

Freeform illumination design in optical systems with partial differential equations

Dissertation

zur Erlangung des akademischen Grades doctor rerum naturalium (Dr.
rer. nat.)



VORGELEGT DEM RAT DER PHYSIKALISCHE-ASTRONOMISCHEN
FAKULTÄT DER FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

von Christoph Bösel
geboren am 08.09.1987 in Erfurt

1. Gutachter: Prof. Dr. Herbert Gross (Friedrich-Schiller-Universität Jena)
2. Gutachter: Prof. Dr. Karl-Heinz Brenner (Universität Heidelberg)
3. Gutachter: Prof. Dr. Fabian Duerr (Vrije Universiteit Brussels)

Tag der Disputation: 26.11.2019

Freiformbeleuchtungsdesign in optischen System mit partiellen Differentialgleichungen

Zahlreiche Anwendungen in der Beleuchtung und Messtechnik erfordern das Design kompakter, energieeffizienter, nicht-abbildender optischer Systeme zur Generierung nichttrivialer Zielintensitätsverteilungen. Eine moderne Möglichkeit dieses Anforderungsprofil zu erfüllen bieten refraktive oder reflektive optische Flächen ohne jegliche Symmetrien, sogenannte Freiformflächen. Im Gegensatz zu klassischen Projektionsmethoden wie zum Beispiel der Durchlichtprojektion bieten Freiformen durch die geeignete Wahl ihrer lokalen Oberflächenkrümmung theoretisch die Möglichkeit beliebige Beleuchtungsmuster nahezu verlustfrei zu erzeugen. Für eine gegebene Lichtquelle und ein gewünschtes Muster besteht die Hauptschwierigkeit hierbei in der Berechnung der entsprechenden Freiformflächen, welche die Energieumverteilung realisieren. Dieses sogenannte inverse Problem der nicht-abbildenden Optik erfordert zum einen dessen mathematische Modellierung und zum anderen die numerische Lösung des Modells. Das Ziel dieser Arbeit ist demzufolge die Entwicklung einer allgemeinen mathematischen Beschreibung des inversen Problems und dessen numerischer Lösung, sowie die Entwicklung anwendungsorientierter Freiformbeleuchtungskonzepte.

Im ersten Teil dieser Arbeit werden dazu zunächst die zur Entwicklung der mathematischen Modelle notwendigen Grundlagen der geometrischen Optik ausgehend vom Fermatschen Prinzip eingeführt. Aufbauend auf den physikalischen und geometrischen Gesetzmäßigkeiten wird eine Beschreibung des inversen Problems in Form nichtlinearer, partieller Differentialgleichungen für Null-Étendue Lichtquellen, welche durch eine einzelne Wellenfront beschrieben werden, hergeleitet. Das Modell beschreibt hierbei sowohl die gezielte Intensitätsumverteilung mittels einer einzelnen refraktiven oder reflektiven Freiformfläche als auch die Umverteilung der Intensität und Phase mit zwei gekoppelten refraktiven und/oder reflektiven Freiformflächen. Im Gegensatz zu vorherigen Modellen ist die Beschreibung hierbei nicht eingeschränkt auf planare oder sphärische Eingangswellenfronten oder auf ein paraxiales Regime. Zudem beschreibt das Modell das Beleuchtungsdesign von Freiformflächen, welche in einem optischen System platziert werden.

Im zweiten Teil der Arbeit wird eine numerische Lösungsstrategie für das mathematische Modell entwickelt. Die Basis bildet dabei die optimale Transporttheorie mit einer quadratischen Kostenfunktion für welche gezeigt wird, dass sie die Freiformbeleuchtung für planare Wellenfronten in einer paraxialen Approximation beschreibt und somit als Startwert für ein Strahlen-Mapping zur numerischen Lösung der partiellen Differentialgleichungen dient. Auf die detaillierten Darstellung der numerischen Lösungsstrategie folgend wird dann eine Evaluation der numerischen Methodik anhand verschiedener Designbeispiele durchgeführt.

Im abschließenden Teil der Arbeit wird ein Designkonzept zur Berechnung kompakter Freiformbeleuchtungssysteme zur gezielten Intensitätsumverteilung realer, ausgedehnter Lichtquellen entwickelt. Dieses Konzept besteht aus der geeigneten Kombination einer dünnen Doppelfreiformlinse, welche für planare Wellenfronten berechnet wird, und eines abbildenden optischen Systems mit objektseitig telezentrischem Strahlengang. Wie gezeigt wird, führt diese Kombination der

beiden Elemente zur Minimierung des typischen Verschmierungseffektes des Intensitätsmusters, welches bei der Freiformbeleuchtung mit realen, ausgedehnten Lichtquellen auftritt. Durch das entwickelte Designkonzept lassen sich vormals nicht realisierbare ultrakompakte, hocheffiziente Projektionssysteme zur Musterprojektion berechnen, welche zum Beispiel zur Entwicklung kompakter 3D Messsysteme verwendet werden können.

Freeform illumination design in optical systems with partial differential equations

Numerous applications in illumination and metrology require the design of compact, energy-efficient nonimaging optical systems for nontrivial irradiance or intensity distributions. A modern way to fulfill the profile of requirements are freeform surfaces, meaning refractive or reflective surfaces without any symmetries. In contrast to classical projection methods, for instance transmitted-light illuminators, freeform surfaces offer the possibility to generate nearly arbitrary target distributions by an appropriate choice of the local surface curvature. For a given light source and a desired target distribution the main difficulty is thereby the computation of the freeform surfaces, which realize the required energy redistribution. This so-called inverse problem of nonimaging optics necessitates on the one hand a mathematical description and on the other hand the numerical solving of the corresponding model. Therefore, the goal of this thesis is to develop a general mathematical description of the inverse problem, the numerical solving of the corresponding model as well as the development of application oriented freeform illumination design concepts.

In the first part of this thesis, starting from Fermat's principle, the basics of geometrical optics which are required for the development of the mathematical models, are presented. Based on the corresponding physical and geometrical laws, a description of the inverse problem in terms of nonlinear, partial differential equations for zero-étendue light sources, which are described by a single wavefront, is introduced. The model thereby describes the specified irradiance redistribution with a single refractive or reflective freeform surface as well as the irradiance and phase redistribution with two coupled refractive and/or reflective freeform surfaces. In contrast to former models, the description is not restricted to planar or spherical input wavefronts or to a paraxial regime. Additionally, the model describes the illumination design for freeform surfaces which are placed in an optical system.

In the second part of the thesis a numerical solving strategy for the design model is developed. The basis builds the optimal transport theory with a quadratic cost function for which it is shown that it describes the freeform illumination design with planar wavefronts in a paraxial approximation. Consequently, it serves as an initial iterate for a ray mapping for the numerical solution of the partial differential equations. After the detailed description of the numerical design strategy follows an evaluation of the design methodology for several design examples.

In the final part of the thesis a design concept for the calculation of freeform illumination systems for irradiance control with real, extended light sources is developed. The concept consists of an appropriate combination of a thin zero-étendue double freeform for collimated beam shaping and an imaging system with a telecentric object space. As it is shown, this combination of both elements leads to a minimization of the typical irradiance pattern blurring effect which occurs in freeform illumination design with extended light sources. This allows the design of compact, energy efficient projection systems for pattern generation, which could not be realized before and which are useful for the development of compact 3D measurement systems.

Contents

1	Introduction	3
1.1	State of the art and contributions of this work	4
2	Fundamentals	9
2.1	Geometrical optics	9
2.1.1	Fermat's principle and deflection laws	9
2.1.2	Wavefronts and optical path length	11
2.1.3	Étendue conservation	12
2.2	Energy conservation	14
2.2.1	Coordinate transformations and local energy redistribution	14
2.2.2	Optimal mass transport and freeform illumination design	16
3	Mathematical modelling of freeform illumination design with partial differential equations (PDEs)	18
3.1	Single freeform design: irradiance control	18
3.1.1	Design geometry	19
3.1.2	PDE system and Monge-Ampère equation	22
3.1.3	Predefined exit surfaces	23
3.2	Double freeform design: irradiance and phase control	25
3.2.1	Generalization to double freeform surfaces	25
3.2.2	Constant optical path length condition	27
3.2.3	Wavefront Mapping Coordinates and PDE System	28
3.3	Formulation of freeform illumination design problem in optical systems	28
3.3.1	Describing a ray-path in an optical system	29
3.3.2	Freeform illumination design in optical systems	30
3.4	Summary	32
4	L_2 optimal mass transport and freeform illumination design	34
4.1	Single freeform surface	35
4.2	Double freeform surface	38
4.3	Summary	40
5	Design method: numerical solution of freeform illumination PDEs	42
5.1	PDE discretization and nonlinear equation systems	43
5.2	Initial mapping calculation	46
5.2.1	Solving the Monge-Ampère equation with finite difference schemes: the problem with nonlinear PDEs	46
5.2.2	Sulman's L_2 OMT algorithm	47
5.3	Initial surface construction	49
5.3.1	Single freeform surface	50
5.3.2	Double freeform surface	51
5.4	Design strategies	52

5.4.1	Freeform illumination design in optical systems	56
5.5	Preparation of input and output data for design process	57
5.5.1	Ray-aiming	57
5.5.2	L_2 OMT mapping inversion	58
5.5.3	Noise-free irradiance computation in (sequential) optical systems	59
5.5.4	Wavefronts and mapping projection	60
5.5.5	Wavefronts and ray directions	60
5.6	Summary	61
6	Design examples	62
6.1	Single freeform design	63
6.2	Double freeform design	67
7	Freeform illumination design for pattern generation with extended sources	71
7.1	Zero-étendue single freeform illumination design for structured illumination with extended sources	73
7.2	Reducing the pattern blurring of freeform projectors	75
7.2.1	Combining a zero-étendue single freeform and an imaging system	75
7.2.2	Combining a zero-étendue double freeform and an imaging system	80
7.3	Design examples	84
7.4	Design procedure	90
7.5	Non-planar input and output wavefronts	91
7.6	Summary	94
8	Conclusion and outlook	96
	Bibliography	100
	Appendices	106
A	Additional content	106
A.1	Local energy conservation for nonplanar target surface	106
A.2	Single FID with general target surfaces	106
A.2.1	Generalized PDE model	106
A.2.2	Numerics	107
A.3	Double lens with two freeform surfaces	107
A.4	Coupling of input wavefronts by freeform systems	108
A.4.1	Example: $N=2$	109
A.5	Derivation of Equation (4.5)	110
A.6	Integrating the linear advection equation (4.9)	111
A.7	Remarks on manufacturability of freeform surfaces	112
B	Publications	115
C	Ehrenwörtliche Erklärung	117

1 Introduction

While the world per capita energy consumption remained nearly constant between the 1970's and 2000's, it has shown a significant growth in the past two decades [1]. This important problem is addressed by, inter alia, increased research in renewable energies to provide environmentally friendly energy production. However, not only the energy production but also the efficient energy usage is the focus of a heightened interest in academic research. An important step towards a more efficient energy utilization in illumination was the development of the light emitting diode (LED), which was first reported in 1927 by Oleg Vladimirovich Losev [2] and which showed a tremendous evolution towards high luminous efficacies in past decades [3]. Due to numerous other advantages, for instance their directional emission, good lifetime and comparably small physical dimensions [4], LEDs might nowadays be the main contributor to an efficient illumination in applications like street lighting, automotive lighting or 3D measurement.

These applications also illustrate another important energy efficiency-determining aspect, which is the light transfer from the light source to the illuminated area. The academic research field, concerned with the efficient light transfer between a light source and a target, and the design of corresponding optical systems, is the nonimaging or illumination optics [5–7]. While many design concepts of nonimaging optics like concentrators focus on the energy transfer from an emitter to a receiver without considering the concrete energy distribution at the target, other applications have requirements that are more stringent. For instance, the above-mentioned street- and automotive lighting and 3D measurement [Fig. 1.1] require the generation of a predefined intensity or irradiance pattern, which is prescribed by legal standards or in case of 3D measurement by the desired measurement precision [8, 9]. In case of complex target patterns with high spatial frequency structures and without any symmetries, this will consequently necessitate optical systems of equal complexity, in the sense that optical surfaces with complicated curvatures are needed and/or a large number of optical surfaces to realize the energy redistribution. These optical surfaces without any symmetries, so-called freeform surfaces, gain more and more interest within the imaging as well as nonimaging optics community. This is caused by recent progresses in high precision manufacturing methods like diamond turning, injection molding or 3D printing [10, 11], and optical design theory and methods [12, 13]. The design and manufacturing of freeform surfaces enables the development of compact imaging and illumination systems [14], which were unthinkable a few years ago. Especially the application of refractive or reflective freeform surfaces in nonimaging optics potentially provides the possibility to design illumination systems for arbitrary illumination patterns, while omitting the reduction of the

energy throughput of classical projection approaches like slide projectors [15].

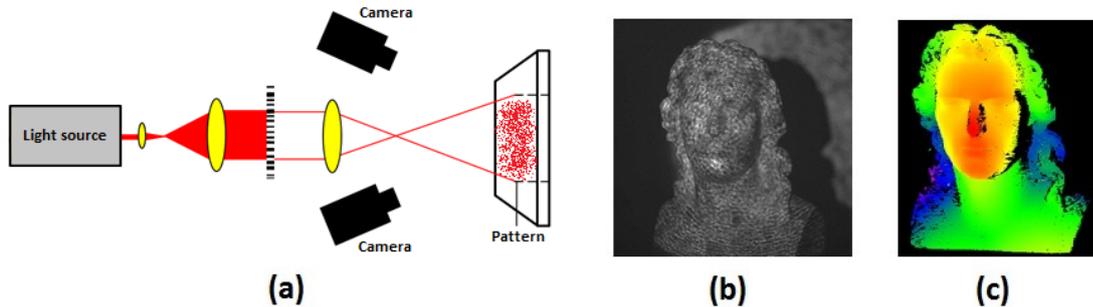


Figure 1.1: (a) Typical 3D measurement setup consisting of a light source, a pattern generating beam shaping optic, two cameras and the target volume. (b) The object is illuminated by a predefined speckle-like pattern, generated by a freeform surface. (c) The three-dimensional object shape is recreated by a triangulation with the two cameras. [9]

An efficient design of these kind of illumination systems requires the formal mathematical connection of the structure of the illumination pattern to the local curvature of the freeform surface, which leads to the central problem statement considered in this thesis: the inverse problem of nonimaging optics. It states that for a given light source and predefined irradiance pattern and/or wavefront, one or more surfaces need to be calculated, which realize the required energy redistribution.

Hereafter, it will be distinguished between the inverse problem for zero-étendue light sources, which are described by a single wavefront, and extended light sources, which are described by a superposition of an infinite number of wavefronts. While the inverse problem for zero-étendue light sources can be solved exactly due to an one-to-one correspondence between input rays and freeform surface points, extended light sources only allow for approximate solutions. Thus, also in literature, there is a clear distinction between published freeform illumination design (FID) methods for both types of light sources. The main focus of this thesis will be the development of design models and numerical algorithms for the calculation of freeform surfaces without any symmetries for zero-étendue light sources in a geometrical optics approximation. Despite the focus on zero-étendue light sources, in the last part of the thesis it will be shown that the developed algorithms can also be applied *without modifications* directly to the design of compact freeform illumination systems for extended light sources.

1.1 State of the art and contributions of this work

Numerous attempts have been made in recent years to model and numerically solve the inverse problem of nonimaging optics for extended sources (e.g. [16–26, 28–35]) as well as for zero-étendue light sources. In the following, a short literature review of design models and methods for the latter will be given. These design methods for zero-étendue light sources can be divided into three main groups, all of which have different advantages and disadvantages: the support-

ing quadratics method (SQM), ray-mapping methods and partial differential equation (PDE) methods.

The SQM is a commonly used method for FID [36–41]. It is based on the construction of the desired freeform surface by utilizing special properties of quadric surfaces. For instance, if a freeform mirror design for a spherical input wavefront with a target plane in the near field is considered, the freeform can be constructed from a unification of ellipsoidal surfaces. Thereby, the point light source is placed in one common focal point of all ellipsoids, which redirect the light to the other focal points placed at different positions in the target plane. The freeform can then be built by an appropriate unification of patches of all ellipsoids to obtain the desired irradiances [36]. Similar techniques can be used for other design geometries with lenses instead of mirrors or for irradiance and wavefront shaping with two freeform surfaces. So far, the SQM has been demonstrated for the shaping of collimated beams [41] or point light sources [37–41] with single freeform surfaces as well as for the collimated beam shaping with double freeform surfaces [42]. Positive aspects of the SQM are the rather simple handling of complicated boundaries of the illumination patterns, the ability to handle high contrasts [41, 42], which is sufficient for many applications, and a comparably profound mathematical understanding and justification of aspects like numerical convergence. This contrasts with rather low possible resolutions of irradiance distributions¹ and low computational performances², respectively. Furthermore, negative aspects are the differentiability of the designed freeform surfaces, which might be critical regarding manufacturing, and the lack of publications with detailed, efficient algorithms for complex irradiance distributions.

Another important group of FID methods are the ray-mapping methods [46–58]. The basis of these methods builds the establishment of a relationship between the input ray coordinates of the light source and the target coordinates of the rays after the reflection or the refraction at the freeform surface. This relationship needs to be calculated in a way that the input energy distribution is redistributed into the required target illumination pattern. After the calculation of this ray mapping, the freeform surfaces is then constructed from the geometrical deflection laws by defining the surface normal vector field through the ray mapping. The question of the integrability of the corresponding ray mapping is related to this. This means that an arbitrary mapping relationship, which redistributes a given input distribution into the target distribution does not necessarily guarantee the existence of a corresponding *continuous* freeform. Only for a few specific cases so far, the design problem has been formulated in terms of a ray-mapping like for instance the double reflector design [43], the double lens design for collimated beam shaping [44] or the single freeform lens design for collimated beam shaping [55]. A commonly used ray-mapping in FID was first applied by Bruneton et al. [47] and later also by other authors [49, 50]. For a given emittance and irradiance distribution, this mapping is calculated from optimal mass transport (OMT) with a quadratic cost function (L_2 OMT) and has shown

¹Designed freeform surfaces for irradiances with approximately 38K pixels have been reported so far [41].

²In Ref. [37] the authors designed freeform mirrors for 3025 target pixels and a contrast ratio of 0.33:1 within 15 min. The code was implemented in FORTRAN and run on a Pentium IV processor.

successful applications to the irradiance control [47, 50] as well as the irradiance and phase control [49]. Due to its nonintegrability, the design geometries obtained by a direct integration of the L_2 OMT mapping are thereby restricted to paraxial regimes with small deflection angles caused by the freeform surface as shown in this thesis. The most powerful ray mapping method based on L_2 OMT in single FID so far was published by Schwartzburg et al. [50], which reported the design of irradiance distributions with areas of vanishing irradiance and complicated boundaries with computational times of about 25 min for 66 K pixels and 95 min for 260 K pixels. A general formulation of a design method for single *and* double FID with arbitrary wavefronts does not exist at the moment.

This thesis mainly concentrates on the PDE methods, which model the FID problem by an elliptic, nonlinear PDE, the Monge-Ampère equation (MAE), or systems of nonlinear PDEs. These are solved numerically for the freeform surface or a combination of a freeform surface and the ray mapping. The difficulty of the development of numerical algorithms hereby results from the nonlinearity of the PDEs, which causes numerical instabilities and slow convergence rates for complicated irradiances with large contrasts and require significant computational resources. Furthermore, the implementation of boundary conditions for complicated irradiance boundary shapes is generally nontrivial. On the other hand, the formulation of FID in terms of PDEs directly relates the structure of the irradiance to the curvature of the freeform and guarantees a certain differentiability of the calculated surfaces due to the explicit occurrence of the partial derivatives of the surface. Numerous mathematical models have been presented so far³, beginning with the differential equation model for reflector design with spherical input wavefronts by Boldyrev in 1932 [59], the first MAEs by Komissarov and Boldyrev in 1941 [60] and by Schruben in 1972 [62] for the reflector design, followed by further models for irradiance [45, 63, 64, 67–72, 76] as well as irradiance and phase control [66, 67, 74] with reflective and refractive surfaces. These models are thereby restricted to planar⁴ and spherical input wavefronts with correspondingly shaped entrance surfaces and/or paraxial geometries for single FID, and at least one planar wavefront or paraxiality for double FID.

Consequently, the proposed numerical algorithms for solving the design models, which were first proposed in the 1970's and 1980's in the context of reflector antenna design [65–67], are restricted to these kind of optical configurations [65–67, 70, 71, 73, 74]. This hinders a more efficient utilization of freeform surfaces in illumination design. For instance, from the viewpoint of manufacturability and cost, a plane entrance surface of a freeform lens is beneficial. Due to the restriction to planar and spherical input wavefronts with correspondingly shaped entrance surfaces, a combination of a spherical input wavefront with a plano-freeform lens at a finite distance is not covered by the previous PDE models and PDE methods for irradiance control in literature. Another example is the irradiance control of a laser diode, which is not described

³Design models with symmetry assumptions like e.g. the circular symmetric model in Ref. [61] by Galindo are not considered in this thesis. Literature on these type of design models can be found in the reference list of chapter 7 in the textbook by Winston et al. [5].

⁴“Planar” in this case means also that the wavefront is parallel to the emittance and irradiance plane. For instance this is not the case for the single FID example in Fig. 5.8(a) or the double FID example in Fig. 6.3(c).

by previous models due to the astigmatic wavefront. Hence, due to its high practical relevance, one aim of this work is to overcome the above-mentioned restrictions of the PDE models and numerical PDE methods, and generalize them to general zero-étendue wavefronts. This leads to the major contributions of this thesis⁵:

- (a) The single FID for irradiance control is formulated in terms of nonlinear PDEs for general zero-étendue input wavefronts. This allows the utilization of nonshperical and nonplanar input wavefronts, and/or prescribed entrance surfaces, which was not described by previous PDE models. Additionally, the PDE model is extended to predefined exit surfaces. This describes the placement of a single freeform surface for irradiance control in a system of predefined surfaces. Consequently, pre- or postshaping of input or output beams can be applied to overcome geometrical restrictions posed by a specific design problem. Furthermore the double FID for irradiance and phase control is formulated in terms of nonlinear PDEs for general zero-étendue wavefronts. In contrast to the previous design models in literature, the model describes general input *and* output wavefronts, and is not restricted to paraxial geometries. [Chapter 3]
- (b) A numerical design strategy to solve the *introduced* PDE models is developed with main focus on the single FID and double FID for *general* zero-étendue wavefronts. It is based on the discretization of the nonlinear PDE systems and the subsequent solution of the resulting nonlinear equations systems by corresponding solvers. The key for the numerically stable solution of the nonlinear equation system, especially for complex irradiance distributions, is thereby the construction of an initial iterate which is suited to the structure of the nonlinear PDE system with an *explicitly* stated energy conservation equation and as well as being related to the considered design geometry. It is based on the analytical finding that the L_2 OMT map, which solves the energy conservation equation, is asymptotically integrable in a paraxial regime for collimated beams [Chapter 4] and on the construction of a suitable initial freeform surface from the given wavefronts and the L_2 OMT map. [Chapter 5]
- (c) A design concept for pattern generation with compact freeform illumination systems with *extended* light sources is developed systematically, which is based on a *direct* application of the introduced zero-étendue models and numerical methods. The illumination system consists of a thin double freeform for collimated beam shaping and an imaging system with a telecentric object space. The freeform surfaces are thereby calculated under a zero-étendue assumption and it is shown that the typical pattern blurring of single freeform projectors, caused by the light source extension, is significantly reduced by an appropriate combination of the double freeform surfaces and the imaging system. Consequently, the design concept enables the generation of complex irradiance patterns with high spatial

⁵The three claims are in agreement with the literature overview given in the recent review paper on FID by Wu et al. [110]. The claims also refer to the time when Refs. [82–88] were published.

frequency structures for extended light sources, while maintaining the compactness of the freeform illumination system. This allows the design of highly energy efficient freeform projection systems for pattern generation with a compactness that arguably could not be achieved with previously published FID methods and concepts for extended light sources. [Chapter 7]

A more detailed discussion of the corresponding points in context of the existing body of literature is provided in the introduction of each chapter and subchapter.

After this introduction, in **Chapter 2**, the fundamental physical principles, necessary to understand and develop the FID models and concepts in this thesis are presented. This includes the presentation of the basic equations and laws of geometrical optics and the fundamentals of energy conservation and optimal mass transport theory.

In **Chapter 3**, the PDE models for FID in optical systems with zero-étendue light sources are derived. At first it is shown how the single FID problem can be formulated for general input wavefronts and exit surfaces in terms of a MAE for the freeform surface and an equivalent system of coupled PDEs for the freeform surface and the ray-mapping components. Then, the PDE model is extended to the double FID by an application of the constant optical path length (OPL) condition and a projection of the ray-mapping onto the outgoing wavefront. Subsequently, a general PDE model for FID in optical systems is given. The results of this chapter were published in Refs. [85–88].

Chapter 4 discusses the relation between FID and the L_2 OMT. It is shown analytically that the corresponding transport map represents an integrable ray-mapping for collimated beam shaping in a paraxial regime. The results of this chapter were published in Refs. [82–84].

Based on these results, in **Chapter 5**, a numerical design strategy for the PDE models of single FID and double FID is developed. The results of this chapter were published in Refs. [85–87].

In **Chapter 6**, various design examples are calculated to demonstrate and evaluate the capability of the numerical design strategy. The results of this chapter were published in Refs. [85–87].

Thereafter, in **Chapter 7**, a freeform illumination system design concept for pattern generation with extended light sources is developed, which is based on the application of the proposed design algorithms for zero-étendue light sources. The results of this chapter were published in Refs. [88].

2 Fundamentals

In this chapter, the fundamental physical principles and equations, which are necessary to understand and develop the PDE models and FID concepts in this thesis, are presented. Consequently, in section 2.1, the elementary propagation laws of light in a geometrical optics approximation are derived, based on Fermat's principle, and the law of étendue conservation, which represents a fundamental geometrical limitation of illumination design with extended light sources, is discussed. Furthermore, in section 2.2, energy conservation and the fundamentals of OMT, which in case of FID is connected to the local energy redistribution by the freeform surfaces for given input and output distributions, are discussed.

2.1 Geometrical optics

For the majority of applications in illumination design, the wavelength of the light source is significantly smaller than the scales of the optical system and diffraction effects can be neglected. In case of a wavelength $\lambda \rightarrow 0$ the light propagation can be described as a limiting case of the electromagnetic theory in a geometrical optics approximation, in which the energy is transported along rays, meaning that the ray directions coincide with the Poynting vector. Therefore, to describe the redistribution of energy or redirection of rays, respectively, of a light source by a refracting or reflecting freeform surface, the corresponding geometrical propagation and deflection laws need to be applied. These can be derived from Fermat's principle, which describes the propagation of a ray between two given points \mathbf{P}_1 and \mathbf{P}_2 and will be used to derive the necessary equations utilized in the upcoming chapters. Conventions and discussions follow the textbook by Chaves [6].

2.1.1 Fermat's principle and deflection laws

Assuming an inhomogeneous medium with the refractive index distribution $n(\mathbf{x})$ in which the light travels between the points \mathbf{P}_1 and \mathbf{P}_2 , the OPL between both points is defined by

$$S = \int_{\mathbf{P}_1}^{\mathbf{P}_2} n(\mathbf{x}(s)) ds \quad (2.1)$$

with $\mathbf{x}(s) = (x(s), y(s), z(s))$ and the parameter s . Utilizing this definition, Fermat's principle is formulated as

$$\delta S = \delta \int_{\mathbf{P}_1}^{\mathbf{P}_2} n(\mathbf{x}(s)) ds = 0. \quad (2.2)$$

This means that the variation δS of the light ray path between \mathbf{P}_1 and \mathbf{P}_2 vanishes and the OPL is stationary along the path of the ray.

Considering Eq. (2.2) for FID, the special case in which the light travels between \mathbf{P}_1 and \mathbf{P}_2 in homogeneous media with constant refractive indices, and is reflected or refracted at a surface [Fig. 2.1], is of interest.

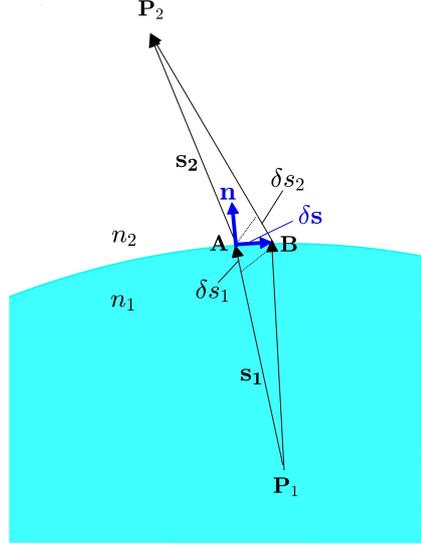


Figure 2.1: An incoming vector \mathbf{s}_1 starting at \mathbf{P}_1 is refracted at the point \mathbf{A} with the surface normal vector \mathbf{n} to the point \mathbf{P}_2 . Additionally, a varied path $\mathbf{P}_1\mathbf{B}\mathbf{P}_2$ is considered. The point \mathbf{B} is shifted relative to \mathbf{A} by an *infinitesimal* vector $\delta\mathbf{s}$ tangential to the surface. The shift leads to the geometrical path differences δs_1 and δs_2 between the paths $\mathbf{P}_1\mathbf{A}\mathbf{P}_2$ and $\mathbf{P}_1\mathbf{B}\mathbf{P}_2$.

If the refraction of the ray at a point \mathbf{A} of the surface is assumed, while propagating in the homogenous media with the constant refractive indices n_1 and n_2 , Eq. (2.1) can be written as a sum of the OPL in both media:

$$S = n_1 s_1 + n_2 s_2 \quad \Rightarrow \quad \delta S = n_1 \delta s_1 + n_2 \delta s_2. \quad (2.3)$$

The ray direction vectors in both media are hereby denoted by \mathbf{s}_1 and \mathbf{s}_2 and the corresponding geometrical path lengths by $s_1 \equiv |\mathbf{s}_1|$ and $s_2 \equiv |\mathbf{s}_2|$. To deduce the law of refraction from (2.3), the variation δS of the OPL needs to be determined. Thus, a second path between \mathbf{P}_1 and \mathbf{P}_2 , going through the surface point \mathbf{B} , is considered [Fig. 2.1]. The points \mathbf{A} and \mathbf{B} are thereby separated by the infinitesimal surface tangential vector $\delta\mathbf{s}$. Denoting the normalized ray direction vector fields by $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$, it follows $\delta s_1 = \hat{\mathbf{s}}_1 \delta\mathbf{s}$ and $\delta s_2 = -\hat{\mathbf{s}}_2 \delta\mathbf{s}$ and thus

$$\delta S = n_1 \hat{\mathbf{s}}_1 \delta\mathbf{s} - n_2 \hat{\mathbf{s}}_2 \delta\mathbf{s} = (n_1 \hat{\mathbf{s}}_1 - n_2 \hat{\mathbf{s}}_2) \delta\mathbf{s} = 0. \quad (2.4)$$

Since the vector $\delta\mathbf{s}$ is tangential to the surface at \mathbf{A} , the vector $n_1 \hat{\mathbf{s}}_1 - n_2 \hat{\mathbf{s}}_2$ is parallel to the

surface normal vector \mathbf{n} and it follows

$$\mathbf{n} = k\hat{\mathbf{n}} = n_1\hat{\mathbf{s}}_1 - n_2\hat{\mathbf{s}}_2 \quad (2.5)$$

with the constant k and the normalized surface normal vector $\hat{\mathbf{n}}$. Thus, the normal vector \mathbf{n} is a linear combination of the incoming and outgoing ray direction vector, and the vectors \mathbf{s}_1 , \mathbf{s}_2 and \mathbf{n} need lie in the same plane. Applying the cross product $(\dots \times \hat{\mathbf{n}})$ to both sides of Eq. (2.5), it follows Snell's law of refraction in the form

$$n_1\hat{\mathbf{s}}_1 \times \hat{\mathbf{n}} = n_2\hat{\mathbf{s}}_2 \times \hat{\mathbf{n}}. \quad (2.6)$$

Additionally, it is helpful to express the outgoing ray direction $\hat{\mathbf{s}}_2$ in terms of $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{n}}$, which are given in many physical situations. To do so, Eq. (2.5) is written as

$$\hat{\mathbf{s}}_2 = \frac{n_1}{n_2}\hat{\mathbf{s}}_1 + k'\hat{\mathbf{n}}. \quad (2.7)$$

with $k' \equiv -k/n_2$. By building the square of both sides (2.7) and using normalization of the vector fields to unity, this leads to a quadratic equation for the constant k' . Solving this equation and choosing the solution with the positive sign leads directly to the ray tracing equation

$$\hat{\mathbf{s}}_2 = n\hat{\mathbf{s}}_1 + \left\{ -n\hat{\mathbf{n}} \cdot \hat{\mathbf{s}}_1 + \sqrt{1 - n^2[1 - (\hat{\mathbf{n}} \cdot \hat{\mathbf{s}}_1)^2]} \right\} \hat{\mathbf{n}}, \quad (2.8)$$

for the refraction at a lens surface. The same process can be applied to derive the corresponding equation for the reflection of a ray at a given surface, which leads to

$$\hat{\mathbf{s}}_2 = \hat{\mathbf{s}}_1 - 2(\hat{\mathbf{n}} \cdot \hat{\mathbf{s}}_1)\hat{\mathbf{n}}. \quad (2.9)$$

While the reflection and refraction is considered at a single point, Eqs. (2.8) and (2.9) are of course valid for every point of a surface. Therefore, instead of considering single vectors, Eqs. (2.8) and (2.9) can also be applied to the vector *fields* $\hat{\mathbf{s}}_1(\mathbf{x})$, $\hat{\mathbf{s}}_2(\mathbf{x})$ and $\hat{\mathbf{n}}(\mathbf{x})$. This will play an important role in chapter 3, when the target coordinates of an input vector field $\hat{\mathbf{s}}_1(\mathbf{x})$, which is deflected by a freeform surface, are expressed through the surface and its gradient.

