the condition to have the system stop on the secondary mirror yields

$$t_1 = -d_1/\Omega_1 \quad (3.19)$$

The eight input parameters of this system are:

$$\begin{cases} d_2 = -d_1, & s'_3 = 1.6820d_1, & f_s = -2.0112d_1, \\ \Omega_1 = 0.5380, & \Omega_2 = 1.5545, \\ \delta_1 = -2.710014, & \delta_2 = -0.415013, & \delta_3 = 0.130922 \end{cases} \quad (3.20)$$

from which we can completely determine the remaining parameters (see Tab. 3.2).

To go into the details of this three-mirror example and others, we redirect the reader to the work of D. Korsch [53].

3.5 ABERRATION THEORY

Stigmatic image In ray optics, we speak of stigmaticity when all rays originating from the point object and reaching the optical system focus at a unique image point after passing through the system. Using wave optics, the image is stigmatic when the incident diverging spherical wavefront centered at the point object is transformed by the optical system into another spherical wavefront converging at the image point. However, geometrical optics only produces a point image of a point object in the paraxial approximation, thus, when rays make small angles with the optical axis. Mathematically, if the first term in the power expansions for the sine and cosine functions

$$\begin{aligned} \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \\ \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \end{aligned} \quad (3.21)$$

is retained, we have the so-called first-order theory. The inclusion of higher-order terms in the derivations, however, predicts increasingly larger aberrations with increasing angles. Similarly, if we approximate the expansions up to the second term, we have the so-called third-order theory.

Tangential and sagittal planes It is important to note that when a surface is used to produce an image of a point object, the cone of rays that falls on the surface has complete rotational symmetry around the optical axis. All of these rays lie in meridional planes, which can be defined as any plane that includes the optical axis. However, this symmetry is maintained only if the object is on-axis.

For off-axis objects, the rotational symmetry is lost, and the cone of rays that falls on the surface lacks this symmetry. However, there is still a plane of symmetry that includes the point object and its paraxial image and also contains the optical axis and the chief ray. This plane of symmetry is called the tangential plane. Assuming that the point object is on the y -axis, and the optical axis overlaps the z -axis, the tangential plane is the vertical plane yz . The rays that originate from an off-axis point object and lie in the tangential plane are called tangential rays (see Fig. 3.6).

Furthermore, the plane that is perpendicular to the tangential plane and contains the chief ray is called the sagittal plane. The only meridional ray in the sagittal plane is the chief ray, and all other rays in this plane are non-meridional or skew rays (see Fig. 3.6). It is crucial to note that the rays in the tangential and sagittal planes constitute only a small subset of the incident cone that originates from an off-axis object point.

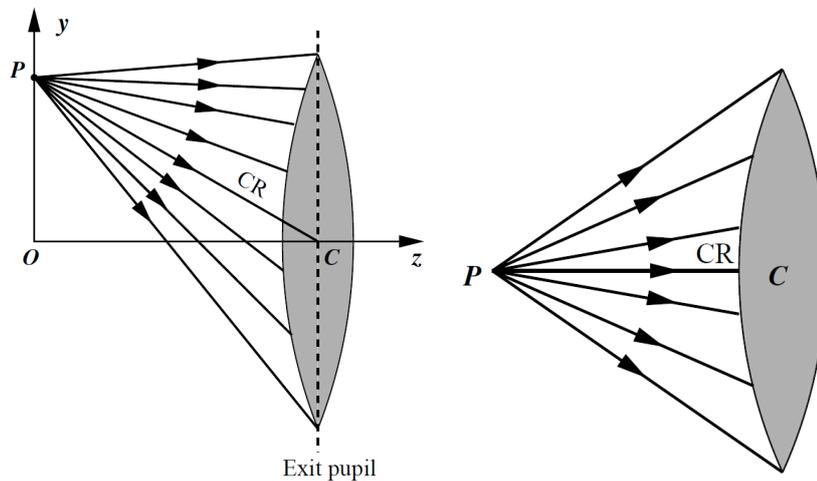


Figure 3.6: Tangential (on the left) and sagittal (on the right) rays, CR represents the chief ray (image source [56]).

Wavefront aberration function In an aberrated system (see Fig. 3.7), rays from adjacent points A and B , being normal to their respective wavefronts, do not

3.5. ABERRATION THEORY

intersect the paraxial image plane at the same point. The “miss” along the optical axis, represented by the distance \overline{LI} , is called the longitudinal aberration, and the miss \overline{IS} , measured in the image plane, is called the transverse, or lateral, aberration. These are ray aberrations.

On the other hand, when evaluating the quality of an optical system, we can also look at the wavefront in its exit pupil. As already stated, if the wavefront is spherical, the image is clear and located at its center of curvature. However, in most cases, the wavefront is not spherical, which leads to image aberrations. To determine the resulting aberrations in the image we can compare the actual wavefront to the ideal wavefront, also called reference sphere.

The wavefront aberration function W is defined as the difference in the optical path length, corresponding to the directed distance $d(= \overline{AB})$, between the actual and ideal wavefront (see Fig. 3.7), multiplied by the refractive index n of the medium in which light propagates

$$W(d) = nd. \quad (3.22)$$

For a precision optical system such as a telescope, the directed distance d must not exceed a fraction of the wavelength of light.

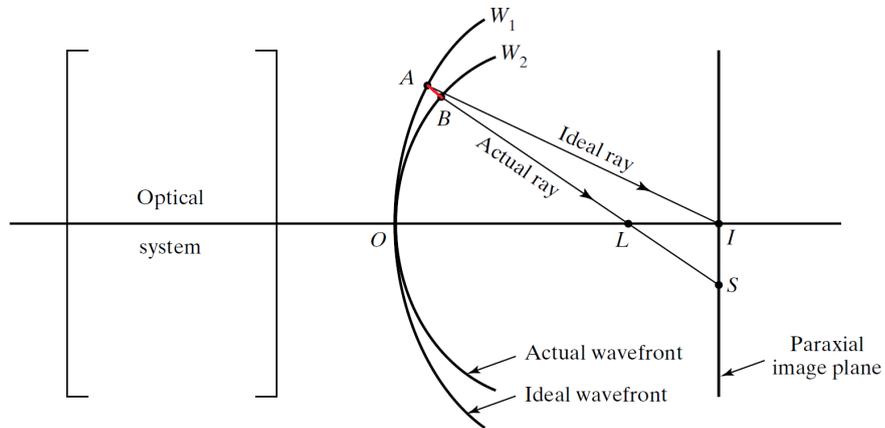


Figure 3.7: Illustration of ray and wave aberrations (image source [51]).

Accordingly to R. V. Shack [57], we can define the wave aberration function for a rotationally symmetric optical system in vector form as

$$W(\vec{H}, \vec{\rho}) = \sum_{k,m,n} W_{2k+n,2m+n,n} (\vec{H} \cdot \vec{H})^k (\vec{\rho} \cdot \vec{\rho})^m (\vec{H} \cdot \vec{\rho})^n \quad (3.23)$$

where H is the field vector located in the image plane, ρ is the aperture vector that defines the intersection of a given ray with the exit pupil plane and the aberration coefficients W embody the magnitude of a given aberration. The subscripts $2k + n, 2m + n, n$ in each coefficient indicate respectively the algebraic power of the field vector, the aperture vector, and the cosine of the azimuthal angle ϕ between these vectors. The sum of all aberration terms and orders produces the actual total wavefront deformation.

Modern optical systems encompass innovative designs that break the rotational symmetry. Several studies have come out in recent years describing classes of non-symmetrical systems. For instance, in a tilted and decentered system, the aberration field for each surface can be centered at a different point in the image. This point can be determined by the vector σ_j , which is obtained by projecting a line connecting the center of the pupil for the surface of interest with the center of curvature of that surface onto the image plane. To determine the contribution of a surface to the overall aberration field, it is necessary to define the effective aberration field height of the surface j as H_{Aj}

$$\vec{H}_{Aj} = \vec{H} - \sigma_j \quad (3.24)$$

The advantage of the vectorial wave aberration expansion with the tilt term is that the nodal points at which the primary aberrations have zero values can be derived. Therefore, it is also called Nodal aberration theory (NAT). This approach, discovered by Shack [58] and developed by Thompson [57] describes the behavior of an optical system with tilted or decentered rotationally symmetric components. Recently, Fuerschbach et al. [59] implemented a generalized approach to NAT where the non-rotationally symmetric components could lie anywhere in the optical system.

Furthermore, Sisian [60] proposed a modified wavefront aberration function to account for plane-symmetric systems

$$W(\vec{H}, \vec{\rho}, \vec{i}) = \sum_{k,m,n,p,q}^{\infty} W_{2k+n+p, 2m+n+q, n, p, q} (\vec{H} \cdot \vec{H})^k (\vec{\rho} \cdot \vec{\rho})^m (\vec{H} \cdot \vec{\rho})^n (\vec{i} \cdot \vec{H})^p (\vec{i} \cdot \vec{\rho})^q \quad (3.25)$$

where \vec{i} is a unit vector that specifies the direction of plane-symmetry. The indices k, m, n, p and q are integer numbers. Note that the first subscript $2k + n + p$

3.5. ABERRATION THEORY

is the algebraic power of the field vector \vec{H} , and the second subscript $2m + n + q$ is the algebraic power of the aperture vector $\vec{\rho}$. To combine several systems, one can use a field displacement term \vec{H}_{Aj} for each of the tilted component systems as Thompson did (see Eq. 3.24). Otherwise, it is possible to integrate several plane-symmetric systems using the vector \vec{i}_j that indicates the relative orientation among each of the j plane-symmetric systems.

Seidel aberrations As mentioned in Section 3.4, any rotationally symmetric system can be described by five coefficients that represent its primary aberrations. These coefficients, known as Seidel aberrations, coincide with the fourth-order terms of the wavefront aberration function (Eq. 3.23). They include spherical aberration, coma, astigmatism, field curvature, and distortion. Since these aberrations are present even in monochromatic light, they are also known as monochromatic aberrations. Hereafter, we will discuss the impact of each aberration on the image when it is present individually.

3.5.1 SPHERICAL ABERRATION

When light parallel to the axis passes through a thin lens, the rays bend towards the axis and intersect at a point, which determines the paraxial focal length f_P of the surface. Beyond the paraxial region, rays that are incident at different heights on the surface hit the axis at different points. The paraxial focus F_P is where the paraxial rays strike the axis, while the marginal focus f_M is where the rays near the periphery intersect (see Fig. 3.8). Spherical aberration on the surface is the measure of the distance between these two foci. Consequently, the image of a point object is not a point, but a circular patch of light. The circle of least confusion is the plane \overline{AB} normal to the z -axis where the circular patch has the least diameter.

3.5.2 COMA

If a point object lies on the axis, the image will only suffer from spherical aberration. However, for off-axis points, the image will also suffer from coma, astigmatism, curvature of field, and distortion.

Due to the effect of coma, the rays that are close to the axis of the surface focus at a different point from that of the marginal rays. As a result, the magnification appears to be different for different areas of the surface.

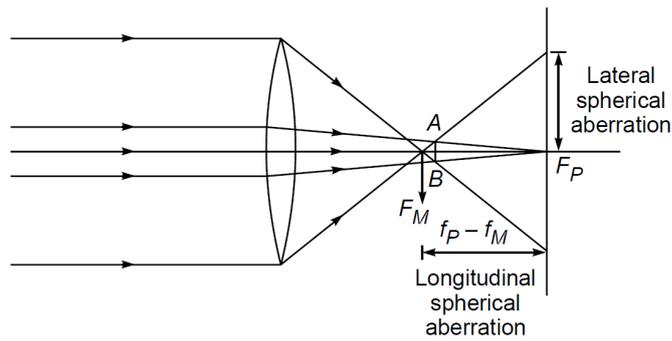


Figure 3.8: The image formation in the presence of spherical aberration (image source [61]).

In Fig. 3.9, we have a three-dimensional view of a set of rays that hit the surface at the same distance from the center. Rays that intersect the surface at diametrically opposite points focus on different points in the image plane, forming a circle. The center of the circle and the distance at which it lies from the ideal image point measure the coma. As the radius (h) of the zone increases, the center of the circle shifts further away from the ideal image. Thus, the composite image will have a comet-like appearance and hence the name coma.

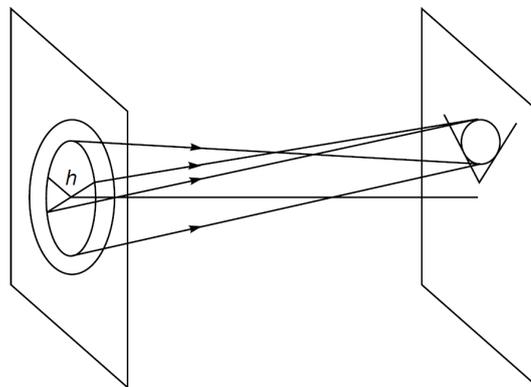


Figure 3.9: The image formation in the presence of coma (image source [61]).

3.5.3 ASTIGMATISM AND FIELD CURVATURE

In Fig. 3.10 left, we can see how astigmatism affects the image formation process in an optical system. When an off-axis point P is imaged by an astigmatic optical system, the rays in the tangential plane converge at a different point than those in the sagittal plane. As an example, rays \overline{PA} and \overline{PB} focus at point T , while rays \overline{PC} and \overline{PD} focus at point S . The distance between these two points

3.5. ABERRATION THEORY

indicates the degree of astigmatism.

In wave optics, aberrations occur when the wavefront emerging from the optical system is not spherical. In the case of astigmatism, the wavefront converges into two distinct lines that are normal to each other and are called the tangential and sagittal focal lines. The image formed somewhere between these two lines is circular and is referred to as the circle of least confusion.

The distance between the tangential and sagittal foci (T and S , respectively) increases as the object point moves away from the axis. As shown in Fig. 3.10 right, the tangential foci and the sagittal foci of points at different distances from the axis lie on two curves. When the T surface falls to the left of the S surface, as shown, the astigmatic difference is taken as positive; otherwise, it is negative. The optical system is considered free from astigmatism when these two surfaces overlap and the resulting surface is called the Petzval surface. Even when these surfaces coincide and point images are formed, it can be shown that the resulting image surface can still be curved (see Eq. 3.14). This curvature is known as the field curvature, and it is a defect that affects the image.

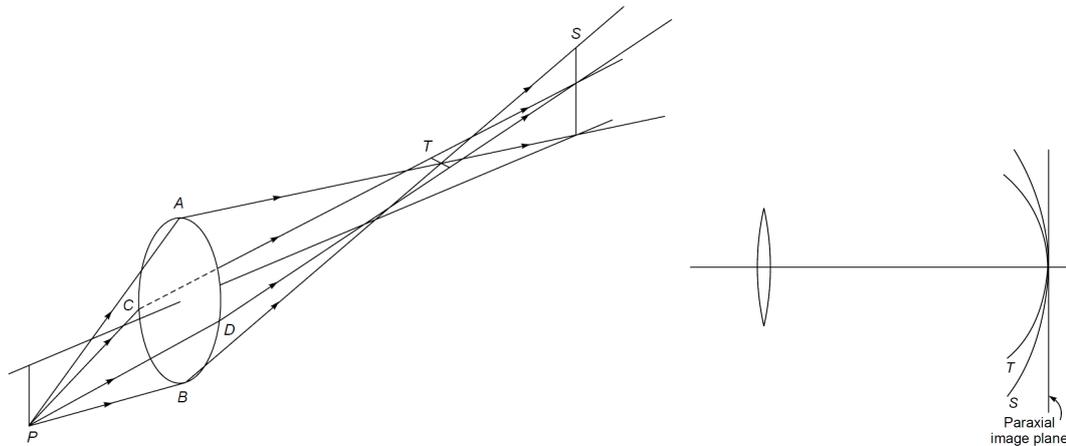


Figure 3.10: On the left, the image formation in the presence of astigmatism. On the right, the tangential and sagittal foci (image source [61]).

3.5.4 DISTORTION

The last of the five monochromatic Seidel aberrations, present even if all the others have been eliminated, is distortion. Even though object points are sharply focused, distortion shows up as a variation in the lateral magnification for object points at different distances from the optical axis. In other words,

distortion arises because different areas of the lens have different focal lengths and different magnifications.

If the magnification increases with distance from the axis, the image appears with pincushion distortion (see Fig. 3.11 right). On the other hand, if magnification decreases with distance from the axis, the image appears with barrel distortion (see Fig. 3.11 left).

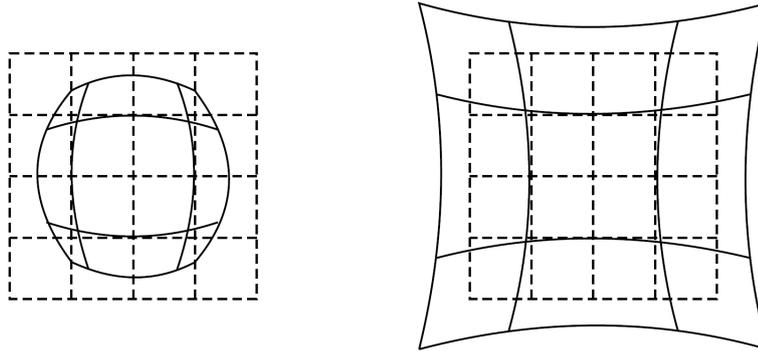


Figure 3.11: On the left, image of a squared grid affected by barrel distortion. On the right, image of a square grid affected by pincushion distortion (image source [61]).

3.6 DESIGN METHODS FOR FREEFORM IMAGING SYSTEMS

A good design method to implement freeform optics in imaging systems should result in initial structures that can effectively utilize the freeform surface's ability to correct aberrations and enhance system design efficiency. However, designing effective freeform imaging systems is challenging due to their non-symmetric configurations, advanced system specifications, and complex aberration theory. Therefore, the generation of favorable starting points is still one of the hot issues in the field of advanced optical design.

One straightforward design method of freeform imaging systems is to find existing patents or other available systems as the starting point. Then, further optimization with optical design software is applied to obtain the final design result. However, as the freeform surfaces are generally used in the off-axis or unobscured forms, there are few existing patents or systems for choice. Therefore, designers may fail to find useful solutions, or they have to spend a long time improving the starting point. Another common design approach is to first

3.6. DESIGN METHODS FOR FREEFORM IMAGING SYSTEMS

create a co-axial spherical or conical starting point based on the primary (or Seidel) aberrations theory. Then an off-axis or unobscured form is created by using an offset aperture and/or a biased input field, or tilting the surfaces until the light clears the mirrors. As the co-axial starting point is generally far from the optimum point, designers may also have complications finding useful solutions.

In this section, we will introduce the most common design methods of freeform imaging optical systems.

3.6.1 ABERRATION-BASED METHOD

A classical idea for designing an obscured reflective system is to first solve the structural parameters for a rotationally symmetric system that corrects third-order aberrations (as we previously did in Section 3.4), and then eliminate the light obstruction in the system by biasing the input FoV and/or decentering the aperture, or tilting the surfaces until the light clears the mirrors [62]. However, this approach does not consider the induced aberrations from newfound asymmetry. In fact, tilting the on-axis solution will break the rotational symmetry of the system and will change where the aberration field zeros (nodes) are located for each aberration type [63]. The shift of the aberration fields will drastically degrade the overall performance of the system. A better approach is to focus on the specific unobscured geometry and its rotationally variant aberrations before correcting the third-order rotationally invariant aberrations [64]. Consequently, by controlling the aberrations with freeform shapes, we are able to restrain the overall freeform sag to the necessary aberration correction, thus decreasing system sensitivity, fabrication cost and testing difficulties.

A comprehensive selection of the possible type Ia geometries for the unobscured three-mirror imager is shown in Fig 3.12.