Also the form (2.5) of the law of reflection/refraction will be important in chapter 4, when the connection of the FID to L_2 OMT is discussed.

2.1.2 Wavefronts and optical path length

Equation (2.1) can be interpreted from the viewpoint of classical mechanics. To do so, the ray path between \mathbf{P}_1 and \mathbf{P}_2 is parameterized by σ . The ray path is then described by the vector $\mathbf{s} = \mathbf{x}(\sigma)$ and $s \equiv |\mathbf{s}|$. Now the OPL S is considered as an action and the Lagrange function $L = L(\mathbf{x}, \mathbf{v}, \sigma)$ with the velocity vector $\mathbf{v} = (v_x, v_y, v_z) \equiv \partial\mathbf{x}/\partial\sigma$ is introduced. The OPL (2.1) then reads

$$S = \int_{\mathbf{P}_1}^{\mathbf{P}_2} n \frac{ds}{d\sigma} d\sigma = \int_{\mathbf{P}_1}^{\mathbf{P}_2} L(\mathbf{x}, \mathbf{v}, \sigma) d\sigma. \quad (2.10)$$

By plugging the Lagrangian into the definition of the generalized momenta

$$\mathbf{p} = \nabla_{\mathbf{v}} L(\mathbf{x}, \mathbf{v}, \sigma) = (\partial_{v_x}, \partial_{v_y}, \partial_{v_z})^T L(\mathbf{x}, \mathbf{v}, \sigma) \quad (2.11)$$

the optical momentum vector

$$\mathbf{p} = n \frac{d\mathbf{s}}{ds}. \quad (2.12)$$

is obtained. This vector describes an infinitesimal displacement ds with the parameter s along the trajectory of the ray and points along the direction of propagation. It is therefore parallel to the direction vectors \mathbf{s}_1 and \mathbf{s}_2 introduced in section 2.1.1.

Furthermore, by applying the Euler-Lagrange equations (ELE) in

$$\nabla S = \nabla \int_{\mathbf{P}_1}^{\mathbf{P}_2} L(\mathbf{x}, \mathbf{v}, \sigma) d\sigma = \int_{\mathbf{P}_1}^{\mathbf{P}_2} \nabla L(\mathbf{x}, \mathbf{v}, \sigma) d\sigma \stackrel{ELE}{=} \int_{\mathbf{P}_1}^{\mathbf{P}_2} \frac{\partial \mathbf{p}}{\partial \sigma} d\sigma = \mathbf{p}, \quad (2.13)$$

it can be concluded that the ray direction vectors are perpendicular to the surfaces S of constant OPLs, which are called wavefronts.

Additionally, Eq. (2.13) and the second fundamental theorem of calculus can be applied to get

$$S = \int_{\mathbf{P}_1}^{\mathbf{P}_2} n(\mathbf{x}) ds \stackrel{(2.12)}{=} \int_{\mathbf{P}_1}^{\mathbf{P}_2} \mathbf{p} ds \stackrel{(2.13)}{=} \int_{\mathbf{P}_1}^{\mathbf{P}_2} \nabla S ds = S(\mathbf{P}_2) - S(\mathbf{P}_1), \quad (2.14)$$

which means that the OPL only depends on the initial and final points of a ray. Two arbitrary neighbouring rays with a common initial and final point (and which remain neighbours during propagation) will therefore have ray paths with an identical OPL. This will play an important role for the description of double freeform systems for irradiance and phase control in section 3.2.

2.1.3 Étendue conservation

The conservation of étendue in a lossless optical system is of major importance since it represents a fundamental geometrical limitation of illumination design and makes a distinction between FID with ideal zero-étendue light sources, and FID with extended light sources inevitable.

A zero-étendue light source can be represented by a single wavefront or ray direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$, respectively, and its characteristic energy distribution. Hence, every input ray of the light source can be redirected by a single point of the freeform surface to a specified target point. This makes it possible to solve the inverse problem of FID with a single freeform for a given output irradiance or intensity distribution.

On the other hand, extended light sources are represented by an infinite number of point light sources over the emitting area of the light source, which leads to a range of ray directions hitting every freeform surface point. Consequently, a single surface cannot generate an arbitrary target

distribution, as discussed in chapter 7.

An extended light source is characterized geometrically by its spatial extension $dA = dxdy$ and its continuous spectrum of ray directions at every source point. These ray directions are limited by the maximum cone angles of every PLS in the area dA . Therefore, the ray directions can be represented by the finite volume of optical momenta $dA_p = dp_x dp_y$. Hence, the areas dA and dA_p build a volume $dU = dAdA_p$ in a four dimensional phase space spanned by a x -, y -, p_x - and p_y -axis. This volume element dU is called étendue. As it can be shown in multiple ways [6, 7], dU is conserved for the lossless propagation of light in an optical system, which means that the phase space volume is invariant:

$$dU = dU^* \quad (2.15)$$

with dU before, and dU^* after the propagation through the system. The étendue conservation is therefore not a physical but a geometrical restriction in the illumination design.

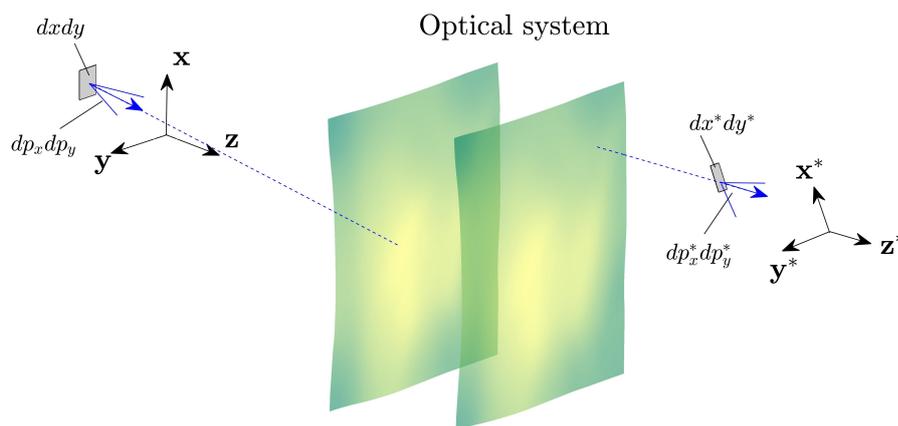


Figure 2.2: The infinitesimal phase space element $dU = dxdy dp_x dp_y$ in the entrance aperture transforms due to the (lossless) propagation in the optical system into $dU^* = dx^*dy^* dp_x^* dp_y^*$ in the exit aperture. While the size of dU and dU^* is equal, due to étendue conservation, the shape of the element might change. A smaller spatial extension of the final element $dx^*dy^* < dxdy$ leads therefore to a larger angular extension $dp_x^* dp_y^* > dp_x dp_y$.

While Eq. (2.15) restricts the size of the phase space volume, the shape of dU can change during propagation through the system [Fig. 2.2]. Thus, the range of ray directions dA_p can be minimized by increasing the spatial extension dA . For instance, if an extended light source with a finite emitting area in the x - y -plane and a finite opening angle is considered, and an observation plane is placed at different distances $z = z_0$ to the extended source, it can be concluded that for larger z_0 the angular range of ray directions dA_p at every point in $z = z_0$ will get smaller.

This will be of key importance in chapter 7, in which different design concepts for FID with extended light sources will be developed and compared. Since the irradiance quality of the final freeform design is inherently limited by the magnitude of the angular range of ray directions dA_p at every freeform surface point (see chapter 7), the lateral extension of the illumination systems

in different design concepts needs to be similar for a valid comparison of the concepts.

2.2 Energy conservation

So far, the propagation of light in a geometrical optics framework was considered, which is important for describing the redirection of light by a freeform surface. Since for FID the light needs to be redistributed by a freeform surface in a specified way to realize the required output distribution, the aspect of energy redistribution will now be discussed.

To do so, a connection between a mathematical coordinate transformation and the local energy distribution in optical systems will be established. As a result follows the Jacobian equation which relates the emittance and irradiance of an optical system through a mapping function between the initial ray coordinates on an input plane and the final ray coordinates on the target surface.

Subsequently the basics of optimal mass transport (OMT), which, roughly speaking, considers the (optimal) mapping of two density functions under certain constraints, are presented. This is of importance since many FID problems can be formulated and solved in an OMT framework by connecting a specific freeform design problem to a corresponding cost function [43, 44, 56]. The density functions are hereby represented e.g. by the given emittance and required irradiance or the intensities of the design problem. In case of FID, the OMT provides a (ray) mapping function, which relates input ray coordinates to the target ray coordinates of the freeform system to realize the required energy redistribution. The main focus will thereby lay on the L_2 OMT, which considers a quadratic cost function and can be represented by an elliptic MAE for a convex potential.

As it will be shown in section 3.1, the single FID problem for zero-étendue wavefronts is represented by a generalization of this MAE. Furthermore, in chapter 4, the L_2 OMT map will be connected to the FID with collimated beams by utilizing its characteristic curl free property. This builds the basis of the numerical FID strategy, developed in chapter 5.

2.2.1 Coordinate transformations and local energy redistribution

In the following a mathematical connection between an emittance $I_S(\mathbf{x})$ on the plane $z = z_0$ and irradiance $I_T(\mathbf{x})$ on the plane $z = z_T$ in an optical system [Fig. 2.3 (a)] is established. The distributions are thereby defined on $\Omega_S := \overline{\{(x, y) \in \mathbb{R}_2 \mid I_S(x, y) \neq 0\}}$ and $\Omega_T := \overline{\{(x, y) \in \mathbb{R}_2 \mid I_T(x, y) \neq 0\}}$ with the boundaries $\partial\Omega_S$ and $\partial\Omega_T$. Assuming a zero-étendue light source and no crossing of rays on $z = z_T$, there is a single ray at every point of the support of $I_S(\mathbf{x})$ and $I_T(\mathbf{x})$. That means that the optical system defines a one-to-one map between every initial ray coordinate $\mathbf{x} = (x, y)$ on the emittance plane and final ray coordinate $\mathbf{u}(\mathbf{x}) = (u_x(x, y), u_y(x, y))$ on the irradiance plane [Fig. 2.3 (b)]. This ray-mapping $\mathbf{u}(\mathbf{x})$ can be interpreted as a coordinate transformation to formally connect the emittance $I_S(\mathbf{x})$ and irradiance $I_T(\mathbf{x})$.

To mathematically describe the local energy redistribution by the optical system, an infinitesimal

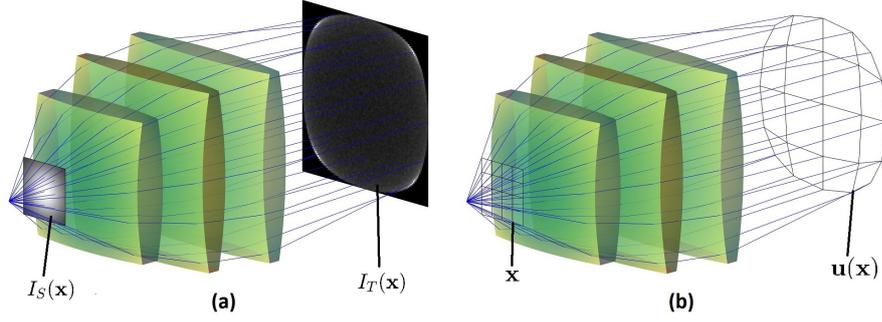


Figure 2.3: (a) An optical system maps an emittance $I_S(\mathbf{x})$ onto an irradiance $I_T(\mathbf{x})$. (b) The optical system connects the initial ray coordinates \mathbf{x} to the final ray coordinates $\mathbf{u}(\mathbf{x})$. This ray mapping $\mathbf{u}(\mathbf{x})$ can be interpreted as a coordinate transformation.

rectangle with the side length dx and dy at \mathbf{x} on the plane $z = z_0$ is considered [Fig. 2.4 (b)]. This rectangle can be interpreted as spanned by initial rays through every corner of the rectangle and is mapped by the optical system onto a quadrilateral at $\mathbf{u}(\mathbf{x})$ on $z = z_T$ [Fig. 2.4 (b)]. The corners of the quadrilateral correspond to the final coordinates of the considered rays. Thus, the variation of the initial rectangular in the optical system can be interpreted as a propagation of a ray tube.

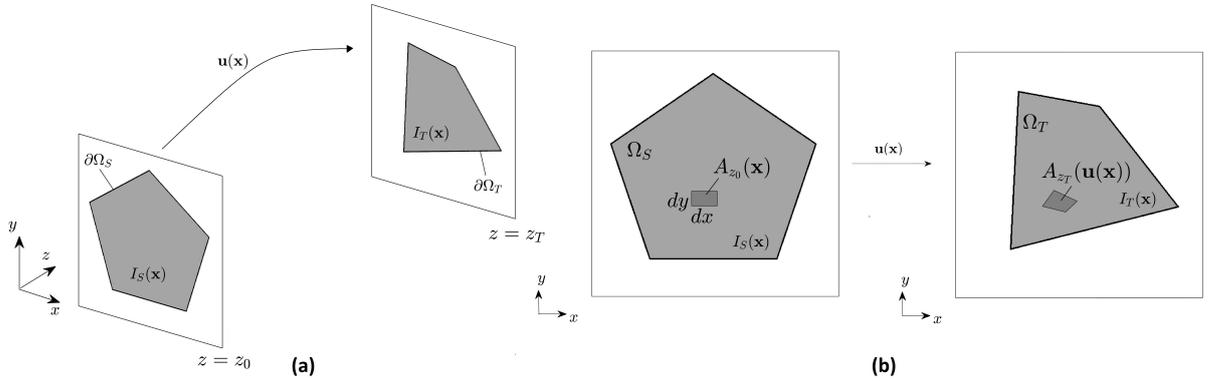


Figure 2.4: (a) The mapping of an emittance $I_S(\mathbf{x})$ on the plane $z = z_0$ onto an irradiance $I_T(\mathbf{x})$ on $z = z_T$ by an optical system is described by the coordinate transformation or ray-mapping $\mathbf{u}(\mathbf{x})$. (b) Under the coordinate transformation an infinitesimal rectangular of the size $A_{z_0}(\mathbf{x}) = dxdy$ is mapped onto the quadrilateral of size $A_{z_T}(\mathbf{u}(\mathbf{x})) = \det \nabla \mathbf{u}(\mathbf{x})dxdy$.

Since the local emittance at \mathbf{x} on $z = z_0$ can be interpreted as inversely proportional to the size $A_{z_0}(\mathbf{x}) = dxdy$ of the initial rectangular, the local irradiance at $\mathbf{u}(\mathbf{x})$ on $z = z_T$ depends on the size $A_{z_T}(\mathbf{u}(\mathbf{x}))$ of the final quadrilateral. To determine the local irradiance $I_T(\mathbf{u}(\mathbf{x}))$, the area change between $A_{z_0}(\mathbf{x})$ and $A_{z_T}(\mathbf{u}(\mathbf{x}))$ needs to be calculated. Therefore, the differential form of the mapping

$$du_x(\mathbf{x}) = \partial_x u_x(\mathbf{x})dx + \partial_y u_x(\mathbf{x})dy, \quad du_y(\mathbf{x}) = \partial_x u_y(\mathbf{x})dx + \partial_y u_y(\mathbf{x})dy \quad (2.16)$$

is considered. By comparing the norm of [90]

$$\begin{pmatrix} dx \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ dy \\ 0 \end{pmatrix} = dx dy \mathbf{e}_z \quad (2.17)$$

with

$$\mathbf{u} \begin{bmatrix} dx \\ 0 \\ 0 \end{bmatrix} \times \mathbf{u} \begin{bmatrix} 0 \\ dy \\ 0 \end{bmatrix} \stackrel{(2.16)}{=} \begin{pmatrix} \partial_x u_x(\mathbf{x}) dx \\ \partial_x u_y(\mathbf{x}) dx \\ 0 \end{pmatrix} \times \begin{pmatrix} \partial_y u_x(\mathbf{x}) dy \\ \partial_y u_y(\mathbf{x}) dy \\ 0 \end{pmatrix} = \det \nabla \mathbf{u}(\mathbf{x}) dx dy \mathbf{e}_z \quad (2.18)$$

it follows that the size of an infinitesimal rectangle changes by the factor $\det(\nabla \mathbf{u}(\mathbf{x}))$, after propagating through the optical system. Consequently, the emittance $I_S(\mathbf{x})$ at \mathbf{x} corresponds to the irradiance $I_T(\mathbf{u}(\mathbf{x}))$ at $\mathbf{u}(\mathbf{x})$, weighted with the local area change $\det(\nabla \mathbf{u}(\mathbf{x}))$ of the initial rectangular. This is described by the Jacobian equation

$$I_S(\mathbf{x}) = \det(\nabla \mathbf{u}(\mathbf{x})) I_T(\mathbf{u}(\mathbf{x})). \quad (2.19)$$

Thus, the ray mapping $\mathbf{u}(\mathbf{x})$ contains the complete information, necessary to describe the local energy redistribution between an emittance and an irradiance in a geometrical optics regime.

The generalization to local energy redistribution for nonplanar target surfaces $z = z_T(\mathbf{x})$ is considered in Appendix A.1.

2.2.2 Optimal mass transport and freeform illumination design

To formulate the problem statement of OMT or the Monge-Kantorovich problem, respectively, two positive density functions $I_S(\mathbf{x})$ and $I_T(\mathbf{x})$ are assumed with

$$\int_{\Omega_S} I_S(\mathbf{x}) d\mathbf{x} = \int_{\Omega_T} I_T(\mathbf{u}) d\mathbf{u} \quad (2.20)$$

with the smooth, bijective mapping $\mathbf{u}(\mathbf{x}) = (u_x(x, y), u_y(x, y))$, $\Omega_S := \overline{\{\mathbf{x} \in \mathbb{R}_2 \mid I_S(\mathbf{x}) \neq 0\}}$ and $\mathbf{u}(\Omega_S) = \Omega_T$ [Fig. 2.4]. As stated above, in FID the density functions correspond, for instance, to the emittance and irradiance of the design problem, which are required to be mapped onto each other by the optical system. Thus, Eq. (2.20) represents the global energy conservation. Hence, in a geometrical optics framework, the map $\mathbf{u}(\mathbf{x})$ can be interpreted as a ray-mapping, which connects the initial coordinates \mathbf{x} of rays in the input plane and the final coordinates $\mathbf{u}(\mathbf{x})$ of the rays, which were redirected by the optical system.

The local energy conservation is expressed by the Jacobian equation (2.19), which describes the local redistribution of the emittance $I_S(\mathbf{x})$ under the mapping $\mathbf{u}(\mathbf{x})$ and represents a PDE for the two mapping components. The integration of Eq. (2.19) over the \mathbb{R}_2 leads to Eq. (2.20). Equation (2.19) with the two unknown functions $u_x(\mathbf{x})$ and $u_y(\mathbf{x})$ is underdetermined. Therefore, in OMT, a specific map $\mathbf{u}(\mathbf{x})$ is fixed by stating an additional constraint in form of a cost function $c(\mathbf{x}, \mathbf{u})$, which e.g. needs to be minimized and characterizes the transformation of the density

functions.

In FID, a certain $c(\mathbf{x}, \mathbf{u})$ therefore defines the redistribution of the emittance by one or more reflective and/or refractive surfaces. So far cost functions for the irradiance control with a single freeform surface and the irradiance and phase control with double freeform surfaces with spherical and planar wavefronts have been proposed [43, 44, 55, 56]. A general formulation for more general wavefronts does currently not exist. The solution of Eq. (2.19) together with the FID cost function provides an integrable map for the design problem. This integrability means that the mapping defines a surface normal vector field that can be integrated to obtain a *continuous*, freeform surface, which redistributes the given input beam correctly into the required output (see chapter 4).

In chapter 4, a specific quadratic cost function $c(\mathbf{x}, \mathbf{u}) = |\mathbf{u}(\mathbf{x}) - \mathbf{x}|^2$ will be considered, which is arguably the standard OMT problem in the mathematical literature. Equation (2.19) then needs to be solved, while simultaneously minimizing the Kantorovich-Wasserstein distance

$$d(I_S, I_T)^2 = \inf_{\mathbf{u} \in M} \int c(\mathbf{x}, \mathbf{u}) I_S(\mathbf{x}) d\mathbf{x} = \inf_{\mathbf{u} \in M} \int |\mathbf{u}(\mathbf{x}) - \mathbf{x}|^2 I_S(\mathbf{x}) d\mathbf{x}, \quad (2.21)$$

with $\inf_{\mathbf{u} \in M}$, representing the infimum of the set M of all energy conserving mappings between $I_S(\mathbf{x})$ and $I_T(\mathbf{x})$. This problem is called the L_2 OMT problem.

One important property of the L_2 OMT map defined by Eqs. (2.19) and (2.21) is that it can be written as the gradient of a *convex* potential $\mathbf{u}(\mathbf{x}) = \nabla u(\mathbf{x})$ [106]. Plugging this relation into the Jacobian equation (2.19) gives the MAE [93]

$$\det(\nabla(\nabla u(\mathbf{x}))) I_T(\partial_x u(\mathbf{x}), \partial_y u(\mathbf{x})) = [\partial_{xx} u(\mathbf{x}) \partial_{yy} u(\mathbf{x}) - \partial_{xy} u(\mathbf{x}) \partial_{yx} u(\mathbf{x})] I_T(\partial_x u(\mathbf{x}), \partial_y u(\mathbf{x})) = I_S(\mathbf{x}), \quad (2.22)$$

for the *unique* convex potential $u(\mathbf{x})$ with the transport boundary condition $\nabla u(\partial\Omega_S) = \partial\Omega_T$. Hence, instead of solving Eqs. (2.19) and (2.21) directly for the map, we can also calculate the potential $u(\mathbf{x})$ and derive the map from its gradient. The numerical solution of Eq. (2.22) is discussed in section 5.2.

Another important property of the L_2 OMT map is its vanishing curl [91]

$$\partial_x u_y(\mathbf{x}) - \partial_y u_x(\mathbf{x}) = 0, \quad (2.23)$$

which will be applied in chapter 4, to show the connection of the L_2 OMT map to the FID design with collimated beams.

3 Mathematical modelling of freeform illumination design with partial differential equations (PDEs)

In this chapter, the modelling of the inverse problem of FID with PDEs for irradiance control with a single freeform surface and for irradiance and phase control with two freeform surfaces is presented. Unlike previous PDE models in literature for irradiance control with single freeform surfaces, it is neither restricted to planar [45, 71, 72] or spherical wavefronts [59, 60, 62–64, 67–70, 76, 80] and correspondingly shaped entrance surfaces nor does it assume paraxial geometries¹. For the irradiance and phase control with double freeform surfaces, it overcomes the restrictions to circular symmetry [61], at least one planar wavefront [66, 67, 74, 78, 84] or paraxiality [49, 54]. Additionally, the PDE model is generalized to predefined exit surfaces. Consequently, the PDE model is valid for freeform surfaces placed in an optical system.

In section 3.1, the single freeform illumination design problem is formulated for general zero-étendue input wavefronts in Cartesian coordinates in terms of a coupled system of nonlinear PDEs or an equivalent MAE, respectively. This, for instance, allows the prescription of arbitrary entrance surfaces between the light source and the freeform surface. Then, by applying additional constraints, the PDE model is generalized to prescribed exit surfaces, leading to the possibility of defining entrance and exit surfaces, simultaneously. In section 3.2, by utilizing the constant OPL condition and a ray-mapping projection onto the prescribed outgoing wavefront, the design model is extended to double freeform systems, enabling therefore the additional control of the outgoing ray directions of the optical system. In section 3.3, the results are summarized in a description of the FID problem in an optical system, which can also be utilized for ray-aiming and noise free irradiance computation.

The results of this chapter were published in references [85], [86], [87] and [88].

3.1 Single freeform design: irradiance control

One of the major drawbacks of the single freeform PDE models for irradiance control in literature is the restriction to planar or spherical wavefronts and correspondingly shaped lens entrance surfaces, respectively and/or the assumption of paraxial geometries. The major difficulty for general input wavefronts is thereby the representation of the unknown ray interception coordinates with the freeform surface and the corresponding surface normal vectors through

¹A MAE design model for general input wavefronts and lens surfaces refracting into the far field was published in Ref. [75]

the initial ray coordinates, which cannot be done straightforwardly with Cartesian or spherical coordinates. Nevertheless, it will be shown that the single FID can be expressed for general zero-étendue wavefronts through the *same* PDE (system) in Cartesian coordinates. This also overcomes problems of some PDE models for spherical wavefronts that are formulated in spherical coordinates [70, 80] and exhibit therefore coordinate singularities, which lead to difficulties regarding the discretization of the PDE's.

To achieve this, the initial coordinates of a ray send out in the input plane and its target coordinates in the output plane are connected to the freeform surface and the surface gradient by applying the laws of refraction and reflection. The resulting ray mapping equations generalize corresponding equations for planar input wavefronts [71, 72] to arbitrary wavefronts and build then together with the Jacobian equation and the transport boundary conditions a system of coupled nonlinear PDE's for the unknown ray mapping components and the freeform surface. This PDE system directly reduces to a Monge-Ampère type equation for the freeform surface only. The specific form of the PDEs is determined through the respective predefined input ray direction vector field and emittance of the incoming beam as well as the required irradiance. Subsequently, it is shown that the PDE model can be extended to predefined exit surfaces, hence allowing the placement of the freeform surface in a system of predefined surfaces.

3.1.1 Design geometry

In the following, a given input beam on the plane $z = z_0$, defined by the normalized ray direction vector field

$$\hat{\mathbf{s}}_1(\mathbf{x}) = \begin{pmatrix} (\hat{\mathbf{s}}_1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_y(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_z(\mathbf{x}) \end{pmatrix}, \quad (3.1)$$

and the emittance $I_S(\mathbf{x})$ with the boundary $\partial\Omega_S$ and $\mathbf{x} = (x, y)$ as defined in section 2.2.1, are considered. This input beam needs to be mapped onto a required irradiance distribution $I_T(\mathbf{x})$ with the boundary $\partial\Omega_T$ on the target plane $z = z_T$ (see Fig. 3.1). Like in the previous chapter, unit vector fields are denoted by a “hat”.

According to the inverse problem of FID, the goal is the calculation of a reflecting or refracting surface, which realizes the mapping of the source onto the target distribution. One important aspect is thereby the desired surface continuity, which simplifies the freeform surface manufacturing process and reduces diffraction effects (depending on the differentiability class) compared to noncontinuous surfaces. This continuity will be implemented through the modelling of the FID problem by PDEs.

In a geometrical optics framework, the inverse problem is fully governed by the energy conservation and the laws of refraction and reflection. The local energy conservation is expressed by the Jacobian equation [Eq. (2.19)]

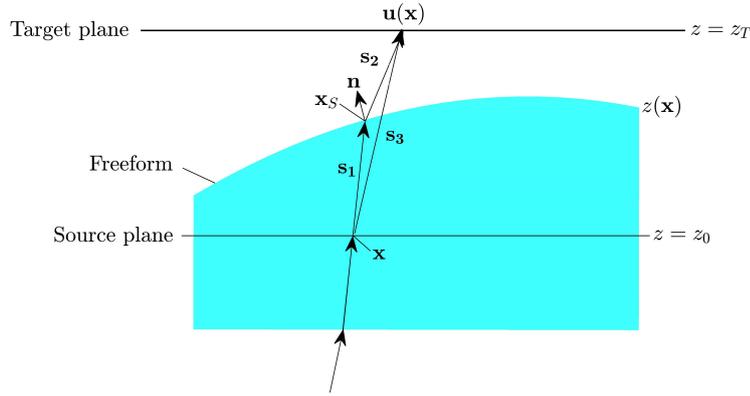


Figure 3.1: The emittance $I_S(\mathbf{x})$ and the input directions $\hat{\mathbf{s}}_1(\mathbf{x})$ on the plane $z = z_0$ as well as the irradiance $I_T(\mathbf{x})$ on the plane $z = z_T$ are given. The input vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ is determined by the given input wavefront and entrance surface(s). To obtain the desired irradiance, the freeform surface $z(\mathbf{x})$ needs to be calculated, which redirects the input beam to the target points $\mathbf{u}(\mathbf{x})$ according to the law of refraction/reflection. The coordinates $\mathbf{x}_S(\mathbf{x})$ represent the intersection points of the input vectors with $z(\mathbf{x})$ and are for noncollimated input beams unknown prior to the freeform calculation. [86]

$$[\partial_x u_x(\mathbf{x}) \partial_y u_y(\mathbf{x}) - \partial_y u_x(\mathbf{x}) \partial_x u_y(\mathbf{x})] I_T(u_x(\mathbf{x}), u_y(\mathbf{x})) = I_S(\mathbf{x}), \quad (3.2)$$

which describes the relation between the emittance and irradiance through the ray mapping or coordinate transformation $\mathbf{u}(\mathbf{x}) = (u_x(\mathbf{x}), u_y(\mathbf{x}))$, respectively. And on the other hand, the ray deflection is expressed through the ray tracing equations for refractive and reflective surfaces [Eqs. (2.8) and (2.9)]

$$\begin{aligned} \hat{\mathbf{s}}_2(\mathbf{x}_S) &= n \hat{\mathbf{s}}_1(\mathbf{x}) + \left\{ -n \hat{\mathbf{n}}(\mathbf{x}_S) \cdot \hat{\mathbf{s}}_1(\mathbf{x}) + \sqrt{1 - n^2 [1 - (\hat{\mathbf{n}}(\mathbf{x}_S) \cdot \hat{\mathbf{s}}_1(\mathbf{x}))^2]} \right\} \hat{\mathbf{n}}(\mathbf{x}_S), \\ \hat{\mathbf{s}}_2(\mathbf{x}_S) &= \hat{\mathbf{s}}_1(\mathbf{x}) - 2 [\hat{\mathbf{n}}(\mathbf{x}_S) \cdot \hat{\mathbf{s}}_1(\mathbf{x})] \hat{\mathbf{n}}(\mathbf{x}_S) \end{aligned} \quad (3.3)$$

with the surface unit normal vector field $\hat{\mathbf{n}}(\mathbf{x})$ at the unknown surface intersection coordinates $\mathbf{x}_S(\mathbf{x}) = (x_S(\mathbf{x}), y_S(\mathbf{x}))$ and the ratio of the refractive indices $n = n_1/n_2$. Equation (3.3) relates the input coordinates \mathbf{x} to the target coordinates $\mathbf{u}(\mathbf{x})$ by the unknown freeform surface $z(\mathbf{x})$ and its gradient. To determine the unknown surface, hereafter, a PDE system for the mapping $\mathbf{u}(\mathbf{x})$ and the surface $z(\mathbf{x})$ is derived. To achieve the desired PDE system for the mapping $\mathbf{u}(\mathbf{x})$ and the surface $z(\mathbf{x})$, the considered geometry is described through the vector fields [Fig. 3.1]

$$\mathbf{s}_1(\mathbf{x}) = \begin{pmatrix} x_S - x \\ y_S - y \\ z(\mathbf{x}_S) - z_0 \end{pmatrix}, \quad \mathbf{s}_3(\mathbf{x}) = \begin{pmatrix} u_x(\mathbf{x}) - x \\ u_y(\mathbf{x}) - y \\ z_T - z_0 \end{pmatrix}, \quad \mathbf{n}(\mathbf{x}_S) = \begin{pmatrix} -\partial_{x_S} z(\mathbf{x}_S) \\ -\partial_{y_S} z(\mathbf{x}_S) \\ 1 \end{pmatrix}, \quad (3.4)$$

and the intersection points $\mathbf{u}(\mathbf{x})$ with the target plane $z = z_T$ are expressed in terms of the given input direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ and the deflected vector field $\hat{\mathbf{s}}_2(\mathbf{x})$ in Eq. (3.3). This is done by applying the geometrical relation

$$\mathbf{s}_1 + \lambda \hat{\mathbf{s}}_2 = \mathbf{s}_3 \quad \Rightarrow \quad \lambda = \frac{(\mathbf{s}_3)_z(\mathbf{x}) - (\mathbf{s}_1)_z(\mathbf{x})}{(\hat{\mathbf{s}}_2)_z(\mathbf{x}_S)} = \frac{z_T - z(\mathbf{x}_S)}{(\hat{\mathbf{s}}_2)_z(\mathbf{x}_S)}, \quad (3.5)$$

in which \mathbf{x}_S and $z(\mathbf{x}_S)$ are the intersection coordinates of the input vector field $\mathbf{s}_1(\mathbf{x})$ with the surface $z(\mathbf{x})$. For the x - and y -component in Eq. (3.5), it then directly follows

$$\mathbf{u}(\mathbf{x}) - \mathbf{x} = \lambda \begin{pmatrix} (\hat{\mathbf{s}}_2)_x(\mathbf{x}_S) \\ (\hat{\mathbf{s}}_2)_y(\mathbf{x}_S) \end{pmatrix} + \begin{pmatrix} (\mathbf{s}_1)_x(\mathbf{x}) \\ (\mathbf{s}_1)_y(\mathbf{x}) \end{pmatrix}. \quad (3.6)$$

This defines a relation between the surface $z(\mathbf{x})$, the surface gradient and the target coordinates. Considering the special case of collimated input beams with vanishing x - and y -components of the input direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$, analogous equations were derived by Wu et al. [71] and Olikier [72]. Consequently, Eq. (3.2), Eq. (3.6) build a PDE system for the unknown functions $u_x(\mathbf{x})$, $u_y(\mathbf{x})$ and $z(\mathbf{x})$.

According to the definition of the surface normal vector field $\mathbf{n}(\mathbf{x})$ in Eq. (3.4), which is apparent in the deflected vector field $\hat{\mathbf{s}}_2(\mathbf{x})$ [Eq. (3.6)], the remaining difficulty is the dependency on the surface intersection points $\mathbf{x}_S(\mathbf{x})$. These are determined by the freeform surface $z(\mathbf{x})$ itself and therefore unknown for noncollimated input beams. In their work on the reconstruction of optical surfaces from experimental ray data [94], a similar problem was solved by the authors using the intercept theorem

$$\frac{\mathbf{x}_S - \mathbf{x}}{z(\mathbf{x}_S) - z_0} = \frac{1}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_y(\mathbf{x}) \end{pmatrix} \quad \Rightarrow \quad \mathbf{x}_S = \frac{z(\mathbf{x}_S) - z_0}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_y(\mathbf{x}) \end{pmatrix} + \mathbf{x}, \quad (3.7)$$

which relates the given ray directions $\hat{\mathbf{s}}_1(\mathbf{x})$ to the vector field $\mathbf{s}_1(\mathbf{x})$ in Eq. (3.4). By utilizing this coordinate transformation for the considered problem of single FID, the gradient $\nabla_S z(\mathbf{x}_S)$ with $\nabla_S \equiv (\partial_{x_S}, \partial_{y_S})^T$ and therefore the surface normal vector field $\hat{\mathbf{n}}(\mathbf{x}_S)$ in Eq. (3.3) can be expressed in terms of the initial coordinates \mathbf{x} on the input plane $z = z_0$. This is straightforwardly done by applying the chain rule to $\nabla_S z(\mathbf{x}_S)$, giving

$$\begin{aligned} \partial_{x_S} z(\mathbf{x}_S) &= \frac{(\partial_{y_S} z)(\mathbf{x}_S) [\partial_x z(\mathbf{x}_S)] - (\partial_x z)(\mathbf{x}_S) [\partial_y z(\mathbf{x}_S)]}{(\partial_x x_S)(\partial_y y_S) - (\partial_x y_S)(\partial_y x_S)} \\ \partial_{y_S} z(\mathbf{x}_S) &= \frac{-(\partial_y x_S) [\partial_x z(\mathbf{x}_S)] + (\partial_x x_S) [\partial_y z(\mathbf{x}_S)]}{(\partial_x x_S)(\partial_y y_S) - (\partial_x y_S)(\partial_y x_S)} \end{aligned} \quad (3.8)$$

The remaining derivatives of the freeform intersection coordinates $\partial_x \mathbf{x}_S$ and $\partial_y \mathbf{x}_S$ can be directly calculated from Eq. (3.7), which also solely includes partial derivatives of $z(\mathbf{x}_S)$ with respect to the initial x - y -coordinates. Consequently, the deflected vector field $\hat{\mathbf{s}}_2(\mathbf{x})$ and normal vector field $\mathbf{n}(\mathbf{x})$ at \mathbf{x}_S in Eq. (3.6) are expressed through $z(\mathbf{x}_S)$, $\partial_x z(\mathbf{x}_S)$ and $\partial_y z(\mathbf{x}_S)$. Hence, by defining

$$z_S(\mathbf{x}) \equiv z(\mathbf{x}_S), \quad (3.9)$$

the ray-mapping equation Eq. (3.6) is fully expressed in terms of the x - y -coordinates. As a result, the Jacobian equation (3.2) and Eq. (3.6) build a system of three PDEs for three

unknown functions $\mathbf{u}(\mathbf{x})$ and $z_S(\mathbf{x})$. After solving the equation system for these functions, the freeform surface $z(\mathbf{x})$, according to Eq. (3.9), is given on the scattered x_S - y_S -grid points, which can be calculated directly from Eq. (3.7). The surface $z(\mathbf{x})$ on the required grid is then derived by a scattered data interpolation.

3.1.2 PDE system and Monge-Ampère equation

To summarize the results of the previous section, it is stated that the single FID problem for zero-étendue sources is governed by the PDE system

$$[(\partial_x u_x)(\partial_y u_y) - (\partial_x u_y)(\partial_y u_x)] I_T(\mathbf{u}) = I_S(\mathbf{x})$$

$$\mathbf{u}(\mathbf{x}) - \mathbf{x} = \frac{z_T - z_S(\mathbf{x})}{(\hat{\mathbf{s}}_2)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_2)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_2)_y(\mathbf{x}) \end{pmatrix} + \frac{z_S(\mathbf{x}) - z_0}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_y(\mathbf{x}) \end{pmatrix}, \quad (3.10)$$

together with the transport boundary conditions $\mathbf{u}(\partial\Omega_S) = \partial\Omega_T$. Thereby $\hat{\mathbf{s}}_1(\mathbf{x})$ is the predefined input direction vector field and $\hat{\mathbf{s}}_2(\mathbf{x})$ is defined as in Eq. (3.3) with the normal vector field $\hat{\mathbf{n}}(\mathbf{x}_S)$ expressed in x - y -coordinates according to Eqs. (3.7) and (3.8).

For a given emittance and irradiance $I_S(\mathbf{x})$ and $I_T(\mathbf{x})$, and a given ray directions $\hat{\mathbf{s}}_1(\mathbf{x})$ the equation system in Eq. (3.10) needs to be solved simultaneously for the surface $z_S(\mathbf{x})$ and the ray mapping components $\mathbf{u}(\mathbf{x})$.

It is important to note that Eq. (3.10) can be further simplified by plugging $u_x(\mathbf{x})$ and $u_y(\mathbf{x})$ into the Jacobian equation. This leads to a nonlinear second order PDE of Monge-Ampère type for the function $z_S(\mathbf{x})$:

$$A\partial_{xx}z_S(\mathbf{x}) + 2B\partial_{xy}z_S(\mathbf{x}) + C\partial_{yy}z_S(\mathbf{x}) + [(\partial_{xx}z_S(\mathbf{x}))(\partial_{yy}z_S(\mathbf{x})) - (\partial_{xy}z_S(\mathbf{x}))^2] = D. \quad (3.11)$$

The coefficients A, B, C, D are hereby functions of $x, y, z_S(\mathbf{x}), \partial_x z_S(\mathbf{x})$ and $\partial_y z_S(\mathbf{x})$, which can be calculated straightforward. The concrete form of the MAE coefficients is thus determined by the input ray direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ and (3.11) consequently represents a unified MAE for zero-étendue wavefronts.

Equation (3.11) is a generalization of the MAE (2.22) for the potential $u(\mathbf{x})$ of the L_2 OMT. In chapter 4 it will be shown that the FID for collimated beams in a paraxial approximation reduces to the L_2 OMT problem, which is equivalent to solving the MAE (3.11) with $A \equiv B \equiv C \equiv 0$ and motivates the proposed numerical design algorithm.

Furthermore it is noteworthy that compared to the Eq. system (3.10), the MAE in (3.11) has numerically the advantage of a lesser number of design variables after its discretization. Despite this advantage, the numerical strategy presented in chapter 5 will nevertheless be based on the PDE system in Eq. (3.10). In contrast to the MAE in Eq. (3.11), the discretization of second order derivatives is therefore avoided. Furthermore, the transport boundary conditions $\mathbf{u}(\partial\Omega_S) = \partial\Omega_T$ can be directly controlled through the mapping $\mathbf{u}(\mathbf{x})$ by the requirement that boundary points of the support of the emittance $I_S(\mathbf{x})$ are mapped onto boundary points of the support of the irradiance $I_T(\mathbf{x})$. For the MAE on the other hand, the same goal has to be

achieved through the surface gradients $\nabla z_S(\mathbf{x})$ at the freeform surface boundary, which is defined by the input direction vector field on the boundary $\partial\Omega_S$ of $I_S(\mathbf{x})$. This leads to the dependency of the boundary conditions on the input field $\hat{\mathbf{s}}_1(\mathbf{x})$. Additionally, the properties of the Jacobian equation and the mapping equations in (3.10) can be controlled separately, which will play a key role for the numerical stability of the design algorithm presented in chapter 5.

3.1.3 Predefined exit surfaces

In the previous sections, a PDE model for the calculation of single freeform surfaces for irradiance control with given input wavefronts of zero-étendue light sources was derived. Consequently, this implies the possibility of designing freeform surfaces for predefined *entrance* surfaces. In the following, the model will be extended to single freeform lenses and mirrors with predefined *exit* surfaces $z_{pre}(\mathbf{x})$. For mirror systems or mirror-lens combinations, the law of refraction thereby simply needs to be replaced with by law of reflection.

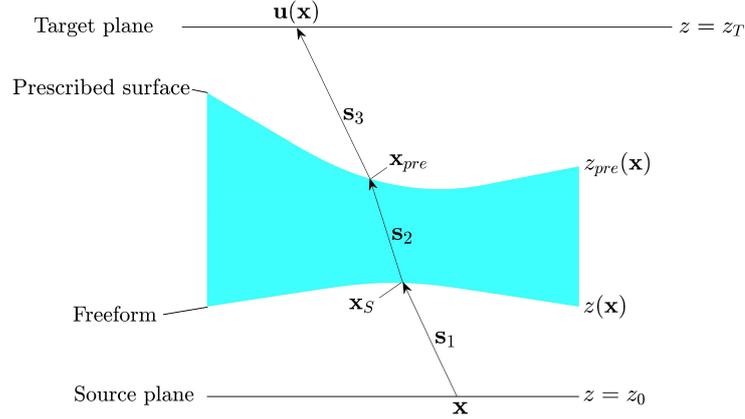


Figure 3.2: Design geometry of a single freeform lens $z(\mathbf{x})$ with a predefined exit surface $z_{pre}(\mathbf{x})$. The input direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ is refracted by the unknown freeform surface and the predefined surface to give the required irradiance distribution on $z = z_T$. Neither the freeform intersection coordinates $\mathbf{x}_S(\mathbf{x})$ nor the intersection coordinates $\mathbf{x}_{pre}(\mathbf{x})$ are known prior to the freeform calculation. [87]

Using the notation and the approach from the previous section, the target coordinates $\mathbf{u}(\mathbf{x})$ on the target plane $z = z_T$ from a ray send out from the coordinate \mathbf{x} on the input plane $z = z_0$ and the local energy conservation can be expressed as [Fig. 3.2]

$$\det(\nabla \mathbf{u}(\mathbf{x})) I_T(\mathbf{u}(\mathbf{x})) = I_S(\mathbf{x})$$

$$\mathbf{u}(\mathbf{x}) = \frac{z_T - z_{pre}(\mathbf{x}_{pre})}{(\hat{\mathbf{s}}_3)_z(\mathbf{x}_{pre})} \begin{pmatrix} (\hat{\mathbf{s}}_3)_x(\mathbf{x}_{pre}) \\ (\hat{\mathbf{s}}_3)_y(\mathbf{x}_{pre}) \end{pmatrix} + \frac{z_{pre}(\mathbf{x}_{pre}) - z_S(\mathbf{x})}{(\hat{\mathbf{s}}_2)_z(\mathbf{x}_S)} \begin{pmatrix} (\hat{\mathbf{s}}_2)_x(\mathbf{x}_S) \\ (\hat{\mathbf{s}}_2)_y(\mathbf{x}_S) \end{pmatrix} + \frac{z_S(\mathbf{x}) - z_0}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_y(\mathbf{x}) \end{pmatrix} + \mathbf{x} \quad (3.12)$$

with $z_S(\mathbf{x})$ defined according to Eq. (3.9). As a consequence, compared to Eq. (3.6) this leads to an extra term in the ray-mapping equation and to the dependency on the ray intersection points $\mathbf{x}_{pre}(\mathbf{x}) = (x_{pre}(\mathbf{x}), y_{pre}(\mathbf{x}))$ of the predefined surface $z_{pre}(\mathbf{x})$. In contrast to section 3.1.1,

therefore not only the normalized ray direction vector field $\hat{\mathbf{s}}_2(\mathbf{x})$ between the freeform and the lens exit surface, but also the vector field $\hat{\mathbf{s}}_3(\mathbf{x})$ between the lens exit surface and the target plane depends on unknown surface intersection points. By applying similar techniques as in section 3.1.1, these can again be expressed in terms of the initial coordinates \mathbf{x} . For the given input ray directions $\hat{\mathbf{s}}_1(\mathbf{x})$ and two refractive surfaces this is done by using

$$\begin{aligned} \hat{\mathbf{s}}_3(\mathbf{x}_{\text{pre}}) &= \frac{1}{n} \hat{\mathbf{s}}_2(\mathbf{x}_S) + \left\{ -\frac{1}{n} \hat{\mathbf{n}}_{\text{pre}}(\mathbf{x}_{\text{pre}}) \cdot \hat{\mathbf{s}}_2(\mathbf{x}_S) + \sqrt{1 - \frac{1}{n^2} [1 - (\hat{\mathbf{n}}_{\text{pre}}(\mathbf{x}_{\text{pre}}) \cdot \hat{\mathbf{s}}_2(\mathbf{x}_S))^2]} \right\} \hat{\mathbf{n}}_{\text{pre}}(\mathbf{x}_{\text{pre}}) \\ \hat{\mathbf{s}}_2(\mathbf{x}_S) &= n \hat{\mathbf{s}}_1(\mathbf{x}) + \left\{ -n \hat{\mathbf{n}}(\mathbf{x}_S) \cdot \hat{\mathbf{s}}_1(\mathbf{x}) + \sqrt{1 - n^2 [1 - (\hat{\mathbf{n}}(\mathbf{x}_S) \cdot \hat{\mathbf{s}}_1(\mathbf{x}))^2]} \right\} \hat{\mathbf{n}}(\mathbf{x}_S) \end{aligned} \quad (3.13)$$

with $n = n_1/n_2$ and $\hat{\mathbf{s}}_2(\mathbf{x}_S) \equiv \hat{\mathbf{s}}_2(\mathbf{x}_{\text{pre}})$. The surface normal vector field $\hat{\mathbf{n}}(\mathbf{x}_S)$ of the freeform surface $z_S(\mathbf{x}) \equiv z(\mathbf{x}_S)$ is hereby written in terms of the initial coordinates \mathbf{x} by the coordinate transformation (3.7) as demonstrated in section 3.1.1.

As pointed out above, compared to the single freeform design problem (3.10), Eq. (3.12) depends on the unknown intermediate coordinates \mathbf{x}_{pre} through $z_{\text{pre}}(\mathbf{x}_{\text{pre}})$ and the surface normal vector field $\hat{\mathbf{n}}_{\text{pre}}(\mathbf{x}_{\text{pre}})$. Therefore, these points \mathbf{x}_{pre} are fixed through the law of refraction $\mathbf{n}(\mathbf{x}_{\text{pre}}) = n_1 \hat{\mathbf{s}}_2(\mathbf{x}) - n_2 \hat{\mathbf{s}}_3(\mathbf{u}(\mathbf{x}))$ [Eq. (2.5)] by applying the constraints

$$\left(\begin{array}{c} \partial_{x_{\text{pre}}} z_{\text{pre}}(\mathbf{x}_{\text{pre}}) \\ \partial_{y_{\text{pre}}} z_{\text{pre}}(\mathbf{x}_{\text{pre}}) \end{array} \right) \stackrel{!}{=} \frac{1}{n_1 (\hat{\mathbf{s}}_2)_z(\mathbf{x}) - n_2 (\hat{\mathbf{s}}_3)_z(\mathbf{u}(\mathbf{x}))} \left(\begin{array}{c} n_1 (\hat{\mathbf{s}}_2)_x(\mathbf{x}) - n_2 (\hat{\mathbf{s}}_3)_x(\mathbf{u}(\mathbf{x})) \\ n_1 (\hat{\mathbf{s}}_2)_y(\mathbf{x}) - n_2 (\hat{\mathbf{s}}_3)_y(\mathbf{u}(\mathbf{x})) \end{array} \right) \quad (3.14)$$

with

$$\mathbf{s}_2(\mathbf{x}) = \left(\begin{array}{c} x_{\text{pre}} - (z_S(\mathbf{x}) - z_0) \frac{(\hat{\mathbf{s}}_1)_x(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} - x \\ y_{\text{pre}} - (z_S(\mathbf{x}) - z_0) \frac{(\hat{\mathbf{s}}_1)_y(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} - y \\ z_{\text{pre}}(\mathbf{x}_{\text{pre}}) - z_S(\mathbf{x}) \end{array} \right), \quad \mathbf{s}_3(\mathbf{u}(\mathbf{x})) = \left(\begin{array}{c} u_x - x_{\text{pre}} \\ u_y - y_{\text{pre}} \\ z_T - z_{\text{pre}}(\mathbf{x}_{\text{pre}}) \end{array} \right). \quad (3.15)$$

Thus, Eq. (3.14) represents a pair of nonlinear equations for every coordinate \mathbf{x}_{pre} or ray, respectively. Hence, the considered design problem is specified by the PDE system in Eq. (3.12) together with the constraints (3.14) and the transport boundary conditions $\mathbf{u}(\partial\Omega_S) = \partial\Omega_T$.

The given formulation can be extended to an arbitrary number of exit surfaces. This can simply be done by adding a corresponding term in Eq. (3.12) for every predefined surface and applying equivalent constraints (3.14) to fix all the unknown predefined surface intersection points. Since only the definition of the input direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$, which is hitting the freeform surface, is required, it of course leaves the possibility to define entrance *and* exit surfaces simultaneously.

Due to its importance, not only for FID, but also for ray aiming and noise-free irradiance computation for extended sources with measurement data, the explicit analytical formulation of the freeform design problem in an optical system will be given in section 3.3.

3.2 Double freeform design: irradiance and phase control

Hereafter, the inverse problem of FID for irradiance and phase control is considered. This means that not only a required irradiance distribution needs to be realized by the optical system but also a predefined output wavefront. The necessary degrees of freedom for fixing the outgoing ray directions are thereby introduced by an additional freeform surface.

As pointed out above, the available double FID models in literature are either restricted to the paraxial regime [49, 54] or require at least one planar wavefront [66, 67, 74, 78, 84]. In the following, it is shown that the single FID model can be extended to double freeform surfaces for irradiance and phase control for arbitrary input *and* output wavefronts [86, 87]. The key steps to this extension are the elimination of one unknown freeform surface in the PDE model through the constant OPL condition and the projection of the ray mapping coordinates onto the output wavefront, thus leading to a PDE system for one of the freeform surfaces and the *projected* mapping coordinates. As it will be seen, in contrast to the single FID, the corresponding PDE system of Eq. (3.10) does not reduce to a MAE for the freeform surface only, if *both* wavefronts are nonplanar.

3.2.1 Generalization to double freeform surfaces

The geometry of the considered double FID problem with the input wavefront $\phi_I(\mathbf{x})$, the output wavefront $\phi_O(\mathbf{x})$ and the freeform surfaces $z_I(\mathbf{x})$ and $z_{II}(\mathbf{x})$ is presented in Fig. 3.3. While the figure shows the design geometry for the example of a double *mirror* system, the derivations presented below are also valid for double freeform lens systems and lens-mirror combinations. The difference is thereby the replacement of the refractive indices for mirrors and the substitution of the deflected vector field $\hat{\mathbf{s}}_2(\mathbf{x})$ according to raytracing equations (3.3).

In accordance with Fig. 3.3, the geometry of the design problem can be expressed through the vector fields

$$\mathbf{s}_I(\mathbf{x}) = \begin{pmatrix} x_I - x \\ y_I - y \\ z_I(\mathbf{x}_I) - z_0 \end{pmatrix}, \quad \mathbf{s}_3(\mathbf{u}) = \begin{pmatrix} u_x(\mathbf{x}) - x_{II} \\ u_y(\mathbf{x}) - y_{II} \\ z_T - z_{II}(\mathbf{x}_{II}) \end{pmatrix}, \quad \mathbf{n}_I(\mathbf{x}_I) = \begin{pmatrix} -\partial_{x_I} z_I(\mathbf{x}_I) \\ -\partial_{y_I} z_I(\mathbf{x}_I) \\ 1 \end{pmatrix}, \quad (3.16)$$

with the surface intersection points

$$\mathbf{x}_I = \frac{z_{I,S}(\mathbf{x}) - z_0}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_y(\mathbf{x}) \end{pmatrix} + \mathbf{x}, \quad \mathbf{x}_{II} = \frac{z_{II,S}(\mathbf{x}) - z_T}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} \begin{pmatrix} (\hat{\mathbf{s}}_3)_x(\mathbf{u}) \\ (\hat{\mathbf{s}}_3)_y(\mathbf{u}) \end{pmatrix} + \mathbf{u}. \quad (3.17)$$

It was thereby defined $z_{I,S}(\mathbf{x}) \equiv z_I(\mathbf{x}_I)$ and $z_{II,S}(\mathbf{x}) \equiv z_{II}(\mathbf{x}_{II})$ in agreement with the definition (3.9).

For given input and output wavefronts $\phi_I(\mathbf{x})$ and $\phi_O(\mathbf{x})$, the corresponding normalized input and output ray direction vector fields $\hat{\mathbf{s}}_1(\mathbf{x})$ and $\hat{\mathbf{s}}_3(\mathbf{x})$ on the input plane $z = z_0$ and target plane

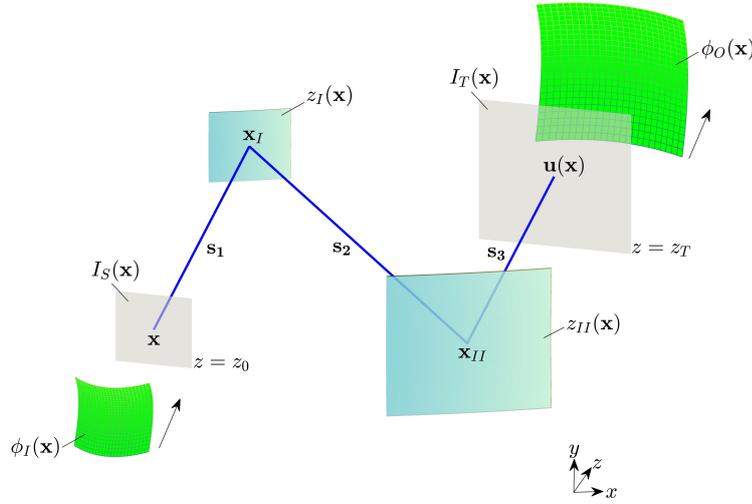


Figure 3.3: Geometry of the design problem for the example of double freeform *mirrors*. The normalized input ray direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ on $z = z_0$ is defined by predefined input wavefront $\phi_I(\mathbf{x})$. This input wavefront and the emittance $I_S(\mathbf{x})$ on $z = z_0$ need to be redistributed by the freeform surfaces $z_I(\mathbf{x})$ and $z_{II}(\mathbf{x})$ to give the required irradiance $I_T(\mathbf{x})$ on $z = z_T$ and output wavefront $\phi_O(\mathbf{x})$. For general, nonplanar wavefronts and unknown surfaces the intermediate surface intersection coordinates \mathbf{x}_I and \mathbf{x}_{II} are a priori unknown. [86]

$z = z_T$ can directly be calculated from the wavefront gradients. With these vector fields, the explicit expression of the analogues for double freeform surfaces of the ray-mapping equations in Eq. (3.10) and (3.12) can be derived straightforwardly, which gives

$$\mathbf{u}(\mathbf{x}) - \mathbf{x} = \frac{z_T - z_{II,S}(\mathbf{x})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} \begin{pmatrix} (\hat{\mathbf{s}}_3)_x(\mathbf{u}) \\ (\hat{\mathbf{s}}_3)_y(\mathbf{u}) \end{pmatrix} + \frac{z_{II,S}(\mathbf{x}) - z_{I,S}(\mathbf{x})}{(\hat{\mathbf{s}}_2)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_2)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_2)_y(\mathbf{x}) \end{pmatrix} + \frac{z_{I,S}(\mathbf{x}) - z_0}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_y(\mathbf{x}) \end{pmatrix}. \quad (3.18)$$

While $\hat{\mathbf{s}}_1(\mathbf{x})$ and $\hat{\mathbf{s}}_3(\mathbf{x})$ are predefined, the deflected vector field $\hat{\mathbf{s}}_2(\mathbf{x})$ is expressed by the ray tracing equations in Eq. (3.3) and the coordinate transformation (3.7) or (3.17) through $\hat{\mathbf{s}}_1(\mathbf{x})$ and freeform surface normals $\hat{\mathbf{n}}_I(\mathbf{x})$ as shown in section 3.1.1.

Hence, the major difference to the single FID case is therefore the presence of the $\hat{\mathbf{s}}_3(\mathbf{x})$ -term, which depends on the ray-mapping $\mathbf{u}(\mathbf{x})$ itself, and moreover, the coupling to the also unknown second freeform surface $z_{II,S}(\mathbf{x})$ in the $\hat{\mathbf{s}}_2(\mathbf{x})$ -term.

It is important to note that Eq. (3.18) is valid for two-mirror system, lens-mirror combinations, single lens systems with two freeform surfaces and two-lens systems with the target plane *within* the medium of the second lens. For a two-lens systems with a finite working distance relative to the exit surface of the second freeform lens, Eq. (3.18) can be generalized straightforwardly, which leads to an additional term in the ray-mapping equation as it is shown in Appendix A.3. Alternatively, it is also possible to use Eq. (3.18) directly for the two-lens system and propagate the predefined irradiance and given output wavefront into the second lens, which is of course possible for an arbitrary number of predefined exit surfaces. The resulting intermediate $I_T(\mathbf{x})$ and $\phi_O(\mathbf{x})$ are then used for the calculations. This on the other hand might yield to addi-

tional difficulties, since depending on the output wavefront, this can lead to more complicated boundary shapes of the intermediate target irradiance and consequently to a more complicated implementation of the transport boundary conditions.

3.2.2 Constant optical path length condition

As noted in section 3.2.1, additional difficulties in the design process compared to the single FID are due to the presence of both freeform surfaces $z_{I,S}(\mathbf{x})$ and $z_{II,S}(\mathbf{x})$ in Eq. (3.18). By applying the constant OPL condition [section 2.1.2], the dependency of Eq. (3.18) on one of the unknown freeform surfaces can be eliminated. Therefore,

$$OPL = n_1|\mathbf{s}_1^I| + n_2|\mathbf{s}_2| + n_1|\mathbf{s}_3^O| \quad (3.19)$$

with the vector fields [Fig. 3.4]

$$\mathbf{s}_1^I = \begin{pmatrix} x_I - x^I \\ y_I - y^I \\ z_{I,S}(\mathbf{x}) - \phi_I(\mathbf{x}^I) \end{pmatrix}, \quad \mathbf{s}_3^O = \begin{pmatrix} u_x^O - x_{II} \\ u_y^O - y_{II} \\ \phi_O(\mathbf{u}^O) - z_{II,S}(\mathbf{x}) \end{pmatrix} \quad (3.20)$$

is considered. The Eq. (3.19) can be solved analytically for the second freeform $z_{II,S}(\mathbf{x})$ using e.g. *Wolfram Mathematica*, which will give several physical and unphysical solutions. These will depend on the respective system type (mirror, lens) and on the ratio of the refractive index n . The appropriate for $z_{II,S}(\mathbf{x})$ is then plugged into (3.18), by which the Jacobian equation (3.2) and the ray-mapping equation (3.18) reduce to a coupled system of three nonlinear PDEs.

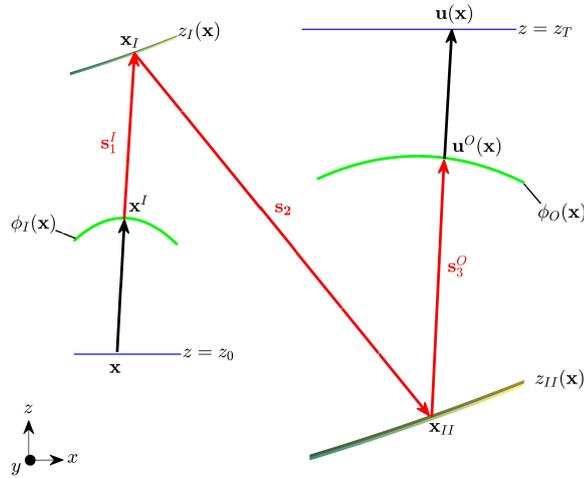


Figure 3.4: The constant OPL in Eq. (3.19) is defined through the vector fields \mathbf{s}_1^I and \mathbf{s}_3^O between both wavefronts and the freeform surfaces. The mapping $\mathbf{u}(\mathbf{x})$ on $z = z_T$ can be projected onto the output wavefront by using the corresponding ray directions in the target plane. [86]

3.2.3 Wavefront Mapping Coordinates and PDE System

In the previous section, the second freeform $z_{II,S}(\mathbf{x})$ was eliminated from the ray-mapping equation (3.18). Thereby, the dependency on the projected mapping coordinates $\mathbf{u}^O(\mathbf{x})$ through the output wavefront $\phi_O(\mathbf{u}^O)$ was introduced (see Eq. (3.20) and Fig. 3.4), which represents the major difference of the double FID problem compared to the single FID. Hence, the PDE system in Eqs. (3.2) and (3.18) can not be solved directly for the mapping $\mathbf{u}(\mathbf{x})$ and surface $z_{I,S}(\mathbf{x})$. However, by using the relation [Fig. 3.4]

$$\mathbf{u} = \frac{z_T - \phi_O(\mathbf{u}^O)}{(\hat{\mathbf{s}}_3)_z(\mathbf{u}^O)} \begin{pmatrix} (\hat{\mathbf{s}}_3)_x(\mathbf{u}^O) \\ (\hat{\mathbf{s}}_3)_y(\mathbf{u}^O) \end{pmatrix} + \mathbf{u}^O, \quad (3.21)$$

which depicts a one-to-one correspondence between the mapping $\mathbf{u}(\mathbf{x})$ and $\mathbf{u}^O(\mathbf{x})$ a PDE system for the projected mapping coordinates $u_x^O(\mathbf{x}), u_y^O(\mathbf{x})$ and $z_{I,S}(\mathbf{x})$ can be derived instead. To achieve this, Eq. (3.21) is plugged into Eqs. (3.2) and (3.18). As a result, it follows a PDE system of the form

$$\begin{aligned} f(\mathbf{u}^O, \nabla u_x^O, \nabla u_y^O) I_T^O(\mathbf{u}^O) &= I_S(\mathbf{x}), \\ \mathbf{u}^O(\mathbf{x}) - \mathbf{x} &= \mathbf{f}(z_{I,S}, \nabla z_{I,S}, \phi_O(\mathbf{u}^O), \hat{\mathbf{s}}_3(\mathbf{u}^O)). \end{aligned} \quad (3.22)$$

This PDE system together with the boundary conditions for $\mathbf{u}^O(\mathbf{x})$, which follow directly from $\mathbf{u}(\partial\Omega_S) = \partial\Omega_T$ and Eq. (3.21) then needs to be solved for the unknown functions. Thereby, the irradiance $I_T(\mathbf{u})$ was redefined through the projected mapping by $I_T^O(\mathbf{u}^O)$.