3.6.2 DIRECT DESIGN METHOD

The direct design method for freeform surfaces involves employing ideal imaging principles, such as geometric optics propagation laws (e.g., the laws of refraction and reflection). This method aims to establish a set of equations connecting the light rays and the surface, based on the object-image relationship within the design index of the imaging system. The subsequent step is to solve these equations, resulting in surface discrete point data information that is then fitted to generate the freeform surfaces.

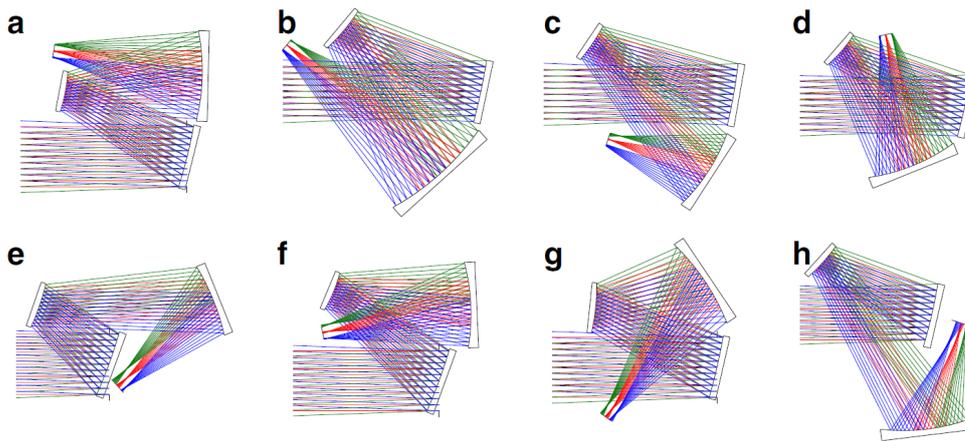


Figure 3.12: A comprehensive selection of type Ia geometries for unobscured three-mirror imager (image source [64]).

Common methods include the partial differential equation (PDE) method, simultaneous multiple surface method (SMS) design method and point-by-point construction-iteration (CI) method. These methods provide diverse avenues for translating ideal imaging principles into practical, mathematically defined freeform surfaces.

Partial differential equation (PDE) method The Wasserman-Wolf differential equation was derived in 1949 to design two aspheric surfaces for a centered system [65]. Based on the Abbe sine condition that solves a first-order differential equation, the equation could obtain axial stigmatism. In 2002, Knapp developed a new approach to design correctors for non-rotationally symmetric optical systems [66]. This approach was similar to that of Wasserman-Wolf's and was implemented in a commercial software macro called the Generalized Aspheric Design Program (GAP). While Wasserman and Wolf's method had limited applications since it calculates discrete light, Knapp's approach provided new design options. However, it had its own set of disadvantages, such as a long time to arrive at a solution and an inability to address chromatic aberrations. In 2007, Andrew Hicks created a freeform surface by setting up a system of partial differential equations for the correction of aberrations in large field-of-view imaging systems [67]. In 2010, D.Cheng et al. completed the design of an off-axis freeform prismatic head-mounted display system by using partial differential equations [68].

The process of designing freeform surfaces using partial differential equa-

3.6. DESIGN METHODS FOR FREEFORM IMAGING SYSTEMS

tions is based on the object-image relationship within the imaging system, on the premise of knowing the direction vectors of the incoming and outgoing light rays. By using vector orientation, a system of differential equations is constructed based on the law of refraction to determine the coordinates of the freeform surface vectors and their corresponding normal vector coordinates. Finally, the discrete points are fitted by the freeform surface representation function to obtain the initial shape of the freeform surface.

Simultaneous multiple surface (SMS) method In 1990, the SMS method was first proposed for the design of 2D non-imaging optical systems by J. C. Minano and P. Benitez [69]. In 2009, it was improved and expanded to be used in the design of 2D imaging optical systems [70]. In 2012, the SMS method was first generalized to 3D geometry [71]. The method utilizes the edge principle and the Fermat theorem of light to deduce functional differential equations that describe the entire optical system. It can calculate the point and extended light sources, taking into account the size and angle of the source. The SMS method uses the principle of *bundle-coupling* and *prescribed-irradiance*, where input and output bundles are coupled, and one bundle must be included in the other. The power of the SMS method lies in the fact that the incoming and outgoing wavefronts can themselves be freeform, giving the method great flexibility.

Point-by-point construction-iteration (CI) method In 2013, J. Zhu's group at Tsinghua University introduced a 2D freeform surface design method based on a point-by-point construction and iterative approach [72].

According to the object-image relationship and design requirements, each initial surface type, be it a plane or a conic, is selected to determine the preliminary spatial layout and relative position of the components. Feature light is then specified, and the spatial layout and positional relationship of each component are defined in polar coordinates. Next, light rays from multiple fields of view and different pupil coordinates are sampled and traced in polar coordinates. Following this, the refraction and reflection laws are employed to solve the surface problem based on the relationship between incident and outgoing rays. A set of feature data points on the surface and their corresponding normal vectors are obtained, and a fitting process is applied to join these points together, completing the point-by-point construction of the freeform surface. Subsequently, through iterative optimization, the freeform surface is refined to establish the

initial structure of the imaging optical system.

In 2014, Zhu Jun's group extended the point-by-point construction and iteration design concept to 3D free-form imaging systems [73]. They improved the computational method for solving discrete points and proposed a high-precision freeform surface fitting method, considering both coordinates and normal vectors of the discrete points on the surface [74].

Building upon this method, the group successfully designed several off-axis reflective free-form surface imaging optical systems [73] [75]. In 2017, they integrated the point-by-point construction-iteration design idea with neural network machine learning, enabling the rapid construction of optical systems with minimal input [76].

4

Freeform Off-axis Three-mirror Telescope Design: A Deep Learning-Enabled Approach

With the rapid development of science and technology, the demands placed on imaging optics are progressively increasing. In order to achieve superior imaging performance, imaging optical systems are being designed with larger FoV, larger apertures, smaller f -numbers ($F/\#$), reduced volume, and fewer optical elements. Furthermore, non-rotationally symmetric system configurations are generally the preferred choice to meet specific system structure requirements while avoiding light obscuration.

Imaging systems have used spherical and aspherical elements for a long time because of their rotational shape and ease of fabrication. However, they hardly achieve high-level design requirements as they present low degrees of design freedom, nevertheless, their ability to correct aberrations is limited especially in non-symmetric systems (see Section 3.5). To overcome these limitations, freeform optical surfaces can be used. In fact, non-rotational symmetric surfaces can improve imaging performance and specifications while reducing the volume and number of elements.

Advanced freeform surfaces improve imaging optics but also increase significantly design difficulty, production time, and cost because of their complex shape and non-symmetric system structure.

Traditional optical design methods often rely on finding a starting point

4.1. FRAMEWORK

through literature or databases, which may not be feasible for freeform optical systems. A design process without good starting points is time-consuming and requires a significant amount of human effort. Thus, Nodal aberration theory and direct/point-by-point design methods have been proposed (see Section 3.6); however, they present limitations in efficiency, simplicity, and generality. When new system parameters and configurations are input, the designer must repeat the design process - which usually requires complex mathematical derivations, - and even the optimization strategy may need to be adjusted accordingly, wasting time and effort. Deep Learning (DL) can be considered a solution to these issues, as it can effectively summarize design knowledge and apply it to design tasks with a wide range of system and structure parameters. DL has been proven successful in many areas of scientific research and engineering, including the field of optics [77][78][79][80]. In 2019, Côté et al. used DL to obtain lens design databases to produce high-quality starting points for coaxial spherical objectives [81]. Then, this was improved by the introduction of more design forms [82][83]. However, the above framework is limited to coaxial spherical systems. In 2019, Yang et al. [84] proposed the preliminary neural network-based design framework of freeform reflective imaging systems, which Chen et al. [85] improved by increasing the range of system specifications. In 2023, Mao et al. [86] presented a deep learning framework for the ultrafast generation of multiple-solution for generalized off-axis reflective, refractive, and catadioptric systems with multiple freeform surfaces.

4.1 FRAMEWORK

In this thesis, we propose a framework that utilizes neural network-based Supervised Learning (SL) [87] for generating starting points for freeform off-axis three-mirror telescopes. The framework builds on the work of Yang et al. [84] and involves obtaining a set of base systems to train a FFNN [87]. These base systems are generated through system evolution, starting from one initial base system. A training dataset consisting of system parameters as the input and surface parameters as the target is then obtained.

Consequently, a FFNN, also known as Multilayer Perceptron (MLP), is trained using the provided dataset. The fully-connected network enables the generation of good starting points with various system specifications for further optimization.

Designers no longer have to manage the starting point exploration or analytical/numerical design process, thus reducing the amount of time and human effort required and the dependence on advanced design skills. Beginners in optical design can also generate a good starting point using the obtained network.

We validated the feasibility of this process by designing the Wetherell-configuration freeform off-axis reflective triplet [9]. The diagram of the design framework is shown in Fig. 4.1. For a given three-mirror system geometry, the entire design framework can be divided into the following main steps:

1. Generate a number of base systems that have the same configuration within a given range of system specifications. The system parameters, such as full-FoV, $F/\#$ and Entrance-Pupil Diameter (ENPD), and the corresponding surface parameters (which can fully describe the location and shape information of surfaces) are taken as the input and target parts, respectively, in the dataset.
2. The FFNN network can be trained using the obtained dataset. Accordingly, the network goes through a testing and validating phase.
3. For a given design requirement, the system parameters are selected and provided to the FFNN, which can output the corresponding surface parameters directly. The resulting telescope can be taken as a good starting point for further optimization.

Development environment: the algorithm and data processing for automatic dataset generation and network training are entirely implemented in MATLAB environment. The initial system configuration and optimization are achieved using Zemax OPTICSTUDIO, an optical design software powered by Ansys. In particular, we used an Application Programming Interface (API), ZOS-API, to communicate directly with OPTICSTUDIO through MATLAB. In fact, ZOS-API relies on a Component Object Model (COM) interface and it is rooted in .NET libraries.

4.1. FRAMEWORK

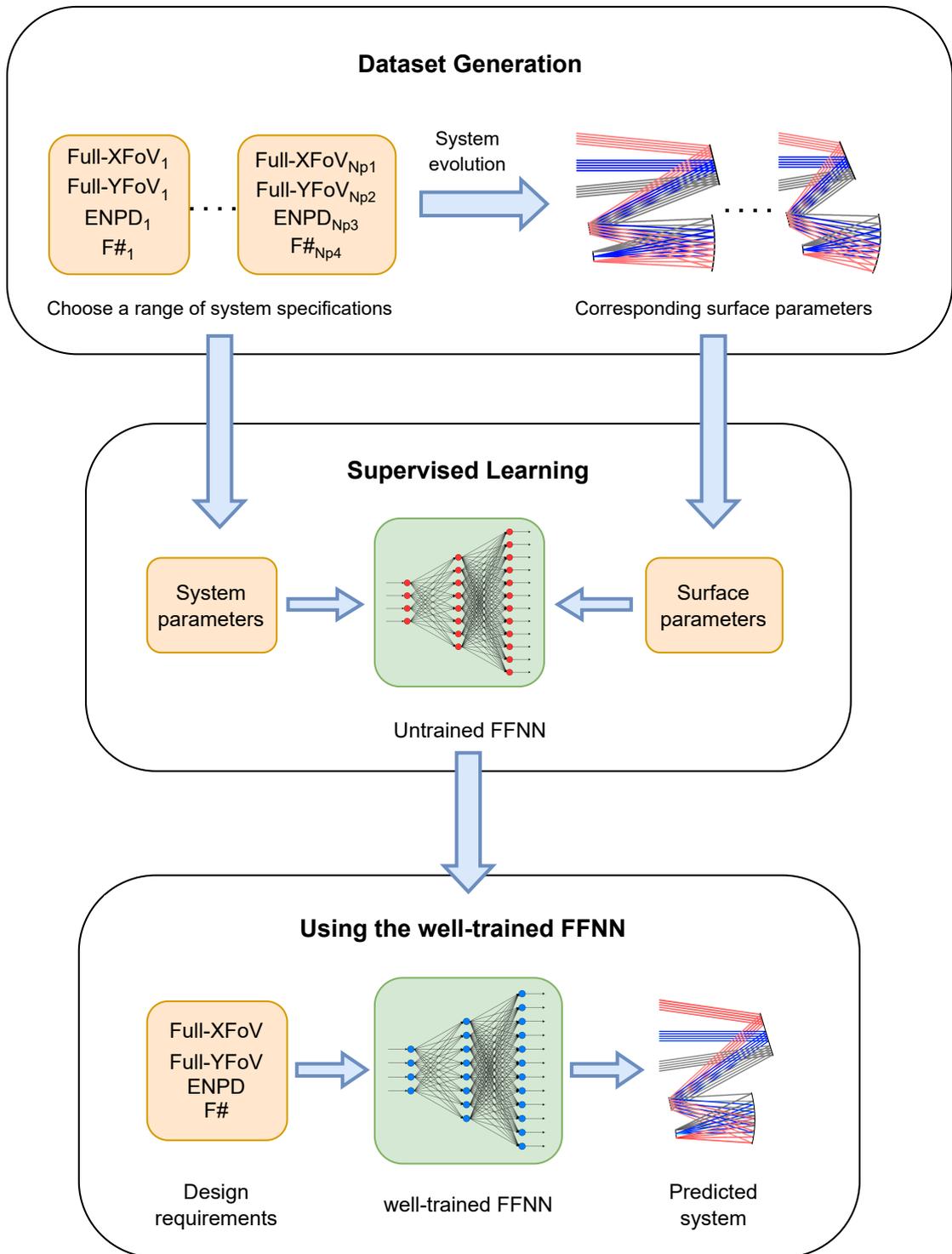


Figure 4.1: Illustration of the design framework.

4.2 DATASET

In our design framework, we use an array, **SYSP**, to describe the system specification. The array contains four representative parameters:

$$\mathbf{SYSP} = [\text{Full-XFoV}, \text{Full-YFoV}, \text{ENPD}, f\#]$$

we define

Full-XFoV: the full FoV of the system in the x -direction.

Full-YFoV: the full FoV of the system in the y -direction.

ENPD: the entrance-pupil diameter

$f\#$: the f -number

The surface parameter array, **SURP**, includes the surface locations and surface coefficients. In this study, only the common case in which the system is symmetric about the YOZ plane is considered. Thus, the local position for each surface can be generally described using the global decenter (y, z) and tilt values α relative to a predefined global coordinate system.

The surface coefficients are the parameters to describe the surface shapes. The proposed design framework can be applied to many different types of freeform surfaces, including the XY polynomial, the Zernike polynomial, the Q-type polynomial, and the radial basis function (see Section 2.1 for a complete review). However, given a specific design specification or geometry, some specific surface types may lead to better results than others. We assume a total number of Ω surface parameters to be recorded in the target array **SURP**.

The next step is to define, according to the final application or function, the range of each system parameter. There are two approaches available for creating representative base systems that fall within the specified ranges

1. Uniform sampling: sample N_{pi} ($1 \leq i \leq 4$) different parameters for each parameter range $\mathbf{SYSP}_{i,min} \leq \mathbf{SYSP}_i \leq \mathbf{SYSP}_{i,max}$ at equal intervals. After full combinations, there is a total of $N = N_{p1} \times N_{p2} \times N_{p3} \times N_{p4}$ different system specifications **SYSP**.
2. Random sampling: randomly sample N different system specifications within the entire range.

4.2. DATASET

According to Yang [84], the equal-interval sampling of system parameters is found to be more stable to get a satisfactory network.

For each \mathbf{SYSP}_i ($1 \leq i \leq N$), the corresponding base system $\mathbf{BaseSys}_i$ with good imaging performance must be generated to obtain the surface parameters \mathbf{SURP}_i . One important step in the design process is finding base systems that match the required folding geometry and deliver high-quality imaging performance. The imaging performance of the starting points generated by the FFNN greatly depends on the performance of the systems in the training dataset, thus, the optimization process should be performed carefully. However, it should not be too meticulous, considering the time cost for dataset generation.

We use a special system evolution method to obtain the training dataset. The main concepts of this method are described as follows.

Phase 1: an initial base system, $\mathbf{BaseSys}_1$, must be generated first. This system can be designed by non-specialists or beginners with little experience. The initial configuration is achieved by exploring starting systems from patents or using some other design methods (see Section 3.6) based on design specifications and folding geometry. Then optimization is conducted to get good imaging performance. Basic constraints for optimization should control the system parameters, the system structure (ensuring geometric constraints while avoiding light obscuration), and aberrations (such as image distortion and others). Note that the above constraints are used for the whole system evolution. It is important to note that the efficacy of the design framework greatly depends on the quality of the initial base system, thus, choosing a non-trivial \mathbf{SYSP} (i.e., large ENPD, wide FoV and low f -number) as the first system may be paramount. However, this requires advanced design skills and experience.

Phase 2: we define \mathbf{SYSP}^* and \mathbf{SYSP}^{**} as the already used and not used \mathbf{SYSP} , respectively. When $\mathbf{BaseSys}_i$ is obtained, the $N - i$ weighted distances (d_i) between \mathbf{SYSP}_i^* and the $N - i$ remaining \mathbf{SYSP}^{**} are calculated according to

$$d_i = \left\| w \otimes (\mathbf{SYSP}_i^* - \mathbf{SYSP}_j^{**}) \right\|_2 \quad \text{with } 1 \leq j \leq N - i \quad (4.1)$$

where d_i is the i^{th} column of the $(N - i) \times i$ distance matrix D

$$D(j, k) = \left\| w \otimes (\mathbf{SYSP}_k^* - \mathbf{SYSP}_j^{**}) \right\|_2 \quad \text{with } 1 \leq k \leq i \text{ and } 1 \leq j \leq N - i \quad (4.2)$$

$\|\cdot\|_2$ is the 2-norm. $w = [w_{Full-XFoV}, w_{Full-YFoV}, w_{ENPD}, w_{F/\#}]$ is the weight vector that determines the individual weight of each system parameter, chosen to balance the contribution of each system parameter, or to yield a specific hierarchy, to the system evolution. \otimes represents element-wise vector multiplication.

Phase 3: before generating $\mathbf{BaseSys}_{i+1}$, note that the next base system $\mathbf{BaseSys}_{i+1}$ may not be evolved from $\mathbf{BaseSys}_i$. Finally, the matrix coordinates (h, q) corresponding to the minimum $D(h, q)$ are evaluated. Therefore, $\mathbf{BaseSys}_h^{**}$ ($1 \leq h \leq N - i$) is chosen to be the next system to be optimized from $\mathbf{BaseSys}_q^*$ ($1 \leq q \leq i$). Hence, the $i + 1$ system specification will be $\mathbf{SYSP}_{i+1}^* = \mathbf{SYSP}_h^{**}$.

Phase 4: the optimization is conducted starting from $\mathbf{BaseSys}_q^*$ to obtain $\mathbf{BaseSys}_{i+1}^{**}$. To boost the system evolution, if Full-XFoV and Full-YFoV of q and $i + 1$ base systems are the same, then is best to scale $\mathbf{BaseSys}_q^*$ by a scale factor $ENPD_{i+1}/ENPD_q$ before optimization. In case the optimization constraints are not satisfied, the algorithm may go back to Phase 3, and the subsequent nearest $\mathbf{BaseSys}_h^{**}$ from $\mathbf{BaseSys}_q^*$ is selected to be further optimized to obtain $\mathbf{BaseSys}_{i+1}^{**}$ again. This step is at the discretion of the developer and can be repeated a number of times. Eventually, if the last generated base system does not meet the requirements, then it is discarded.

Phase 5: Phase (2)-(4) are repeated until all the base systems have been generated and all the surface parameters have been recorded.

4.3 SUPERVISED LEARNING

4.3.1 NETWORK ARCHITECTURE

After the fundamental dataset is obtained, the neural network can be trained using supervised training. Before that, we need to define the network architecture. For any given labeled training set, we aim to learn a function F that maps a

4.3. SUPERVISED LEARNING

sequence of real-valued input parameters X to a sequence of real-valued output parameters Y

$$F : X \rightarrow Y$$

SL is a task-driven approach, hence, it is carried out when certain goals are identified to be accomplished from a certain set of inputs [88]. In our design framework, the supervised task is inferred as a regression, i.e., a proper fit of the training dataset with a continuous function. For this reason, we choose a FFNN: a fully-connected network consisting of an input layer, one or more hidden layers, and an output layer (see Fig. 4.2). By "Feed-Forward" we mean that the information moves unidirectionally from the input to the output without making cycles.

Except the first layer, the inputs to the nodes are weighted sums of the outputs from the nodes in the previous layer. For instance, the k^{th} node in the first hidden layer (i.e., second layer) will have as input

$$I_k^1 = \sum_{\psi=1}^4 w_{k,\psi}^1 \mathbf{SYSP}_{\psi} + b_k^1 \quad (4.3)$$

where $w_{k,\psi}^1$ is the weight of the link between the ψ^{th} ($1 \leq \psi \leq 4$) node in the input layer to the k^{th} node in the first hidden layer, while b_k^1 is the bias term. Each node in the hidden layers expresses a nonlinear activation function (e.g., Sigmoid-type function) that is the same for all nodes. Conversely, the whole nodes in the output layer convey a type of linear activation function. O_k^l is the output of the k^{th} node in the l^{th} ($2 \leq l \leq L$) layer that will be part of the input for the nodes in the $(l+1)^{\text{th}}$ layer. O_k^l can be represented as

$$O_k^l = A\left(\sum_{\psi=1}^{\Psi} w_{k,\psi}^{l-1} O_{\psi}^{l-1} + b_k^l\right) \quad (4.4)$$

where L is the total number of layers in the FFNN, A is the activation function, $w_{k,\psi}^{l-1}$ is the weight of the link between the ψ^{th} node in the $(l-1)^{\text{th}}$ layer to the k^{th} node in the l^{th} , while b_k^l is the bias term. The "deepness" of the network (i.e., the number of hidden layers) should be selected carefully: more complex architectures allow stronger expressive abilities at the cost of a higher time

complexity. Two hidden layers are found to be a good trade-off, thus, we propose a 4-layer FFNN.

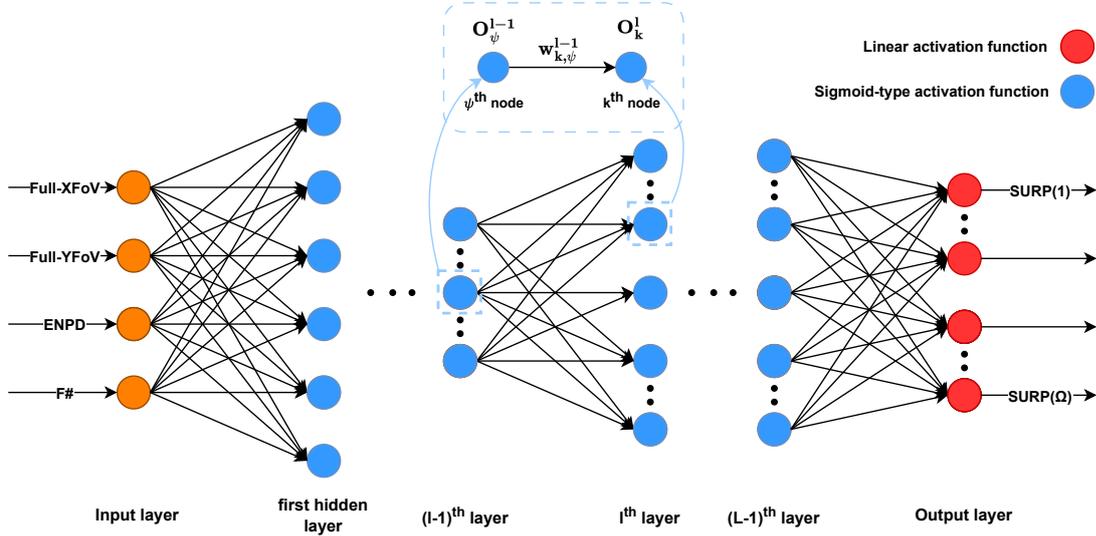


Figure 4.2: Illustration of a generic FFNN.

4.3.2 NETWORK TRAINING AND VALIDATION

To improve the efficiency and convergence of training, all input data with the same type are normalized to the range of $[-1, 1]$. Since the values and units of different parameters can vary significantly, normalization helps to make them comparable and easier to work with. This is done using a linear preprocessing method that is similar to min-max scaling. For instance, if $\overline{\mathbf{SYSP}}_a^\tau$ is the preprocessed system parameter of \mathbf{SYSP}_a^τ for each system a ($1 \leq a \leq N$) and parameter τ ($1 \leq \tau \leq 4$), it holds that

$$\overline{\mathbf{SYSP}}_a^\tau = \frac{\mathbf{SYSP}_a^\tau - 0.5 \times [\max(\mathbf{SYSP}_{TOT}^\tau) + \min(\mathbf{SYSP}_{TOT}^\tau)]}{0.5 \times [\max(\mathbf{SYSP}_{TOT}^\tau) - \min(\mathbf{SYSP}_{TOT}^\tau)]} \quad (4.5)$$

where, $\max(\mathbf{SYSP}_{TOT}^\tau)$ and $\min(\mathbf{SYSP}_{TOT}^\tau)$ are the maximum and minimum values of the τ^{th} system parameter in the training dataset, respectively.

Similarly, if $\overline{\mathbf{SURP}}_a^\tau$ is the preprocessed surface parameter of \mathbf{SURP}_a^τ for each system a ($1 \leq a \leq N$) and parameter τ ($1 \leq \tau \leq 4$), it holds that

$$\overline{\mathbf{SURP}}_a^\tau = \frac{\mathbf{SURP}_a^\tau - 0.5 \times [\max(\mathbf{SURP}_{TOT}^\tau) + \min(\mathbf{SURP}_{TOT}^\tau)]}{0.5 \times [\max(\mathbf{SURP}_{TOT}^\tau) - \min(\mathbf{SURP}_{TOT}^\tau)]} \quad (4.6)$$

4.3. SUPERVISED LEARNING

where, $\max(\mathbf{SYSP}_{TOT}^{\tau})$ and $\min(\mathbf{SYSP}_{TOT}^{\tau})$ are the maximum and minimum values of the τ^{th} surface parameter in the training dataset, respectively. The choice of the preprocessing technique employed in this context is not definitive, and alternative approaches may also be suitable. Note that to extract the exact surface parameters from the trained network, we need to perform reverse processing.

The training parameters, including the loss function, the optimization algorithm, the learning rate, and weight initialization, must be chosen carefully.

The loss function can be defined as the Mean Squared Error (MSE) between the target and actual predicted output (**SURP**). Once the loss is calculated by the feed-forward process, the derivative or slope of the loss function across all the parameters in the neural network has to be computed. This is done using a backpropagation algorithm [89], which propagates the derivative backward through the network. The derivative is then utilized by the optimizer to adjust the weights and biases, reducing the loss function. This helps the neural network to generate predictions that are closer to the target values.

Many backpropagation training algorithms exist, here we present the two most common classes. The first employs the Jacobian derivatives (e.g., Levenberg-Marquardt and Bayesian Regulation algorithms), known to be fast but memory-starving. The other class uses the Gradient derivatives (e.g., Scaled conjugate gradient and Gradient descent algorithms), which usually are better at finding the global minimum but require more time. Since the low complexity of the proposed FFNN, any backpropagation algorithm should be able to perform greatly (provided that the input dataset is of decent quality).

The learning rate is crucial when training an FFNN. If too high, the model may not converge; if too low, training time increases. Optimal performance requires finding the right balance. Also, at the beginning of the training, the weights and biases should be initialized to some values. The standard method for initializing weights in network layers that use Sigmoid-type activation functions is known as Xavier initialization [90].

When training a neural network, there are two common obstacles to overcome: underfitting and overfitting. Underfitting happens when the model is too simple to capture the data's underlying patterns. As a result, it performs poorly on both the training and unseen data. Overfitting, on the other hand, occurs

when the model becomes too complex, fitting the training data too closely and performing poorly on new, unseen data. To achieve a well-generalized model, it is crucial to mitigate the overfitting problem. One effective strategy is to use a validation set and the early stopping method. A validation set that is distinct from the training set allows for monitoring the model's performance on data it has not seen during training. By regularly evaluating the model on the validation set, it is possible to detect when overfitting starts, as the performance on the validation set may plateau or degrade. The early stopping method involves halting the training process once the model's performance on the validation set ceases to improve. This prevents it from becoming too specialized to the training data and enhances its ability to generalize to new, unseen data. This approach aims to balance between model complexity and generalization, creating a more robust and effective neural network.

5

Example Demonstration

In this chapter, we aim to demonstrate the feasibility of our design framework by designing a freeform implementation of the Wetherell-configuration reflective triplet as described in Section 3.4. This imaging system is a freeform off-axis TMA telescope that presents symmetry about the YOZ plane, a traditional zig-zag folding geometry, and no intermediate images (see Fig. 5.1). M_1 , M_2 , M_3 , and IMG denote the primary mirror, secondary mirror, tertiary mirror, and image plane, respectively. The aperture stop is located at M_2 . To eliminate the light obstruction, the system is tilted until the light clears the mirrors.

Each surface in the optical system is represented by the XY-polynomial, which is a type of freeform surface (see Section 2.1 for a complete review). This surface type is used up to the fourth-term order (i.e., 8th polynomial order) and has a conic base. Since the optical system is symmetric about the YOZ plane, only even components of x are utilized:

$$z(x, y) = \frac{c(x^2 + y^2)}{1 + \sqrt{1 - (1 + k)(x^2 + y^2)}} + A_2y + A_3x^2 + A_5y^2 + A_7x^2y + A_9y^3 + A_{10}x^4 + A_{12}x^2y^2 + A_{14}y^4 \quad (5.1)$$

where c is the curvature, k is the conic constant and A_i is the coefficient of the polynomial terms (see Tab. 2.2 for the polynomial terms and corresponding interpretation).

Due to computational constraints, we limited the range of system parameters.

After taking NanoSat and CubeSat instrumentation as a case study [91], we determine the appropriate range of specifications, or, system parameter space:

$$\begin{cases} 2^\circ \leq \text{Full-XFoV} \leq 8^\circ \\ 1^\circ \leq \text{Full-YFoV} \leq 3^\circ \\ 40 \text{ mm} \leq \text{ENPD} \leq 100 \text{ mm} \\ 2 \leq F/\# \leq 4 \end{cases} \quad (5.2)$$

Then, we perform uniform sampling within the above range, in particular, we take $N_{\text{Full-XFoV}} = 7$, $N_{\text{Full-YFoV}} = 3$, $N_{\text{ENPD}} = 13$ and $N_{F/\#} = 9$. After full combinations, we obtain $N = N_{\text{Full-XFoV}} \times N_{\text{Full-YFoV}} \times N_{\text{ENPD}} \times N_{F/\#} = 2457$. Furthermore, we include a new constraint that takes into account the actual design cases. We choose not to consider telescopes that have a Full-YFoV greater than Full-XFoV, as such configurations are unlikely to occur in YOZ -symmetric systems [91]. Eventually, the total number of base systems becomes $N = 2106$.

The surface parameters comprise the surface locations and surface coefficients. To determine the surface locations, the vertex of M_2 is set as the reference of the global coordinate system. These locations include the global z -decenter of M_1 , M_3 , and IMG, in addition to the α -tilt, which is the system rotation required to remove the light obstruction. The surface coefficients for each freeform surface are the values given in Eq. 5.1, and comprehend the curvature c , the conic constant k and the eight polynomial coefficients. A total of $\Omega = 34$ surface parameters are recorded in **SURP** for each system.

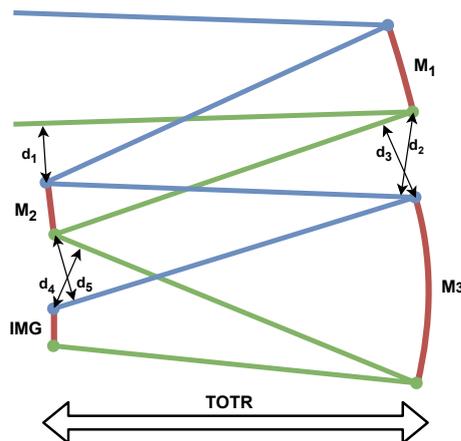


Figure 5.1: Ziz-zag geometry of freeform off-axis three-mirror telescope and its structure constraints.

5.1 BASE SYSTEM DESIGN

The system parameters of the initial base system **BaseSys₁** are

$$\mathbf{SYSP}_1 = [2^\circ, 1^\circ, 50 \text{ mm}, 4]$$

which leads to an easily designed system, since the smallest FoV and the largest $F/\#$ are used. To generate this system, we first employ the third-order correction method (see Section 3.4) to get an on-axis, rotationally symmetric TMA Wetherell-configuration. Consequently, we eliminate the light obscuration by tilting the surfaces and we pursue a gradual optimization strategy.

First phase: On-axis rotationally symmetric TMA The eight input parameters of the Wetherell reflective triplet are (see Section 3.4):

$$\begin{cases} d_2 = -d_1, & s'_3 = 1.6820d_1, & f_s = -2.0112d_1, \\ \Omega_1 = 0.5380, & \Omega_2 = 1.5545, \\ k_1 = -2.710014, & k_2 = -0.415013, & k_3 = 0.130922 \end{cases} \quad (5.3)$$

where d_i is the mirror separation between M_{i+1} and M_i , s'_3 is the image distance from M_3 to the image plane IMG, f_s is the image focal length of the system, Ω_i is the paraxial ray-height ratio and k_i is the conic constant of mirror M_i (corresponding to δ_i in Section 3.4).

Now, based on the information presented in Tab. 3.2, we can derive the following relationships

$$\begin{aligned} c_1 &= \frac{1}{2} \left(\frac{1 - \Omega_1}{d_1} \right) \\ c_2 &= \frac{1}{2} \left(\frac{1 - \Omega_1}{\Omega_1 d_1} + \frac{1 - \Omega_2}{d_2} \right) \\ c_3 &= \frac{1}{2} \left(\frac{1 - \Omega_2}{\Omega_2 d_2} + \frac{1}{s'_3} \right) \end{aligned}$$

where c_i is the mirror curvature of M_i .

By imposing $\text{ENPD} = 50 \text{ mm}$ and $F/4$, we set $f_s = \text{ENPD} \cdot F/\# = 200 \text{ mm}$. The surface data are evaluated from Eq. 5.3 and reported in Tab. 5.1 in mm-unit. The on-axis optical layout is presented in Fig. 5.2.

	M ₁	M ₂	M ₃
Curvature	-0.002323	-0.007105	-0.004782
Conic	-2.710014	-0.415013	0.130922
Separation	-99.453000	99.453000	-167.279946

Table 5.1: Surface data of $F/4$ on-axis rotationally symmetric TMA with ENPD = 50 mm.

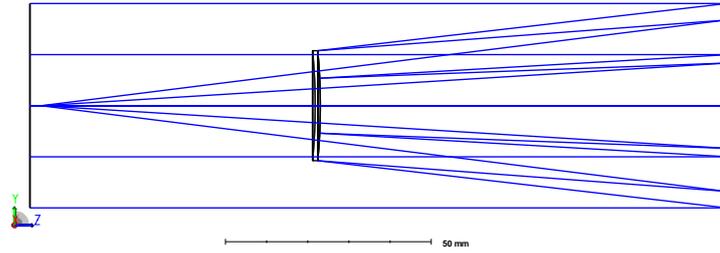


Figure 5.2: On-axis rotationally symmetric TMA layout.

Second phase: Off-axis freeform TMA Now we are ready to enter the remaining system parameters. First, we choose a rectangular biased-FoV, centered on $(0^\circ, -0.5^\circ)$. Because the system is symmetric about the YOZ -plane, only half of the 2° full-XFoV needs to be considered during the design. Thus, we select the following 9 fields: $(0^\circ, 0^\circ)$, $(0^\circ, -0.5^\circ)$, $(0^\circ, -1^\circ)$, $(0.5^\circ, 0^\circ)$, $(0.5^\circ, -0.5^\circ)$, $(0.5^\circ, -1^\circ)$, $(1^\circ, 0^\circ)$, $(1^\circ, -0.5^\circ)$, $(1^\circ, -1^\circ)$. Then, we set the operation wavelength $\lambda = 0.55 \mu\text{m}$.

Initially, the light clearance is achieved by manually tilting the system through a coordinate break surface [92] placed before the first mirror (α -tilt). Then, a progressive strategy is adopted to introduce and optimize each freeform surface. To assess the imaging performance in the optimization process, we choose to minimize the RMS spot size using the Zemax OPTICSTUDIO *Optimization Wizard* and control the Seidel aberrations through Zemax OPTICSTUDIO built-in operands (SPHA, COMA, ASTI, DIMX) [92]. To control the Total Track Length (TOTR) and the telescope volume, it is crucial to bound $d_1 + d_2$ and the mirror diameters.

Before the optimization, proper constraints should be selected. To handle the configuration geometry and structure constraints (see Fig. 5.1), we implement a computationally efficient raytracing method through a Zemax Programming Language (ZPL) macro (briefly, ZPLM). The ZPLM individually controls the

distances (d_1, d_2, d_3, d_4 , and d_5) between marginal rays and mirror edges and is suitable for any off-axis reflective systems. The general mathematical principle is described in Subsection 5.1.1.

In the first step of the optimization, the system structure needs to be maintained, avoiding any obscuration and overlap between mirrors. The α -tilt, the curvatures, and conic constants are set as variables. The $F/\#$ is set to 4. The IMG y -decenter and tilt about x -axis are evaluated through the Chief Ray solver [92] from the $(0^\circ, 0^\circ)$ field and $\lambda = 0.55 \mu\text{m}$ wavelength. In the second step, all the mirrors are replaced with XY-polynomial surfaces (known as Extended Polynomial in OPTICSTUDIO). Thus, XY-terms up to the 2nd order and mirror separations are set as variables. As the optimization progresses, the XY-polynomial order is increased until the 4th term-order, for a total of eight polynomial coefficients per surface. After the final stage of the optimization, a diffraction-limited freeform off-axis three-mirror telescope is obtained (see Fig. 5.3). The surface parameters are reported in Tab. 5.2 in mm-unit. The RMS spots and the Seidel aberrations are plotted in Fig. 5.4 and Fig. 5.5, respectively. The TOTR is 142.35 mm.

	M₁	M₂	M₃
Curvature	-0.002410	-0.006999	-0.004920
Conic	-1.319729	0.238910	0.273655
Separation	-105.860491	106.319803	-142.354031
α-tilt	-8.160475°		
A₂	-3.160755e - 03	5.664674e - 02	-7.579894e - 03
A₃	1.172562e - 05	3.864781e - 05	-2.600642e - 05
A₅	-1.575939e - 05	-6.570647e - 05	-1.733906e - 05
A₇	2.099685e - 07	4.303813e - 06	2.021158e - 07
A₉	3.197684e - 08	1.721260e - 06	-6.418369e - 08
A₁₀	9.725825e - 10	1.887585e - 08	2.006808e - 10
A₁₂	1.513949e - 09	1.808823e - 08	-5.293407e - 10
A₁₄	5.197476e - 10	5.048701e - 09	8.217985e - 10

Table 5.2: BaseSys₁ surface data.