It is important to note that Eq. (3.22) is not stated explicitly here due to its length and due to the simplicity of its derivation: this is done by plugging at first the analytical solution for $z_{II,S}(\mathbf{x})$ from the OPL condition into Eq. (3.18) and then Eq. (3.21) into the Jacobian equation (3.2) and the ray-mapping equations (3.18).

The PDE model for a double lens with two freeform surfaces is described in Appendix A.3.

3.3 Formulation of freeform illumination design problem in optical systems

As pointed out in section 3.1.3, the presented PDE model is not only applicable to FID but also to ray aiming and irradiance computation. Due to its diverse application possibilities, a summary of the results of the previous sections will be given and the FID problem will be stated explicitly for freeform surfaces in optical systems. To do so, at first an analytical description of a (sequential) optical ray path in an optical system of N predefined surfaces is given. The model is thereby characterized through a simultaneous description of every surface intersection point of a ray through nonlinear equations. Secondly, the description is used to model the FID for irradiance or irradiance and phase control in a system of predefined surfaces. The formulation

of sections 3.1 and 3.2 is thereby slightly modified so that every surface in the system refers to the same input aperture instead of an aperture directly between the last entrance surface and the freeform. Additionally, in Appendix A.4 it is argued that the model also directly describes the coupling of N input wavefronts through N surfaces.

3.3.1 Describing a ray-path in an optical system

In the following, a description scheme for a single (sequential) ray-path in an optical system by a system of nonlinear equations is presented, which relates the ray intersection points of every considered surface simultaneously. It represents a generalization of the ray-mapping equations (3.6), (3.12) and (3.18) and the application of the constraints in Eqs. (3.14) (or (A.5)) to every predefined exit surface.

It is assumed that an optical system of N lens and/or mirror surfaces $z_i(\mathbf{x})$, $i = 1, \dots, N$ is given together with the input surface $z_0(\mathbf{x})$ at which the input distributions of the system are defined and the target surface $z_{N+1}(\mathbf{x})$ (see Fig. 3.5).

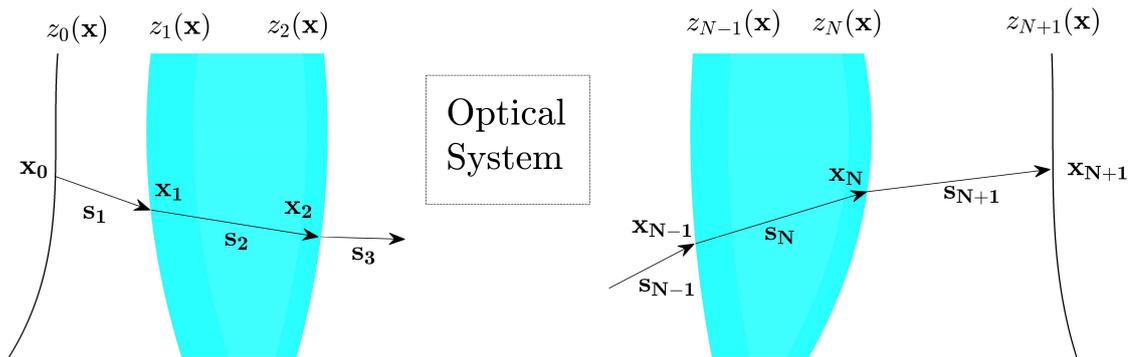


Figure 3.5: Ray path in an optical system between the surfaces $z_0(\mathbf{x})$ and $z_{N+1}(\mathbf{x})$. The ray path is defined by the given input direction $\mathbf{s}_1(\mathbf{x}_0)$ and the surface intersection coordinates $\mathbf{x}_i(\mathbf{x})$, $i = 1, \dots, N+1$. [88]

As shown in the previous sections, the ray path of a ray between the input coordinate $\mathbf{x} = \mathbf{x}_0$ and the respective target coordinate $\mathbf{x}_{N+1}(\mathbf{x})$ can be described by the law of refraction/reflection and the application of the intercept theorem. For a predefined normalized input ray direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ on the input surface $z_0(\mathbf{x})$, the ray intersection coordinate \mathbf{x}_{N+1} on the target surface $z_{N+1}(\mathbf{x})$ of a ray send out from $(\mathbf{x}_0, z_0(\mathbf{x}_0))$ can thereby be represented as

$$\mathbf{x}_{N+1}(\mathbf{x}_0) = \sum_{i=0}^N \frac{z_{i+1}(\mathbf{x}_{i+1}) - z_i(\mathbf{x}_i)}{(\hat{\mathbf{s}}_{i+1})_z(\mathbf{x}_i)} \begin{pmatrix} (\hat{\mathbf{s}}_{i+1})_x(\mathbf{x}_i) \\ (\hat{\mathbf{s}}_{i+1})_y(\mathbf{x}_i) \end{pmatrix} + \mathbf{x}_0. \quad (3.23)$$

with the surface intersection coordinates $\mathbf{x}_i(\mathbf{x}_0)$, $i = 1, \dots, N$. The deflected vector fields $\hat{\mathbf{s}}_{i+1}(\mathbf{x}_i)$, $i = 2, \dots, N$ can then be written iteratively in terms of the predefined input vector field $\hat{\mathbf{s}}_1(\mathbf{x})$, the initial ray coordinates $(\mathbf{x}_0, z_0(\mathbf{x}_0))$ and the predefined surfaces $z_i(\mathbf{x})$, $i = 1, \dots, N$ through the ray tracing equations

$$\hat{\mathbf{s}}_{i+1}(\mathbf{x}_i) = \frac{n_i}{n_{i-1}}\hat{\mathbf{s}}_i(\mathbf{x}_{i-1}) + \left\{ -\frac{n_i}{n_{i-1}} \cdot \hat{\mathbf{n}}_i(\mathbf{x}_i) \cdot \hat{\mathbf{s}}_i(\mathbf{x}_{i-1}) + \sqrt{1 - \left(\frac{n_i}{n_{i-1}}\right)^2 [1 - (\hat{\mathbf{n}}_i(\mathbf{x}_i) \cdot \hat{\mathbf{s}}_i(\mathbf{x}_{i-1}))^2]} \right\} \hat{\mathbf{n}}_i(\mathbf{x}_i),$$

$$\hat{\mathbf{s}}_{i+1}(\mathbf{x}_i) = \hat{\mathbf{s}}_i(\mathbf{x}_{i-1}) - 2[\hat{\mathbf{n}}_i(\mathbf{x}_i) \cdot \hat{\mathbf{s}}_i(\mathbf{x}_{i-1})]\hat{\mathbf{n}}_i(\mathbf{x}_i), \quad (3.24)$$

with the surface normal vector fields

$$\mathbf{n}_i(\mathbf{x}_i) = \begin{pmatrix} -\nabla_{\mathbf{x}_i} z_i(\mathbf{x}_i) \\ 1 \end{pmatrix}, \quad \nabla_{\mathbf{x}_i} \equiv (\partial_{x_i}, \partial_{y_i})^T \quad (3.25)$$

and the refractive indices $n_i, i = 1, \dots, N$. Hence Eq. (3.23), expresses the target coordinate in terms of the initial coordinate \mathbf{x}_0 of the given input ray direction $\hat{\mathbf{s}}_1(\mathbf{x}_0)$ and the surface intersection coordinates $\mathbf{x}_i(\mathbf{x}_0), i = 1, \dots, N$. Since the surface intersection coordinates $\mathbf{x}_i(\mathbf{x}_0)$ themselves are unknown, Eq. (3.23) is not sufficient to determine the complete ray path in an optical system. Analogous to Eq. (3.14), therefore additional constraints need to be applied to determine the surface intersection coordinates. Assuming a pair of given intersection coordinates $\mathbf{x}_{i+1}(\mathbf{x}_0)$ and $\mathbf{x}_{i-1}(\mathbf{x}_0)$, the coordinate $\mathbf{x}_i(\mathbf{x}_0)$ is determined by the law of refraction/reflection by applying the surface normal constraints

$$\nabla_{\mathbf{x}_i} z_i(\mathbf{x}_i) = \frac{1}{n_{i-1}(\hat{\mathbf{s}}_i)_z(\mathbf{x}_{i-1}) - n_i(\hat{\mathbf{s}}_{i+1})_z(\mathbf{x}_i)} \cdot \begin{pmatrix} n_{i-1}(\hat{\mathbf{s}}_i)_x(\mathbf{x}_{i-1}) - n_i(\hat{\mathbf{s}}_{i+1})_x(\mathbf{x}_i) \\ n_{i-1}(\hat{\mathbf{s}}_i)_y(\mathbf{x}_{i-1}) - n_i(\hat{\mathbf{s}}_{i+1})_y(\mathbf{x}_i) \end{pmatrix} \quad (3.26)$$

with the vector fields

$$\mathbf{s}_{i+1}(\mathbf{x}_i) = \begin{pmatrix} x_{i+1} - x_i \\ y_{i+1} - y_i \\ z_{i+1}(\mathbf{x}_{i+1}) - z_i(\mathbf{x}_i) \end{pmatrix}, \quad i = 1, \dots, N \quad (3.27)$$

to every surface. Equation (3.26), represents a pair of nonlinear equations for every coordinate $\mathbf{x}_i(\mathbf{x}_0), i = 1, \dots, N$. Since every ray in an optical system is described through the Eqs. (3.23) and (3.26), these equations induce coordinate transformations between the rays on the input and target surface or object and image space, respectively [section 2.2.1].

It is important to note that the equation system (3.23) and (3.26) not only applies to FID but that it also serves as an elegant and efficient way for doing ray-aiming and noise free irradiance computation as shown in section 5.5. This can be used to generate the emittance $I_S(\mathbf{x})$ and the input ray directions for the single FID problem (3.10) and double FID problem (3.22) with a nontrivial entrance surface and is applied for the design examples in chapter 6.

3.3.2 Freeform illumination design in optical systems

Hereafter, the conventions from sections 3.1 and 3.2 are used, the initial and target points are denoted by $\mathbf{x} \equiv \mathbf{x}_0$ and $\mathbf{u}(\mathbf{x}) \equiv \mathbf{x}_{N+1}(\mathbf{x})$, and plane source and target surfaces $z_0(\mathbf{x}) \equiv z_0$ and $z_{N+1}(\mathbf{x}) \equiv z_T$ are considered.

From the viewpoint of FID, the Eqs. (3.2) and (3.23) together with the constraints Eq. (3.26) build a PDE system for the irradiance control with a single freeform surface $z_{FF}(\mathbf{x}_j) \equiv z_j(\mathbf{x}_j)$ at an arbitrary position $j = 1, \dots, N$ in a system of predefined surfaces. In contrast the previous formulations from sections 3.1 and 3.2, in which the input aperture is placed directly between the last entrance surface and the freeform surface, the PDE system here refers to the same aperture and input vector field $\hat{\mathbf{s}}_1(\mathbf{x})$, independent of the chosen position of the freeform in the optical system.

Assuming a given emittance $I_S(\mathbf{x})$ and vector field $\hat{\mathbf{s}}_1(\mathbf{x})$, and a required irradiance distribution $I_T(\mathbf{x})$, the PDE system (3.2) and (3.23) together with the constraints (3.26) needs to be solved simultaneously for the surface $z_{FF}(\mathbf{x}_j)$, the mapping $\mathbf{u}(\mathbf{x})$ and the surface intersection coordinates $\mathbf{x}_i(\mathbf{x})$ of the predefined surfaces $z_i(\mathbf{x})$ with $i = 1, \dots, N; i \neq j$.

In the design models from sections 3.1 and 3.2, an important point was the dependency of the unknown freeform surface $z_{FF}(\mathbf{x})$ and its normal vector field $\mathbf{n}_{FF}(\mathbf{x}) \equiv \mathbf{n}_j(\mathbf{x})$, apparent in the deflection vector field $\hat{\mathbf{s}}_{j+1}(\mathbf{x}_j)$, on the unknown freeform intersection point $\mathbf{x}_j(\mathbf{x})$. As described before, this freeform normal vector field, according to Eq. (3.25), needs to be evaluated at these intersection points $\mathbf{x}_j(\mathbf{x})$. This dependency can be eliminated by representing $\mathbf{x}_j(\mathbf{x})$ analogous to Eq. (3.7) in terms of the initial coordinates \mathbf{x} through

$$\mathbf{x}_j(\mathbf{x}) = \frac{z_{FF}(\mathbf{x}_j) - z_{j-1}(\mathbf{x}_{j-1})}{(\hat{\mathbf{s}}_j)_z(\mathbf{x}_{j-1})} \begin{pmatrix} (\hat{\mathbf{s}}_j)_x(\mathbf{x}_{j-1}) \\ (\hat{\mathbf{s}}_j)_y(\mathbf{x}_{j-1}) \end{pmatrix} + \sum_{i=0}^{j-2} \frac{z_{i+1}(\mathbf{x}_{i+1}) - z_i(\mathbf{x}_i)}{(\hat{\mathbf{s}}_{i+1})_z(\mathbf{x}_i)} \begin{pmatrix} (\hat{\mathbf{s}}_{i+1})_x(\mathbf{x}_i) \\ (\hat{\mathbf{s}}_{i+1})_y(\mathbf{x}_i) \end{pmatrix} + \mathbf{x}. \quad (3.28)$$

Following the steps from Eq. (3.7) to Eq. (3.9), the chain rule is applied to this coordinate transformation and $z_{FF}(\mathbf{x}_j) \equiv z_{FF,S}(\mathbf{x})$ defined, so that x and y component of the normal vector field $\mathbf{n}_{FF}(\mathbf{x}_j)$ can be written as

$$\nabla_{\mathbf{x}_j} z_{FF,S}(\mathbf{x}) = \mathbf{f}(\partial_x z_{FF,S}(\mathbf{x}), \partial_y z_{FF,S}(\mathbf{x})). \quad (3.29)$$

The intersection coordinates $\mathbf{x}_i(\mathbf{x})$, $i = 1, \dots, j-1$ before the freeform surface can be determined by a simple forward raytracing from the initial ray coordinates \mathbf{x} . This leads to the PDE system of the form

$$\begin{aligned} \mathbf{u}(\mathbf{x}) &= \mathbf{f}(z_{FF,S}(\mathbf{x}), \partial_x z_{FF,S}(\mathbf{x}), \partial_y z_{FF,S}(\mathbf{x}), \mathbf{x}_{j+1}(\mathbf{x}), \dots, \mathbf{x}_N(\mathbf{x})) \\ \det(\nabla \mathbf{u}(\mathbf{x})) I_T(\mathbf{u}(\mathbf{x})) &= I_S(\mathbf{x}) \\ BC : \mathbf{u}(\partial \Omega_S) &= \partial \Omega_T \end{aligned} \quad (3.30)$$

with the constraints

$$\nabla_{\mathbf{x}_i} z_i(\mathbf{x}_i) = \frac{n_{i-1} \hat{\mathbf{s}}_i^\perp(\mathbf{x}_{i-1}) - n_i \hat{\mathbf{s}}_{i+1}^\perp(\mathbf{x}_i)}{(n_{i-1} \hat{\mathbf{s}}_i(\mathbf{x}_{i-1}) - n_i \hat{\mathbf{s}}_{i+1}(\mathbf{x}_i))_z}, \quad i = j+1, \dots, N. \quad (3.31)$$

with \perp denoting the x - y -component of the vector fields. As in section 3.1, the *unknown* freeform surface intersection coordinate $\mathbf{x}_j(\mathbf{x})$ were eliminated in favor of the *known* initial coordinates

\mathbf{x} . The resulting PDE system therefore describes the single FID problem for placing a freeform at the j 's position in an optical system of N surfaces and has to be solved simultaneously for the surface intersection points $\mathbf{x}_{j+1}(\mathbf{x}), \dots, \mathbf{x}_N(\mathbf{x})$, the freeform surface $z_{FF,S}(\mathbf{x})$ and the mapping coordinates $\mathbf{u}(\mathbf{x})$. From a numerical viewpoint, the structure of the PDE system is therefore equivalent to Eqs. (3.10) and (3.22), which makes a numerical treatment possible as shown in chapter 5.

In case that not only the control of the irradiance but also the phase is required, the introduction of an additional freeform surface $z_{FF,II}(\mathbf{x}) \equiv z_k(\mathbf{x}_k)$ is necessary to provide the degrees of freedom for controlling also the outgoing ray directions of the optical system. A system equivalent to Eq. (3.30) is then straightforwardly derived by applying the constant optical path length condition

$$OPL = n_1|\mathbf{s}_1| + \dots + n_j|\mathbf{s}_j| + \dots + n_k|\mathbf{s}_k| + \dots + n_{N+1}|\mathbf{s}_{N+1}|. \quad (3.32)$$

This condition can always be solved analytically for one of the freeform surfaces, which then can be eliminated from the PDE system according to section 3.2.2. Since this introduces a dependency on the projected wavefront mapping coordinates, analogously to section 3.2.3 the mapping $\mathbf{u}(\mathbf{x})$ is then projected onto the predefined outgoing wavefront and eliminated as well from the corresponding PDE system.

For double freeform systems with predefined exit surfaces, the irradiance on $z = z_T$ can be propagated by the given output wavefront to a plane between the second freeform and the first exit surface to specify the intermediate irradiance and intermediate output wavefront. Alternatively, the surface intersection points are fixed through (3.31) and for the last exit surface the constraint is replaced by Eq. (A.5) from Appendix A.3.

Hence, it follows a PDE system for the unknown freeform surface and the wavefront mapping coordinates equivalent to (3.22). Its numerical solving process is described in chapter 5.

3.4 Summary

The FID problem for irradiance control with a single freeform and irradiance and phase control with a double freeform was formulated for zero-étendue light sources in terms of PDEs. In contrast to the single FID, the PDE system for the double FID does not reduce to a MAE for the freeform surface only if the input and output beam are both non-collimated. In both cases the specific form of the PDE system is thereby defined through the given ray direction vector fields. Consequently, the PDE systems (3.10) and (3.22) represent a unification and generalization of the available PDE models in literature, since all possible wavefronts are described by the same PDE system².

Furthermore it was shown that the design models can be generalized to the single FID and double FID in optical systems. Therefore, the freeform surfaces can be placed at an arbitrary position in the optical system, which makes a pre- and post beam shaping with predefined entrance and

²The extension of the single FID to non-planar target surfaces is described in Appendix A.2

exit surfaces possible.

It is important to note that the formulations (3.10), (3.22) and (3.30) are structurally similar in the sense that they consist of three first-order PDE's for three unknown functions: two mapping components and one freeform surface. This structural similarity will make it possible to solve the single FID and double FID problem with the similar numerical techniques as discussed in chapter 5.

Also the description of (sequential) ray-paths in optical systems, given in section 3.3.1, will be of importance for the generation of the input data of the design algorithms, like emittances and ray directions, as discussed in chapter 5.5.

4 L_2 optimal mass transport and freeform illumination design

In this chapter, a connection between the FID and L_2 OMT [section 2.2.2] is established by analytically showing that the L_2 OMT solves the FID problem for collimated beams and infinite freeform-target or freeform-freeform distances. This is of special importance since the efficient and stable numerical solving of PDEs or PDE systems, like in Eqs (3.10), (3.11), (3.22) or (3.30), requires a suitable initial iterate close to a root to ensure numerical convergence, as discussed in chapter 5. In contrast to the MAE (3.11) for single freeform surfaces, PDE systems like in Eq. (3.10) not only need be solved for the freeform surface but also simultaneously for the mapping components. While this, after the discretization, leads to a higher number of design variables compared to the MAE, it has the major advantage that the energy conservation equation in (3.10) is decoupled from the freeform surface and is independent of geometrical constraints like ray directions. Hence, by constructing a suitable mapping, which on the one hand already fulfills the energy conservation equation (3.2) and on the other hand has a physical meaningful relation to the FID problem, the numerical convergence can be significantly stabilized compared to solving the MAE model (3.11).

As stated above, one promising candidate for such a ray mapping can be calculated from L_2 OMT [section 2.2.2] due to several reasons. Firstly, this specific mapping is connected to the paraxial propagation of rays in a homogeneous medium [95]. Secondly, it is arguably the standard problem in numerical OMT and consequently provides us with various available numerical algorithms for the mapping calculation. And thirdly, it was successfully applied in publications on the FID for irradiance control [25, 46, 47, 50, 77] and for irradiance and phase control [49, 53], which inspired the following investigations. The focus of these publications thereby was the practical application of the L_2 OMT to numerical design of freeform surfaces, rather than the *theoretical* justification of the application of L_2 OMT to FID, which will be done in this chapter. This is of major importance as it is a priori not clear if and under which conditions for a given map, which fulfills the energy conservation equation (2.19), a corresponding freeform surface exists that maps $I_S(\mathbf{x})$ onto $I_T(\mathbf{x})$.

Therefore, in sections 4.1 and 4.2 it will be shown that the OMT mapping solves the single FID and double FID problem for collimated beams approximately in a paraxial regime, which is realized through a large distance between the freeform surface and the target distribution for single freeforms [82] or a large distance between the freeform surfaces for double FID [83, 84]. It is important to note that these results are in agreement with the findings by Doskolovich et al.

[55, 56] through a variational approach. Earlier attempts in Ref. [98] to clarify the question under which conditions the L_2 OMT mapping is integrable lack of correct mathematical derivations.¹ Furthermore, in section 4.2, for the case of double FID with collimated beams, a PDE system for the ray-mapping components only is derived, which consists of the energy conservation equation and a ray-mapping condition. Due to the explicitly present energy conservation equation and the independence from the surfaces it allows a stable numerical solution process with a lower number of design variables compared to Eq. (3.22) and hence results in a faster numerical convergence, but is only applicable to collimated beams. Additionally, it is shown that the OMT mapping solves the double freeform *mirror* design problem with collimated beams exactly, which is in agreement with the formal mathematical proof by Glimm and Olikier [43]. In section 4.3, the application of the L_2 OMT to more general design geometries in FID is discussed. The results of this chapter were published in references [82], [83] and [84].

4.1 Single freeform surface

In the following, the surface integrability condition [81], which is frequently utilized in FID [19, 47, 58, 76], and the laws of reflection and refraction will be applied to derive a condition, which an energy conserving ray mapping $\mathbf{u}(\mathbf{x})$ and a freeform surface $z(\mathbf{x})$ for collimated input beams need to fulfill to ensure surface continuity. Thereby, the propagation of the input beam is described through the direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ and the reflected/refracted vector field is denoted by $\hat{\mathbf{s}}_2(\mathbf{x})$. In that case, the deflection laws can be described through the freeform surface normal vector field [section 2.1.1]

$$\mathbf{n} = n_1 \hat{\mathbf{s}}_1 - n_2 \hat{\mathbf{s}}_2 \quad (4.1)$$

with the refractive indices n_1 of the lens and n_2 of the surrounding medium. For a freeform mirror it is $n_1 = n_2 = -1$. Since surface continuity is required for the freeform $z(\mathbf{x})$, the integrability condition [81]

$$\mathbf{n} \cdot (\nabla \times \mathbf{n}) = 0 \quad (4.2)$$

needs to be applied to the surface normal vector field. By using the explicit expressions for the vector fields

$$\mathbf{s}_1 = \begin{pmatrix} 0 \\ 0 \\ z(\mathbf{x}) - z_0 \end{pmatrix}, \quad \mathbf{s}_3 = \begin{pmatrix} u_x(\mathbf{x}) - x \\ u_y(\mathbf{x}) - y \\ z_T - z_0 \end{pmatrix}, \quad \mathbf{s}_2 = \mathbf{s}_3 - \mathbf{s}_1, \quad (4.3)$$

¹In Ref. [98] the authors consider a collimated input beam and describe the freeform surface by the vector $\mathbf{S} = (x_0, y_0, z_0)$ with x_0 and y_0 and z_0 being constants. From this, the authors conclude that the curl of this vector vanishes and neglect it in their derivations, which is incorrect. The correct description of the freeform surface is the vector field $\mathbf{S} = (0, 0, z(\mathbf{x}))$, whose curl does not vanish and therefore cannot be neglected in the derivations. The conclusions, given in Ref. [98], are therefore unjustified.

which describe the considered design geometry with a collimated input beam [Fig. 4.1 (a)], according to Eq. (4.1), the normal vector field $\mathbf{n}(\mathbf{x})$ can be written in terms of the ray mapping and the surface.

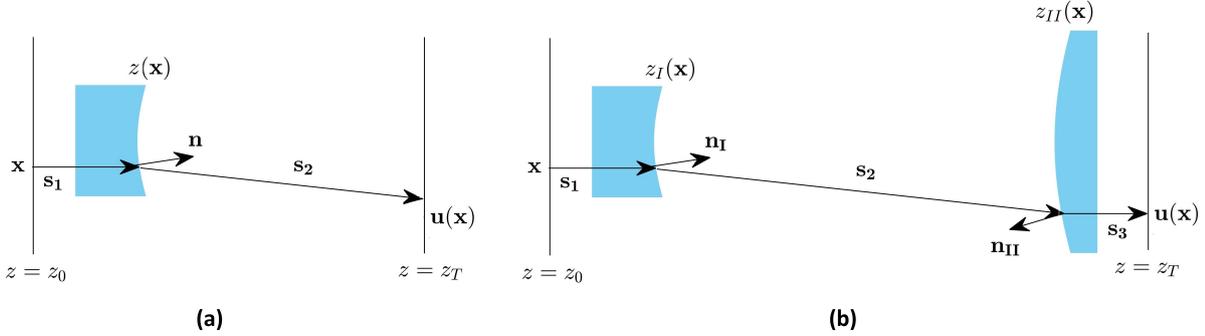


Figure 4.1: (a) Single freeform lens system. The collimated input beam is redistributed by the freeform surfaces $z(\mathbf{x})$ into the required irradiance $I_T(\mathbf{x})$ on the target plane $z = z_T$. (b) Design geometry of a double lens system. The collimated input beam is redistributed by the freeform surfaces $z_I(\mathbf{x})$ and $z_{II}(\mathbf{x})$ to realize the required irradiance $I_T(\mathbf{x})$ and a plane output wavefront.

Hence, by plugging Eq. (4.1) into the integrability condition (4.2), the differential equation (see Appendix A.5)

$$\mathbf{s}_2 \cdot (\nabla \times \mathbf{s}_1) = n_1 \frac{\{\mathbf{s}_2 \times [(\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2]\}_z}{\mathbf{n} \cdot \mathbf{s}_2} + \mathbf{s}_2 \cdot (\nabla \times \mathbf{s}_3), \quad (4.4)$$

can be derived, which couples the ray mapping and the surface. After inserting the explicit expressions for the vector fields (4.3)

$$\mathbf{v} \cdot \nabla z(\mathbf{x}) = n_1 \frac{\mathbf{v} \cdot [(\mathbf{v}^\perp \cdot \nabla) \mathbf{v}^\perp]}{\mathbf{n} \cdot \mathbf{s}_2} - (z_T - z(\mathbf{x})) \nabla \mathbf{v}, \quad (4.5)$$

is obtained, in which $\mathbf{v} \equiv (\mathbf{u}(\mathbf{x}) - \mathbf{Id})^\perp$, the identity vector $\mathbf{Id} \equiv (x, y)^T$ and $\nabla^\perp \equiv (-\partial_y, \partial_x)$ were defined. Equation (4.5) is organized in a way that only the left hand side (LHS) depends on the freeform surface derivatives. Consequently, it can be concluded that for the considered design problem the integrability condition corresponds to a semilinear two dimensional advection equation in which the freeform surface $z(\mathbf{x})$ represents the conserved transport quantity, the ray-mapping and the right hand side (RHS) of Eq. (4.5) represent a source term and the orthogonal of the mapping corresponds to the velocity field.

As it will be shown in the following by a straightforward application of the law of reflection/refraction, Eq. (4.5) itself does not describe a physically valid situation. This is can be seen directly by applying the condition that the gradient of the surface needs to be equal to the surface normal vector field defined by the ray deflection laws [Eq. (4.1)]

$$\nabla(z - z(\mathbf{x})) \stackrel{!}{=} \frac{\mathbf{n}}{(\mathbf{n})_z} \Leftrightarrow \begin{pmatrix} -\partial_x z(\mathbf{x}) \\ -\partial_y z(\mathbf{x}) \\ 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} -\frac{n_2}{|\mathbf{s}_2| \cdot (\mathbf{n})_z} v_y \\ \frac{n_2}{|\mathbf{s}_2| \cdot (\mathbf{n})_z} v_x \\ 1 \end{pmatrix}, \quad (4.6)$$

which relates the surface $z(\mathbf{x})$ to its derivatives. Consequently, $\nabla z(\mathbf{x}) \propto \mathbf{v}^\perp$ needs to hold, which

means that the LHS of Eq.(4.5) gives $\mathbf{v} \cdot \nabla z(\mathbf{x}) \equiv 0$. This can only be fulfilled if the RHS of Eq. (4.5) is identical to zero:

$$n_1 \frac{\mathbf{v} \cdot [(\mathbf{v}^\perp \cdot \nabla) \mathbf{v}^\perp]}{\mathbf{n} \cdot \mathbf{s}_2} - (z_T - z(\mathbf{x})) \nabla \mathbf{v} \stackrel{!}{\equiv} 0. \quad (4.7)$$

Thus, for every integrable ray mapping and continuous surface the condition (4.7) needs to be valid and represents a manifestation of the laws of refraction and reflection, which state that the vector fields $\mathbf{s}_1(\mathbf{x})$, $\mathbf{s}_2(\mathbf{x})$ and $\mathbf{n}(\mathbf{x})$ need to lie in the same plane.

Having found an energy conserving mapping $\mathbf{u}(\mathbf{x})$ fulfilling the Eq. (4.7), one is then left with the linear advection equation $\mathbf{v} \nabla z(\mathbf{x}) = 0$, which can be solved for the surface $z(\mathbf{x})$ by applying appropriate boundary conditions (see Appendix A.6). Due to the dependence of Eq. (4.7) on the surface itself, finding such a mapping for general design geometries is a nontrivial task.

Nevertheless, there is a special case in which Eq. (4.7) can be fulfilled without any *a priori* knowledge of the freeform surface. By observing that the second term of the LHS in Eq. (4.7) is proportional to $\nabla \mathbf{v}$, it can be seen that this term vanishes for non-trivial cases if and only if the curl of the ray-mapping $\mathbf{u}(\mathbf{x})$ vanishes. As pointed out in section 2.2.2, this vanishing curl $\nabla \times \mathbf{s}_3 = 0$ is equivalent to the quadratic cost function condition (2.21), meaning that this condition represents the only mapping for which $\nabla \mathbf{v} \equiv 0$, besides the mirrored L_2 OMT map $-\mathbf{u}(\mathbf{x})$.