5.1. BASE SYSTEM DESIGN

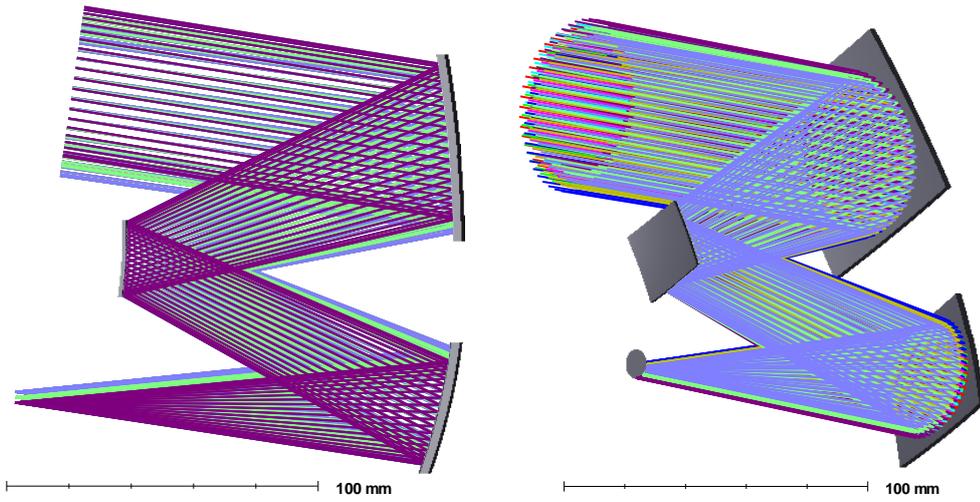


Figure 5.3: BaseSys₁ layout.

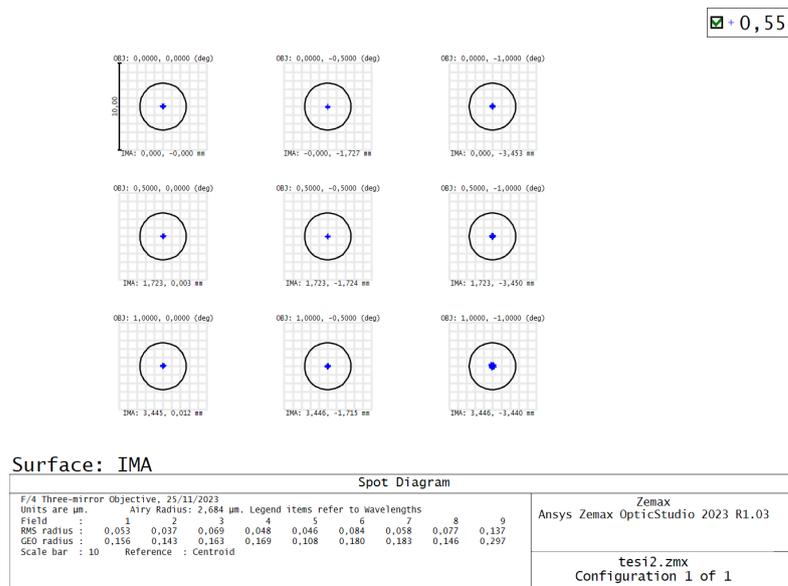


Figure 5.4: BaseSys₁ RMS spots.

5.1.1 CONTROL OF STRUCTURE CONSTRAINTS: THE 3D POINT-TO-LINE DISTANCE

Referring to Fig. 5.6, the equation of the light ray \vec{s} can be described as

$$s(\vec{t}) = a + t\vec{n} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} + t \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \quad (5.4)$$

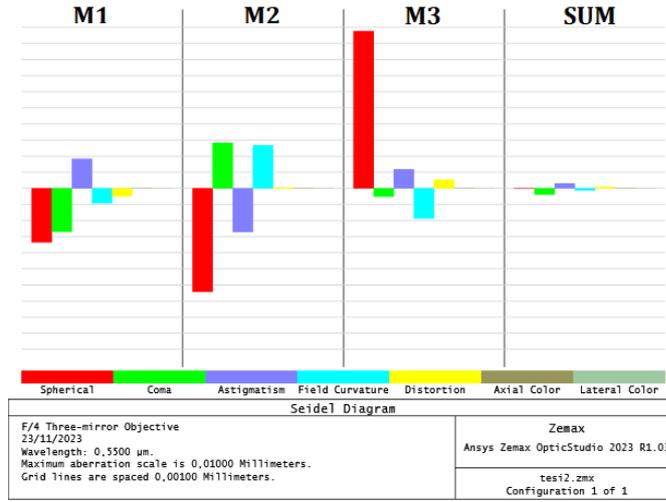
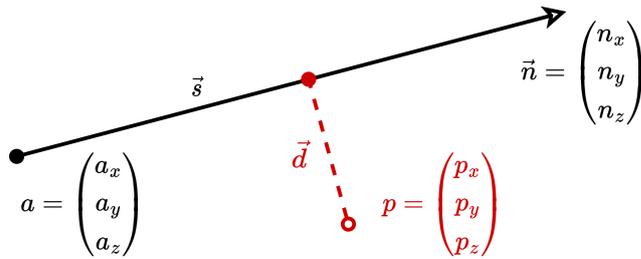

 Figure 5.5: BaseSys₁ Seidel aberrations diagram.


Figure 5.6: The point-to-line distance in three-dimensional space.

where a is a point on the light ray and \vec{n} is a unit vector in the direction of the ray. \vec{s} varies with t and represents any point on that ray.

The vector \vec{d} links the ray \vec{s} to the mirror edge p . In general, for any t , it holds

$$\vec{d} = \begin{pmatrix} a_x + tn_x \\ a_y + tn_y \\ a_z + tn_z \end{pmatrix} - \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} \quad (5.5)$$

We want to solve t to minimize $\|\vec{d}\|$. Practically, the vectors \vec{n} and \vec{d} must be orthogonal, thus, their scalar product is

$$\begin{cases} \vec{n} \cdot \vec{d} = \|\vec{n}\| \|\vec{d}\| \cos \theta|_{\theta=90^\circ} = 0 \\ \vec{n} \cdot \vec{d} = n_x(a_x - p_x + tn_x) + n_y(a_y - p_y + tn_y) + n_z(a_z - p_z + tn_z) \end{cases} \quad (5.6)$$

5.2. DATASET GENERATION AND ANALYSIS

Following easy calculations, it can be found that

$$t^* = n_x(p_x - a_x) + n_y(p_y - a_y) + n_z(p_z - a_z) \quad (5.7)$$

The shortest distance from the light ray \vec{s} to the mirror edge p , corresponding to the norm of \vec{d} , is found by substituting t^* back into \vec{d}

$$\begin{aligned} \text{distance}(\vec{s}, p) &= \left\| \vec{d}(t^*) \right\| = \\ &= \sqrt{(a_x - p_x + t^*n_x)^2 + (a_y - p_y + t^*n_y)^2 + (a_z - p_z + t^*n_z)^2} \quad (5.8) \end{aligned}$$

Ultimately, by examining the sign of vector \vec{d} , it is possible to determine whether the mirror edge p is above (> 0) or below (< 0) the ray.

5.2 DATASET GENERATION AND ANALYSIS

The dataset generation is achieved through a system evolution algorithm (see A.1 for the full MATLAB code). The algorithm evaluates the minimum weighted distance from the already used and not used **SYSP** at every cycle, according to the definition given in Eq.4.2. In particular, we choose the weight vector to balance the contribution of each system parameter:

$$w = [0.25, 0.25, 0.05, 1].$$

For each subsequent base system optimization, we choose to minimize the RMS spot size while imposing the above structure constraints and controlling the maximum distortion. The system optimization is fully automated and doesn't need the supervision of a human operator. In particular, we decided to follow an approach that optimizes radii, curvatures, conics, and surface separations while gradually incrementing the polynomial order of the XY-polynomial surfaces. It took a total of 58.5 hours to generate the dataset on a computer using an AMD Ryzen 5 5500 U central processing unit @2.1 GHz and 8 GB of internal memory.

Tab 5.3 presents nine base systems generated through the system evolution algorithm proposed by us. For each system, we report the layout, the **SYSP**, the average RMS spot size over the fields (RMS Spot), the maximum relative distortion (Dist), and the TOTR. Moreover, we point out in red all the RMS spots

whose size is above the diffraction limit DL, defined as [52]

$$DL = 1.22 \lambda F/\# \quad \text{with } \lambda = 0.55 \mu\text{m} \quad (5.9)$$

We present nine 3D plots to assess the quality of the dataset generated with the system evolution method (see Fig. 5.7, 5.8, 5.9, 5.10, 5.11). The plots depict 234 systems per each $F/\#$, for a total of 2106 systems. The axes are defined according to the system parameter space (see Eq. 5.2); in particular, the x -axis, y -axis, and z -axis take the Full-XFoV, the ENPD and the Full-YFoV values, respectively. Furthermore, the color of the sphere represents the average RMS spot size.

By looking at the figures, we can state that the average RMS spot size increases with Full-XFoV, Full-YFoV, and ENPD, while it decreases with $F/\#$. Moreover, by means of Eq. 5.9, we find that the diffraction-limited systems per each $F/\#$ are 39.7% of $F/2$, 67.5% of $F/2.25$, 85.9% of $F/2.5$, and 100% of $F/2.75$, $F/3$, $F/3.25$, $F/3.5$, $F/3.75$ and $F/4$.

In Fig. 5.12 right, we display the maximum relative distortion and the TOTR of each generated system. The system number is determined according to the minimum weighted distance from \mathbf{SYSP}_1 .

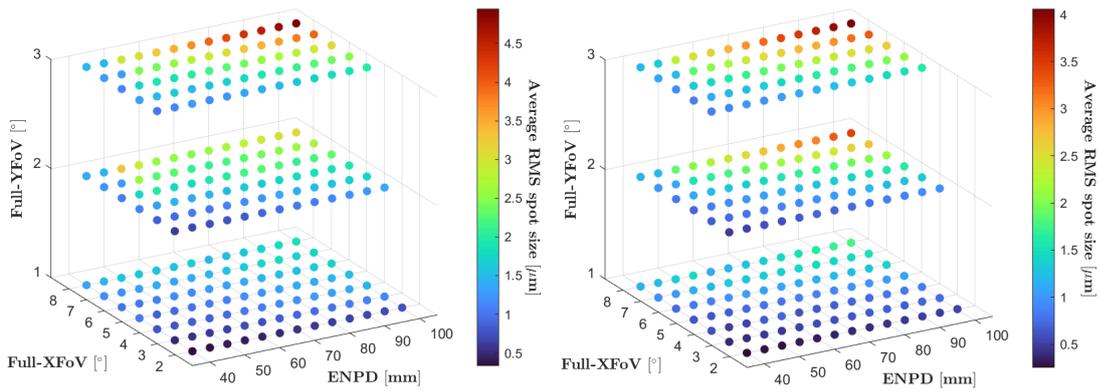


Figure 5.7: The average RMS spot size of each $F/2$ (on the left) and $F/2.25$ (on the right) imaging system as a function of Full-XFov, Full-YFoV, and ENPD.

5.2. DATASET GENERATION AND ANALYSIS

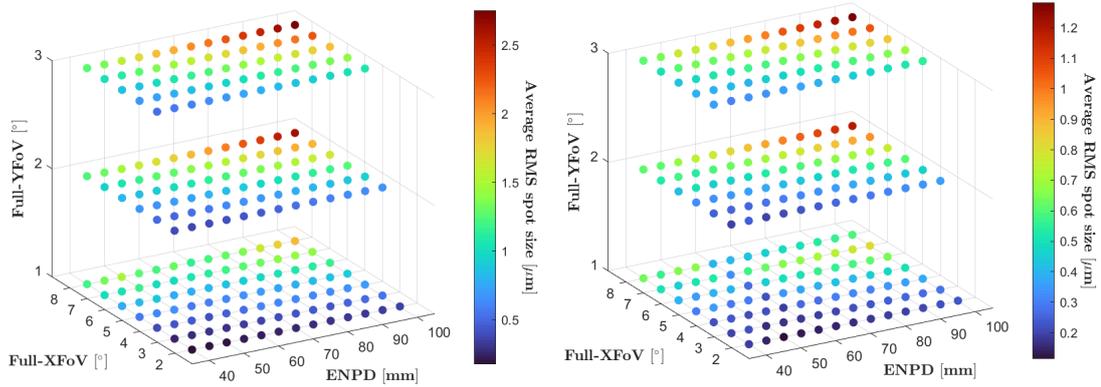


Figure 5.8: The average RMS spot size of each $F/2.5$ (on the left) and $F/2.75$ (on the right) imaging system as a function of Full-XFoV, Full-YFoV, and ENPD.

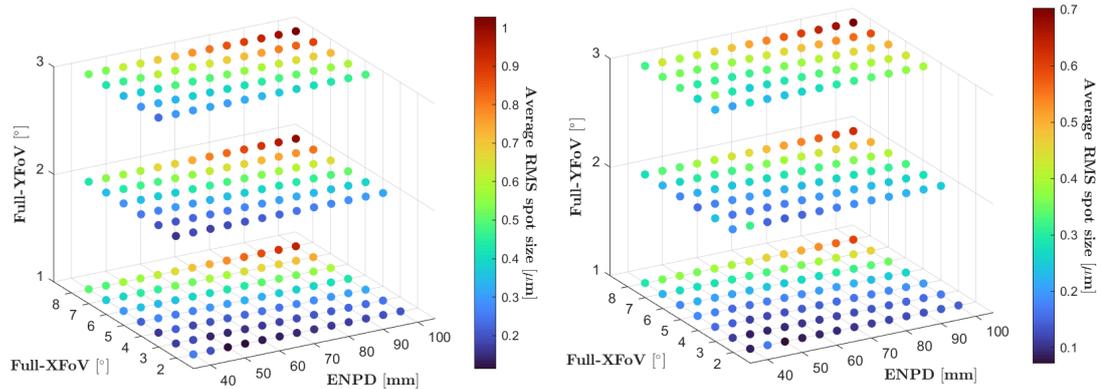


Figure 5.9: The average RMS spot size of each $F/3$ (on the left) and $F/3.25$ (on the right) imaging system as a function of Full-XFoV, Full-YFoV, and ENPD.

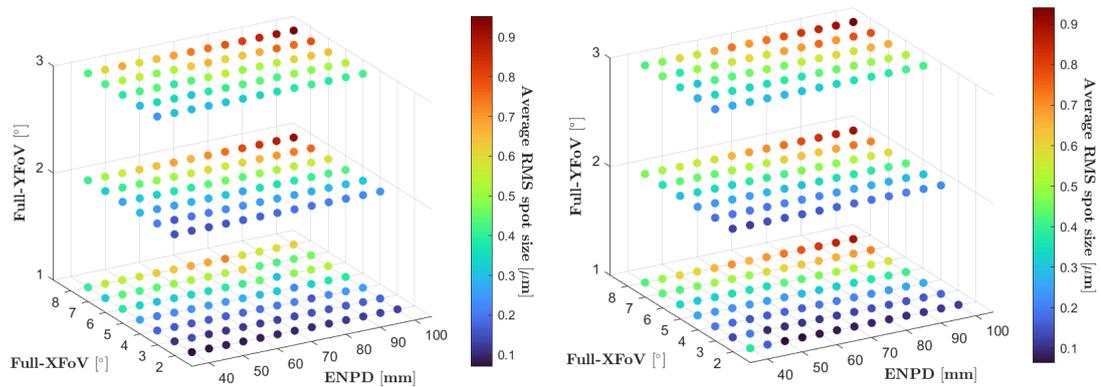


Figure 5.10: The average RMS spot size of each $F/3.5$ (on the left) and $F/3.75$ (on the right) imaging system as a function of Full-XFoV, Full-YFoV, and ENPD.

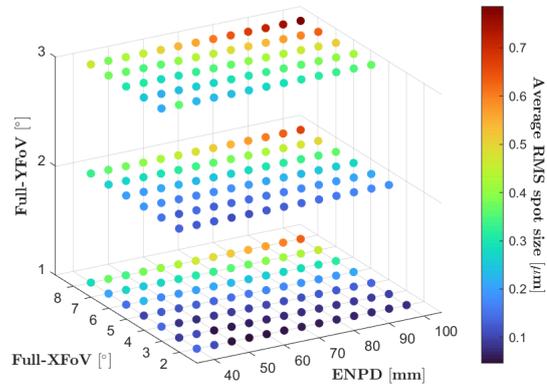


Figure 5.11: The average RMS spot size of each $F/4$ imaging system as a function of Full-XFoV, Full-YFoV, and ENPD.

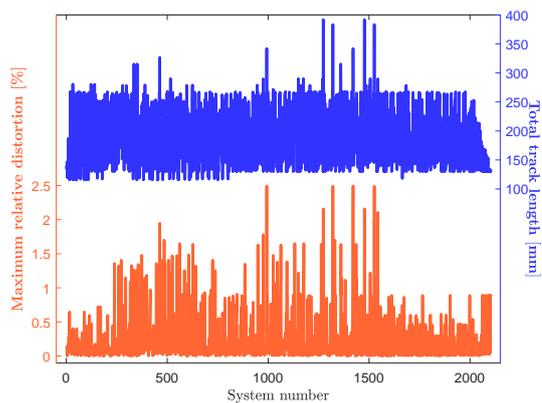


Figure 5.12: The maximum relative distortion and the total track length of each imaging system.

5.2. DATASET GENERATION AND ANALYSIS

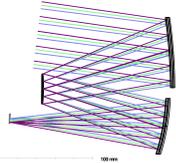
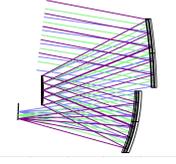
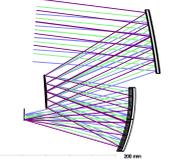
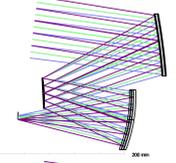
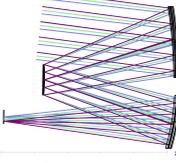
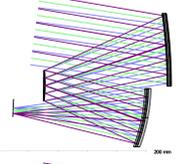
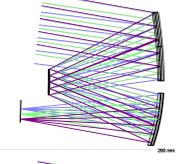
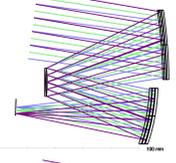
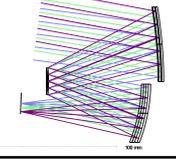
Layout	FoV	ENPD [mm]	F/#	RMS Spot [μm]	Dist [%]	TOTR [mm]
	$2^\circ \times 1^\circ$	65	3.5	0.091	0.11	173.25
	$4^\circ \times 3^\circ$	60	2.5	0.684	0.25	148.19
	$4^\circ \times 3^\circ$	90	2	1.803	1.35	231.02
	$6^\circ \times 2^\circ$	85	2.25	1.846	1.09	233.69
	$6^\circ \times 1^\circ$	75	4	0.246	0.07	209.78
	$5^\circ \times 2^\circ$	80	2.75	0.492	0.09	211.12
	$8^\circ \times 3^\circ$	70	3.75	0.693	0.28	187.10
	$7^\circ \times 2^\circ$	40	3.25	0.225	0.14	130
	$5^\circ \times 3^\circ$	45	3	0.318	0.06	118.95

Table 5.3: Typical base systems obtained through system evolution.

5.3 NETWORK DESIGN AND ANALYSIS

In this design framework, the FFNN has an input layer, two hidden layers, and an output layer, having 4, 30, 40, and 34 nodes, respectively. The network design is displayed in Fig. 5.13 and is trained using the generated dataset.

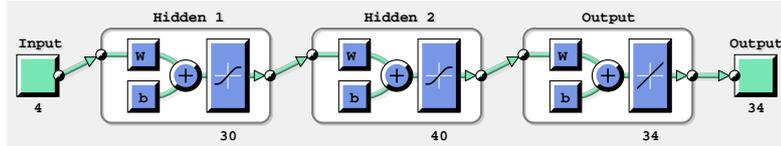


Figure 5.13: The Feed-Forward Neural Network design.

Hereafter, we briefly present the MATLAB code reported in A.2. We set a division by randomly sampling the dataset for training (70%), validation (15%), and testing (15%). The Levenberg-Marquardt algorithm (i.e., *trainlm* [93]) and the MSE are selected as the optimization algorithm and the loss function, respectively. We employ a built-in function-fitting neural network (i.e., *fitnet* [94]), and the training parameters are left unchanged. The total number of training epochs is 110, but the validation stops at 104. The train, validation, and test loss values are $4.47e - 04$, $6.82e - 04$, and $9.35e - 04$, respectively. The training time is approximately 35 minutes on a computer using an AMD Ryzen 5 5500 U central processing unit @2.1 GHz and 8 GB of internal memory. In Fig. 5.14, we plot the train, validation, and test performances, the training state values, the error histogram, and the linear regression (with $R = 0.99937$).

To better evaluate the feasibility of the trained FFNN, we utilize it to predict surface parameters using a new, different set of system parameters within the system parameter space defined by Eq. 5.2. Thus, we perform uniform sampling and we take $N_{\text{Full-XFoV}} = 6$, $N_{\text{Full-YFoV}} = 4$, $N_{\text{ENPD}} = 6$ and $N_{F/\#} = 6$. After full combinations, we obtain $N = N_{\text{Full-XFoV}} \times N_{\text{Full-YFoV}} \times N_{\text{ENPD}} \times N_{F/\#} = 864$. As before, we keep only the systems having Full-XFoV greater than Full-YFoV. Eventually, the total number of predicted systems becomes $N = 648$. The subsequent step is to evaluate the performance of the predicted systems that correspond to these output surface data.

For each system, it is required to quickly optimize by minimizing the RMS spot size, determine whether or not there are ray tracing errors and/or light obscuration issues, and then evaluate the system imaging performance as before. Tab 5.4 presents nine imaging systems predicted by the FFNN. For each system,

5.3. NETWORK DESIGN AND ANALYSIS

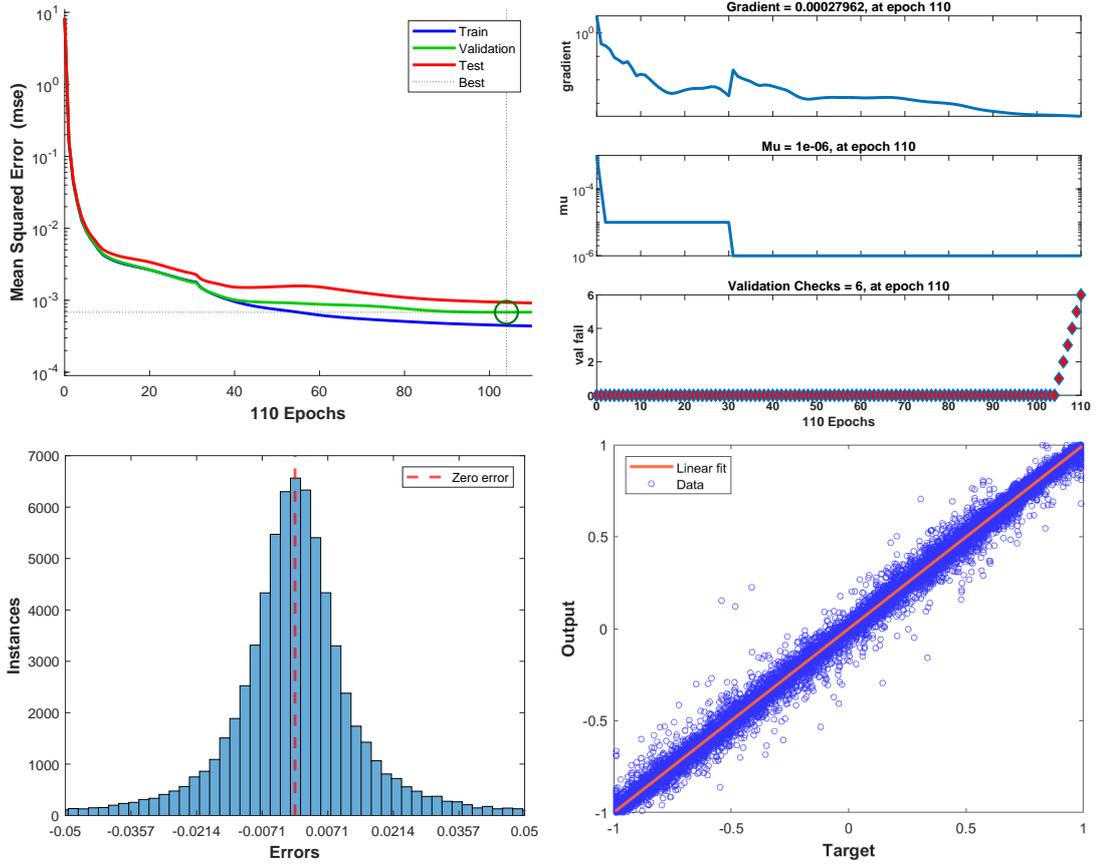


Figure 5.14: From the top to the bottom and from the left to the right. The network performance, the training state values, the error histogram, and the linear regression.

we report the layout, the **SYSP**, the average RMS spot size over the fields (RMS Spot), the maximum relative distortion (Dist), and the TOTR. Moreover, we point out in red all the RMS spots whose size is above the diffraction limit DL.

Finally, we present six 3D plots to assess the quality of the systems predicted by the FFNN (see Fig. 5.15, 5.16, 5.17). Similarly to before, the plots depict 108 systems per each $F/\#$, for a total of 648 systems. The axes are defined according to the system parameter space (see Eq. 5.2); in particular, the x -axis, y -axis, and z -axis take the Full-XFoV, the ENPD and the Full-YFoV values, respectively. Furthermore, the color of the sphere represents the average RMS spot size.

We can observe that the average RMS spot size no longer depends on Full-XFoV, Full-YFoV, and ENPD, but instead presents several outliers. However, it generally decreases with $F/\#$. Furthermore, by means of Eq. 5.9, we find that the

diffraction-limited systems per each $F/\#$ are 26.9% of $F/2$, 66.7% of $F/2.4$, 88.0% of $F/2.8$, 99.1% of $F/3.2$, 92.6% of $F/3.6$, and 96.3% of $F/4$.

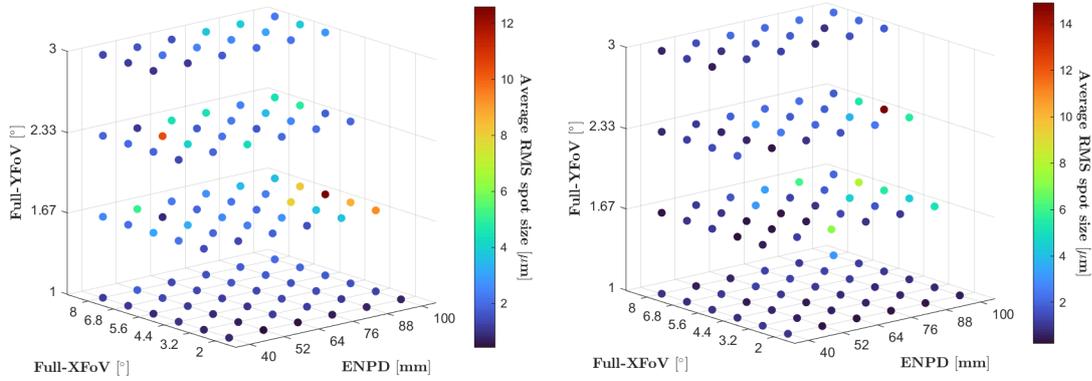


Figure 5.15: The average RMS spot size of each $F/2$ (on the left) and $F/2.4$ (on the right) imaging system as a function of Full-XFoV, Full-YFoV, and ENPD.

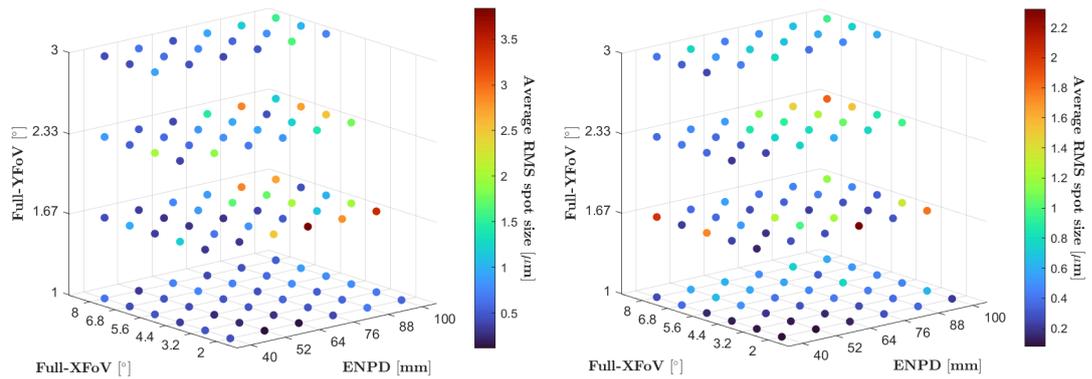


Figure 5.16: The average RMS spot size of each $F/2.8$ (on the left) and $F/3.2$ (on the right) imaging system as a function of Full-XFoV, Full-YFoV, and ENPD.

In Fig. 5.18, we display the maximum relative distortion and the TOTR of each predicted system. The system number is determined according to the minimum weighted distance from \mathbf{SYSP}_1 .

It is worth noting that the predicted system undergoes rapid and automated optimization. Therefore, we are optimistic that if a human operator were to use the network to predict and optimize a single system, they could achieve significantly better performance.

5.3. NETWORK DESIGN AND ANALYSIS

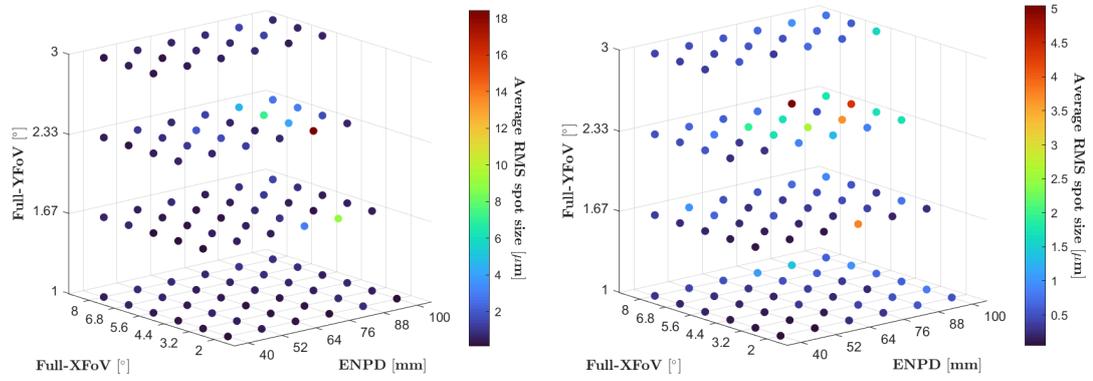


Figure 5.17: The average RMS spot size of each $F/3.6$ (on the left) and $F/4$ (on the right) imaging system as a function of Full-XFoV, Full-YFoV, and ENPD.

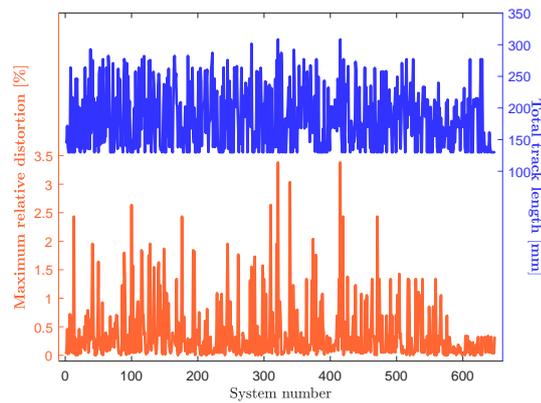


Figure 5.18: The maximum relative distortion and the total track length of each imaging system.

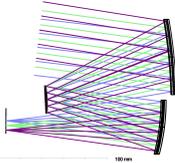
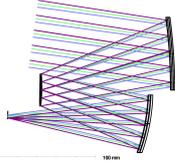
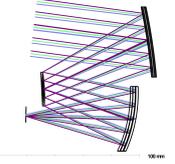
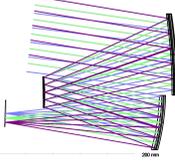
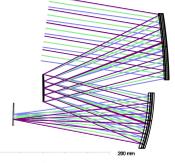
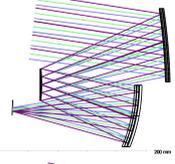
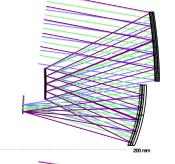
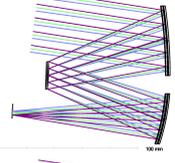
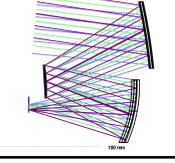
Layout	FoV	ENPD [mm]	F/#	RMS Spot [μm]	Dist [%]	TOTR [mm]
	$4.4^\circ \times 2.33^\circ$	52	4	0.272	0.68	154.26
	$2^\circ \times 1^\circ$	64	2.8	0.137	0.05	170.05
	$8^\circ \times 1^\circ$	40	2	1.147	1.29	130
	$8^\circ \times 2.33^\circ$	88	3.6	4.587	0.07	249.48
	$6.8^\circ \times 1.67^\circ$	100	3.2	0.465	0.29	274.73
	$5.6^\circ \times 1.67^\circ$	76	2.4	0.876	0.40	209.56
	$5.6^\circ \times 3^\circ$	88	2.4	1.587	0.25	210.90
	$3.2^\circ \times 1^\circ$	40	4	0.0876	0.08	130.44
	$6.8^\circ \times 3^\circ$	52	2	2.133	1.26	130

Table 5.4: Typical predicted systems obtained through our FFNN.



Conclusions and Future Works

In this thesis, we present a neural network-based design framework for off-axis three-mirror telescopes. This project aims to enhance the design of these freeform optical systems for multiple applications by leveraging machine learning and system evolution techniques. Our approach provides an automated and scalable solution that can generate starting points for the design process.

In the first chapter, we discuss the limitations of traditional telescopes and introduce freeform optics as a groundbreaking solution. We provide a historical context, highlight applications, and outline the transformative impact of freeform optics. The chapter emphasizes the need for innovative optical designs in the evolving landscape of space exploration and other industries. The second chapter provides insights into the mathematical descriptions, manufacturing processes, and metrology techniques involved in freeform optics. The third chapter covers the design and analysis of all-reflective unobscured TMA telescopes. In particular, we explore the fundamental characteristics of imaging systems, the design principles of TMA telescopes, the aberration theory, and the current design methods for freeform imaging systems as discussed in scientific literature. In the fourth chapter, we introduce a deep learning approach for designing freeform off-axis three-mirror telescopes. This method is needed to meet the growing demand for better imaging systems, as traditional designs have limitations. The framework uses supervised learning with a neural network to generate starting points for freeform telescope designs, reducing manual effort. The dataset is created through a system evolution method, and training involves normalizing input data, defining a multi-layer neural network, and

addressing underfitting and overfitting. ZEMAX OPTICSTUDIO and MATLAB are used for initial configuration and training, respectively. In the fifth chapter, we demonstrate the feasibility of the design framework by implementing a freeform version of the Wetherell-configuration reflective triplet, a freeform off-axis TMA telescope. The system parameters are constrained based on a case study, and a total of 2106 base systems are generated through a system evolution algorithm. The initial base system, **BaseSys₁**, achieves diffraction-limited performance. A dataset is generated, and a Feed-Forward Neural Network is trained to predict surface parameters based on system parameters. The FFNN achieved promising results, indicating its ability to predict systems with acceptable performance.

In future developments, this design framework holds the potential to advance through the exploration of deeper neural networks capable of accommodating broader system parameter ranges and various telescope configurations. Additionally, there is an opportunity to integrate unsupervised learning methods with deep neural networks, enhancing the framework's adaptability. This combination could empower the model to uncover inherent patterns independently, providing a more comprehensive understanding of the optical design space. Such advancements may facilitate the discovery of novel configurations and optimal solutions, making the framework more versatile and effective in optical system optimization.



Appendix

A.1 SYSTEM EVOLUTION ALGORITHM

```
1 function [ r ] = DatasetGeneration( args )
2
3 if ~exist('args', 'var')
4     args = [];
5 end
6
7 % Initialize the OpticStudio connection
8 TheApplication = InitConnection();
9 if isempty(TheApplication)
10     % failed to initialize a connection
11     r = [];
12 else
13     try
14         r = BeginApplication(TheApplication, args);
15         CleanupConnection(TheApplication);
16     catch err
17         CleanupConnection(TheApplication);
18         rethrow(err);
19     end
20 end
21 end
22
23 function [r] = BeginApplication(TheApplication, args)
24
```

A.1. SYSTEM EVOLUTION ALGORITHM

```
25 import ZOSAPI.*;
26
27 sampleDir = 'C:\Users\loren\Documents\Zemax\ZOS-API Projects\
    MATLABStandaloneApplication1\base systems\';
28
29 % weight vector
30 w = [0.25, 0.25, 0.05, 1]; % w = [Full-XFOV, Full-YFOV, ENPD, F#]
31
32 % parameters range
33 FullXFoV_range = [2, 8];
34 FullyFoV_range = [1, 3];
35 ENPD_range = [40, 100];
36 Fnum_range = [2, 4];
37
38 % number of samples per parameter
39 N_FullXFoV = 7;
40 N_FullyFoV = 3;
41 N_ENPD = 13;
42 N_Fnum = 9;
43
44 % freeform coefficients having even components of x
45 paramNumbers = [27 28 30 32 34 35 37 39];
46 nPar = length(paramNumbers); % nPar=8 coefficients correspond to 4th-
    term order
47
48 % initialize structure array that collects the base systems generated
    through system evolution
49 systems = struct('name', {}, 'SYSP', {}, 'SURP', {}, 'SpotSize', {}, 'dst'
    , {}, 'totr', {}, 'genf', {});
50 % name: name of the .zmx file
51 % SYSP: system parameters (Full-XFoV, Full-YFoV, ENPD, F#)
52 % SURP: surface parameters (surface locations and surface
    coefficients)
53 % SpotSize: array of spot sizes (one for each field)
54 % dst: maximum relative distortion (%)
55 % totr: total track length (mm)
56 % genf: encircled energy
57
58 % uniform sampling
59 FullXFoV_vec = linspace(FullXFoV_range(1), FullXFoV_range(2),
    N_FullXFoV);
60 FullyFoV_vec = linspace(FullyFoV_range(1), FullyFoV_range(2),
    N_FullyFoV);
```

```

61 ENPD_vec = linspace(ENPD_range(1), ENPD_range(2), N_ENPD);
62 Fnum_vec = linspace(Fnum_range(1), Fnum_range(2), N_Fnum);
63
64 % create the list of base systems to be optimized
65 notUsedSYSPlist = [];
66 for a=1:N_FullXFoV
67     for b=1:N_FullyFoV
68         for c=1:N_ENPD
69             for d=1:N_Fnum
70                 % keep only the systems having XFov > YFoV
71                 if(FullXFoV_vec(a)>=FullXFoV_vec(b))
72
73                     notUsedSYSPlist = [notUsedSYSPlist; round(
FullXFoV_vec(a),4), round(FullyFoV_vec(b),4), round(ENPD_vec(c),4),
round(Fnum_vec(d),4)];
74
75                 end
76             end
77         end
78     end
79 end
80
81 % load the 'systems' struct containing BaseSys1 or a list of
optimized base systems
82 load('InitialSystems.mat','systems')
83
84 % initialize the list of previously optimized base systems
85 optimizedSYSPlist = zeros(length(systems),4);
86
87 for i=1:length(systems)
88
89     fprintf('Load base system number %d/%i\n', i, length(systems));
90     optimizedSYSPlist(i,:) = systems(i).SYSP;
91     % remove all the previously optimized base systems contained by
systems
92     % struct
93     notUsedSYSPlist(ismember(notUsedSYSPlist,systems(i).SYSP,'rows')
,:) = []; % remove
94
95 end
96
97 for i=1:length(systems)
98