Thus, the remaining term in Eq. (4.7) is the first term. Since for a prescribed emittance $I_S(\mathbf{x})$ and irradiance $I_T(\mathbf{x})$, the mapping $\mathbf{u}(\mathbf{x})$ and therefore the numerator is fixed, the remaining term vanishes if $\mathbf{n} \cdot \mathbf{s}_2$ is equal to infinity. By using Eqs. (4.1) and (4.3), it follows that $\mathbf{n} \cdot \mathbf{s}_2 \sim (z_T - z(\mathbf{x}))$ for z_T approaching infinity. Therefore, the quadratic cost function mapping is exactly integrable for infinite distances between the freeform surface and the target distribution. Hence, we can conclude that an *approximate* integrability can be reached in the paraxial regime defined by

$$\mathbf{n} \cdot \mathbf{s}_2 \gg n_1 \mathbf{v} \cdot [(\mathbf{v}^\perp \cdot \nabla) \mathbf{v}^\perp]. \quad (4.8)$$

After calculating the corresponding mapping the surface can then either be calculated by solving the linear advection equation

$$\mathbf{v} \cdot \nabla z(\mathbf{x}) = \nabla(\mathbf{v}z(\mathbf{x})) = 0, \quad \mathbf{v} = (\mathbf{u}(\mathbf{x}) - \mathbf{Id})^\perp \quad (4.9)$$

with appropriate boundary conditions or by integrating Eq. (4.6) on the support of $I_S(\mathbf{x})$ [Appendix A.6].

Consequently, it has been shown that the L_2 OMT map, defined by Eqs. (2.19) and (2.21), describes a concrete single FID geometry for which it is integrable and can be calculated independent from the freeform surface itself. The L_2 OMT map and the mirrored L_2 OMT map, both for which $\mathbf{v}(\mathbf{x}) \equiv 0$ holds, then provide two possible approximate freeform solutions in a paraxial regime for collimated input beams: a noncrossing geometry and a crossing geometry [Fig. 4.2].

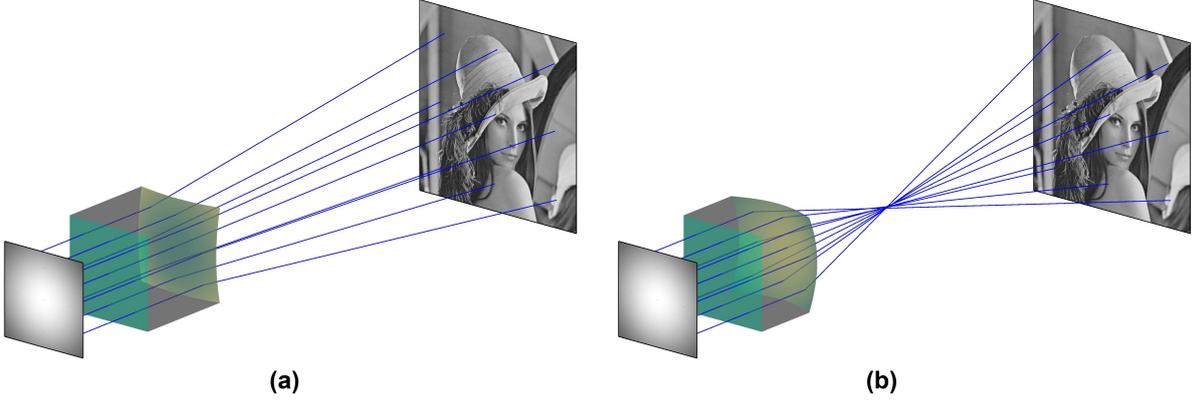


Figure 4.2: The L_2 OMT map offers two solutions for the FID with collimated beams in a paraxial regime: (a) a noncrossing geometry for $\mathbf{u}(\mathbf{x})$ and (b) a crossing geometry for $-\mathbf{u}(\mathbf{x})$. In both cases the curl of map vanishes and the map is asymptotically integrable.

In section 4.2, the application of the OMT mapping to the double FID for collimated beam shaping will be discussed. As it will turn out, in contrast to the single FID, the equivalent of the ray mapping condition (4.7) for the double FID can be reduced to a PDE for the mapping components only, which allows for an efficient numerical solution of the collimated beam shaping problem beyond the paraxial approximation.

4.2 Double freeform surface

The major difference compared to the single FID is the coupling of two freeform surfaces $z_I(\mathbf{x})$ and $z_{II}(\mathbf{x})$. As it is shown below, the design task generally cannot be separated into the calculation of an energy redistributing freeform and the calculation of a second phase correcting freeform surface. For the considered design geometry the freeform surfaces are *both* responsible for obtaining the required irradiance *and* phase redistribution, simultaneously.

According to Fig. 4.1 (b), the design geometry is expressed through the vector fields

$$\mathbf{s}_1 = \begin{pmatrix} 0 \\ 0 \\ z_I(\mathbf{x}) - z_0 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} u_x(\mathbf{x}) - x \\ u_y(\mathbf{x}) - y \\ z_{II}(\mathbf{u}) - z_I(\mathbf{x}) \end{pmatrix}, \quad \mathbf{s}_3 = \begin{pmatrix} 0 \\ 0 \\ z_T - z_{II}(\mathbf{u}) \end{pmatrix}, \quad \mathbf{s}_4 = \begin{pmatrix} u_x(\mathbf{x}) - x \\ u_y(\mathbf{x}) - y \\ z_T - z_0 \end{pmatrix}. \quad (4.10)$$

For the considered double FID for collimated beam shaping, the surface continuity condition (4.2) needs to be required for both freeform surfaces. Applying the definition Eq. (4.1) to the normal vector field $\mathbf{n}_I(\mathbf{x})$ of surface $z_I(\mathbf{x})$, it follows the equivalent of Eq. (4.4)

$$\mathbf{s}_2 \cdot (\nabla \times \mathbf{s}_1) = n_1 \frac{\{\mathbf{s}_2 \times [(\mathbf{s}_2 \nabla) \mathbf{s}_2]\}_z}{\mathbf{n}_I \cdot \mathbf{s}_2} + \mathbf{s}_2 \cdot (\nabla \times \mathbf{s}_4) - \mathbf{s}_2 \cdot (\nabla \times \mathbf{s}_3) \quad (4.11)$$

and Eq. (4.5)

$$\mathbf{v} \cdot \nabla z_I(\mathbf{x}) = n_1 \frac{\mathbf{v} \cdot [(\mathbf{v}^\perp \cdot \nabla) \mathbf{v}^\perp]}{\mathbf{n}_I \cdot \mathbf{s}_2} - (z_{II}(\mathbf{u}) - z_I(\mathbf{x})) \nabla \mathbf{v} + \mathbf{v} \cdot \nabla z_{II}(\mathbf{u}) \quad , \quad (4.12)$$

respectively. The difference to the single freeform case is therefore on the one hand the additional third term on the RHS's of Eqs. (4.11) and (4.12), which are of the same form as the LHS terms and reflect the symmetry of the design problem, and on the other hand the dependence on the second freeform surface. Thus, using the analogue of condition Eq. (4.6) leads to $\nabla z_I(\mathbf{x}) \propto \mathbf{v}^\perp$, which again shows that a nonvanishing RHS of Eq. (4.12) contradicts the laws of reflection or refraction.

To rewrite the gradient of the second surface $\nabla_{\mathbf{x}} z_{II}(\mathbf{u})$ in the third term of Eq. (4.12), the chain rule is applied to $\nabla_{\mathbf{u}} z_{II}(\mathbf{u}) = (\partial_{u_x} z_{II}(\mathbf{u}), \partial_{u_y} z_{II}(\mathbf{u})) = \mathbf{n}_{II}(\mathbf{u})/(\mathbf{n}_{II})_z$. Hence, it follows the linear advection equation together with the ray-mapping condition for double FID with collimated beams in the form

$$\mathbf{v} \cdot \nabla z_I(\mathbf{x}) = 0, \quad (4.13)$$

$$n_1 \frac{\mathbf{v} \cdot [(\mathbf{v}^\perp \cdot \nabla) \mathbf{v}^\perp]}{\mathbf{n}_I \cdot \mathbf{s}_2} + n_2 \frac{g(\mathbf{u})}{(\mathbf{n}_{II})_z |\mathbf{s}_2|} - (z_{II}(\mathbf{u}) - z_I(\mathbf{x})) \nabla \mathbf{v} = 0. \quad (4.14)$$

Thereby the function $g(\mathbf{u}) := -v_x^2 \partial_x u_y + v_y^2 \partial_y u_x + v_x v_y (\partial_x u_x - \partial_y u_y)$ was defined.

Equation (4.14) again raises the question, if and under what kind of geometrical conditions a ray-mapping $\mathbf{u}(\mathbf{x})$ exists, which gives continuous freeform surfaces allowing a redistribution of the given emittance into the required irradiance without any a priori knowledge. By using the definition of the vector fields (4.10) and the deflection laws (4.1) it follows that both denominators in Eq. (4.14) are asymptotically proportional to the difference between both freeform surfaces $(z_{II}(\mathbf{u}) - z_I(\mathbf{x}))$, whereas the numerators are invariant under this geometrical variation for a fixed mapping $\mathbf{u}(\mathbf{x})$.

Thus, in that approximation, the first two terms in Eq. (4.14) and the last term only vanish for the L_2 OMT map due to $\nabla \mathbf{v} = 0$, which again results in two possible approximate FID solutions for collimated beams in a paraxial regime. Alternatively, this can also be shown by expanding the Rubinstein-Wolansky cost function [44], which describes the double FID for lenses and collimated beams, for small refraction angles.

Hereafter the case of double freeform mirrors for collimated beam shaping will be discussed shortly. Hence, $n_1 = n_2 = -1$ is used and $\mathbf{n}_I \mathbf{s}_2 = -(\mathbf{n}_{II})_z |\mathbf{s}_2|$ follows. Equation (4.14) then reduces to

$$\left\{ \frac{|\mathbf{v}|^2}{\mathbf{n}_I \cdot \mathbf{s}_2} + [z_{II}(\mathbf{u}) - z_I(\mathbf{x})] \right\} \nabla \mathbf{v} = 0. \quad (4.15)$$

Therefore, without symmetry assumptions, this condition can only be fulfilled by the L_2 OMT map. This is in agreement with the results by Glimm and Olikier, which gave a rigorous mathematical proof in Ref. [43]. Consequently, in contrast to the other considered design cases, the

L_2 OMT map solves the double mirror design problem for collimated beam shaping independent from the freeform-freeform distance within physical bounds.

Equation (4.14) for lenses can be further simplified by eliminating the dependency on both freeform surfaces. To do so, the constant OPL condition (3.19) with collimated beams is solved for

$$z_{II}(\mathbf{u}) - z_I(\mathbf{x}) = -\frac{n}{n^2-1}OPL_{red} \mp \frac{1}{n^2-1}\sqrt{OPL_{red}^2 + (n^2-1)|\mathbf{u}(\mathbf{x}) - \mathbf{Id}|^2}, \quad (4.16)$$

with $n := n_1/n_2$ and the reduced optical path length $OPL_{red} := (OPL - n_1 z_T)/n_2$ between the first and second surface. The sign before the second term of Eq. (4.16) thereby depends on the chosen system type. For single lens systems ($OPL_{red} > 0$; $n < 1$) a negative sign is needed and for double lens systems ($OPL_{red} < 0$; $n > 1$) a positive sign. According to the square root in Eq. (4.16), the possible parameter space is restricted by $|\mathbf{u}(\mathbf{x}) - \mathbf{Id}|^2 < OPL_{red}^2/(n^2-1)$.

Equation (4.16) can be plugged into Eq. (4.14), which gives the PDE system

$$\begin{aligned} \det(\nabla \mathbf{u}(\mathbf{x}))I_T(\mathbf{u}(\mathbf{x})) - I_S(\mathbf{x}) &= 0, \\ \partial_y u_x - \partial_x u_y - \frac{(n^2-1)[(u_x-x)^2 \partial_y u_x - (u_y-y)^2 \partial_x u_y + (u_x-x)(u_y-y)(\partial_y u_y - \partial_x u_x)]}{OPL_{red}^2 + (n^2-1)|\mathbf{u}(x,y) - \mathbf{Id}|} &= 0, \end{aligned} \quad (4.17)$$

for the mapping $\mathbf{u}(\mathbf{x})$. After solving the PDE system [84], the mapping can then be integrated to give the surface $z_I(\mathbf{x})$ [Appendix A.6], from which the $z_{II}(\mathbf{u})$ follows by Eq. (4.16). The second surface, defined at the points $\mathbf{u}(\mathbf{x})$, is then interpolated onto the required grid.

4.3 Summary

In sections 4.1 and 4.2, it was shown that the L_2 OMT map for collimated beams is asymptotically integrable due to the curl-free property (2.23). This plays a significant role for the numerical solving of the PDE models (3.10), (3.22) and (3.30) in chapter 5 with a finite difference discretization. In chapter 5, the L_2 OMT map and a surface, constructed from the map, will be used as an initial iterate for the root-finding of the resulting nonlinear equation system. A fast and reliable root-finding thereby requires a convergent discretization scheme (see section 5.1) and/or an initial iterate close to a root of the nonlinear equation system.

Considering the single FID with a collimated input beam and applying the findings of this chapter, the latter can be guaranteed by calculating the L_2 OMT map and the corresponding freeform surface for a sufficiently large freeform-to-target distance ($z_T - z(\mathbf{x})$). The L_2 OMT map and the freeform are then used as an initial iterate to solve the PDE system (3.10). By an iterative repetition of the process by a stepwise reduction $z_T \rightarrow (z_T - \Delta z_T)$ with a sufficiently small Δz_T , it can be ensured that the initial iterate for solving process is always close to the solution of the PDE system (3.10), even in a nonparaxial regime. The same argument holds for the double FID for collimated beam shaping by a stepwise reduction of the freeform-freeform

distance.

It is important to note that this iterative procedure is a thought experiment and purely theoretical. In concrete design examples [chapter 6], an application of the iterative procedure was *never* necessary and the numerical solver [chapter 5] converged stably even for highly nonparaxial configurations with noncollimated beams. Additionally, it is noteworthy that for these kind of more general wavefronts, one can extend the argument given above by not only reducing distances iteratively, but by also changing the wavefront stepwise. Starting from a collimated beam, the transformation $\phi(\mathbf{x}) \rightarrow (\phi^{coll}(\mathbf{x}) + \Delta\phi(\mathbf{x}))$ is done until the required wavefront is reached. It is hereby assumed that a small change of the wavefront leads to a small step in the solution space or small variation of the freeform surface, respectively.

While the L_2 OMT map and its mirror map allow the calculation of two numerical solutions to the FID problem for a large space of geometrical parameters, it is important to emphasize that the PDE systems (3.10), (3.22) and (3.30) might have solutions outside of the accessible geometrical parameter space, like for instance saddle-shaped freeform solutions. A necessary mathematical analysis for a better understanding of this problem lies outside the scope this thesis.

5 Design method: numerical solution of freeform illumination PDEs

In this chapter, a numerical strategy for the solution of the nonlinear PDE systems (3.10), (3.22) and (3.30) is proposed. Since previous PDE models are restricted to spherical or planar wavefronts and correspondingly shaped entrance surfaces, a numerically reliable solution method for *general* zero-étendue wavefronts is required. A common technique for solving nonlinear PDEs is the discretization by finite differences (FD) and the subsequent solution of the resulting coupled nonlinear equation systems by iterative techniques like e.g. the Newton method (see e.g. [65, 70, 80, 93]). In the context of reflector antenna design, this was done first in 1976 by Norris and Westcott, which computed reflector surfaces by solving a linearized MAE [65]. For complex-structured irradiance distributions with high resolutions the simple application of standard FDs and the utilization of nonlinear equation solvers will lead to numerical instabilities and likely fail [93]. Consequently, for MAE methods in literature that are based on standard FDs [70, 71], it can be observed that design examples are shown with irradiance distributions of low complexity, low resolutions and/or low contrasts. Since the development of a provably convergent FD discretization scheme for nonlinear PDEs is nontrivial [93], the construction of an appropriate initial iterate for the nonlinear equation system solver is of major importance to ensure a reliable numerical approach. In case of FID this means the requirement of a suitable initial ray mapping and/or initial freeform surface.

To achieve this for the single FID, the direct solution of the PDE system (3.10) instead of the MAE (3.11) is proposed¹. This has the advantage that the energy conservation equation is stated explicitly. Thus, as discussed in chapter 4.3, it is possible to construct an initial mapping by L_2 OMT, which already fulfills the energy conservation equation *and* is related to the optical design problem. This greatly benefits the stability of the numerical calculations with standard FDs. Consequently, the main difficulty of the numerical computations is shifted to the stable calculation of the L_2 OMT map, which is independent of geometrical considerations like for instance ray directions and for which available numerical algorithms in literature can be utilized (e.g. [91–93]).

Furthermore, an initial surface construction approach for general zero-étendue wavefronts is proposed, which incorporates properties of the corresponding L_2 OMT map and design geometry. To do so, a differential equation model from Rubinstein and Wolansky [94] for the reconstruction

¹A method based on numerically solving a PDE system was published in 1987 by Galindo-Israel et al. in the context of reflector antenna design [67].

of optical surfaces from experimental ray data is adapted to integrate suitable initial surfaces directly from the L_2 OMT map and the given input and output ray directions by the Runge-Kutta method. This reduces the number of necessary iterations for the solution of the nonlinear equation system greatly compared to the often utilized simple initial surface shapes, like e.g. spheres, and is also applicable to non-planar and non-spherical zero étendue wavefronts.

Moreover, the design approach is extended to the double FID [Eq. (3.22)]. The major difference to the single FID is thereby the numerical projection of the L_2 OMT map onto the output wavefront to achieve a suitable initial iterate for a stable solution of the nonlinear equation systems resulting from the discretization of Eq. (3.22).

Additionally, the generalization of the numerical design approach to the placement of freeform surfaces in an optical system of predefined surfaces is discussed. Since, in particular for predefined surfaces, the initialization of the numerical FID process requires the preparation of input data like emittances, ray directions and/or surface interception coordinates, numerical procedures for ray aiming and noise-free irradiance computation are proposed. These procedures are based on the description of optical systems from chapter 3.3.

The chapter is structured as follows. In section 5.1, the discretization of the PDE systems (3.10), (3.22) and (3.30) by standard FDs and the solving of the resulting equation systems by Newton-type nonlinear equation solvers is explained. This is followed in section 5.2, by the presentation of two numerical L_2 OMT map calculation approaches from literature and a discussion of their advantages and disadvantages. Based on this, the numerical solution of the MAE (3.11) is discussed. In section 5.3, it is shown how to construct the initial freeform surface for single and double freeform systems from the L_2 OMT map by ordinary differential equations with the Runge-Kutta method. Then, in section 5.4, a summary and discussion of the design strategies for single and double freeform surfaces and their application to optical systems is given. The preparation of the input data of the design algorithm like emittances, ray directions and wavefronts is discussed in section 5.5.

The results of this chapter were published in references [85], [86] and [87].

5.1 PDE discretization and nonlinear equation systems

In chapter 3, it was shown that the FID problem for single and double freeform lenses can be formulated in terms of a PDE systems of the form (3.10) and (3.22) with three unknown functions: the freeform surface and two (wavefront) mapping components.² Therefore, these coupled PDEs need to be solved simultaneously for all three functions. In the following, this is done by a FD discretization of the PDEs and the application of a nonlinear equation system solver. As mentioned above, the L_2 OMT map will thereby be used as the initial iterate for the mapping components [section 5.2] and to construct the initial surface [section 5.3]. These initial iterates will be denoted by the superscript “ ∞ ”. Instead of directly solving the PDE system for

²The solution of Eq. (3.30) with the additional surface intersection points of predefined exit surfaces will be discussed in section 5.4.1

the unknown functions, the corrections $\Delta z_S(\mathbf{x})$ and $\Delta \mathbf{u}(\mathbf{x})$ are formally introduced through

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}^\infty(\mathbf{x}) + \Delta \mathbf{u}(\mathbf{x}), \quad z_S(\mathbf{x}) = z_S^\infty(\mathbf{x}) + \Delta z_S(\mathbf{x}). \quad (5.1)$$

and the initial iterate is redefined by $\Delta z_S(\mathbf{x}) = 0$ and $\Delta \mathbf{u}(\mathbf{x}) = 0$.

In the following, the discretization of the PDE systems is discussed. It is assumed that the emittance $I_S(\mathbf{x}) \geq 0$ and the input ray directions $\hat{\mathbf{s}}_1(\mathbf{x})$ are defined on the rectangular aperture $\Omega_S := \overline{\{\mathbf{x} \in \mathbb{R}_2 \mid x_{min} < x < x_{max}, y_{min} < y < y_{min}\}}$ [Fig. 5.1] and that the functions are discretized on a two dimensional cartesian grid

$$(x, y) \rightarrow (x_i, y_j) = (x_{min} + (i - 1)dx, y_{min} + (j - 1)dy) \quad (5.2)$$

with the spacing $(dx, dy) = ((x_{max} - x_{min})/(M - 1), (y_{max} - y_{min})/(N - 1))$ and $i = 1, \dots, M; j = 1, \dots, N$.

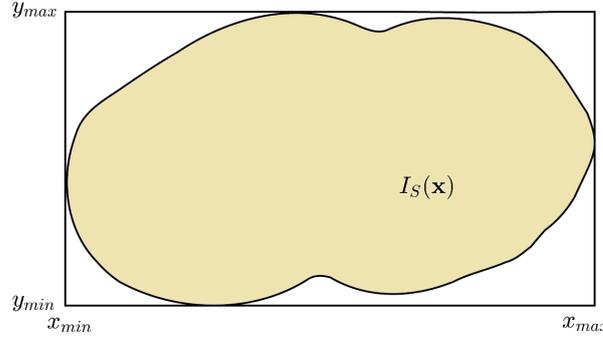


Figure 5.1: The emittance $I_S(\mathbf{x}) \geq 0$ is defined on a rectangular aperture.

Hence, it follow the discretized functions

$$\Delta \mathbf{u}(\mathbf{x}) \rightarrow \Delta \mathbf{u}(x_i, y_j) \equiv (\Delta \mathbf{u})_{i;j}, \quad \Delta z_S(\mathbf{x}) \rightarrow \Delta z_S(x_i, y_j) \equiv (\Delta z_S)_{i;j}. \quad (5.3)$$

Now the discretization of the partial derivatives is considered. As it is known, the application of standard FD's for the discretization of general nonlinear PDE's is in general insufficient for a stable numerical convergence of root-finders [93]. Especially, for the case of FID, the more complicated structure of the MAE (3.11), compared to Eq. (2.22), which is considered in Ref. [93], makes the development of convergent schemes nontrivial (see 5.2.1 below). Due to the lack of an in-depth numerical analysis of the first-order PDE systems in Eqs. (3.10) and (3.22), standard FDs will nevertheless be applied for the discretization of the partial derivatives of $\Delta z_S(\mathbf{x})$ and $\Delta \mathbf{u}(\mathbf{x})$. Instead, the focus will be put on the calculation of an appropriate initial iterate $\mathbf{u}^\infty(\mathbf{x})$ and $z^\infty(\mathbf{x})$. Consequently, for the interior points of Ω_S , the central FDs [Fig. 5.2 (a)]

$$\partial_x(\Delta z_S) \rightarrow \frac{1}{2dx} [(\Delta z_S)_{i;j+1} - (\Delta z_S)_{i;j-1}], \quad \partial_y(\Delta z_S) \rightarrow \frac{1}{2dy} [(\Delta z_S)_{i+1;j} - (\Delta z_S)_{i-1;j}] \quad (5.4)$$

are applied with $i = 2, \dots, M - 1; j = 2, \dots, N - 1$, and on the boundary the second-order FDs [Fig.

5.2 (b)]

$$\begin{aligned}\partial_x(\Delta z_S) &\rightarrow -\frac{1}{2dx} [3(\Delta z_S)_{i;j+2} - 4(\Delta z_S)_{i;j+1} + (\Delta z_S)_{i;j}] \\ \partial_y(\Delta z_S) &\rightarrow -\frac{1}{2dy} [3(\Delta z_S)_{i+2;j} - 4(\Delta z_S)_{i+1;j} + (\Delta z_S)_{i;j}]\end{aligned}\quad (5.5)$$

are utilized.

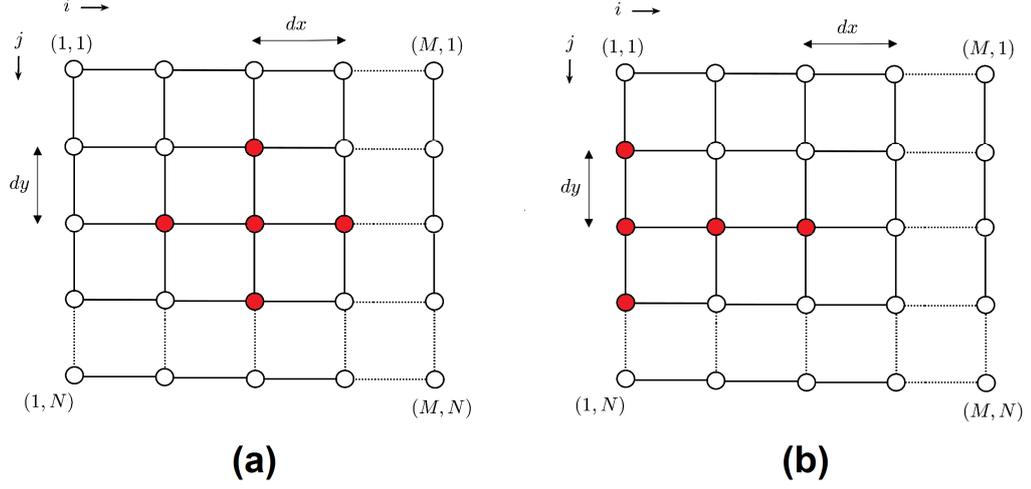


Figure 5.2: The partial derivatives of $\Delta z_S(\mathbf{x})$ and the mapping components $\Delta \mathbf{u}(\mathbf{x})$ are discretized on a cartesian grid with the spacing (dx, dy) . Stencils for the FD approximation of the derivatives for (a) interior points and (b) boundary points.

So far, the discretization of the PDEs in Eqs. (3.10) and (3.22) was discussed. Additionally, the transport boundary condition $\mathbf{u}(\partial\Omega_S) = \partial\Omega_T$ needs to be implemented, meaning that the boundary of the input aperture is required to be mapped onto the boundary of the irradiance distribution. For source and target distributions defined on a unit square, it follows

$$\begin{aligned}\Delta u_x(\pm 0.5, y) = 0, \quad y \in [-0.5, 0.5] &\rightarrow (\Delta \mathbf{u})_{1;j} = (\Delta \mathbf{u})_{M;j} = 0, \quad j = 1, \dots, N \\ \Delta u_y(x, \pm 0.5) = 0, \quad x \in [-0.5, 0.5] &\rightarrow (\Delta \mathbf{u})_{i;1} = (\Delta \mathbf{u})_{i;N} = 0, \quad i = 1, \dots, N\end{aligned}\quad (5.6)$$

Hence, due to the explicit presence of the mapping components in Eqs. (3.10) and (3.22), the transport boundary conditions can be controlled directly through the mapping itself. In case of the MAE (3.11), the same condition $\mathbf{u}(\partial\Omega_S) = \partial\Omega_T$ has to be achieved through the surface gradients $\nabla z_S(\mathbf{x})$ at $\partial\Omega_S$, which leads to a derivative-dependency and thus to more complicated equations than Eq. (5.6).

After applying the discretization schemes, we are left with a system of nonlinear equations of the form $\mathbf{F}(\mathbf{X}) = 0$ for the $3 \cdot M \cdot N$ unknowns $(\Delta \mathbf{u})_{i;j}$ and $(\Delta z_S)_{i;j}$. Typically such complex nonlinear equation systems with a large number of unknowns are solved by the Newton method [65, 70, 93]. For the design examples in the following chapters, the *trust-region reflective* solver from MATLAB 2015b's optimization toolbox is applied, which is based on the interior-reflective Newton method by Coleman and Li [100, 101]. Thereby, it is of special importance for a fast

numerical calculation to provide the corresponding MATLAB function $fsolve()$ with the structure of the non-zero elements of the Jacobian matrix of the nonlinear equation system. Otherwise $fsolve()$ will apply a FD discretization to the full Jacobian matrix, which significantly slows down the root-finding process and increases the memory consumption. The sparse structure of the Jacobian matrix follows directly from the structure of the corresponding PDE system and the applied FD discretization scheme.

5.2 Initial mapping calculation

In this section, algorithms for the numerical calculation of the L_2 OMT map $\mathbf{u}^\infty(\mathbf{x})$, which will serve as an initial iterate for the nonlinear equation system solver, are discussed. Due to the applicability of this specific mapping to a variety of scientific areas, the L_2 OMT map has drawn large interest in numerical mathematics. Thus, the mathematical literature provides numerous available methods for the mapping calculation (e.g. [91–93]), which differ in regard to the capability of being able to handle complex boundaries of the input and output distribution or in regard to their numerical stability, especially for complex input and output distributions with high contrasts. For the FID strategy, which will be given in section 5.4, the application of a specific numerical L_2 OMT method is of minor importance.

In the following, at first the solution of the elliptic MAE (2.22) with standard FD schemes is discussed, which allows the calculation of maps for emittances and irradiances with complicated boundary shapes but leads numerical instabilities for complex distributions with high contrasts. This illustrates numerical disadvantages of solving the single FID problem with the MAE (3.11) and supports the utilization of the PDE system (3.10).

Afterwards, the numerical L_2 OMT algorithm by Sulman et al. [92] will be presented and discussed, which represents a good compromise between computational speed, implementation effort and numerical robustness and will be used for the numerical design examples in the following chapters. A disadvantage of Sulman’s algorithm is the restriction to emittances and irradiances with square boundaries.

5.2.1 Solving the Monge-Ampère equation with finite difference schemes: the problem with nonlinear PDEs

In section 3.1, it was shown that the single FID for zero-étendue wavefronts can be formulated in terms of the MAE (3.11), which is a generalization of the MAE (2.22) for the L_2 OMT problem. To solve the MAE (2.22), the second order partial derivatives of the mapping potential need to be discretized. Assuming a cartesian grid as in section 5.1, the standard FD scheme is defined by

$$\begin{aligned}
\partial_x u &\rightarrow \frac{1}{2dx} [u_{i+1,j} - u_{i-1,j}], & \partial_y u &\rightarrow \frac{1}{2dy} [u_{i,j+1} - u_{i,j-1}], \\
\partial_{xx} u &\rightarrow \frac{1}{dx^2} [u_{i+1,j} - 2u_{i,j} + u_{i-1,j}], & \partial_{yy} u &\rightarrow \frac{1}{dy^2} [u_{i,j+1} - 2u_{i,j} + u_{i,j-1}] \\
\partial_{xy} u &\rightarrow \frac{1}{4dxdy} [u_{i+1,j+1} - u_{i+1,j-1} - u_{i-1,j+1} + u_{i-1,j-1}].
\end{aligned} \tag{5.7}$$

on interior grid points and

$$\partial_x u \rightarrow -\frac{1}{2dx} [3u_{i;j+2} - 4u_{i;j+1} + u_{i;j}], \quad \partial_y u \rightarrow -\frac{1}{2dy} [3u_{i+2;j} - 4u_{i+1;j} + u_{i;j}]. \tag{5.8}$$

on boundary grid points. This discretization leads to a system of nonlinear equations, which can be solved by the Newton method. The numerically *stable* solution is in general nontrivial. For instance, by applying Eqs. (5.7) and (5.8) to Eq. (2.22), it can be seen directly that the resulting nonlinear equation system allows non-convex solutions, which contradicts the convexity constraint of the potential $u(\mathbf{x})$. Also, due to the mixed derivatives in Eq. (2.22), the standard FD scheme does not fulfill simultaneously the monotonicity, stability and consistency conditions [93] as required for convergence of approximation schemes of fully nonlinear second order PDEs [104]. The resulting numerical instabilities are especially problematic for large contrasts in the input and output distributions $I_S(\mathbf{x})$ and $I_T(\mathbf{x})$.