```

A.1. SYSTEM EVOLUTION ALGORITHM

```

99     % determine the weighted distance between the base systems to
      optimize and
100     % the previously optimized base systems
101     distance(:,i) = vecnorm(w.*(notUsedSYSPlist-optimizedSYSPlist(i
      ,:)),2,2);
102
103 end
104
105 fileList = dir(fullfile(sampleDir, '*.zmx'));
106
107 % generate the j-th system from the (j-1) previously optimized
      systems
108 j=1+length(systems);
109
110 while (j<=length(notUsedSYSPlist)+length(systems)+1)
111
112     % distance matrix: rows=notUsedSYSPlist columns=optimizedSYSPlist
113     [~,nearestSYSPindex] = min(min(distance,[],2));
114     % select the base system to optimize according to the distance
      matrix
115     nearestSYSP = notUsedSYSPlist(nearestSYSPindex,:);
116     % the chosen base system will evolve from one of these 2
      previously optimized base systems
117     [~,systemIndeces] = mink(distance(nearestSYSPindex,:),2);
118
119     i = 0; % select system index
120     % save systems temporarily
121     SaveSystem = struct('system',{},'SpotSize',{},'meanSpotSize',{},'
      SYSPindex',{},'dst',{},'totr',{},'genf',{});
122     WFNO = zeros(1,2); % F-number initialization
123     EFFL = zeros(1,2); % effective focal length initialization
124
125     % this cycle will only run once if the first system is
      diffraction-limited
126     while 1
127
128         i=i+1;
129
130         fprintf('Generate system number %d/%i version %d\n', j,
      length(notUsedSYSPlist)+length(systems), i);
131         % select the base system from which the new one should be
      evolving
132         systemIndex = systemIndeces(i);

```

```

133     initialSYSP = optimizedSYSPlist(systemIndex,:);
134
135     % select SYSP, SURP from systems structure array
136     if systems(systemIndex).SYSP == initialSYSP
137         SURP = systems(systemIndex).SURP;
138     else
139         HandleError('The initial SYSP arrays do not match!');
140     end
141
142     % extract global surface data
143     alpha = SURP(3*nPar+1);
144     Z3 = SURP(3*nPar+2);
145     Z5 = SURP(3*nPar+3);
146     Z7 = SURP(3*nPar+4);
147     radius3 = SURP(3*nPar+5);
148     conic3 = SURP(3*nPar+6);
149     radius4 = SURP(3*nPar+7);
150     conic4 = SURP(3*nPar+8);
151     radius5 = SURP(3*nPar+9);
152     conic5 = SURP(3*nPar+10);
153
154     % create baseSys_i, nearest SYSP, new SYSP
155     TheSystem = TheApplication.LoadNewSystem(strcat(sampleDir,
fileList(1).name));
156     TheSystemData = TheSystem.SystemData;
157     TheSystem.UpdateMode = ZOSAPI.LensUpdateMode.None;
158
159     % change Field
160     SysFields = TheSystemData.Fields;
161
162     field2 = SysFields.GetField(2);
163     field2.X = 0.0;
164     field2.Y = -nearestSYSP(2)/2;
165
166     field3 = SysFields.GetField(3);
167     field3.X = 0.0;
168     field3.Y = -nearestSYSP(2);
169
170     field4 = SysFields.GetField(4);
171     field4.X = nearestSYSP(1)/4;
172     field4.Y = 0.0;
173
174     field5 = SysFields.GetField(5);

```

A.1. SYSTEM EVOLUTION ALGORITHM

```
175     field5.X = nearestSYSP(1)/4;
176     field5.Y = -nearestSYSP(2)/2;
177
178     field6 = SysFields.GetField(6);
179     field6.X = nearestSYSP(1)/4;
180     field6.Y = -nearestSYSP(2);
181
182     field7 = SysFields.GetField(7);
183     field7.X = nearestSYSP(1)/2;
184     field7.Y = 0.0;
185
186     field8 = SysFields.GetField(8);
187     field8.X = nearestSYSP(1)/2;
188     field8.Y = -nearestSYSP(2)/2;
189
190     field9 = SysFields.GetField(9);
191     field9.X = nearestSYSP(1)/2;
192     field9.Y = -nearestSYSP(2);
193
194     % change Aperture
195     TheSystemData.Aperture.ApertureValue = nearestSYSP(3);
196
197     % assign values by converting from global to local
coordinates and
198     % make the solve variable for alpha, radii, conics and
thicknesses
199     Surface1 = TheSystem.LDE.GetSurfaceAt(1); % 'dummy' surface
200     Surface2 = TheSystem.LDE.GetSurfaceAt(2); % coordinate break
201     Surface3 = TheSystem.LDE.GetSurfaceAt(3); % M1
202     Surface4 = TheSystem.LDE.GetSurfaceAt(4); % M2
203     Surface5 = TheSystem.LDE.GetSurfaceAt(5); % M2
204
205     % surface 1
206     Surface1.Thickness = 130;
207     % surface 2
208     Surface2.GetCellAt(14).DoubleValue = -alpha;
209     Surface2.GetCellAt(14).MakeSolveVariable();
210     % Surface 3
211     Surface3.Radius = radius3;
212     Surface3.Conic = conic3;
213     Surface3.Thickness = -Z3;
214     Surface3.RadiusCell.MakeSolveVariable();
215     Surface3.ConicCell.MakeSolveVariable();
```

```

216     Surface3.ThicknessCell.MakeSolveVariable();
217     % Surface 4
218     Surface4.Radius = radius4;
219     Surface4.Conic = conic4;
220     Surface4.Thickness = Z5;
221     Surface4.RadiusCell.MakeSolveVariable();
222     Surface4.ConicCell.MakeSolveVariable();
223     Surface4.ThicknessCell.MakeSolveVariable();
224     % Surface 5
225     Surface5.Radius = radius5;
226     Surface5.Conic = conic5;
227     Surface5.Thickness = -Z5 + Z7;
228     Surface5.RadiusCell.MakeSolveVariable();
229     Surface5.ConicCell.MakeSolveVariable();
230     Surface5.ThicknessCell.MakeSolveVariable();
231
232     % assign surface coefficients
233     for k=1:1:nPar
234
235         Surface3.GetCellAt(paramNumbers(k)).DoubleValue = SURP(k)
;
236         Surface4.GetCellAt(paramNumbers(k)).DoubleValue = SURP(
nPar+k);
237         Surface5.GetCellAt(paramNumbers(k)).DoubleValue = SURP(2*
nPar+k);
238
239     end
240
241     % merit function definition
242     TheMFE = TheSystem.MFE;
243     OptWizard = TheMFE.SEQOptimizationWizard2;
244     % optimize for smallest RMS Spot
245     OptWizard.Criterion = ZOSAPI.Wizards.CriterionTypes.Spot;
246     OptWizard.OverallWeight = 1;
247     OptWizard.UseAllFields = true;
248     OptWizard.Reference = ZOSAPI.Wizards.ReferenceTypes.Centroid;
249     OptWizard.Arms = ZOSAPI.Wizards.PupilArmsCount.Arms_6;
250     OptWizard.Type = ZOSAPI.Wizards.OptimizationTypes.RMS;
251     OptWizard.UseGaussianQuadrature = true;
252     % gaussian quadrature with 3 rings
253     OptWizard.Rings = 3;
254     OptWizard.StartAt = 2;
255     OptWizard.AssumeAxialSymmetry = false;

```

A.1. SYSTEM EVOLUTION ALGORITHM

```
256     OptWizard.Apply();
257     mf_filename = 'C:\Users\loren\Documents\Zemax\MeritFunction\
RMS_Spot_Radius.mf';
258     TheMFE.SaveMeritFunction(mf_filename);
259     TheMFE.LoadMeritFunction(mf_filename);
260
261     % upload custom merit functions for TMA optimization and to
control structure constrains
262     TMAFile = 'C:\Users\loren\Documents\Zemax\MeritFunction\tma.
MF';
263     TheMFE.InsertMeritFunction(TMAFile,1);
264     % impose the new F-number (WFNO)
265     Operand_Fnum = TheMFE.InsertNewOperandAt(2);
266     Operand_Fnum.ChangeType(ZOSAPI.Editors.MFE.MeritOperandType.
WFNO);
267     Operand_Fnum.Target = nearestSYSP(4);
268     Operand_Fnum.Weight = 1.0;
269     % control the effective focal length (EFFL)
270     Operand_EFFL = TheMFE.InsertNewOperandAt(3);
271     Operand_EFFL.ChangeType(ZOSAPI.Editors.MFE.MeritOperandType.
EFFL);
272     Operand_EFFL.Target = 0;
273     Operand_EFFL.Weight = 0;
274
275     % spot diagram analysis
276     spot = TheSystem.Analyses.New_Analysis(ZOSAPI.Analysis.
AnalysisIDM.StandardSpot);
277     spot_setting = spot.GetSettings();
278     spot_setting.Field.SetFieldNumber(0);
279     spot_setting.Wavelength.SetWavelengthNumber(0);
280     spot_setting.ReferTo = ZOSAPI.Analysis.Settings.Spot.
Reference.Centroid;
281
282     % scale system if XFoV and YFoV of the 'initial' and 'nearest
' base systems are equal
283     if(all(initialSYSP(1:2)==nearestSYSP(1:2)))
284
285         ScaleFactor = nearestSYSP(3)/initialSYSP(3);
286         TheSystemData.Aperture.ApertureValue = initialSYSP(3);
287         ScaleSystem = TheSystem.Tools.OpenScale();
288         ScaleSystem.ScaleFactor = ScaleFactor;
289         ScaleSystem.ScaleByFactor = true;
290         ScaleSystem.RunAndWaitForCompletion();
```

```

291     ScaleSystem.Close();
292
293     Surface3.GetCellAt(25).DoubleValue = 1;
294     Surface4.GetCellAt(25).DoubleValue = 1;
295     Surface5.GetCellAt(25).DoubleValue = 1;
296
297     exp = [1,2,2,3,3,4,4,4];
298
299     % rescale the coefficients to maintain Norm Radius = 1
300     for k=1:1:nPar
301
302         Surface3.GetCellAt(paramNumbers(k)).DoubleValue =
Surface3.GetCellAt(paramNumbers(k)).DoubleValue*(1/ScaleFactor)^(
exp(k));
303         Surface4.GetCellAt(paramNumbers(k)).DoubleValue =
Surface4.GetCellAt(paramNumbers(k)).DoubleValue*(1/ScaleFactor)^(
exp(k));
304         Surface5.GetCellAt(paramNumbers(k)).DoubleValue =
Surface5.GetCellAt(paramNumbers(k)).DoubleValue*(1/ScaleFactor)^(
exp(k));
305
306     end
307
308 end
309
310 % optimize alpha, radii, conics, thicknesses and freeform
311 % coefficients up to 2nd-term order for 30 sec
312 for k=1:1:nPar-5
313     pause(0.1)
314     Surface3.GetCellAt(paramNumbers(k)).MakeSolveVariable();
315     Surface4.GetCellAt(paramNumbers(k)).MakeSolveVariable();
316     Surface5.GetCellAt(paramNumbers(k)).MakeSolveVariable();
317 end
318
319 tic;
320
321 LocalOptimTimeInSeconds = 10;
322 LocalOpt = TheSystem.Tools.OpenLocalOptimization();
323 pause(0.1)
324 if ~isempty(LocalOpt)
325     LocalOpt.Algorithm = ZOSAPI.Tools.Optimization.
OptimizationAlgorithm.DampedLeastSquares;
326     LocalOpt.Cycles = ZOSAPI.Tools.Optimization.

```

A.1. SYSTEM EVOLUTION ALGORITHM

```
OptimizationCycles.Automatic;
327     LocalOpt.NumberOfCores = 12;
328     fprintf('Local Optimization...\n');
329     pause(0.1)
330     fprintf('Initial Merit Function %8.6f\n', LocalOpt.
InitialMeritFunction);
331     LocalOpt.RunAndWaitWithTimeout(LocalOptimTimeInSeconds);
332     fprintf('Final Merit Function %8.6f\n', LocalOpt.
CurrentMeritFunction);
333     pause(0.1)
334
335     LocalOpt.Cancel();
336     pause(0.1)
337     LocalOpt.WaitForCompletion();
338     pause(0.1)
339     LocalOpt.Close();
340     pause(0.1)
341     end
342
343     toc;
344
345     % optimize alpha, radii, conics, thicknesses and freeform
346     % coefficients up to 3rd-term order for 40 sec
347     for k=1:1:nPar-3
348         pause(0.1)
349         Surface3.GetCellAt(paramNumbers(k)).MakeSolveVariable();
350         Surface4.GetCellAt(paramNumbers(k)).MakeSolveVariable();
351         Surface5.GetCellAt(paramNumbers(k)).MakeSolveVariable();
352     end
353
354     tic;
355
356     LocalOptimTimeInSeconds = 30;
357     LocalOpt = TheSystem.Tools.OpenLocalOptimization();
358     pause(0.1)
359     if ~isempty(LocalOpt)
360         LocalOpt.Algorithm = ZOSAPI.Tools.Optimization.
OptimizationAlgorithm.DampedLeastSquares;
361         LocalOpt.Cycles = ZOSAPI.Tools.Optimization.
OptimizationCycles.Automatic;
362         LocalOpt.NumberOfCores = 12;
363         fprintf('Local Optimization...\n');
364         pause(0.1)
```

```

365         fprintf('Initial Merit Function %8.6f\n', LocalOpt.
InitialMeritFunction);
366         LocalOpt.RunAndWaitWithTimeout(LocalOptimTimeInSeconds);
367
368         fprintf('Final Merit Function %8.6f\n', LocalOpt.
CurrentMeritFunction);
369         pause(0.1)
370
371         LocalOpt.Cancel();
372         pause(0.1)
373         LocalOpt.WaitForCompletion();
374         pause(0.1)
375
376         Fnum = TheMFE.GetOperandAt(2);
377         WFNO(i) = Fnum.Value;
378         EFFL(i) = TheMFE.GetOperandAt(3).Value;
379         SaveSystem(i).dst = TheMFE.GetOperandAt(4).Value;
380         SaveSystem(i).totr = TheMFE.GetOperandAt(6).Value;
381         SaveSystem(i).genf = zeros(1,7);
382         for h=7:1:13
383             SaveSystem(i).genf(1,h-6) =
384                 TheMFE.GetOperandAt(h).Value;
385         end
386
387         LocalOpt.Close();
388         pause(0.1)
389     end
390
391     toc;
392
393     % extract RMS spot size for field points
394     spot.ApplyAndWaitForCompletion();
395     spot_results = spot.GetResults();
396     SaveSystem(i).SpotSize = [spot_results.SpotData.
GetRMSSpotSizeFor(1,1), spot_results.SpotData.GetRMSSpotSizeFor
(2,1), spot_results.SpotData.GetRMSSpotSizeFor(3,1), spot_results.
SpotData.GetRMSSpotSizeFor(4,1), spot_results.SpotData.
GetRMSSpotSizeFor(5,1), spot_results.SpotData.GetRMSSpotSizeFor
(6,1), spot_results.SpotData.GetRMSSpotSizeFor(7,1), spot_results.
SpotData.GetRMSSpotSizeFor(8,1), spot_results.SpotData.
GetRMSSpotSizeFor(9,1)];
397
398     fprintf('Full-FOV: %2.1f x %2.1f, ENPD: %6.3f, F#: %6.4f\n',

```

A.1. SYSTEM EVOLUTION ALGORITHM

```

nearestSYSP(1), nearestSYSP(2), nearestSYSP(3), nearestSYSP(4))
399     fprintf('RMS radius: %6.3f %6.3f %6.3f %6.3f %6.3f %6.3f
%6.3f\n', SaveSystem(i).SpotSize(1), SaveSystem(i).SpotSize(2),
SaveSystem(i).SpotSize(3), SaveSystem(i).SpotSize(4), SaveSystem(i).
SpotSize(5), SaveSystem(i).SpotSize(6), SaveSystem(i).SpotSize(7))
400     TheSystem.Analyses.CloseAnalysis(spot);
401     pause(0.1)
402
403     SaveSystem(i).system = TheSystem.CopySystem();
404     SaveSystem(i).meanSpotSize = mean(SaveSystem(i).SpotSize);
405     SaveSystem(i).SYSPindex = nearestSYSPindex;
406     SaveSystem(i).DL = 1.22*nearestSYSP(4)*0.55; % diffraction
limit for 550nm
407
408     % if the system is diffraction-limited, F-number is correct
409     % and the focal length is positive then breaks the cycle
410     if (all(SaveSystem(i).SpotSize<=1) && abs(WFNO(i)-nearestSYSP
(4))<1e-2 && EFFL(i)>0)
411
412         TheSystem = SaveSystem(i).system.CopySystem();
413         nearestSYSPindex = SaveSystem(i).SYSPindex;
414         nearestSYSP = notUsedSYSPlist(nearestSYSPindex,:);
415         trueSystemIndex = i;
416
417         pause(0.1)
418         break;
419
420     % if both initial systems are used, optimize one last time
and the cycle ends
421     elseif(i>=length(systemIndeces))
422
423         [~,trueSystemIndex] = min([SaveSystem.meanSpotSize]./[
SaveSystem.DL]);
424         TheSystem = SaveSystem(trueSystemIndex).system.
CopySystem();
425         nearestSYSPindex = SaveSystem(trueSystemIndex).SYSPindex;
426         nearestSYSP = notUsedSYSPlist(nearestSYSPindex,:);
427
428
429     % optimize alpha, radii, conics, thicknesses and freeform
430     % coefficients up to 4th-term order for 60 sec
431     for k=1:1:nPar
432
433         pause(0.1)

```

```

434     Surface3.GetCellAt(paramNumbers(k)).
435     MakeSolveVariable();
436     Surface4.GetCellAt(paramNumbers(k)).
437     MakeSolveVariable();
438     Surface5.GetCellAt(paramNumbers(k)).
439     MakeSolveVariable();
440
441     end
442
443     tic;
444     pause(0.1)
445     LocalOptimTimeInSeconds = 60;
446     LocalOpt = TheSystem.Tools.OpenLocalOptimization();
447     pause(0.2)
448
449     if ~isempty(LocalOpt)
450         LocalOpt.Algorithm = ZOSAPI.Tools.Optimization.
OptimizationAlgorithm.DampedLeastSquares;
451         LocalOpt.Cycles = ZOSAPI.Tools.Optimization.
OptimizationCycles.Automatic;
452         LocalOpt.NumberOfCores = 12;
453         fprintf('Local Optimization...\n');
454         pause(0.1)
455         fprintf('Initial Merit Function %8.6f\n', LocalOpt.
InitialMeritFunction);
456         LocalOpt.RunAndWaitWithTimeout(
LocalOptimTimeInSeconds);
457         fprintf('Final Merit Function %8.6f\n', LocalOpt.
CurrentMeritFunction);
458         pause(0.1)
459         LocalOpt.Cancel();
460         pause(0.1)
461         LocalOpt.WaitForCompletion();
462         pause(0.2)
463
464         Fnum = TheMFE.GetOperandAt(2);
465         WFNO(i) = Fnum.Value;
466         EFFL(i) = TheMFE.GetOperandAt(3).Value;
467         SaveSystem(i).dst = TheMFE.GetOperandAt(4).Value;
468         SaveSystem(i).totr = TheMFE.GetOperandAt(6).Value;
469
470         LocalOpt.Close();
471         pause(0.1)

```

A.1. SYSTEM EVOLUTION ALGORITHM

```
472     end
473     toc;
474
475     spot = TheSystem.Analyses.New_Analysis(ZOSAPI.Analysis.
AnalysisIDM.StandardSpot);
476     pause(0.1)
477     spot_setting = spot.GetSettings();
478     spot_setting.Field.SetFieldNumber(0);
479     spot_setting.Wavelength.SetWavelengthNumber(0);
480     spot_setting.ReferTo = ZOSAPI.Analysis.Settings.Spot.
Reference.Centroid;
481     spot.ApplyAndWaitForCompletion();
482     spot_results = spot.GetResults();
483     SaveSystem(trueSystemIndex).SpotSize = [spot_results.
SpotData.GetRMSSpotSizeFor(1,1), spot_results.SpotData.
GetRMSSpotSizeFor(2,1), spot_results.SpotData.GetRMSSpotSizeFor
(3,1), spot_results.SpotData.GetRMSSpotSizeFor(4,1), spot_results.
SpotData.GetRMSSpotSizeFor(5,1), spot_results.SpotData.
GetRMSSpotSizeFor(6,1), spot_results.SpotData.GetRMSSpotSizeFor
(7,1), spot_results.SpotData.GetRMSSpotSizeFor(8,1), spot_results.
SpotData.GetRMSSpotSizeFor(9,1)];
484
485     fprintf('Full-FOV: %2.1f x %2.1f, ENPD: %6.3f, F#: %6.4f\
n', nearestSYSP(1), nearestSYSP(2), nearestSYSP(3), nearestSYSP(4)
)
486     fprintf('RMS radius: %6.3f %6.3f %6.3f %6.3f %6.3f %6.3
f %6.3f\n', SaveSystem(trueSystemIndex).SpotSize(1), SaveSystem(
trueSystemIndex).SpotSize(2), SaveSystem(trueSystemIndex).SpotSize
(3), SaveSystem(trueSystemIndex).SpotSize(4), SaveSystem(
trueSystemIndex).SpotSize(5), SaveSystem(trueSystemIndex).SpotSize
(6), SaveSystem(trueSystemIndex).SpotSize(7))
487
488     pause(0.1)
489     break;
490
491     end
492
493     pause(0.1)
494
495     end
496
497     % remove the base system from the list of base systems to
optimize
```

```

498     notUsedSYSPlist(nearestSYSPindex,:) = [];
499     distance(nearestSYSPindex,:) = [];
500
501     % assign SURP values
502     SURP = zeros(3*nPar+10,1);
503     % get global surface locations for M1, M3 and IMG
504     [~,~,~,~,~,~, R23,~,~, R33,~,~,~] = TheSystem.LDE.
GetGlobalMatrix(1);
505     SURP(3*nPar+1,1) = atan2d(-R23,R33);
506     [~,~,~,~,~,~,~,~,~,~,~,~,~, Z0] = TheSystem.LDE.
GetGlobalMatrix(3);
507     SURP(3*nPar+2,1) = Z0;
508     [~,~,~,~,~,~,~,~,~,~,~,~,~, Z0] = TheSystem.LDE.
GetGlobalMatrix(5);
509     SURP(3*nPar+3,1) = Z0;
510     [~,~,~,~,~,~,~,~,~,~,~,~,~, Z0] = TheSystem.LDE.
GetGlobalMatrix(7);
511     SURP(3*nPar+4,1) = Z0;
512     % get the surface coefficients
513     SURP(3*nPar+5,1) = TheSystem.LDE.GetSurfaceAt(3).Radius;
514     SURP(3*nPar+6,1) = TheSystem.LDE.GetSurfaceAt(3).Conic;
515     SURP(3*nPar+7,1) = TheSystem.LDE.GetSurfaceAt(4).Radius;
516     SURP(3*nPar+8,1) = TheSystem.LDE.GetSurfaceAt(4).Conic;
517     SURP(3*nPar+9,1) = TheSystem.LDE.GetSurfaceAt(5).Radius;
518     SURP(3*nPar+10,1) = TheSystem.LDE.GetSurfaceAt(5).Conic;
519
520     for k=1:1:nPar
521
522         SURP(k,1) = TheSystem.LDE.GetSurfaceAt(3).GetCellAt(
paramNumbers(k)).DoubleValue();
523         SURP(nPar+k,1) = TheSystem.LDE.GetSurfaceAt(4).GetCellAt(
paramNumbers(k)).DoubleValue();
524         SURP(2*nPar+k,1) = TheSystem.LDE.GetSurfaceAt(5).GetCellAt(
paramNumbers(k)).DoubleValue();
525
526     end
527
528     % if the F-number is correct and the focal length is positive
then saves the base system
529     if(abs(WFNO(trueSystemIndex)-nearestSYSP(4))<1e-2 && EFFL(
trueSystemIndex)>0)
530
531         % insert the new optimized base system and evaluate the

```

A.1. SYSTEM EVOLUTION ALGORITHM

```
weighted distances
532     distance = [distance, vecnorm(w.*(notUsedSYSPlist-nearestSYSP
),2,2)];
533     optimizedSYSPlist = [optimizedSYSPlist; nearestSYSP];
534
535     % save the base system data in 'systems' struct
536     systems(j).SURP = SURP';
537     clear SURP;
538     systems(j).name = strcat(num2str(nearestSYSP(1)),'$',num2str(
nearestSYSP(2)),'$',num2str(nearestSYSP(3)),'$',num2str(
nearestSYSP(4)),'$.zmx');
539     systems(j).SYSP = nearestSYSP;
540     systems(j).SpotSize = SaveSystem(trueSystemIndex).SpotSize;
541     systems(j).dst = SaveSystem(trueSystemIndex).dst;
542     systems(j).totr = SaveSystem(trueSystemIndex).totr;
543     systems(j).genf = SaveSystem(trueSystemIndex).genf;
544     clear nearestSYSP;
545     clear SaveSystem;
546
547     save('InitialSystems.mat','systems')
548     j=j+1;
549
550     end
551
552     TheSystem.Close(false);
553     pause(0.1)
554 end
555
556 r = [];
557
558 end
559
560 function app = InitConnection()
561
562 import System.Reflection.*;
563
564 % Find the installed version of OpticStudio.
565 zemaxData = winqueryreg('HKEY_CURRENT_USER', 'Software\Zemax', '
ZemaxRoot');
566 NetHelper = strcat(zemaxData, '\ZOS-API\Libraries\ZOSAPI_NetHelper.
dll');
567 % Note -- uncomment the following line to use a custom NetHelper path
568 % NetHelper = 'C:\Users\loren\Documents\Zemax\ZOS-API\Libraries\
```

```

    ZOSAPI_NetHelper.dll';
569 % This is the path to OpticStudio
570 NET.addAssembly(NetHelper);
571
572 success = ZOSAPI_NetHelper.ZOSAPI_Initializer.Initialize();
573 % Note -- uncomment the following line to use a custom initialization
    path
574 % success = ZOSAPI_NetHelper.ZOSAPI_Initializer.Initialize('C:\
    Program Files\OpticStudio\');
575 if success == 1
576     LogMessage(strcat('Found OpticStudio at: ', char(ZOSAPI_NetHelper
        .ZOSAPI_Initializer.GetZemaxDirectory())));
577 else
578     app = [];
579     return;
580 end
581
582 % Now load the ZOS-API assemblies
583 NET.addAssembly(AssemblyName('ZOSAPI_Interfaces'));
584 NET.addAssembly(AssemblyName('ZOSAPI'));
585
586 % Create the initial connection class
587 TheConnection = ZOSAPI.ZOSAPI_Connection();
588
589 % Attempt to create a Standalone connection
590
591 % NOTE - if this fails with a message like 'Unable to load one or
    more of
592 % the requested types', it is usually caused by try to connect to a
    32-bit
593 % version of OpticStudio from a 64-bit version of MATLAB (or vice-
    versa).
594 % This is an issue with how MATLAB interfaces with .NET, and the only
595 % current workaround is to use 32- or 64-bit versions of both
    applications.
596 app = TheConnection.CreateNewApplication();
597 if isempty(app)
598     HandleError('An unknown connection error occurred!');
599 end
600 if ~app.IsValidLicenseForAPI
601     HandleError('License check failed!');
602     app = [];
603 end

```

A.2. NEURAL NETWORK TRAINING CODE

```
604
605 end
606
607 function LogMessage(msg)
608 disp(msg);
609 end
610
611 function HandleError(error)
612 ME = MException('zosapi:HandleError', error);
613 throw(ME);
614 end
615
616 function CleanupConnection(TheApplication)
617 % Note - this will close down the connection.
618
619 % If you want to keep the application open, you should skip this step
620 % and store the instance somewhere instead.
621 TheApplication.CloseApplication();
622 end
```

A.2 NEURAL NETWORK TRAINING CODE

```
1 % solve an Input-Output Fitting problem with a Neural Network
2 %
3 %   inputs - input data.
4 %   targets - target data.
5
6 x = inputs';
7 t = targets';
8
9 % create a Fitting Network
10 % takes a row vector of N hidden layer sizes, and a backpropagation
    training function, and returns a feed-forward neural network with
    N+1 layers.
11 hiddenLayerSize = [30 40];
12 % choose a Training Function
13 trainFcn = 'trainlm'; % Levenberg-Marquardt backpropagation.
14
15 net = fitnet(hiddenLayerSize, trainFcn);
16
17 % choose Input and Output Pre/Post-Processing Functions
```

```

18 net.input.processFcns = {'removeconstantrows','mapminmax'};
19 net.output.processFcns = {'removeconstantrows','mapminmax'};
20
21 % setup Division of Data for Training, Validation, Testing
22 net.divideFcn = 'dividerand'; % divide data randomly
23 net.divideMode = 'sample'; % divide up every sample
24 net.divideParam.trainRatio = 70/100;
25 net.divideParam.valRatio = 15/100;
26 net.divideParam.testRatio = 15/100;
27
28 % choose a Performance Function
29 net.performFcn = 'mse'; % MSE - Mean Squared Error
30
31 % choose Plot Functions
32 net.plotFcns = {'plotperform','plottrainstate','ploterrhist',
                 'plotregression','plotfit'};
33
34 % train the Network
35 [net,tr] = train(net,x,t);
36
37 % test the Network
38 y = net(x);
39 e = gsubtract(t,y);
40 performance = perform(net,t,y);
41
42 % recalculate Training, Validation and Test Performance
43 trainTargets = t .* tr.trainMask{1};
44 valTargets = t .* tr.valMask{1};
45 testTargets = t .* tr.testMask{1};
46 trainPerformance = perform(net,trainTargets,y);
47 valPerformance = perform(net,valTargets,y);
48 testPerformance = perform(net,testTargets,y);
49
50 % view the Network
51 view(net)
52
53 % plots
54 figure, plotperform(tr)
55 figure, plottrainstate(tr)
56 figure, ploterrhist(e)
57 figure, plotregression(t,y)
58 figure, plotfit(net,x,t)

```


References

- [1] Jannick P. Rolland et al. "Freeform optics for imaging". In: *Optica* 8.2 (Feb. 2021), pp. 161–176. DOI: 10.1364/OPTICA.413762. URL: <https://opg.optica.org/optica/abstract.cfm?URI=optica-8-2-161>.
- [2] X. Jiang, P. Scott, and D. Whitehouse. "Freeform Surface Characterisation - A Fresh Strategy". In: *CIRP Annals* 56.1 (2007), pp. 553–556. ISSN: 0007-8506. DOI: <https://doi.org/10.1016/j.cirp.2007.05.132>. URL: <https://www.sciencedirect.com/science/article/pii/S0007850607001333>.
- [3] Guohao Ju, Hongcai Ma, and Changxiang Yan. "Aberration fields of off-axis astronomical telescopes induced by rotational misalignments". In: *Opt. Express* 26.19 (Sept. 2018), pp. 24816–24834. DOI: 10.1364/OE.26.024816. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-19-24816>.
- [4] Haodong Shi et al. "Analysis of nodal aberration properties in off-axis freeform system design". In: *Appl. Opt.* 55.24 (Aug. 2016), pp. 6782–6790. DOI: 10.1364/AO.55.006782. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-55-24-6782>.
- [5] "Two-element variable-power spherical lens". US3305294A. Luis W Alvarez. Feb. 21, 1967.
- [6] William T. Plummer. "Free-form optical components in some early commercial products". In: *Tribute to Warren Smith: A Legacy in Lens Design and Optical Engineering*. Ed. by Robert E. Fischer. Vol. 5865. International Society for Optics and Photonics. SPIE, 2005, p. 586509. DOI: 10.1117/12.623875. URL: <https://doi.org/10.1117/12.623875>.
- [7] "Catoptric anastigmatic afocal optical system". US3674334A. Abe Offner. July 4, 1972.

REFERENCES

- [8] “Three mirror anastigmatic optical system”. US4265510A. Lacy G. Cook. May 5, 1981.
- [9] “All-reflective three element objective”. US4240707A. William Wetherell and David Womble. Dec. 23, 1980.
- [10] Berge Tatian. “Aberration balancing in rotationally symmetric lenses*”. In: *J. Opt. Soc. Am.* 64.8 (Aug. 1974), pp. 1083–1091. DOI: 10.1364/JOSA.64.001083. URL: <https://opg.optica.org/abstract.cfm?URI=josa-64-8-1083>.
- [11] David Shafer. “Design With Two-Axis Aspheric Surfaces”. In: *Computer-Aided Optical Design*. Ed. by Robert E. Fischer. Vol. 0147. International Society for Optics and Photonics. SPIE, 1978, pp. 171–176. DOI: 10.1117/12.956638. URL: <https://doi.org/10.1117/12.956638>.
- [12] H. Philip Stahl et al. “Survey of interferometric techniques used to test JWST optical components”. In: *Interferometry XV: Techniques and Analysis*. Ed. by Catherine E. Towers, Joanna Schmit, and Katherine Creath. Vol. 7790. International Society for Optics and Photonics. SPIE, 2010, p. 779002. DOI: 10.1117/12.862234. URL: <https://doi.org/10.1117/12.862234>.
- [13] Joseph A. Connelly et al. “Imaging performance and modeling of the infrared multi-object spectrometer focal reducer”. In: *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*. Ed. by Masanori Iye and Alan F. M. Moorwood. Vol. 4841. International Society for Optics and Photonics. SPIE, 2003, pp. 702–714. DOI: 10.1117/12.458974. URL: <https://doi.org/10.1117/12.458974>.
- [14] Eli Atad-Ettedgui et al. “Opto-mechanical design of SCUBA-2”. In: *Optomechanical Technologies for Astronomy*. Ed. by Eli Atad-Ettedgui, Joseph Antebi, and Dietrich Lemke. Vol. 6273. International Society for Optics and Photonics. SPIE, 2006, 62732H. DOI: 10.1117/12.669231. URL: <https://doi.org/10.1117/12.669231>.
- [15] Jon Ranson. *MiniSpec: Miniaturized Imaging Spectrometer to Measure Vegetation Structure and Function*. Accessed: 16-11-2023.

- [16] J.P. Veefkind et al. "TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications". In: *Remote Sensing of Environment* 120 (2012). The Sentinel Missions - New Opportunities for Science, pp. 70–83. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2011.09.027>. URL: <https://www.sciencedirect.com/science/article/pii/S0034425712000661>.
- [17] Jess Köhler et al. "Sentinel 5 instrument and UV1 spectrometer subsystem optical design and development". In: *International Conference on Space Optics — ICSO 2020*. Ed. by Bruno Cugny, Zoran Sodnik, and Nikos Karafolas. Vol. 11852. International Society for Optics and Photonics. SPIE, 2021, p. 118522M. DOI: [10.1117/12.2599430](https://doi.org/10.1117/12.2599430). URL: <https://doi.org/10.1117/12.2599430>.
- [18] William H. Swartz. *CHAPS-D: The Compact Hyperspectral Air Pollution Sensor–Demonstrator*. Accessed: 16-11-2023.
- [19] Jerome Caron and Stefan Bäumer. "Progress in freeform mirror design for space applications". In: *International Conference on Space Optics — ICSO 2020*. Ed. by Bruno Cugny, Zoran Sodnik, and Nikos Karafolas. Vol. 11852. International Society for Optics and Photonics. SPIE, 2021, 118521S. DOI: [10.1117/12.2599322](https://doi.org/10.1117/12.2599322). URL: <https://doi.org/10.1117/12.2599322>.
- [20] T. Peschel et al. "Design of an imaging spectrometer for earth observation using freeform mirrors". In: *International Conference on Space Optics — ICSO 2016*. Ed. by Bruno Cugny, Nikos Karafolas, and Zoran Sodnik. Vol. 10562. International Society for Optics and Photonics. SPIE, 2017, p. 1056237. DOI: [10.1117/12.2296156](https://doi.org/10.1117/12.2296156). URL: <https://doi.org/10.1117/12.2296156>.
- [21] Benqi Zhang et al. "Simultaneous improvement of field-of-view and resolution in an imaging optical system". In: *Opt. Express* 29.6 (Mar. 2021), pp. 9346–9362. DOI: [10.1364/OE.420222](https://doi.org/10.1364/OE.420222). URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-29-6-9346>.
- [22] Jun Zhu et al. "Design of a low F-number freeform off-axis three-mirror system with rectangular field-of-view". In: *Journal of Optics* 17.1 (Dec. 2014), p. 015605. DOI: [10.1088/2040-8978/17/1/015605](https://doi.org/10.1088/2040-8978/17/1/015605). URL: <https://dx.doi.org/10.1088/2040-8978/17/1/015605>.

REFERENCES

- [23] Roland Geyl et al. “Freeform optics design, fabrication and testing technologies for Space applications”. In: *International Conference on Space Optics — ICSO 2018*. Ed. by Zoran Sodnik, Nikos Karafolas, and Bruno Cugny. Vol. 11180. International Society for Optics and Photonics. SPIE, 2019, 111800P. DOI: 10.1117/12.2535944. URL: <https://doi.org/10.1117/12.2535944>.
- [24] Weichen Wu, Guofan Jin, and Jun Zhu. “Optical design of the freeform reflective imaging system with wide rectangular FOV and low F-number”. In: *Results in Physics* 15 (2019), p. 102688. ISSN: 2211-3797. DOI: <https://doi.org/10.1016/j.rinp.2019.102688>. URL: <https://www.sciencedirect.com/science/article/pii/S2211379719323551>.
- [25] Anika Broemel, Uwe Lippmann, and Herbert Gross. In: *Advanced Optical Technologies* 6.5 (2017), pp. 327–336. DOI: doi:10.1515/aot-2017-0030. URL: <https://doi.org/10.1515/aot-2017-0030>.
- [26] H. Gross et al. “Overview on surface representations for freeform surfaces”. In: *Optical Systems Design 2015: Optical Design and Engineering VI*. Ed. by Laurent Mazuray, Rolf Wartmann, and Andrew P. Wood. Vol. 9626. International Society for Optics and Photonics. SPIE, 2015, 96260U. DOI: 10.1117/12.2191255. URL: <https://doi.org/10.1117/12.2191255>.
- [27] Jingfei Ye et al. “Review of optical freeform surface representation technique and its application”. In: *Optical Engineering* 56.11 (2017), p. 110901. DOI: 10.1117/1.OE.56.11.110901. URL: <https://doi.org/10.1117/1.OE.56.11.110901>.
- [28] Nick Takaki, Aaron Bauer, and Jannick P. Rolland. “Degeneracy in freeform surfaces described with orthogonal polynomials”. In: *Appl. Opt.* 57.35 (Dec. 2018), pp. 10348–10354. DOI: 10.1364/AO.57.010348. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-57-35-10348>.
- [29] Vasudevan Lakshminarayanan and Andre Fleck. “Zernike polynomials: a guide”. In: *Journal of Modern Optics* 58.18 (2011), pp. 1678–1678. DOI: 10.1080/09500340.2011.633763. eprint: <https://doi.org/10.1080/09500340.2011.633763>. URL: <https://doi.org/10.1080/09500340.2011.633763>.