To overcome these problems, some authors propose to represent the determinant in (2.22) as a product of its smallest and largest eigenvalue [93]. This allows to rewrite the determinant in the terms of a set of eigenvectors to eliminate the mixed derivatives and write Eq. (2.22) in the form [45, 93]

$$\min_{(\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2) \in V} \left\{ \left(\frac{\partial^2 u}{\partial v_1^2} \right)_\delta^+ \left(\frac{\partial^2 u}{\partial v_2^2} \right)_\delta^+ + \left(\frac{\partial^2 u}{\partial v_1^2} \right)_\delta^- \left(\frac{\partial^2 u}{\partial v_2^2} \right)_\delta^- \right\} I_T(\nabla u) = I_S(\mathbf{x}) \tag{5.9}$$

with the set V of orthonormal bases of \mathbb{R}_2 , the direction derivative $\frac{\partial}{\partial v_i}$ in the direction of $\hat{\mathbf{v}}_i; i = 1, 2$, as well as $(z)_\delta^+ := \max(z, \delta)$ and $(z)_\delta^- := \min(z, \delta)$ with $\delta > 0$. This formulation allows the application of a FD discretization scheme, which enforces convexity and fulfills the monotonicity, stability and consistency condition [93].

The ability to rewrite the MAE (2.22) in the form of Eq. (5.9) illustrates an important difference of the L_2 OMT compared to the MAE (3.11) in single FID. Consequently, it can be concluded that the MAE methods for FID, based on standard FD schemes, are not numerically stable for complex emittances and irradiances with strong gradients.

Due to these numerical instabilities, the standard FDs (5.7) and (5.8) will only be applied to calculate $\mathbf{u}^\infty(\mathbf{x})$ for simple *uniform* distributions to demonstrate the capabilities of the numerical FID strategies to handle non-rectangular boundary shapes.

5.2.2 Sulman's L_2 OMT algorithm

For the design examples with complex irradiances with high spatial frequency features Sulman's algorithm [92] will be used, which is numerically more stable but can only handle distributions

with square boundaries.

Instead of solving the elliptic MAE (2.22) directly for the potential $u(\mathbf{x})$, Sulman et al. solve the parabolic MAE [92]

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} = \log \left[\frac{I_T(\nabla u(\mathbf{x}, t)) \det(\nabla^2 u(\mathbf{x}, t))}{I_S(\mathbf{x})} \right] \quad (5.10)$$

with the initial conditions

$$u(\mathbf{x}, 0) = \frac{1}{2}(\mathbf{x} \cdot \mathbf{x}^T) \quad (5.11)$$

and the Neumann boundary conditions

$$\nabla u(\mathbf{x}, t) \cdot \mathbf{n} = \mathbf{x} \cdot \mathbf{n}, \quad \mathbf{n} \in \partial\Omega_S \quad (5.12)$$

with the vector \mathbf{n} representing the outward boundary normal of Ω_S . These boundary conditions guarantee that the boundary $\partial\Omega_S$ is mapped onto $\partial\Omega_T$ of the target distribution.

The calculation of the mapping potential is done by a search for the steady state solution $u_\infty(\mathbf{x})$, as Eq. (5.10) for $t \rightarrow \infty$ converges to the solution of the elliptical MAE (2.22). The logarithm on the RHS of eq. (5.10) hereby preserves the convexity [section 2.2.1] of the potential $u_\infty(\mathbf{x})$. The L_2 OMT map is then simply calculated by the gradient of the mapping potential $\mathbf{u}^\infty(\mathbf{x}) = \nabla u_\infty(\mathbf{x})$. The numerical procedure for solving eq. (5.10) is summarized in Alg. 5.1.

Algorithm 5.1 Sulman's algorithm

- 1: Initialize potential $u^n = \frac{1}{2}(x^2 + y^2)$
 - 2: Set tolerance Tol
 - 3: Set step size dt
 - 4: *loop*:
 - 5: **while** $maxNorm \geq Tol$ **do**
 - 6: Gradient of potential $(u_x^n, u_y^n) = \nabla u^n$.
 - 7: Boundary conditions $\mathbf{u}(\partial\Omega_S) = \partial\Omega_T$.
 - 8: Hesse determinant $det(H^n) = |u_{xx}^n u_{yy}^n - u_{yx}^n u_{xy}^n|$
 - 9: Set $F^n = \log \left[\frac{I_T(u_x^n, u_y^n) det(H^n)}{I_S(x, y)} \right]$
 - 10: Update potential $u^n = u^n + dt F^n$
 - 11: $maxNorm = \|F^n\|_2$
 - 12: stop or set $n = n + 1$
 - 13: $(u_x^n, u_y^n) = \nabla u^n$
-

Existence, uniqueness and convergence results of the algorithm can be found in [92].

As it was discussed in section 2.2.1 and chapter 4, $\mathbf{u}^\infty(\mathbf{x})$ represents a ray-mapping between the input rays on the input plane and the target plane. Consequently, the transformation between the emittance and irradiance can be visualized by a grid transformation as shown in Fig. 5.3.

The important parameters of the algorithm, which are defined in Alg. 5.1, are the tolerance Tol and the stepsize dt . Therefore, Tol determines the precision of the solution $\mathbf{u}(\mathbf{x})$ of the energy conservation equation according to the definition in Alg. 5.1. This means that the algorithm

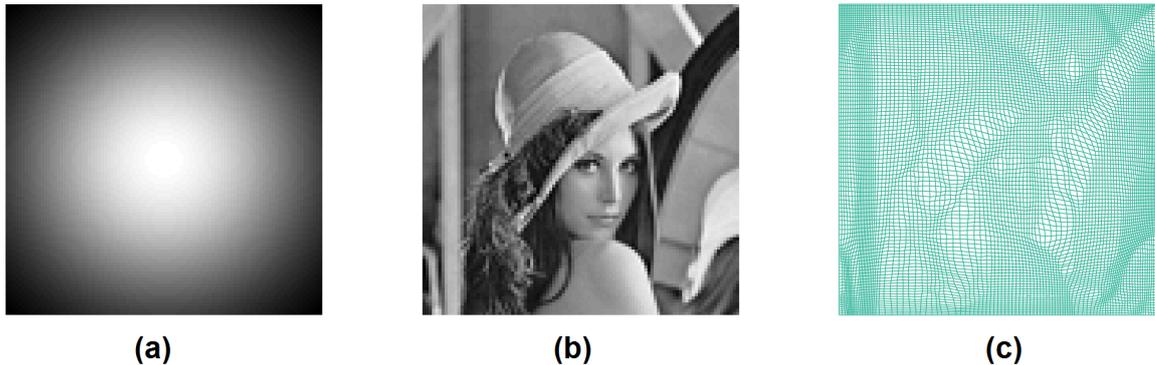


Figure 5.3: (a) Input distribution. (b) Output distribution. (c) Grid deformation according to the L_2 OMT map $\mathbf{u}^\infty(\mathbf{x})$, in which the input distribution is represented by a cartesian grid. The local energy redistribution is therefore proportional to the change in size of the quadrilaterals [section 2.2.1].

stops if the logarithm of the ratio between the discretized mapped irradiance $I_T(\mathbf{x})$ and the emittance $I_S(\mathbf{x})$ is sufficiently small. The stepsize dt on the other hand determines the speed of convergence but also has influence on the numerical stability. As the L_2 OMT map solves the FID for planar wavefronts and infinite freeform-target distances [chapter 4], Tol is connected to the quality of the energy redistribution for these kind of design geometries. Since the L_2 OMT map also serves as an initial iterate for the solving of the PDE systems (3.10) and (3.22), the FID configuration with a planar wavefront and large freeform-target distances might be utilized as a benchmark in a ray-tracing simulation to evaluate quality of the freeform solutions of the PDE systems.

Appropriate values for Tol and dt for a stable numerical convergence thereby depend on the resolution and the individual structure of the considered emittance and irradiance. Typical (empirically determined) values for distributions like “Lena” (Fig. 5.3) with a resolution of 250 pixels \times 250 pixels are $Tol \approx 10$ and $dt \approx 5 \cdot 10^{-6}$. Additional values for different resolutions are listed in Table 5.1 below.

Strong gradients in the distributions might lead to instabilities, which manifest themselves in a locally nonconvex potential $u(\mathbf{x}, t)$. This will require smaller values for dt . Since the numerical mapping $\mathbf{u}(\mathbf{x})$, which is calculated by Alg. 5.1, usually shows oscillations, an interpolation of the mapping components by a spline interpolation is applied after Tol is reached. Thus, the precision of the mapping, defined by Tol , is therefore not identical to the final map.

5.3 Initial surface construction

After calculating the initial map $\mathbf{u}^\infty(\mathbf{x})$, the initial surface $z_S^\infty(\mathbf{x})$ needs to be constructed. Since the PDE systems (3.10) and (3.22) are formulated for general zero-étendue wavefronts, the initial surface construction approach should take the nature of the chosen wavefronts into account.

To do so, an approach by Rubinstein and Wolansky [94] for the reconstruction of optical surfaces

from ray data will be adapted. In this work the authors argue that one or two unknown optical surfaces can be reconstructed by measuring the input and output ray directions and target points e.g. by a Hartmann-Shack sensor and solving ordinary differential equations. For the calculation of the initial single and double freeform surfaces, this means that predefined input and output ray directions will be applied and the L_2 OMT map $\mathbf{u}^\infty(\mathbf{x})$ will be used for the target points.

It is important to note that in contrast to Ref. [94], the integration of the differential equations will in general not lead to surfaces which accurately generate the required irradiance. While it is reasonable to assume that (ideal) experimental data will generate an integrable relation between the input and target coordinates, this is not true for the mapping $\mathbf{u}^\infty(\mathbf{x})$ as it was shown in chapter 4. Since $\mathbf{u}^\infty(\mathbf{x})$ already correctly describes the energy redistribution between the emittance and irradiance, it is nevertheless reasonable to assume that the initial surface construction approach is more suitable for the considered FID task than any other approach proposed so far.

The differential equations given in the following are applicable to the standard single and double freeform design geometries [Figs. 3.1 and 3.3]. A discussion of their applicability to more general optical systems with predefined entrance and exit surfaces will be given in section 5.4.1.

5.3.1 Single freeform surface

At first the case of a single freeform surface is considered and the given input ray direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ on the input plane $z = z_0$ and the mapping $\mathbf{u}^\infty(\mathbf{x})$ between the input and the target plane $z = z_T$ are assumed. The design geometry corresponds to Fig. 3.1 and Eq. (3.4), respectively. Similar to chapter 3, the law for refraction/reflection

$$\frac{\mathbf{n}(\mathbf{x}_S)}{(\mathbf{n}(\mathbf{x}_S))_z} = \frac{n_1 \hat{\mathbf{s}}_1 - n_2 \hat{\mathbf{s}}_2}{(n_1 \hat{\mathbf{s}}_1 - n_2 \hat{\mathbf{s}}_2)_z} = \nabla_S(z - z_S^\infty(\mathbf{x})) \quad (5.13)$$

is applied at the freeform surface intersection points \mathbf{x}_S and the gradient

$$\partial_{x_S} z_S^\infty(\mathbf{x}) = -\frac{n_1(\hat{\mathbf{s}}_1)_x - n_2(\hat{\mathbf{s}}_2)_x}{n_1(\hat{\mathbf{s}}_1)_z - n_2(\hat{\mathbf{s}}_2)_z} \equiv n_x(\mathbf{x}), \quad \partial_{y_S} z_S^\infty(\mathbf{x}) = -\frac{n_1(\hat{\mathbf{s}}_1)_y - n_2(\hat{\mathbf{s}}_2)_y}{n_1(\hat{\mathbf{s}}_1)_z - n_2(\hat{\mathbf{s}}_2)_z} \equiv n_y(\mathbf{x}). \quad (5.14)$$

is inverted by using the coordinate transformation in Eq. (3.7). By application of the chain rule to the LHS of Eq. (5.14), the gradient of the surface can then be written in terms of the \mathbf{x} -coordinates to give the ordinary differential equations [94]

$$\begin{aligned} \partial_x z_S^\infty(\mathbf{x}) &= \frac{n_x(\mathbf{x}) + \left[n_x(\mathbf{x}) \partial_x \left(\frac{(\hat{\mathbf{s}}_1)_x}{(\hat{\mathbf{s}}_1)_z} \right) + n_y(\mathbf{x}) \partial_x \left(\frac{(\hat{\mathbf{s}}_1)_y}{(\hat{\mathbf{s}}_1)_z} \right) \right] (z_S^\infty(\mathbf{x}) - z_0)}{1 - n_x(\mathbf{x}) \frac{(\hat{\mathbf{s}}_1)_x}{(\hat{\mathbf{s}}_1)_z} - n_y(\mathbf{x}) \frac{(\hat{\mathbf{s}}_1)_y}{(\hat{\mathbf{s}}_1)_z}}, \\ \partial_y z_S^\infty(\mathbf{x}) &= \frac{n_y(\mathbf{x}) + \left[n_x(\mathbf{x}) \partial_y \left(\frac{(\hat{\mathbf{s}}_1)_x}{(\hat{\mathbf{s}}_1)_z} \right) + n_y(\mathbf{x}) \partial_y \left(\frac{(\hat{\mathbf{s}}_1)_y}{(\hat{\mathbf{s}}_1)_z} \right) \right] (z_S^\infty(\mathbf{x}) - z_0)}{1 - n_x(\mathbf{x}) \frac{(\hat{\mathbf{s}}_1)_x}{(\hat{\mathbf{s}}_1)_z} - n_y(\mathbf{x}) \frac{(\hat{\mathbf{s}}_1)_y}{(\hat{\mathbf{s}}_1)_z}}. \end{aligned} \quad (5.15)$$

This differential equations can be integrated along an arbitrary path in the input aperture Ω_S to

give the freeform surface $z_S(\mathbf{x})$ for an integrable map. For the nonintegrable L_2 OMT map $\mathbf{u}^\infty(\mathbf{x})$ the initial surface $z_S^\infty(\mathbf{x})$ will be path dependent and the initial surface will vary slightly with the chosen path, which has no significant influence on the convergence of the subsequent nonlinear equation system solving process. A possible integration path is shown in Fig. 5.4. Hereby, the integration constant of the surface is fixed at the center of the freeform. The surface is then integrated first along the x-direction and then along the y-direction. Due to the non-integrability of $\mathbf{u}^\infty(\mathbf{x})$ the structure of the chosen path is clearly visible on $z_S^\infty(\mathbf{x})$ [Fig. 5.4 (b)].

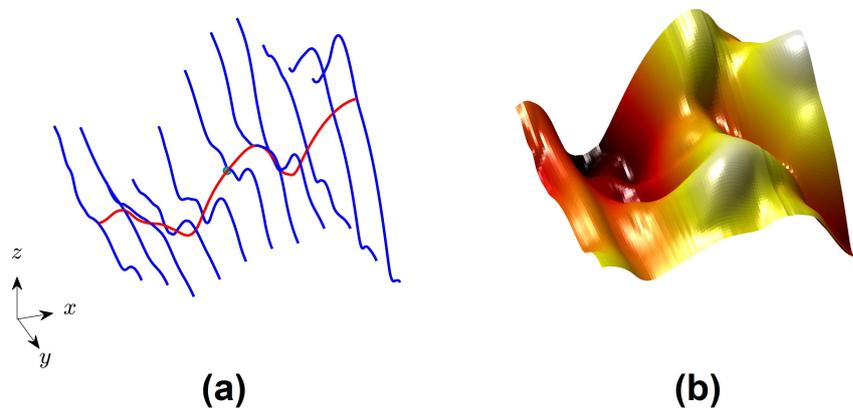


Figure 5.4: (a) The initial surface is calculated by a path integration with the integration constant of the surface at center of the freeform (green dot). The surface is integrated first along the x-direction (red curve) and then along the y-direction (blue curves) to give the (b) initial surface. A line structure along the y-direction is visible on the freeform surface. This is a result of the error accumulation along the x-direction and the subsequent error accumulation along the y-direction. Since the L_2 OMT map is not integrable, the initial surface will be different for different integration paths.

For the numerical computation of the solution of the differential equations (5.15), MATLABs ode45 solver, which is based on the Runge-Kutta method [102, 103], is applied.

To give an impression of the difference in quality of an initial and a final single freeform lens for a concrete design example, Fig. 5.6 (c) and 5.6 (d) below show the simulated irradiance distributions for both cases.

5.3.2 Double freeform surface

Similar to single freeform surfaces, also differential equations for double freeform surfaces for irradiance and phase control can be derived. Therefore, not only predefined input ray directions $\hat{\mathbf{s}}_1(\mathbf{x})$ and the initial mapping $\mathbf{u}^\infty(\mathbf{x})$ are required but also predefined outgoing ray directions $\hat{\mathbf{s}}_3(\mathbf{x})$ (see section 3.2 and Fig. 3.3). Consequently, both freeform surfaces need to be determined simultaneously from the given data. This then leads to a system of *coupled* ordinary differential equations. By utilizing the same procedure as for single freeforms, it follows [94]

$$\partial_x z_{I,S}^\infty(\mathbf{x}) = \frac{n_{I,x}(\mathbf{x}) + h_{I,x}(\mathbf{x})[z_{I,S}^\infty(\mathbf{x}) - z_0]}{1 - n_{I,x}(\mathbf{x}) \frac{(\hat{\mathbf{s}}_1)_x(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} - n_{I,y}(\mathbf{x}) \frac{(\hat{\mathbf{s}}_1)_y(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})}}, \quad \partial_y z_{I,S}^\infty(\mathbf{x}) = \frac{n_{I,y}(\mathbf{x}) + h_{I,y}(\mathbf{x})[z_{I,S}^\infty(\mathbf{x}) - z_0]}{1 - n_{I,x}(\mathbf{x}) \frac{(\hat{\mathbf{s}}_1)_x(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} - n_{I,y}(\mathbf{x}) \frac{(\hat{\mathbf{s}}_1)_y(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})}}, \quad (5.16)$$

with the coefficients

$$\begin{aligned} h_{I,x}(\mathbf{x}) &= n_{I,x}(\mathbf{x}) \partial_x \left(\frac{(\hat{\mathbf{s}}_1)_x(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \right) + n_{I,y}(\mathbf{x}) \partial_x \left(\frac{(\hat{\mathbf{s}}_1)_y(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \right), \\ h_{I,y}(\mathbf{x}) &= n_{I,x}(\mathbf{x}) \partial_y \left(\frac{(\hat{\mathbf{s}}_1)_x(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \right) + n_{I,y}(\mathbf{x}) \partial_y \left(\frac{(\hat{\mathbf{s}}_1)_y(\mathbf{x})}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \right), \\ n_{I,x}(\mathbf{x}) &\equiv -\frac{n_1(\hat{\mathbf{s}}_1)_x(\mathbf{x}) - n_2(\hat{\mathbf{s}}_2)_x(\mathbf{x})}{n_1(\hat{\mathbf{s}}_1)_z(\mathbf{x}) - n_2(\hat{\mathbf{s}}_2)_z(\mathbf{x})}, \quad n_{I,y}(\mathbf{x}) \equiv -\frac{n_1(\hat{\mathbf{s}}_1)_y(\mathbf{x}) - n_2(\hat{\mathbf{s}}_2)_y(\mathbf{x})}{n_1(\hat{\mathbf{s}}_1)_z(\mathbf{x}) - n_2(\hat{\mathbf{s}}_2)_z(\mathbf{x})}, \end{aligned} \quad (5.17)$$

which represent the generalization of Eq. (5.15) with z_T replaced by $z_{II,S}^\infty(\mathbf{x})$.

Additionally, by applying Eq. (5.13) to the normal vector field of the second surface, an extra pair of equations

$$\begin{aligned} \partial_x z_{II,S}^\infty(\mathbf{x}) &= \frac{n_{II,x}(\mathbf{x}) \partial_x u_x + n_{II,y}(\mathbf{x}) \partial_x u_y + h_{II,x}(\mathbf{x})[z_{II,S}^\infty(\mathbf{x}) - z_T]}{1 - n_{II,x}(\mathbf{x}) \frac{(\hat{\mathbf{s}}_3)_x(\mathbf{u})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} - n_{II,y}(\mathbf{x}) \frac{(\hat{\mathbf{s}}_3)_y(\mathbf{u})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})}}, \\ \partial_y z_{II,S}^\infty(\mathbf{x}) &= \frac{n_{II,x}(\mathbf{x}) \partial_y u_x + n_{II,y}(\mathbf{x}) \partial_y u_y + h_{II,y}(\mathbf{x})[z_{II,S}^\infty(\mathbf{x}) - z_T]}{1 - n_{II,x}(\mathbf{x}) \frac{(\hat{\mathbf{s}}_3)_x(\mathbf{u})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} - n_{II,y}(\mathbf{x}) \frac{(\hat{\mathbf{s}}_3)_y(\mathbf{u})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})}} \end{aligned} \quad (5.18)$$

with the coefficients

$$\begin{aligned} h_{II,x}(\mathbf{x}) &= n_{II,x}(\mathbf{x}) \partial_x \left(\frac{(\hat{\mathbf{s}}_3)_x(\mathbf{u})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} \right) + n_{II,y}(\mathbf{x}) \partial_x \left(\frac{(\hat{\mathbf{s}}_3)_y(\mathbf{u})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} \right), \\ h_{II,y}(\mathbf{x}) &= n_{II,x}(\mathbf{x}) \partial_y \left(\frac{(\hat{\mathbf{s}}_3)_x(\mathbf{u})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} \right) + n_{II,y}(\mathbf{x}) \partial_y \left(\frac{(\hat{\mathbf{s}}_3)_y(\mathbf{u})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} \right), \\ n_{II,x}(\mathbf{x}) &\equiv -\frac{n_1(\hat{\mathbf{s}}_3)_x(\mathbf{u}) - n_2(\hat{\mathbf{s}}_2)_x(\mathbf{x})}{n_1(\hat{\mathbf{s}}_3)_z(\mathbf{u}) - n_2(\hat{\mathbf{s}}_2)_z(\mathbf{x})}, \quad n_{II,y}(\mathbf{x}) \equiv -\frac{n_1(\hat{\mathbf{s}}_3)_y(\mathbf{u}) - n_2(\hat{\mathbf{s}}_2)_y(\mathbf{x})}{n_1(\hat{\mathbf{s}}_3)_z(\mathbf{u}) - n_2(\hat{\mathbf{s}}_2)_z(\mathbf{x})} \end{aligned} \quad (5.19)$$

can be obtained. The system of differential equations in Eqs. (5.16) and (5.18) can again be solved by MATLABs ode45 solver. As an additional complication, the output vector field $\hat{\mathbf{s}}_3(\mathbf{x})$ thereby depends on the mapping $\mathbf{u}(\mathbf{x})$, which from a numerical viewpoint requires the repeated two dimensional interpolation of $\hat{\mathbf{s}}_3(\mathbf{x})$ during the ODE solving. This leads to an initial surface construction process, which is significantly slower than for single freeform surfaces (see Table 5.1 in section 5.4.).

Alternatively, it is also possible to eliminate the dependency on the second surface by solving the constant OPL condition (3.29). Due to the decoupling of Eq. (5.16) from Eq. (5.18), this presumably leads to a faster computation.

It is also important to note that the initial surface construction approach can be applied to the FID with a MAE like (3.11). Since this initial surface will not fulfill the energy conservation equation to the same degree as the L_2 OMT map, the numerical stability of solving process will be critical for complex irradiance distributions.

5.4 Design strategies

Hereafter, the design strategies for the single FID and the double FID problem will be summarized. It is assumed that the corresponding input data like incoming ray directions/wavefronts $\hat{\mathbf{s}}_1(\mathbf{x})/\phi^I(\mathbf{x})$, outgoing ray directions/wavefronts $\hat{\mathbf{s}}_3(\mathbf{x})/\phi^O(\mathbf{x})$, emittance and irradiance $I_S(\mathbf{x})$ and $I_T(\mathbf{x})$, refractive indices (n_1, n_2 , etc.) and the geometrical parameters of the freeform system (z_0, z_T , integration constants) are given. The input and output distributions can be either experimental data or synthetic data from ray tracing simulations. Then the workflows given in Fig. 5.5 can be applied directly.

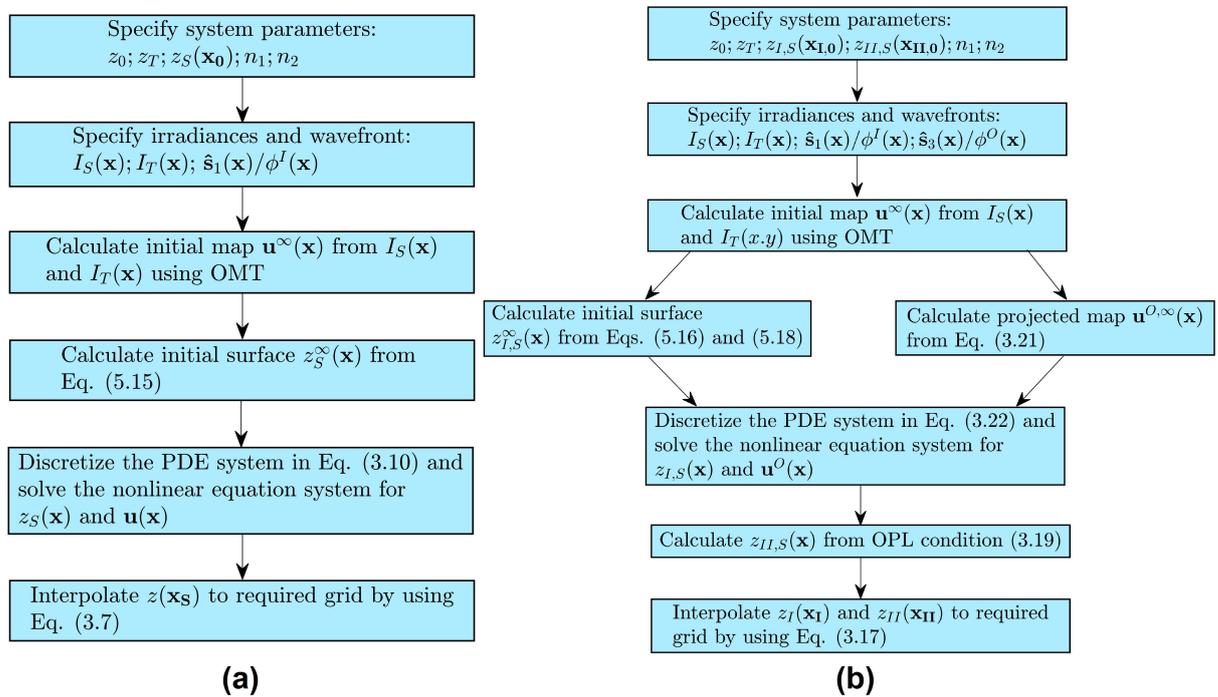


Figure 5.5: Workflow for (a) single freeform surfaces and (b) double freeform surfaces. [86]

In both cases the initial map is calculated from a numerical L_2 OMT algorithm [section 5.2], the initial surface/s is/are constructed from the ordinary differential equations [section 5.3] by e.g. the Runge-Kutta method and these functions serve as an initial iterate for the nonlinear equation system solving with a Newton-type method [section 5.1]. The different stages of the single FID process are illustrated in Fig. 5.6.

The major difference between the single FID and the double FID is the necessity to project the initial mapping coordinates $\mathbf{u}^\infty(\mathbf{x})$ onto the output wavefront $\phi^O(\mathbf{x})$. Numerically this projection can straightforwardly be done by applying Eq. (3.21), which represents a nonlinear equation for every coordinate $\mathbf{u}^\infty(\mathbf{x})$ (see section 5.5.4).

Furthermore, for the double FID, the second freeform surface $z_{II,S}(\mathbf{x})$ is calculated from $z_{I,S}(\mathbf{x})$ by the constant OPL condition (3.19).

After solving the discretized PDE systems, the numerical design approach will provide the dis-

cretized surfaces $z_S(\mathbf{x})$ or $z_{I,S}(\mathbf{x})$ and $z_{II,S}(\mathbf{x})$. According to Eqs. (3.7), (3.9) and (3.17), the freeform surfaces are given on the scattered grid points \mathbf{x}_S or \mathbf{x}_I and \mathbf{x}_{II} , respectively. To validate the calculated surfaces by a ray tracing simulation, e.g. in a commercial software like Zemax OpticStudio, it is necessary to interpolate the scattered freeform data onto the required grid points. In general, one can choose between different interpolation schemes, whose quality is determined *a posteriori* by a ray tracing simulation, or by experimental results after a manufacturing process. For the design examples in chapters 6 and 7, the bicubic interpolation of the function *RegularizeData3D*, available on the MATLAB file exchange website [105], with a smoothness to fidelity ratio of 0.00001, will be applied. This generates freeform data without oscillations in the simulated irradiances like it can be observed for the interpolation with MATLAB 2015b's built-in *scatteredInterpolant()* function.

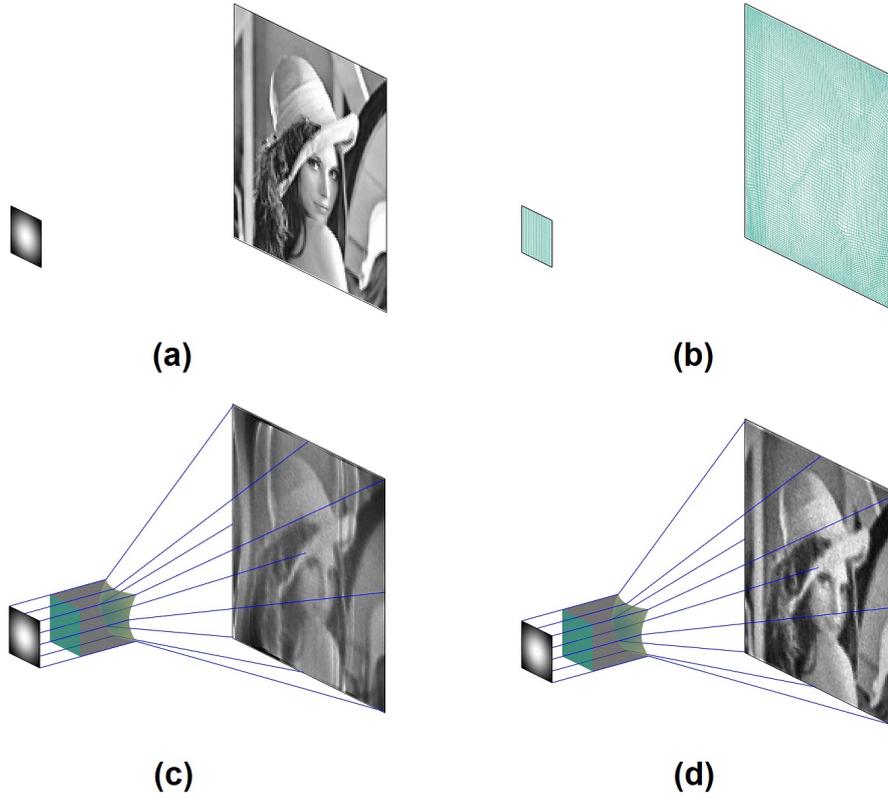


Figure 5.6: Illustration of the workflow for the single FID. (a) Predefined emittance and irradiance. (b) Calculation of the L_2 OMT map $\mathbf{u}^\infty(\mathbf{x})$ between the emittance and irradiance. (c) Construction of the initial surface. Due to the non-paraxial design geometry, $\mathbf{u}^\infty(\mathbf{x})$ is not integrable and the simulated irradiance shows strong deviations from the predefined irradiance. (d) Final surface with simulated irradiance distribution. In case of the double FID there is an additional step in which the ray mapping $\mathbf{u}^\infty(\mathbf{x})$ is projected onto the output wavefront.

In Table 5.1, to illustrate the scaling of the time consumption, typical computational times of every design step are shown as a function of the irradiance resolution for a single freeform lens with a spherical input wavefront [Fig. 5.7(a)] and a double freeform lens with a spherical input and output wavefront [Fig. 5.7(b)]. Since both designs use a lambertian intensity with the same

square input aperture and the same irradiance, the initial maps $\mathbf{u}^\infty(\mathbf{x})$ for both systems are identical.

For the calculation of $\mathbf{u}^\infty(\mathbf{x})$, a tolerance of $Tol_{256} = 10$ (defined in Alg. 5.1) for 256×256 pixels was chosen and the other tolerances were defined by $Tol_N = Tol_{256} \cdot (N/256)^2$. These, according to experience, correspond to values, which lead to a good compromise between the achievable irradiance pattern quality and the computational time. As pointed out in section 5.2.2, due to the interpolation of the mapping components after applying Alg. 5.1, Tol correlates with quality of the mapping \mathbf{u}^∞ , but does not directly represent the quality.