- [30] G. W. Forbes. “Shape specification for axially symmetric optical surfaces”. In: *Opt. Express* 15.8 (Apr. 2007), pp. 5218–5226. DOI: 10.1364/OE.15.005218. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-15-8-5218>.
- [31] G. W. Forbes. “Robust and fast computation for the polynomials of optics”. In: *Opt. Express* 18.13 (June 2010), pp. 13851–13862. DOI: 10.1364/OE.18.013851. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-18-13-13851>.
- [32] G.W. Forbes. “Robust, efficient computational methods for axially symmetric optical aspheres”. In: *Opt. Express* 18.19 (Sept. 2010), pp. 19700–19712. DOI: 10.1364/OE.18.019700. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-18-19-19700>.
- [33] G.W. Forbes. “Manufacturability estimates for optical aspheres”. In: *Opt. Express* 19.10 (May 2011), pp. 9923–9942. DOI: 10.1364/OE.19.009923. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-19-10-9923>.
- [34] G. W. Forbes. “Characterizing the shape of freeform optics”. In: *Opt. Express* 20.3 (Jan. 2012), pp. 2483–2499. DOI: 10.1364/OE.20.002483. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-20-3-2483>.
- [35] Anika Brömel. “Development and evaluation of freeform surface descriptions”. en. Dissertation, Friedrich-Schiller-Universität Jena, 2018. PhD thesis. Jena, 2018. DOI: 10.22032/dbt.38072. URL: <https://doi.org/10.22032/dbt.38072>.
- [36] Michaela Bray. “Orthogonal polynomials: a set for square areas”. In: *SPIE Optical Systems Design*. 2004. URL: <https://api.semanticscholar.org/CorpusID:120141642>.
- [37] Virendra N. Mahajan and Guang-ming Dai. “Orthonormal polynomials in wavefront analysis: analytical solution”. In: *J. Opt. Soc. Am. A* 24.9 (Sept. 2007), pp. 2994–3016. DOI: 10.1364/JOSAA.24.002994. URL: <https://opg.optica.org/josaa/abstract.cfm?URI=josaa-24-9-2994>.
- [38] Peter Ott. “Optic design of head-up displays with freeform surfaces specified by NURBS”. In: *Optical Design and Engineering III*. Ed. by Laurent Mazuray et al. Vol. 7100. International Society for Optics and Photonics. SPIE, 2008, 71000Y. DOI: 10.1117/12.797761. URL: <https://doi.org/10.1117/12.797761>.

REFERENCES

- [39] Michael P. Chrisp. "Three Mirror Anastigmat Designed with NURBS Freeform Surfaces". In: *Renewable Energy and the Environment: Postdeadline Papers*. Optica Publishing Group, 2013, FM4B.3. DOI: 10.1364/FREEFORM.2013.FM4B.3. URL: <https://opg.optica.org/abstract.cfm?URI=Freeform-2013-FM4B.3>.
- [40] Kainian Tong et al. "Model of radial basis functions based on surface slope for optical freeform surfaces". In: *Opt. Express* 26.11 (May 2018), pp. 14010–14023. DOI: 10.1364/OE.26.014010. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-11-14010>.
- [41] Jingfei Ye et al. "Modal wavefront reconstruction over general shaped aperture by numerical orthogonal polynomials". In: *Optical Engineering* 54.3 (2015), p. 034105. DOI: 10.1117/1.OE.54.3.034105. URL: <https://doi.org/10.1117/1.OE.54.3.034105>.
- [42] Sumit Kumar, Zhen Tong, and Xiangqian Jiang. "Advances in the design and manufacturing of novel freeform optics". In: *International Journal of Extreme Manufacturing* 4.3 (July 2022), p. 032004. DOI: 10.1088/2631-7990/ac7617. URL: <https://dx.doi.org/10.1088/2631-7990/ac7617>.
- [43] F.Z. Fang et al. "Manufacturing and measurement of freeform optics". In: *CIRP Annals* 62.2 (2013), pp. 823–846. ISSN: 0007-8506. DOI: <https://doi.org/10.1016/j.cirp.2013.05.003>. URL: <https://www.sciencedirect.com/science/article/pii/S0007850613001935>.
- [44] Johannes Stock et al. "Description and reimplementation of real freeform surfaces". In: *Appl. Opt.* 56.3 (Jan. 2017), pp. 391–396. DOI: 10.1364/AO.56.000391. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-56-3-391>.
- [45] Tom Pertermann et al. "Angular resolved power spectral density analysis for improving mirror manufacturing". In: *Appl. Opt.* 57.29 (Oct. 2018), pp. 8692–8698. DOI: 10.1364/AO.57.008692. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-57-29-8692>.
- [46] Danyun Cai. "Development of freeform optical systems". en. PhD thesis. Jena, 2022. URL: <http://uri.gbv.de/document/gvk:ppn:1822254582>.
- [47] Christina Reynolds and David Walker. "Pseudo-Random Tool Paths for CNC Sub-Aperture Polishing and Other Applications". In: *Optics express* 16 (Dec. 2008), pp. 18942–9. DOI: 10.1364/OE.16.018942.

- [48] Zahra Hosseinimakarem, Angela D. Davies, and Chris J. Evans. "Zernike polynomials for mid-spatial frequency representation on optical surfaces". In: *Reflection, Scattering, and Diffraction from Surfaces V*. Ed. by Leonard M. Hanssen. Vol. 9961. International Society for Optics and Photonics. SPIE, 2016, 99610P. DOI: 10.1117/12.2236520. URL: <https://doi.org/10.1117/12.2236520>.
- [49] G. W. Forbes. "Fitting freeform shapes with orthogonal bases". In: *Opt. Express* 21.16 (Aug. 2013), pp. 19061–19081. DOI: 10.1364/OE.21.019061. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-21-16-19061>.
- [50] İlhan Kaya and Jannick P. Rolland. In: *Advanced Optical Technologies 2.1* (2013), pp. 81–88. DOI: doi:10.1515/aot-2012-0075. URL: <https://doi.org/10.1515/aot-2012-0075>.
- [51] F.L. Pedrotti, L.M. Pedrotti, and L.S. Pedrotti. *Introduction to Optics*. Cambridge University Press, 2017. ISBN: 9781108428262. URL: https://books.google.it/books?id=06o_DwAAQBAJ.
- [52] E. Hecht. *Optics*. Pearson, 2012. ISBN: 9788131718070. URL: <https://books.google.it/books?id=wcMWPBMMzIkC>.
- [53] Dietrich Korsch. "CHAPTER NINE - THIRD-ORDER CORRECTION OF THREE-MIRROR SYSTEMS". In: *Reflective Optics*. Ed. by Dietrich Korsch. San Diego: Academic Press, 1991, pp. 207–259. ISBN: 978-0-12-421170-4. DOI: <https://doi.org/10.1016/B978-0-12-421170-4.50013-3>. URL: <https://www.sciencedirect.com/science/article/pii/B9780124211704500133>.
- [54] R. Barry Johnson. "Historical perspective on understanding optical aberrations". In: *Lens Design: A Critical Review*. Ed. by Warren J. Smith. Vol. 10263. International Society for Optics and Photonics. SPIE, 1992, p. 1026303. DOI: 10.1117/12.131966. URL: <https://doi.org/10.1117/12.131966>.
- [55] Dietrich Korsch. "CHAPTER FIVE - THIRD-ORDER OPTICS". In: *Reflective Optics*. Ed. by Dietrich Korsch. San Diego: Academic Press, 1991, pp. 85–104. ISBN: 978-0-12-421170-4. DOI: <https://doi.org/10.1016/B978-0-12-421170-4.50009-1>. URL: <https://www.sciencedirect.com/science/article/pii/B9780124211704500091>.

REFERENCES

- [56] K.K. Sharma. *Optics: Principles and Applications*. Academic Press, 2006. ISBN: 9780123706119. URL: <https://books.google.it/books?id=RHHslAEACAAJ>.
- [57] Kevin Thompson. "Description of the third-order optical aberrations of near-circular pupil optical systems without symmetry". In: *J. Opt. Soc. Am. A* 22.7 (July 2005), pp. 1389–1401. DOI: 10.1364/JOSAA.22.001389. URL: <https://opg.optica.org/josaa/abstract.cfm?URI=josaa-22-7-1389>.
- [58] Roland V. Shack and Kevin Thompson. "Influence Of Alignment Errors Of A Telescope System On Its Aberration Field". In: *Optical Alignment I*. Ed. by Richard N. Shagam and William C. Sweatt. Vol. 0251. International Society for Optics and Photonics. SPIE, 1980, pp. 146–153. DOI: 10.1117/12.959464. URL: <https://doi.org/10.1117/12.959464>.
- [59] Kyle Fuerschbach, Jannick P. Rolland, and Kevin P. Thompson. "Extending Nodal Aberration Theory to include mount-induced aberrations with application to freeform surfaces". In: *Opt. Express* 20.18 (Aug. 2012), pp. 20139–20155. DOI: 10.1364/OE.20.020139. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-20-18-20139>.
- [60] Lori B. Moore, Anastacia M. Hvisc, and Jose Sasian. "Aberration fields of a combination of plane symmetric systems". In: *Opt. Express* 16.20 (Sept. 2008), pp. 15655–15670. DOI: 10.1364/OE.16.015655. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-16-20-15655>.
- [61] A.K. Ghatak. *Optics*. Tata McGraw-Hill Publishing Company Limited, 2009. ISBN: 9780070262157. URL: <https://books.google.it/books?id=fZSVXt0e8WAC>.
- [62] John W. Figoski. "Aberration Characteristics Of Nonsymmetric Systems". In: *1985 International Lens Design Conference*. Ed. by Duncan T. Moore and William H. Taylor. Vol. 0554. International Society for Optics and Photonics. SPIE, 1986, pp. 104–111. DOI: 10.1117/12.949203. URL: <https://doi.org/10.1117/12.949203>.
- [63] Kyle Fuerschbach, Jannick P. Rolland, and Kevin P. Thompson. "A new family of optical systems employing ϕ -polynomial surfaces". In: *Opt. Express* 19.22 (Oct. 2011), pp. 21919–21928. DOI: 10.1364/OE.19.021919. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-19-22-21919>.

- [64] Aaron Bauer, Eric M. Schiesser, and Jannick P. Rolland. "Starting geometry creation and design method for freeform optics". In: *Nature Communications* 9.1 (May 2018), p. 1756. ISSN: 2041-1723. DOI: 10.1038/s41467-018-04186-9. URL: <https://doi.org/10.1038/s41467-018-04186-9>.
- [65] G D Wassermann and E Wolf. "On the Theory of Aplanatic Aspheric Systems". In: *Proceedings of the Physical Society. Section B* 62.1 (Jan. 1949), p. 2. DOI: 10.1088/0370-1301/62/1/302. URL: <https://dx.doi.org/10.1088/0370-1301/62/1/302>.
- [66] David James Knapp. *Conformal optical design*. 2002. URL: <http://hdl.handle.net/10150/289852>.
- [67] R. Andrew Hicks. "Direct methods for freeform surface design". In: *Novel Optical Systems Design and Optimization X*. Ed. by R. John Koshel and G. Groot Gregory. Vol. 6668. International Society for Optics and Photonics. SPIE, 2007, p. 666802. DOI: 10.1117/12.731453. URL: <https://doi.org/10.1117/12.731453>.
- [68] Dewen Cheng, Yongtian Wang, and Hong Hua. "Free form optical system design with differential equations". In: *Optical Design and Testing IV*. Ed. by Yongtian Wang et al. Vol. 7849. International Society for Optics and Photonics. SPIE, 2010, 78490Q. DOI: 10.1117/12.869690. URL: <https://doi.org/10.1117/12.869690>.
- [69] Juan C. Miñano, Pablo Benítez, and Juan C. González. "RX: a nonimaging concentrator". In: *Appl. Opt.* 34.13 (May 1995), pp. 2226–2235. DOI: 10.1364/AO.34.002226. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-34-13-2226>.
- [70] Juan C. Miñano et al. "An application of the SMS method for imaging designs". In: *Opt. Express* 17.26 (Dec. 2009), pp. 24036–24044. DOI: 10.1364/OE.17.024036. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-17-26-24036>.
- [71] Fabian Duerr et al. "Analytic free-form lens design in 3D: coupling three ray sets using two lens surfaces". In: *Opt. Express* 20.10 (May 2012), pp. 10839–10846. DOI: 10.1364/OE.20.010839. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-20-10-10839>.

REFERENCES

- [72] Jun Zhu, Tong Yang, and Guofan Jin. "Design method of surface contour for a freeform lens with wide linear field-of-view". In: *Opt. Express* 21.22 (Nov. 2013), pp. 26080–26092. DOI: 10.1364/OE.21.026080. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-21-22-26080>.
- [73] Tong Yang et al. "Design method of freeform off-axis reflective imaging systems with a direct construction process". In: *Opt. Express* 22.8 (Apr. 2014), pp. 9193–9205. DOI: 10.1364/OE.22.009193. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-22-8-9193>.
- [74] Jun Zhu et al. "Generating optical freeform surfaces considering both coordinates and normals of discrete data points." In: *Journal of the Optical Society of America. A, Optics, image science, and vision* 31 11 (2014), pp. 2401–8. URL: <https://api.semanticscholar.org/CorpusID:207311860>.
- [75] Tong Yang et al. "Direct design of freeform surfaces and freeform imaging systems with a point-by-point three-dimensional construction-iteration method". In: *Opt. Express* 23.8 (Apr. 2015), pp. 10233–10246. DOI: 10.1364/OE.23.010233. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-23-8-10233>.
- [76] Tong Yang, Guo-Fan Jin, and Jun Zhu. "Automated design of freeform imaging systems". In: *Light: Science & Applications* 6.10 (Oct. 2017), e17081–e17081. ISSN: 2047-7538. DOI: 10.1038/lsa.2017.81. URL: <https://doi.org/10.1038/lsa.2017.81>.
- [77] Kyong Hwan Jin et al. "Deep Convolutional Neural Network for Inverse Problems in Imaging". In: *IEEE Transactions on Image Processing* 26.9 (2017), pp. 4509–4522. DOI: 10.1109/TIP.2017.2713099.
- [78] Jiaqi Jiang, Mingkun Chen, and Jonathan Fan. "Deep neural networks for the evaluation and design of photonic devices". In: *Nature Reviews Materials* 6 (Dec. 2020). DOI: 10.1038/s41578-020-00260-1.
- [79] Ryoichi Horisaki, Ryosuke Takagi, and Jun Tanida. "Deep-learning-generated holography". In: *Appl. Opt.* 57.14 (May 2018), pp. 3859–3863. DOI: 10.1364/AO.57.003859. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-57-14-3859>.

- [80] Zhiyuan Cao et al. “Back propagation neural network based signal acquisition for Brillouin distributed optical fiber sensors”. In: *Opt. Express* 27.4 (Feb. 2019), pp. 4549–4561. DOI: 10.1364/OE.27.004549. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-4-4549>.
- [81] Geoffroi Côté, Jean-François Lalonde, and Simon Thibault. “Extrapolating from lens design databases using deep learning”. In: *Opt. Express* 27.20 (Sept. 2019), pp. 28279–28292. DOI: 10.1364/OE.27.028279. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-20-28279>.
- [82] Geoffroi Côté, Jean-François Lalonde, and Simon Thibault. “Deep learning-enabled framework for automatic lens design starting point generation”. In: *Opt. Express* 29.3 (Feb. 2021), pp. 3841–3854. DOI: 10.1364/OE.401590. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-29-3-3841>.
- [83] Geoffroi Côté et al. “Inferring the solution space of microscope objective lenses using deep learning”. In: *Opt. Express* 30.5 (Feb. 2022), pp. 6531–6545. DOI: 10.1364/OE.451327. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-30-5-6531>.
- [84] Tong Yang, Dwen Cheng, and Yongtian Wang. “Direct generation of starting points for freeform off-axis three-mirror imaging system design using neural network based deep-learning”. In: *Opt. Express* 27.12 (June 2019), pp. 17228–17238. DOI: 10.1364/OE.27.017228. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-12-17228>.
- [85] Wenchen Chen et al. “Generating starting points for designing freeform imaging optical systems based on deep learning”. In: *Opt. Express* 29.17 (Aug. 2021), pp. 27845–27870. DOI: 10.1364/OE.432745. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-29-17-27845>.
- [86] Boyu Mao et al. “FreeformNet: fast and automatic generation of multiple-resolution freeform imaging systems enabled by deep learning”. In: *Photon. Res.* 11.8 (Aug. 2023), pp. 1408–1422. DOI: 10.1364/PRJ.492938. URL: <https://opg.optica.org/prj/abstract.cfm?URI=prj-11-8-1408>.
- [87] Iqbal H. Sarker. “Machine Learning: Algorithms, Real-World Applications and Research Directions”. In: *SN Computer Science* 2.3 (Mar. 2021), p. 160. ISSN: 2661-8907. DOI: 10.1007/s42979-021-00592-x. URL: <https://doi.org/10.1007/s42979-021-00592-x>.

REFERENCES

- [88] Iqbal H Sarker et al. "Cybersecurity data science: an overview from machine learning perspective". In: *Journal of Big data* 7 (2020), pp. 1–29.
- [89] Raúl Rojas. "The Backpropagation Algorithm". In: *Neural Networks: A Systematic Introduction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996, pp. 149–182. ISBN: 978-3-642-61068-4. DOI: 10.1007/978-3-642-61068-4_7. URL: https://doi.org/10.1007/978-3-642-61068-4_7.
- [90] Xavier Glorot and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks". In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Ed. by Yee Whye Teh and Mike Titterton. Vol. 9. Proceedings of Machine Learning Research. Chia Laguna Resort, Sardinia, Italy: PMLR, May 2010, pp. 249–256. URL: <https://proceedings.mlr.press/v9/glorot10a.html>.
- [91] *Nanosats Database*. <https://www.nanosats.eu/>. Accessed: 2023-11-22.
- [92] *Ansys Zemax OpticStudio 2023 R1 User Manual*. Ansys Co. Canonsburg, Pennsylvania, Jan. 2023.
- [93] Inc. The MathWorks. *Levenberg-Marquardt backpropagation*. Accessed: 29-11-2023.
- [94] Inc. The MathWorks. *Function fitting neural network*. Accessed: 29-11-2023.

Ringraziamenti

Raggiungere la Laurea Magistrale in Ingegneria per le Comunicazioni Multimediali e Internet è, ad oggi, il mio più grande successo personale e il coronamento di anni di *studio matto e disperatissimo*. Per questi motivi, ci tengo a ringraziare tutti coloro che mi hanno accompagnato e sostenuto in questo lungo percorso. Un ringraziamento speciale va al mio Relatore di tesi, Giampiero Naletto, che mi ha permesso di conoscere il nuovo e stimolante campo delle ottiche a forma libera.

Venendo ai ringraziamenti più informali, vorrei fare un plauso ai temerari che hanno letto tutta la tesi, ammesso che ce ne siano. Grazie ai miei insostituibili compagni di corso, chiacchiere e serate: Enrico e Matteo. Grazie a Davide, fonte inesauribile di consigli e pozzo infinito di scienza, che ha reso la Magistrale una realtà meno paurosa. Grazie ai coinquilini che hanno sopportato il Lorenzo in sessione esami. Grazie ai miei genitori per il supporto economico e a mia sorella per avermi guidato nelle scelte più ardue. Grazie ai miei nonni, per l'amore e il cibo. Grazie a Grammarly, messo a dura prova dai miei strafalcioni grammaticali. Grazie alla scena trap genovese, al lo-fi hip hop e all'indie rock, che hanno reso lo studio meno oberante.

Infine, è il turno di colei che da sei anni è la persona più importante. Grazie a Chiara, che mi sprona ogni giorno ad essere la versione migliore di me stesso. Senza di lei nulla di questo sarebbe stato possibile.