For a stable convergence on the other hand the stepsizes $dt_{256} = 5 \cdot 10^{-5}$ [Alg. 5.1] and $dt_N = dt_{256} \cdot (256/N)^2$ were utilized, which causes an increase in computational time with N . Further increase in time consumption mainly results from the interpolation of $I_T(\mathbf{x})$ at $\mathbf{u}(\mathbf{x})$ in the algorithm 5.1.

To calculate the initial surfaces with MATLABs *ode45* solver, tolerances of 10^{-8} were applied. These tolerances [108] determine the precision of the solution of the differential equations (5.15) and (5.16). Hence, they define how accurate the initial surface $z_S^\infty(\mathbf{x})$ maps the incoming rays onto the points $\mathbf{u}^\infty(\mathbf{x})$ along the one dimensional integration path in the x - or y direction. Since $\mathbf{u}^\infty(\mathbf{x})$ is not integrable [chapter 4] and the initial surfaces $z_S^\infty(\mathbf{x})$ is integration-path dependent [section 5.3.1], the tolerances of *ode45* only have limited informative value on the quality of $z_S^\infty(\mathbf{x})$ regarding the intended energy redistribution. Higher tolerances than the utilized 10^{-8} will lead to improved simulated irradiances of the initial surfaces (e.g. tolerances 10^{-10} are utilized in section 6.1), but usually do not have a significant effect on the convergence of the root finding of the PDE systems.

For the root-finding with MATLABs *fsolve()*, the value of the function tolerance $TolFun$ determines the precision of the approximation of the root [109]. An equation system $\mathbf{F}(\mathbf{X}) = 0$ is then considered as solved if $|\mathbf{F}(\mathbf{X})|^2 < TolFun^2$. Hence, $TolFun$ describes the deviation of the approximate discrete solution of the PDE systems (3.10) and (3.22) from its root. In case of FID, the equation system $\mathbf{F}(\mathbf{X}) = 0$ consists of the discretized energy conservation equation and both ray-mapping equations. Hence, in case of the energy conservation equation $TolFun$ describes the numerical deviation of the mapped irradiance $I_T(\mathbf{x})$ from the emittance $I_S(\mathbf{x})$, and for the ray-mapping equations $TolFun$ describes the deviation of the ray target points, which are generated by the freeform surface, from the mapping $\mathbf{u}(\mathbf{x})$ that is defined by the energy conservation equation. In the numerical experiments the function tolerance is chosen to be $TolFun = 10^{-4}$. This is approximately of the same order of magnitude as the precision of the initial solution $\mathbf{u}^\infty(\mathbf{x})$ of the energy conservation equation in $\mathbf{F}(\mathbf{X}) = 0$ and ensures a similar precision of the solution of the ray mapping equations in the PDE systems. This means that the summed quadratic difference between the ray target points and the mapping $\mathbf{u}(\mathbf{x})$ is approximately $TolFun^2$. The deviation should thereby be sufficiently small to be negligible in the energy conservation equation, whose discretization error of the FDs [section 5.1] scales quadratically with the grid spacing, which e.g. for $N = 512$ is approximately of order 10^{-6} . For the chosen $TolFun$, the number of necessary iterations for the root-findings varied between two and three.

Typical achievable accuracies of the calculated freeform systems with regard to the difference between the predefined and simulated irradiances, or predefined and simulated wavefronts are provided for the design examples in chapter 6.³ Since the energy conservation equation in (3.10) and (3.22) is already fulfilled by the initial map the difference between the predefined and simulated irradiances is usually in the range of the precision of the L_2 OMT map defined by Tol .

Table 5.1: Typical computational times for the single and double FID with $N \times N$ pixels.

N	$\mathbf{u}^\infty(\mathbf{x})$	$z_S^\infty(\mathbf{x})$	$z_S(\mathbf{x})$	$\mathbf{u}^{O,\infty}(\mathbf{x})$	$z_{I,S}^\infty(\mathbf{x})$	$z_{I,S}(\mathbf{x})$
128	50 s	4.7 s	22.6 s	1.8 s	49.7 s	26.5 s
192	162 s	7.2 s	61.0 s	8.6 s	101.8 s	70.5 s
256	366 s	9.9 s	231.3 s	58.6 s	170.5 s	248.8 s
384	1240 s	16.0 s	1197.6 s	335.1 s	369.8 s	1243.7 s
512	2728 s	23.1 s	3887.9 s	1058.1 s	696.2 s	3302.8 s

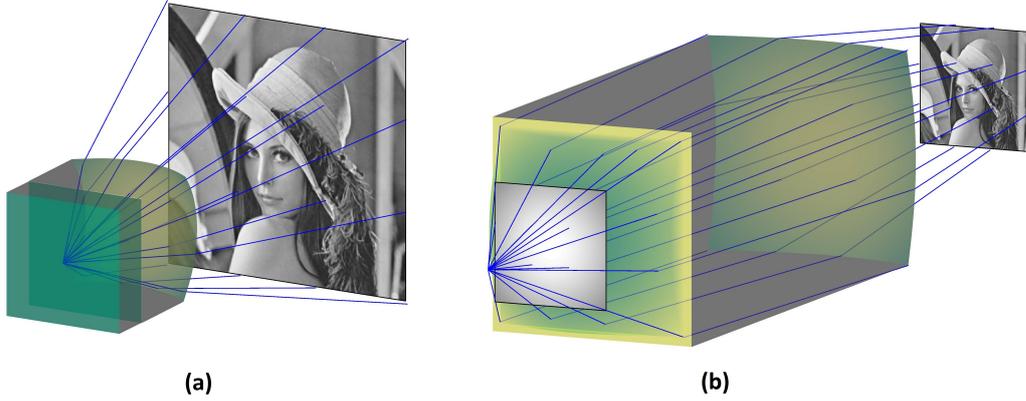


Figure 5.7: (a) Single freeform lens design with a spherical input wavefront. (b) Double freeform lens design with a spherical input and output wavefront.

5.4.1 Freeform illumination design in optical systems

In the following, the numerical design strategies for freeform surfaces in optical systems with predefined exit and/or entrance surfaces before and after the freeform surfaces, are discussed.

For FID in optical systems not only the mapping coordinates and the surface(s) need to be determined, but also the unknown intersection points of the rays with the predefined surfaces. For the root finding of the corresponding PDE system (3.30), the surface intersection coordinates $\mathbf{x}_i(\mathbf{x}), i = j + 1, \dots, N$ and the constraints (3.14)/(3.31)/(A.5) need to be discretized by the FD scheme (5.4) and (5.5). This leads to two additional equations for every unknown discretized surface intersection point $\mathbf{x}_i(\mathbf{x})$, which in contrast to the discretized PDEs are not coupled directly to neighbouring points. Hence, MATLABs trust-region reflective solver can still be applied.

³In chapter 6, rms values for the difference between the predefined and simulated irradiance for design examples with 250×250 pixels are provided. The rms values for other resolutions scale approximately with $(250/N)^2$

The initial surface construction approach from section 5.3 on the other hand cannot be applied directly and initial iterates $\mathbf{x}_i^\infty(\mathbf{x})$ for the unknown ray intersection points need to be determined. In the following, the generalization of the initial surface construction process from section 5.3 is discussed. Hereafter, it is distinguished between ray intersection points with *entrance* surfaces and with *exit* surfaces.

Assuming predefined *entrance* surfaces only, the input ray direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ in Eqs. (5.15) and (5.16) can be determined by a ray aiming [section 5.5.1] onto a cartesian grid on a plane between the last entrance surface and the (first) freeform surface. The Eqs. (5.15) and (5.16) can therefore still be applied.

For predefined *exit* surfaces, the initial surface construction for single freeforms and a double freeforms is considered separately. At first, a pairwise presence of double freeforms in the optical system is assumed, meaning that there is no predefined surface *between* both freeforms. Then, the initial map $\mathbf{u}^\infty(\mathbf{x})$ and the predefined output wavefront $\phi^O(\mathbf{x})$ can be applied to propagate the map to an intermediate plane between the second freeform and first exit surface by a ray tracing. The intermediate map $\mathbf{u}_I^\infty(\mathbf{x})$ instead of $\mathbf{u}^\infty(\mathbf{x})$ is then used in Eqs. (5.16) and (5.18), which can be integrated as before. If the double freeform are not pairwise appearing, Eqs. (5.16) and (5.18) cannot be applied. In that case, simple initial surface shapes (e.g. spheres or planes) can be applied and the initial surface intersection points $\mathbf{x}_i^\infty(\mathbf{x})$ are determined by a ray aiming [section 5.5.1] onto $\mathbf{u}^\infty(\mathbf{x})$. Consequently, the initial surface shapes should be chosen so that the resulting irradiance covers Ω_T

For the initial surface construction with single freeform surfaces additional difficulties arise due to the dependence of the ray directions after the freeform on the freeform surface itself. Considering Eq. (5.14), this means that $\hat{\mathbf{s}}_2(\mathbf{x})$ needs to be expressed in terms of the unknown surface intersection points after the freeform, rather than $\mathbf{u}^\infty(\mathbf{x})$. One approach to overcome this, was applied in Ref. [87]. There, at first a simple surface shape, like a plane surface, was chosen for the freeform. Then, a ray aiming at the $\mathbf{u}^\infty(\mathbf{x})$ was applied and the resulting surface intersection points were used to define $\hat{\mathbf{s}}_2(\mathbf{x})$ and Eq. (5.14) integrated to get $z_S^\infty(\mathbf{x})$. For a better approximation of $z_S^\infty(\mathbf{x})$, the process was repeated iteratively with $z_S^\infty(\mathbf{x})$ replacing the simple surface shape from the initial step. As pointed out before, the application of simple initial surface shapes in the nonlinear equation system solving process in general leads to a higher number of iterations with the trust-region-reflective solver. While this leads to increased computational times, by experience, the stability of the numerical convergence is not influenced.

5.5 Preparation of input and output data for design process

While the design algorithm given in section 5.5 can readily be applied, the input and output data like ray directions, emittances and wavefronts still need to be prepared for the design algorithm itself. Especially in optical systems this might not be trivial. Fortunately, efficient numerical techniques based on the description of ray paths in optical systems from section 3.3 can be applied as it will be shown in the following.

5.5.1 Ray-aiming

As pointed out above, the input ray direction vector field $\hat{\mathbf{s}}_{In}(\mathbf{x})$ needs to be determined on a cartesian grid to utilize the design strategies from section 5.4. To determine these vector fields, the description of ray paths in Eqs. (3.23) and (3.26) can be applied directly. Therefore, a zero-étendue light source, N predefined entrance surfaces and a cartesian target grid defined by \mathbf{x} is assumed [Fig. 3.5]. Then, $\mathbf{x}_{N+1}(\mathbf{x}_0) \stackrel{!}{=} \mathbf{x}$ is required and (3.23) and (3.26) solved simultaneously for all surface intersections points and the initial points \mathbf{x}_0 . From the resulting intersection coordinates \mathbf{x}_N and the target grid \mathbf{x} , the vector field $\hat{\mathbf{s}}_{In}(\mathbf{x})$ can be determined directly.

Similar to the PDEs of the FID, the discretized Eqs. (3.23) and (3.26) can be solved numerically by applying MATLABs trust-region-reflective solver. Even though the ray paths for every grid point \mathbf{x} are independent of each other and can be calculated point by point, it is recommended to solve Eqs. (3.23) and (3.26) simultaneously by providing *fsolve()* with the structure of the Jacobian matrix for a fast calculation. To get an appropriate initial iterate for the nonlinear equation system solving of Eqs. (3.23) and (3.26), \mathbf{x}_0 and surface intersection points $\mathbf{x}_i(\mathbf{x}), i = j + 1, \dots, N$ are estimated by a forward ray tracing on a cartesian input grid. The target points of the forward ray tracing can then be used to determine the support of the irradiance on the target surface, e.g by the MATLAB function *boundary()*. Subsequently, the target grid on the corresponding support can be defined with the MATLAB function *inpolygon()* and the function *knnsearch()* can be applied to determine all surface intersection points to generate the initial iterate for the nonlinear equation system.

5.5.2 L_2 OMT mapping inversion

Most of the L_2 OMT approaches in literature are limited to target distributions with convex boundary shapes [93]. For some practical applications in FID it might be helpful to first calculate the L_2 OMT by interchanging the convex irradiance with the concave emittance and subsequently calculate the inverse map. In Ref. [45], Prins et al. propose an algorithm for the inverse mapping calculation based on the Legendre-transform of the mapping potential.

Here, a straightforward mapping inversion approach based on the numerical ray aiming presented in the section 5.5.1 and the findings of chapter 4 is proposed. To do so, at first the *inverse* mapping between the given emittance with a convex boundary shape and irradiance with a concave boundary shape is calculated. According to the arguments given in chapter 4, the L_2 OMT mapping is asymptotically integrable for collimated input beams and an increasing distance between the freeform and the target plane. Consequently, a single freeform lens for a collimated beam and a large distance to the target $z_T \gg z(\mathbf{x})$ with the design method from section 5.4 is calculated. By using ray-aiming [section 5.5.1] onto a cartesian grid \mathbf{x}_T on the target plane, the inverse map then corresponds to the initial points \mathbf{x}_0 on the input plane: $\mathbf{u}(\mathbf{x}) \equiv \mathbf{x}_0(\mathbf{x}_T)$.

In Fig. 5.8(a), a design example with a concave target is shown.

The inverse L_2 OMT map between the emittance and the irradiance is thereby calculated with the standard FD techniques from section 5.2.1 and the subsequent mapping inversion is done by the

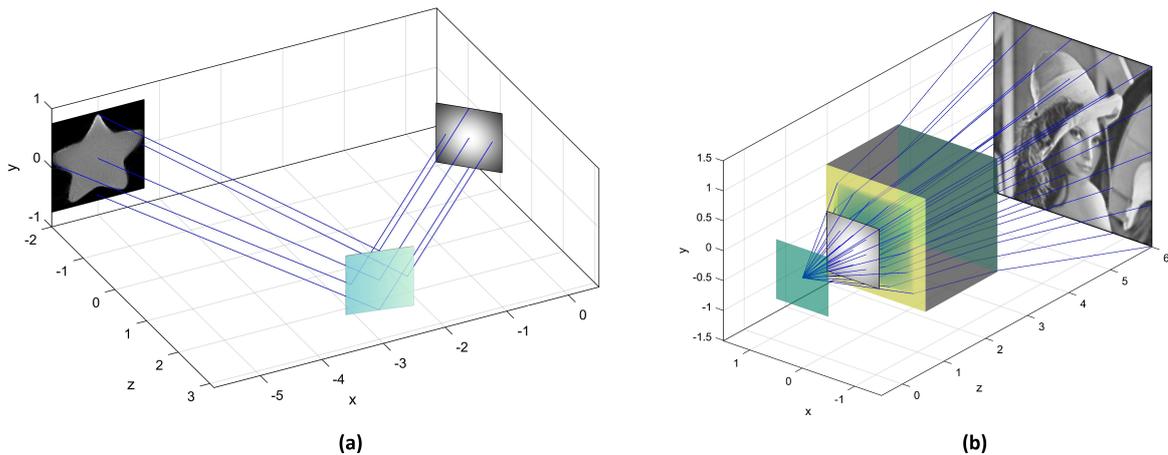


Figure 5.8: (a) Freeform mirror design for a collimated gaussian input beam, which is tilted relative to the detector plane. The irradiance is nonconvex, meaning that the black areas have zero energy. Slight inhomogeneities of the simulated irradiance are due to inaccuracies of the standard FD L_2 OMT algorithm from section 5.2.1. (b) Freeform-plano lens design. The irradiance distribution “Lena” was calculated by applying ray-aiming and using the energy conservation equation. [87]

described algorithm to get the L_2 OMT map. This map is then used for the single FID algorithm [Fig. 5.6 (a)] to give the freeform mirror. Interestingly, a modification of the transport boundary conditions (5.6) to non-square boundary shapes was not necessary for numerical convergence.

5.5.3 Noise-free irradiance computation in (sequential) optical systems

One major disadvantage of conventional Monte-Carlo raytracing is the raytracing noise due to the statistical distribution of input rays. Therefore, to compute accurate noise-free irradiance distributions on detectors with high resolutions, a large number of rays is required if smoothing techniques are prohibited. This is especially critical for FID with feedback optimization methods in which the extension of a light source is taken into account by a repeated freeform design with a subsequent ray tracing [80]. Thus, a fast computation method for a noise free irradiance is of interest.

According to the design model in section 3.3, this can be done by a calculation of the inverse map between the input plane $z = z_0$ and the target plane $z = z_T$ [section 5.5.1] and by applying the Jacobian equation (3.2). This does not only take into account the ray positions in the source and target plane but also their differential change between both planes [section 2.2.1].

To calculate the irradiance for an extended light source, the extended source is separated into a set of N point sources with emittances $I_S^i(\mathbf{x}_0)$ and for every point source the inverse maps $\mathbf{x}_0^i(\mathbf{x}_T)$, $i = 1, \dots, N$ between the input and output plane are determined. The irradiance of the extended light source is then the sum of all point source contributions, which gives⁴

$$I_T^{total}(\mathbf{x}_T) = \sum_{i=1}^N I_T^i(\mathbf{x}_T) = \sum_{i=1}^N \det(\nabla \mathbf{x}_0^i(\mathbf{x}_T)) I_S^i(\mathbf{x}_0^i(\mathbf{x}_T)). \quad (5.20)$$

⁴For a non-planar target surface the Jacobian equation is replaced by Eq. (A.2) from Appendix A.1.

The inverse maps $\mathbf{x}_0^i(\mathbf{x}_T)$, $i = 1, \dots, N$ can hereby either be determined by the ray-aiming approach from section 5.5.1 or, alternatively, by a scattered data interpolation of the mapping components. While the second approach allows much faster computations, it is less accurate since the surface information of the optical elements in the system are not taken into account in the mapping interpolation process.⁵

In Fig. 5.8(b), the described procedure was utilized to compute the irradiance distribution from the given plano-freeform lens and a point light source. For the two predefined surfaces and the target with 250 pixels \times 250 pixels, the ray-aiming and irradiance computation took approximately 78 s.

5.5.4 Wavefronts and mapping projection

In this section, the preparation of the input and output wavefront data, which are necessary for the single and double FID, is discussed.

For the double FID, the projection of the initial mapping $\mathbf{u}^\infty(\mathbf{x})$ onto the output wavefront $\phi^O(\mathbf{x})$ [Fig. 3.4] is needed. This projected mapping $\mathbf{u}^{O,\infty}(\mathbf{x})$ is used for the modified Jacobian equation in Eq. (3.22) and defines the ray directions $\hat{\mathbf{s}}_3^O(\mathbf{u}^{O,\infty})$. On the other hand, according to the solution of the constant OPL condition (3.19), the input and output wavefront $\phi_I(\mathbf{x}^I)$ and $\phi^O(\mathbf{u}^O)$ need to be determined.

The input wavefront $\phi_I(\mathbf{x})$ at the coordinates \mathbf{x}^I [Fig. 3.4] is determined by the constant OPL condition. For a system with predefined entrance surfaces and a collimated input beam or a point light source, a ray aiming [section 5.5.1] to a cartesian grid on a plane between the last entrance surface and the freeform is applied and every ray path is normalized by a constant OPL value which directly gives $\phi_I(\mathbf{x}^I)$.

To calculate the projected mapping coordinates for a given map $\mathbf{u}^\infty(\mathbf{x})$, wavefront $\phi_O(\mathbf{x})$ and ray directions $\nabla\phi_O(\mathbf{x})$, the nonlinear equations (3.21) are applied. These are solved by the trust-region-reflective solver of MATLABs optimization toolbox for every coordinate $\mathbf{u}^\infty(\mathbf{x})$ simultaneously to give $\mathbf{u}^{O,\infty}(\mathbf{x})$ from which $\phi_O(\mathbf{u}^{O,\infty})$ and the ray directions $\hat{\mathbf{s}}_3^O(\mathbf{u}^{O,\infty})$ follow immediately.

5.5.5 Wavefronts and ray directions

Depending on the physical situation, either the wavefront data $\phi(\mathbf{x})$ or the ray directions $\hat{\mathbf{s}}(\mathbf{x})$ on a plane $z = z_{rd}$ might be available. Thus, a conversion between those data might be necessary. If the wavefront $\phi(\mathbf{x})$ is given (consider for instance Fig. 3.4 with $\phi(\mathbf{x}) \equiv \phi_I(\mathbf{x})$ and $z_{rd} \equiv z_0$), the corresponding ray directions on a given Cartesian grid \mathbf{x}_{rd} can then be calculated by solving the pair of nonlinear equations

$$\mathbf{x}_{rd} \stackrel{!}{=} [\phi(\mathbf{x}_0) - z_{rd}] \nabla^\perp \phi(\mathbf{x})|_{\mathbf{x}_0} + \mathbf{x}_0 \quad (5.21)$$

for every \mathbf{x}_0 . This gives $\mathbf{s}(\mathbf{x}_{rd}) = \mathbf{x}_{rd} - \mathbf{x}_0$, which needs to be normalized.

⁵An example of such a superposition of point source contributions can be seen in Fig. 7.5 of chapter 7

For given ray directions $\hat{\mathbf{s}}(\mathbf{x})$, Eq. (5.15) can be integrated for $\phi(\mathbf{x}_S)$ by reducing Eq. (5.13) to

$$\frac{\mathbf{n}(\mathbf{x}_S)}{(\mathbf{n}(\mathbf{x}_S))_z} = \frac{\hat{\mathbf{s}}(\mathbf{x})}{(\hat{\mathbf{s}}(\mathbf{x}))_z} = \nabla_S(z - \phi(\mathbf{x}_S)); \quad \mathbf{x}_S = \frac{\phi(\mathbf{x}_S) - z_{rd}}{(\hat{\mathbf{s}})_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}})_x(\mathbf{x}) \\ (\hat{\mathbf{s}})_y(\mathbf{x}) \end{pmatrix} + \mathbf{x}. \quad (5.22)$$

The integration leads to $\phi(\mathbf{x}_S)$ with the scattered grid points \mathbf{x}_S , which is then interpolated onto the Cartesian grid \mathbf{x} .

5.6 Summary

A FID strategy for single and double freeform surfaces in optical systems with general zero-étendue input wavefronts was developed. It is based on the numerical solving of the PDE models from chapter 3 and motivated by the findings of chapter 4 on the connection of L_2 OMT and FID. The PDE models are thereby discretized by standard FDs to obtain systems of coupled nonlinear equations, which are solved by a Newton-like solver [section 5.1]. Of great importance for the stable numerical convergence to a root of the nonlinear equation system is the calculation of a suitable initial iterate for the mapping by L_2 OMT [section 5.2] and the construction of an initial surface by ordinary differential equations [section 5.3]. The explicit appearance of the energy conservation equation in the PDE systems (3.10), (3.22) and (3.30) is thereby the keypoint for the numerical stability for complex emittances and irradiances and represents a major advantage compared to the direct solving of a MAE like (3.11).

It is noteworthy that the L_2 OMT algorithms, discussed in section 5.2, are not essential to the presented design strategy and can be replaced by other available L_2 OMT algorithms in literature. In case that e.g. Sulman's algorithm [section 5.2.2] fails to converge, which might happen for strong gradients in the irradiance distribution, the algorithm can be replaced by a provably convergent L_2 OMT algorithm [93].

Furthermore, it was shown how the description of optical ray path from section 3.3.1 can be applied to ray-aiming and noise-free irradiance computation. These techniques will be applied for preparation of the input data of the design examples in chapters 6 and 7.

6 Design examples

In this chapter, the developed numerical design strategy from chapter 5 will be evaluated and its capabilities demonstrated. Design examples, which were not possible to calculate with previous design models are thereby shown. This means that single freeform surfaces for non-spherical and non-planar wavefronts and planar or spherical entrance surfaces, respectively, and double freeform surfaces with nonplanar input *and* output wavefronts without the restriction to paraxiality will be designed.

All the design examples are calculated on an Intel Core i3 at 2×2.4 Ghz with 16GB RAM. The design algorithms were implemented in MATLAB 2015b. For the initial surface construction MATLABs *ode45* solver is used. For the root-finding of the PDE systems MATLABs *fsolve* function with the trust-region-reflective solver is applied. The surface interpolation of the scattered freeform data is done with the function *RegularizeData3D* [105] using cubic interpolation and a smoothness to fidelity ratio of 0.00001. The design examples are evaluated in a ray-tracing simulation with a MATLAB ray tracing toolbox “MatLightTracer”, which is based on the ray-tracing procedure by Spencer and Murty [107]. The detector resolution thereby corresponds to the grid resolution of the computed freeform surfaces ($250 \text{ pixels} \times 250 \text{ pixels}$) and $200 \cdot 10^6$ rays are used for the irradiance computation. All freeform surfaces are interpolated onto a grid with $500 \text{ pixels} \times 500 \text{ pixels}$ to minimize errors at the freeform boundary, which for non-planar wavefronts is non-square. As irradiance distributions for the design, the test images “lake” and “boat” [Fig. 6.1] are used, which show smooth distributions of gray values as well as steep gradients and are interpolated bilinear.

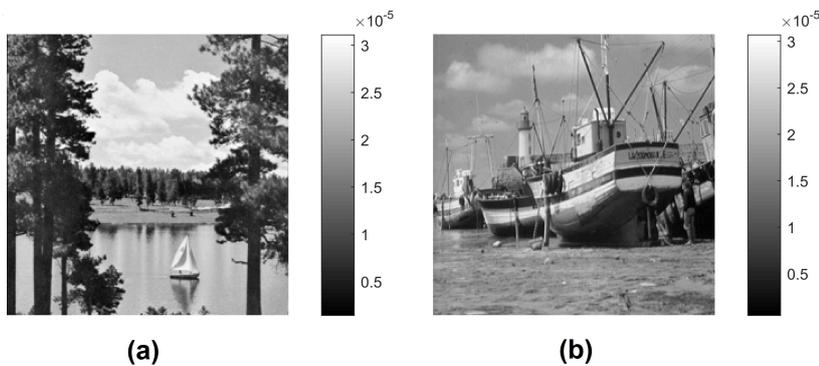


Figure 6.1: Test images (a) “lake” with $\langle I_T \rangle = 3.1094 \cdot 10^{-5}$ and $\max(I_T) = 1.6110 \cdot 10^{-5}$ and (b) “boat” with $\langle I_T \rangle = 3.0594 \cdot 10^{-5}$ and $\max(I_T) = 1.6126 \cdot 10^{-5}$. The minimum irradiance for both images is 0.4% of the maximum value.

The design examples for single freeform surfaces are presented in section 6.1, followed by double freeform surfaces in section 6.2. Remarks on the manufacturability of calculated freeform surfaces can be found in Appendix A.7.

The results of this chapter were published in Refs. [85] and [86].¹

6.1 Single freeform design

An important difference to the other PDE models in literature is the dependence of the PDEs in Eqs. (3.10) and (3.11) on the input vector field $\hat{\mathbf{s}}_1(\mathbf{x})$. Therefore, the FID for zero-étendue wavefronts is described by the same PDE (system) and the implementation of the design algorithm is identical. Thus, there is an infinite number of possible input wavefronts for evaluating the proposed design strategy from section 5.4. Since numerical convergence is not *proven*, but the functioning of the design algorithm should be supported with evidence, several relevant design examples will be presented in this section. The first two design examples consist of shaping a planar and spherical wavefront with a single freeform and acquire their relevance due to their importance for e.g. laser beam shaping. Furthermore, they represent standard examples for the FID in (current) literature. The third design example consists of shaping an astigmatic input wavefront and shows the ability of the design algorithm to handle non-standard input wavefronts, which is, for instance, important for the shaping of laser diodes, but it also demonstrates the possibility of preshaping input beams by predefined entrance surfaces.

In Table 6.1 all the required input data for the examples to initialize the design algorithm are summarized. These are the OMT parameters [section 5.2.2], the geometrical data of the design examples and the tolerances of *ode45* for the initial surface computation [section 5.3] and *fsolve* for the root-finding of the discretized PDE system [section 5.1].² The geometrical data, consisting of the input plane $z = z_0$, the integration constant $z_{S,0}$ of the freeform, the target plane $z = z_T$ and the detector extension and position are hereby measured relative to the extension of the input beam on $z = z_0$, which is defined on a square area of sidelength 1 arb. unit with the center at $(x, y) = (0, 0)$.

At first, the shaping of a collimated beam with a Gaussian emittance by a freeform mirror [Fig. 6.2 (a)] is considered. As shown in section 4.1, this problem is solved asymptotically in a paraxial regime by the L_2 OMT map. This paraxiality in general can be ensured by choosing a large distance between the freeform surface and the target plane, which is fixed by the integration constant $z_{S,0}$ and $z = z_T$. Since applications might have restrictions regarding the spatial extension of the experimental setup, a design geometry within a non-paraxial regime will be considered.

As it can be seen in Fig. 6.2 (a), the non-paraxiality results from the short freeform-detector distance as well as the detector size, which is three times as large as the input beam, and from

¹Additional design examples can be found in Ref. [87].

²While the chosen tolerance for the root-finding is one order of magnitude smaller than for the examples in Table 5.1, the number of necessary iterations and the final precision is the same.

Table 6.1: OMT parameters, design geometries and root-finding parameters for the design examples.

	Collimated	PLS	Astigmatic
Resolution	250×250	250×250	250×250
Emittance	$I_0 \cdot \exp[-4 \cdot (x^2 + y^2)]$	Lambert: $\theta_{max} = 30deg$	[section 5.5.3]
Irradiance	“lake”	“lake”	“lake”
Tolerance Tol [Alg. 5.1]	10	10	10
Stepsize dt [Alg. 5.1]	$5 \cdot 10^{-6}$	$5 \cdot 10^{-6}$	$5 \cdot 10^{-6}$
z_0	0 arb. unit	1.2247 arb. unit	5 arb. unit
$\hat{\mathbf{s}}_1(\mathbf{x})$	$(0, 0, 1)^T$	$\frac{(x, y, z_0)^T}{\sqrt{x^2 + y^2 + z_0^2}}$	[section 5.5.1]
$z_{S,0}$	5 arb. unit	1 arb. unit	6 arb. unit
z_T	0 arb. unit	10 arb. unit	11 arb. unit
Refractive index	-1	1.5	1.5
Detector size	3×3 arb. unit ²	5×5 arb. unit ²	2×2 arb. unit ²
Detector center	$(-5, 0)$ arb. unit	$(0, 0)$ arb. unit	$(0, 0)$ arb. unit
Tol. $ode45()$ for $z_S^\infty(\mathbf{x})$	10^{-10}	10^{-10}	10^{-10}
Tol. $fsolve()$ for $z_S(\mathbf{x})$	10^{-3}	10^{-3}	10^{-3}
Max. iter. $fsolve()$ for $z_S(\mathbf{x})$	3	3	3

the detector decentering relative to the input beam.

In the second example, a PLS with a Lambertian intensity is mapped by a single freeform lens onto “lake” [Fig. 6.2 (f)]. To consider a spherical wavefront, the point light source (PLS) needs to be placed either within the lens or in the center of a spherical entrance surface. To apply the design algorithm, the Lambertian intensity is projected onto the input plane $z = z_0$ to obtain the emittance $I_S(\mathbf{x})$ and to define the input direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$. A maximum opening angle of $\theta_{max} = 30$ deg is chosen and $z = z_0$ is placed so that $\hat{\mathbf{s}}_1(\mathbf{x})$ and $I_S(\mathbf{x})$ are defined on a square with a side length of 1 arb. unit. It is important to note that for a fixed detector size and position as well as a fixed integration constant of the freeform, the maximum possible angle θ_{max} will be limited by total internal reflection. A preshaping with entrance surfaces might therefore be beneficial. It is also important to note that since $\hat{\mathbf{s}}_1(\mathbf{x})$ is defined on a cartesian grid and does not represent a collimated beam, $z_S(\mathbf{x})$ will be defined on a scattered grid \mathbf{x}_S . The same is true for the next design example.

There, a collimated beam with Gaussian emittance (waist 1) is considered, which is first redistributed by the predefined astigmatic surface $z_{pre}(\mathbf{x}) = \frac{1}{9}y^2 - \frac{2}{3}x^2 + 2$ and a plane lens entrance surface at $z = 3$ arb. unit of a plano-freeform lens [Fig. 6.2 (k)]. Consequently, the input direction vector field $\hat{\mathbf{s}}_1(\mathbf{x})$ and the emittance $I_S(\mathbf{x})$ are determined numerically after the entrance surfaces at $z = 5$ arb. unit. In contrast to the first two examples, there is no analytical expression for $\hat{\mathbf{s}}_1(\mathbf{x})$ and the input vector field and emittance need to be determined on a cartesian grid, which is done by the procedures described in sections 5.5.1 and 5.5.3.

After fixing the required input parameters, the OMT map $\mathbf{u}^\infty(\mathbf{x})$ is calculated, the initial surface $z_S^\infty(\mathbf{x})$ is integrated and the discretized PDE system (3.10) solved by the root finding. The computational times of each design step are summarized in Table 6.2.

Since the irradiance distribution for the three design examples are equal, the different computa-

Table 6.2: Computational time.

	Collimated	PLS	Astigmatic
$\mathbf{u}^\infty(\mathbf{x})$	320 s	378 s	417 s
$z_S^\infty(\mathbf{x})$	68 s	183 s	145 s
$z_S(\mathbf{x})$	208 s	213 s	215 s

tional times for determining $\mathbf{u}^\infty(\mathbf{x})$ are due to the individual structures of the emittances. While the time consumption of the root-finding with the nonlinear equation system solver is nearly the same for the design examples, the integration of the initial surface $z_S^\infty(\mathbf{x})$ is significantly influenced by the structure of the input vector field. For instance, in case of the design example with the collimated input beam, the x - and y -component of the vector field vanish and the corresponding components are not interpolated during the Runge-Kutta method, which leads to the lowest computational time of $z_S^\infty(\mathbf{x})$. This is in contrast to the root-finding of the nonlinear equation system, which does not require the interpolation of the vector fields and shows similar computational times for each design example.

It is noteworthy that while smaller tolerances of *ode45* will allow faster calculations of $z_S^\infty(\mathbf{x})$ and also generate equally sufficient initial surfaces for a fast convergence of the root finding, the precision is required for a meaningful quality comparison between the initial surface and the final surface in a ray-tracing simulation, which can be seen in Fig. 6.2. That means that the tolerance of *ode45* [Table 6.1] was decreased until the simulated irradiance of the initial surface did not significantly improve anymore.

For the evaluation of the quality of the calculated freeform surfaces, the corresponding irradiance distributions $I_{T,RT}(\mathbf{x})$ are calculated in ray-tracing simulations. These are then compared with the predefined irradiance distribution $I_T(\mathbf{x})$ by $\Delta I_T(\mathbf{x}) = I_T(\mathbf{x}) - I_{T,RT}(\mathbf{x})$. To give an impression on the improvement of the initial surface by the root-finding process, the ray-tracing results with the initial surface and the final surface as well as a comparison with the predefined irradiance are plotted. This can be seen in Fig. 6.2 for the collimated input beam [Fig. 6.2 (b)-(e)], for the PLS [Fig. 6.2 (g)-(j)] and for the astigmatic input wavefront [Fig. 6.2 (l)-(o)].

It is important to note that many publications in FID literature forego a quality assessment of the simulated irradiances and that there is no commonly utilized quality metric for comparison of the prescribed irradiance and simulated irradiance for complex patterns. This makes a general cross comparison between different publications and design methods difficult. Nevertheless, in Table 6.3, metrics for the quality assessment of the simulated irradiance distributions are provided. This includes the rms of $\Delta I_T(\mathbf{x})$, the correlation coefficient $corr_{I_T}$ between $I_T(\mathbf{x})$ and $I_{T,RT}(\mathbf{x})$, and the energy efficiency η . For a prescribed irradiance $I_T(\mathbf{x})$, which is defined on a uniform cartesian grid with $N \times M$ grid points, the rms of $\Delta I_T(\mathbf{x})$ is given by³

³In Ref. [47], the authors propose to either normalize the rms value by an arithmical average or the maximum value of the prescribed irradiance. These values are given in the caption of Fig. 6.1 and can be used to normalize the rms values in Tables 6.3 and 6.6.

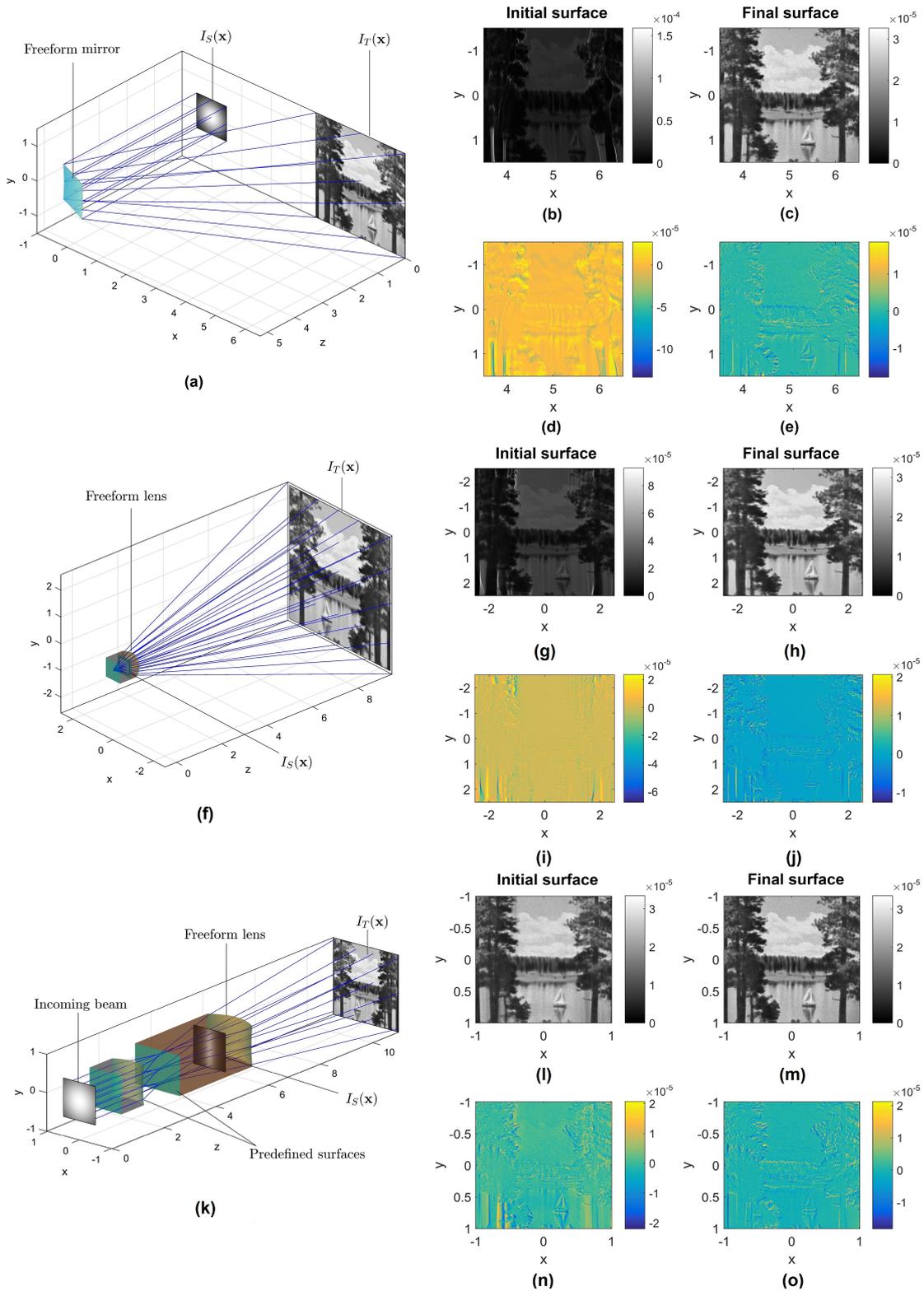


Figure 6.2: (a),(f),(k) Geometries of the design examples. (b), (g), (l) Simulated irradiance for the initial surface. (c), (h), (m) Simulated irradiance for the final surface. (d), (i), (n) Difference between predefined irradiance and simulated irradiance for the initial surface. (e), (j), (o) Difference between predefined irradiance and simulated irradiance for the final surface. [85]

$$rms_{\Delta I_T} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (\Delta I_T)_{i;j}^2}. \quad (6.1)$$

The correlation coefficient is defined as

$$corr_{I_T} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \left(\frac{\langle (I_T)_{i;j} - \mu_{I_T} \rangle}{\sigma_{I_T}} \right) \left(\frac{\langle (I_{T,RT})_{i;j} - \mu_{I_{T,RT}} \rangle}{\sigma_{I_{T,RT}}} \right) \quad (6.2)$$

with the mean deviations μ_I and standard deviations σ_I of the respective quantities, and the energy efficiency η gives the percentage of energy of the simulated irradiance, which reaches the prescribed detector area.

Especially for the first example there is a significant improvement in pattern quality between the initial and final surface due to strong non-paraxiality of the setup. This example also shows the highest energy efficiency. The energy loss of the other examples is mainly due to the inter- and extrapolation of the scattered freeform data $z_S(\mathbf{x}) \equiv z(\mathbf{x}_S)$ onto a cartesian grid. This is required for the ray-tracing simulation and leads to inaccuracies at the freeform boundary. In particular, the simulated irradiance for the initial freeform surface in the third example [Fig. 6.2 (l)] already shows a comparably high quality. Observing Fig. 6.2 (k), it is reasonable to conclude that this is due to the rather paraxial deflection of incoming rays at the freeform surface. Due to the similarities in the structure of $I_S(\mathbf{x})$ and $I_T(\mathbf{x})$, for which the energy in both cases is focused mostly around the y -axis [Fig. 6.2 (k)], the necessary redistribution of energy by the freeform seems to be already described rather accurately by the L_2 OMT map.

Table 6.3: Comparison of ΔI_T for example “lake” [Fig. 6.1 (a)].

	$z_{coll}^\infty(\mathbf{x})$	$z_{coll}(\mathbf{x})$	$z_{PLS}^\infty(\mathbf{x})$	$z_{PLS}(\mathbf{x})$	$z_{ast}^\infty(\mathbf{x})$	$z_{ast}(\mathbf{x})$
$rms_{\Delta I_T}$	$9.2288 \cdot 10^{-6}$	$2.3472 \cdot 10^{-6}$	$4.1373 \cdot 10^{-6}$	$2.1480 \cdot 10^{-6}$	$3.2072 \cdot 10^{-6}$	$2.7386 \cdot 10^{-6}$
$corr_{I_T}$	0.5750	0.9607	0.8889	0.9668	0.9252	0.9464
η	85.66%	99.98%	99.62%	99.55%	99.23%	99.20%

6.2 Double freeform design

Similar to the single FID, an advantage of the PDE system (3.22) is its design flexibility. For double freeform surfaces not only $\hat{s}_1(\mathbf{x})$ serves as an input for the design algorithm but also the output wavefront $\phi^O(\mathbf{x})$. This allows an universal application of the same numerical computational routines. In the following, two examples will be presented, which will demonstrate the capability of being able to handle non-collimated input and output beams as well as non-paraxiality. The latter aspect will be of special importance for the next chapter in which the shaping of extended light sources by double freeform surfaces is considered.

The input parameters of the design examples are summarized in Table 6.4.

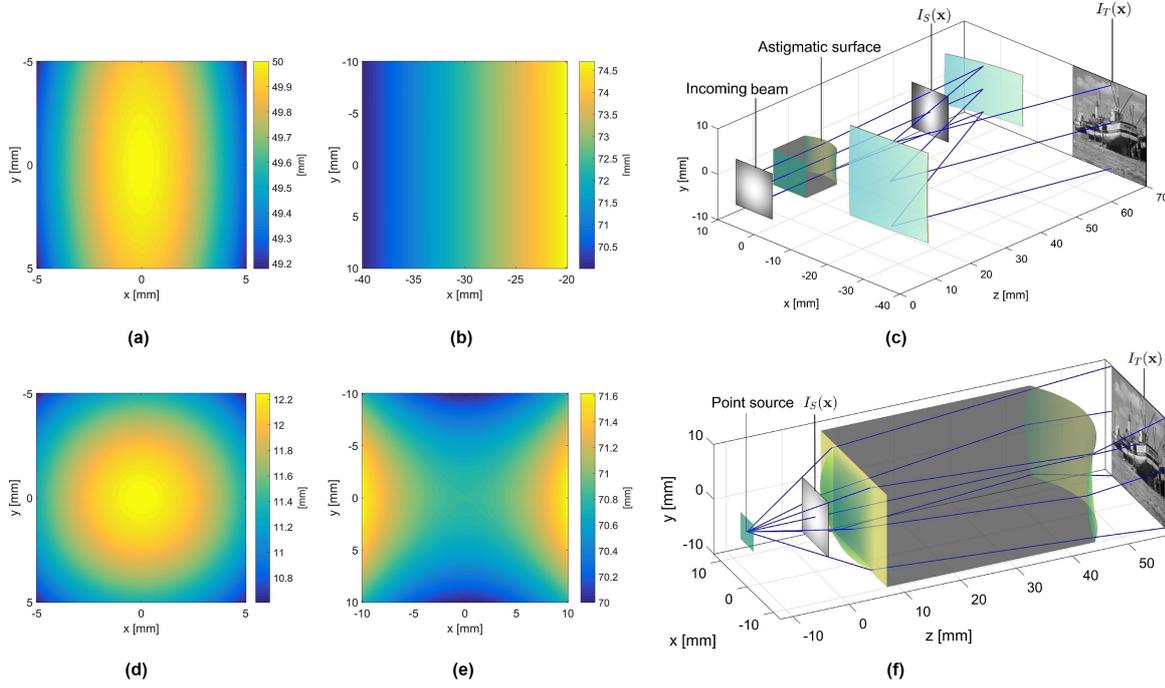


Figure 6.3: Absolute values of the prescribed (a) input wavefront and (b) output wavefront for (c) double mirror system. Absolute values of the prescribed (d) input wavefront and (e) output wavefront for (f) single lens system. [87]

Table 6.4: OMT parameters, design geometries and root-finding parameters for the design examples.

	Mirror system	Lens system
Resolution	250 × 250	250 × 250
Emittance	ray aiming [section 5.5.3]	Lambert: $\theta_{max} = 30deg$
Irradiance	“boat”	“boat”
Tolerance Tol [Alg. 5.1]	8	8
Stepsize dt [Alg. 5.1]	$6 \cdot 10^{-6}$	$6 \cdot 10^{-6}$
z_0	50 mm	12.247 mm
$\hat{\mathbf{s}}_1(\mathbf{x})$	ray aiming [section 5.5.1]	$\frac{(x, y, z_0)^T}{\sqrt{x^2 + y^2 + z_0^2}}$
$z_{I,S,0}$	65 mm	15 mm
$z_{II,S,0}$	20 mm	60 mm
z_T	70 mm	70 mm
Refractive index	-1	1.5
Detector size	20 mm × 20 mm	20 mm × 20 mm
Detector center	(-30, 0) mm	(0, 0) mm
$\phi^O(\mathbf{x})$	planar; tilted: 25 deg	astigmatic
Tol. $ode45()$ for $z_S^\infty(\mathbf{x})$	10^{-8}	10^{-8}
Tol. $fsolve()$ for $z_S(\mathbf{x})$	10^{-3}	10^{-3}
Max. iter. $fsolve()$ for $z_S(\mathbf{x})$	3	3

Consequently, compared to the single FID, not only the position of the first freeform in space by the integration constant $z_{I,S,0}$ needs to be specified, but also the position of the second freeform by $z_{II,S,0}$. Furthermore, the definition the required output wavefront $\phi^O(\mathbf{x})$ is required.

The first design example consists of the predefined astigmatic entrance surface of section 6.1 and two freeform mirrors [Fig. 6.3 (c)], which redistribute the emittance and the input wavefront [Fig. 6.3 (a)] at $z = 50$ mm into a planar but tilted output wavefront [Fig. 6.3 (b)] and the irradiance “boat”.

In the second example, a PLS with a Lambertian intensity is redistributed by a single lens with two freeform surfaces [Fig. 6.3 (f)] into “boat” and an astigmatic output wavefront [Fig. 6.3 (e)]. The maximum opening angle of the PLS, referring to the corners of the square input aperture, is 30 deg.

After specifying all required input parameters, the freeform surfaces are calculated by the design approach from Fig. 5.5 (b). The computational times of each design step are presented in Table 6.5.⁴ Additional time consumption, compared to the single FID, is due to the added wavefront mapping projection and due to the initial surface construction. The increased computational time of the latter is caused by the coupling of the two ordinary differential equations (5.16) and (5.18) for the two freeform surfaces and the repeated two dimensional interpolation of the outgoing vector field $\hat{\mathbf{s}}_3(\mathbf{x})$ at $\mathbf{u}^\infty(\mathbf{x})$ in the Runge-Kutta procedure [section 5.3.2]. Slight increases in the root-finding are mainly due to the higher complexity of the PDE system (3.22), which requires additional interpolations due to the additional vector fields and wavefronts.

Table 6.5: Computational time.

	Mirror system	Lens system
$\mathbf{u}^\infty(\mathbf{x})$	458 s	313 s
$\mathbf{u}^{O,\infty}(\mathbf{x})$	< 1 s	16.9 s
$z_{I,S}^\infty(\mathbf{x})$	269 s	251 s
$z_{I,S}(\mathbf{x})$	258 s	234 s

After designing the freeform surfaces, their quality is evaluated by ray-tracing simulations and a comparison with the predefined irradiance [Fig. 6.4 (a), (b), (d) and (e), and Table 6.6]. As it can be seen from Table 6.6 there is a slight deviation in quality between both design examples. Identifying the reason for the differences is in general a nontrivial task due to the complexity of the solving process of the PDE system (3.22), whose structure depends the specific input wavefront, output wavefront, emittance and irradiance. Additionally the ray-tracing simulation requires the interpolation of the numerical freeform data onto a regular Cartesian grid. Since the input wavefront is different in both examples, the scattered grid [Eq. (3.7)] on which the freeform sag values are distributed after solving Eq. (3.22) will be different. Hence, the impact of the interpolation process onto the Cartesian grid will be different for both examples.

Furthermore, the quality of the wavefront redistribution is determined by measuring the optical path difference OPD between the predefined wavefronts and simulated wavefronts [Fig. 6.4 (c) and (f)] and by calculating the corresponding root-mean-square value rms_{OPD} , which is defined analogously to Eq. (6.1), with a reference wavelength of $\lambda = 550$ nm.

It is important to note that, according to section 5.4, the second freeform surface is calculated

⁴The implementation of the numerical wavefront mapping projection was improved compared to Ref. [86].

Table 6.6: Comparison of ΔI_T and OPD for the freeform systems for example “boat” [Fig. 6.1 (b)].

	Mirror System	Lens system
$rms_{\Delta I_T}$	$2.369 \cdot 10^{-6}$	$2.0179 \cdot 10^{-6}$
$corr_{I_T}$	0.9127	0.9349
η	99.77%	99.70%
rms_{OPD}	0.0154λ	0.0670λ

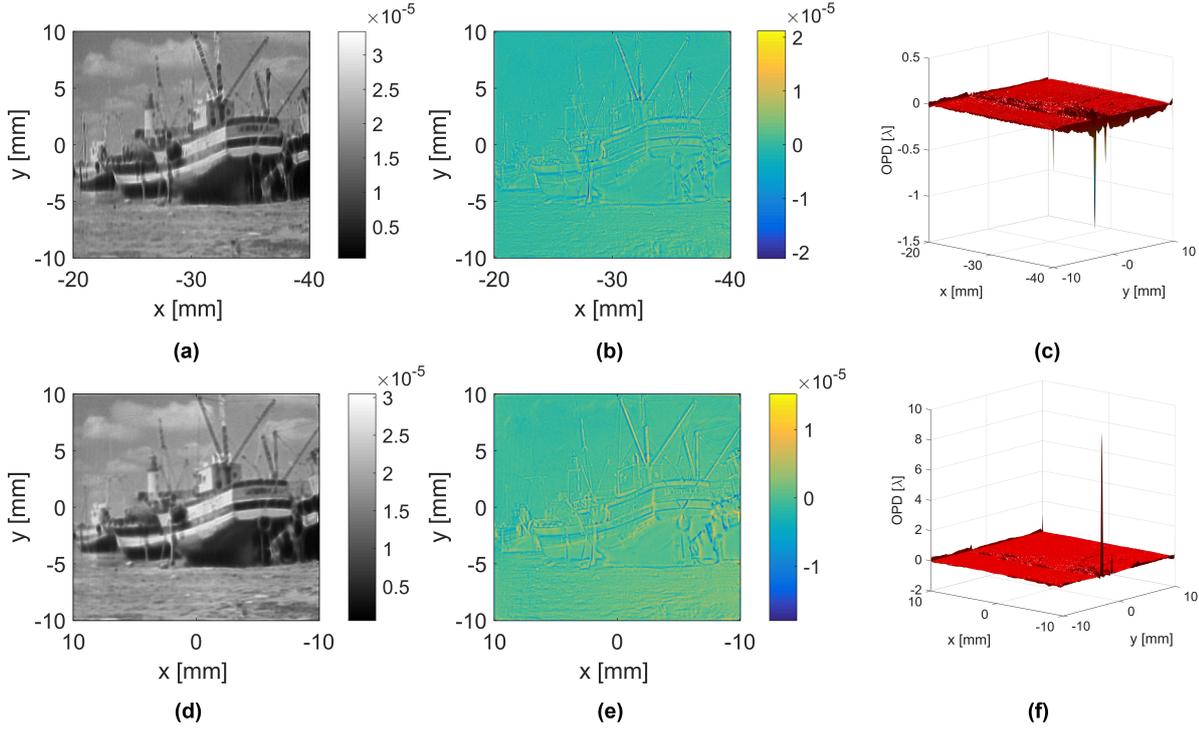


Figure 6.4: (a) Simulated irradiance $I_{RT}(\mathbf{x})$, (b) difference between predefined and simulated irradiance $\Delta I_T(\mathbf{x})$ and (c) OPD between predefined wavefronts of double mirror system. (d) Simulated irradiance $I_{RT}(\mathbf{x})$, (e) difference between predefined and simulated irradiance $\Delta I_T(\mathbf{x})$ and (f) OPD between predefined wavefronts of single lens system. [87]

analytically from the first freeform and the constant OPL condition (3.19). Therefore, the OPD should be zero with regard to numerical precision, when measured along the ray paths defined by $\hat{\mathbf{s}}_1(\mathbf{x})$, which is not the case in Fig. 6.4 (c) and (f). This is due to the interpolation of the scattered freeform data onto a cartesian grid, required for the ray tracing simulation. Consequently, the largest deviations of the OPD are caused by the boundary inter- and extrapolation as observed in Fig. 6.4 (c) and (f).

7 Freeform illumination design for pattern generation with extended sources

In this chapter, the benefits of the developed design model and numerical algorithms for zero-étendue FID are explored. This is especially the minimization of the typical irradiance pattern blurring effect of single freeform projectors due to the extension of the light source [112], which so far got limited attention for complicated irradiances with high spatial frequencies over the complete target distribution [26, 27] and is relevant for applications like 3D measurement [112]. Therefore, a design concept is proposed that allows the design of compact, energy efficient freeform projection systems for generating complex irradiances with large contrasts and minimal blurring. In particular, the high compactness of the projection system, while at the same being able to generate a complex-structured irradiance pattern, was *arguably* not achievable with previous FID methods and concepts for extended light sources in literature. Thereby, especially the overcoming of paraxial restrictions in the single and double FID, and the simultaneous ability to generate complex irradiances plays a crucial role for the blurring reduction.

While for uniform or slowly varying irradiance distributions the design of compact freeform illumination systems is addressed by various methods [16–25, 28–35] as discussed in the review paper on freeform illumination design by Wu et al. [110] and of great importance for applications like for instance street- or automotive lighting [17, 34], the blurring reduction for high spatial frequency patterns usually requires a large lateral extension of the pattern generating single freeform surface if a certain energy throughput is desired. This is a consequence of the étendue conservation for lossless optical systems [section 2.1.3]. Consequently, this prevents a compact projection system and omits the energy efficient utilization of freeform surfaces with extended light sources for pattern generation. The goal of this chapter is therefore the development of a compact freeform illumination system design concept with a high energy efficiency and a reduced pattern blurring *compared to* single freeform projectors with the same lateral extension.

To achieve this goal, in Ref. [26, 27], the authors emulate the functionality of a typical array slide projection system, consisting of a light source, a collimator, arrays of condenser lenses, object slides and projection lenses, by replacing each condenser lens and object slide with double freeform for irradiance and phase control. Instead of focusing the collimated light source by the condensers onto the object slides, the double freeform is applied to generate a spherical converging wavefront and the required target distribution [26, 27]. This configuration thereby generates a virtual object structure, which is imaged similar to the Köhler illumination principle by the subsequent projection lens onto the target. As it turns out, this configuration reduces the

pattern blurring compared to a single freeform projector (array) of the same lateral extension, which directly generates the required pattern on the target. A *coherent* explanation of why and under which conditions this configuration has positive effects on the blurring minimization, and how the projection system should be designed to minimize the blurring was not provided. For instance, instead of “[...] replacing the condenser micro-lens array of the array projector by an appropriate array of two micro-optical freeform surfaces” [26], it is argued below [section 7.5] that the double freeforms should *not* replace the condenser lenses of the slide projector, but only the object slides to achieve the stated objective of minimizing the blurring. Also it is shown below that the double freeform needs to fulfill certain conditions to be advantageous compared to a single freeform.

Furthermore, it is shown in this chapter that the blurring reduction does neither require a converging output wavefront of the double freeform surface element nor might it be the best possible concept to reduce the pattern blurring. As opposed to the results of Ref. [26, 27], it is shown that the potential degree of blurring is entirely determined by the design of the double freeform surface itself, which *then* defines the design constraints of the subsequent imaging system. The most important question is thereby how the double freeform for irradiance and phase control should be designed to minimize the blurring most effectively¹.

To answer this question in this chapter, it is proposed to interpret the pattern blurring effect as a composition of a shift and distortion contribution. This interpretation will thereby be formally connected to the mathematical models from chapter 3. It is shown that both blurring contributions can be minimized simultaneously by a *suitable* combination of a double freeform surface for irradiance and phase control and an imaging system. As it will be argued, this can be done most effectively by a (zero-étendue) collimated freeform beam shaper, which is calculated for a small as possible freeform-freeform distance².

This is demonstrated by the application of a double freeform lens for collimated beam shaping and an imaging system with a telecentric object space and by a comparison of the concept to single freeform projectors with the same lateral extension and energy throughput. Furthermore, estimates of the blurring extension of the different design concepts are given and design rules for the double freeform concept are deduced. Additionally, a design procedure for the concept is proposed, which allows the systematic design of a corresponding freeform projection system for a required pattern and an acceptable blurring extension.

The chapter is structured as follows. In section 7.1, the limitations of the illumination pattern generation with single FID for extended light sources is discussed. Thereafter, in section 7.2, it is explained how the single freeform projection can be improved by a combination with an imaging system. Then, a double FID concept for pattern generation with extended light sources, based on the combination of a thin zero-étendue double freeform lens for collimated beam shaping and

¹An alternative formulation of this question would be, how the double freeform should be designed to realize Köhler illumination in the most accurate way, while simultaneously generating the required pattern.

²The smallest possible distance depends on the given emittance and irradiance and can be determined in the zero-étendue design process [Fig. 5.5] by a variation of the integration constants of both freeform surfaces. A limiting factor for the minimal reachable distance is total internal reflection due to local energy redistribution.

an imaging system with a telecentric object space, is developed. In section 7.3, the benefits of the new design concept are demonstrated by a comparison to single freeform design concepts for different design examples and in section 7.4 the specific design procedure for the design concept is summarized. This is followed in section 7.5, by a discussion of possible variations of the double freeform illumination system concept.

The results of this chapter were published in Refs. [88] and [89].

7.1 Zero-étendue single freeform illumination design for structured illumination with extended sources

One of the major problems in FID is the generation of structured illumination for extended light sources. In case of a single freeform surface, which is calculated for a zero-étendue light source, every source ray is redirected by a single point of the freeform to a single point on the target. For an extended light source on the other hand, the source extension leads to a range of ray directions at every point of the freeform, which results in the typical blurring effect of the required target pattern [Fig. 7.2]. While this problem has been widely studied for uniform illumination patterns [110], the generation of high spatial frequency structures is rarely discussed in literature [26–28]. Such kind of structures are for example interesting for 3D measurement, in which objects are measured by illuminating them with prescribed patterns to obtain the spatial structure of the 3D object by a triangulation [111, 112]. Such kind of a typical pattern can be seen in Fig. 7.1. A fast and accurate measurement thereby benefits in general from a high energy and a large illumination pattern contrast. Classical approaches for the pattern generation like object-slide-based methods or the application of diffractive optical elements are thereby limited by the possible energy-throughput or by the realizable patterns, particularly for extended light sources.

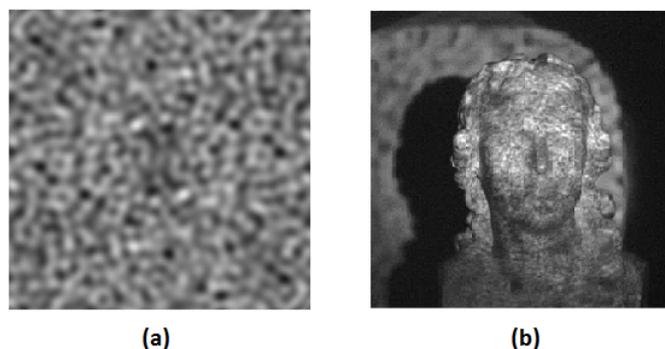


Figure 7.1: (a) Typical illumination pattern for 3D measurement and (b) reconstruction of a 3D dimensional object by utilizing the pattern. [9]

Due to the blurring effect, mentioned above, the producibility of prescribed patterns with a certain complexity, like in Fig. 7.1, with compact freeform projection systems and extended light

sources is nontrivial. In many applications, requiring *slowly* varying irradiances, the pattern can be realized by the application of the so-called “five times rule”. This means that the distance between the freeform-optic and the source should be at least five times as large as the largest source dimension to be able to approximate the emitter by a point light source [Fig. 7.2] [110]. While this rule-of-thumb might be sufficient for most applications, it is only partly applicable to pattern generating freeform surfaces [112], since the blurring extension Δ_T scales with the distance of the target area to the freeform surface [Fig. 7.2]. This can be understood directly by interpreting the illumination pattern as a superposition of pinhole contributions of the extended source with the pinholes placed at every surface point of the freeform. The *spatially variant* blurring is then characterized by the extension Δ_T of a single pinhole contribution. The scaling properties of Δ_T are illustrated for different design examples in Fig. 7.3. In these examples three freeform surfaces are calculated for the same source-freeform distance and a varying freeform-target distance, and all freeform surfaces are illuminated by an identical extended light source.

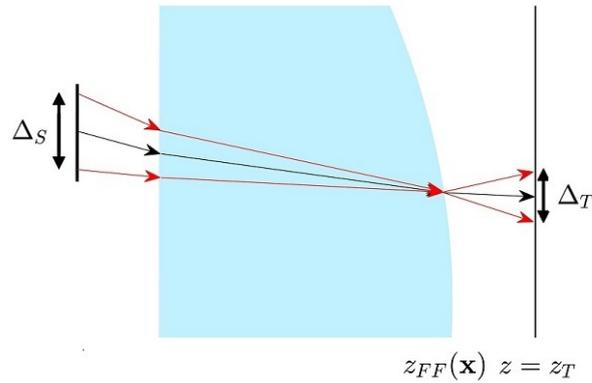


Figure 7.2: The freeform surface $z_{FF}(\mathbf{x})$ is calculated for a point light source. A single input ray (black) is considered and a pinhole placed at the corresponding point of the freeform, which redirects the input ray to the target plane $z = z_T$. Considering an extended light source, the pinhole defines a light cone of rays (red) which is redirected by the freeform towards the target. Due to the light source extension, there is a spreading Δ_T around the (ideal) target point. The spreading scales with Δ_S , the freeform target distance $z_T - z_{FF}(\mathbf{x})$ and is inverse to the freeform-source distance $z_{FF}(\mathbf{x})$. [88]

According to étendue conservation [section 2.1.3], for a fixed light source extension and a required energy throughput, a large lateral extension of the single freeform and a short distance between the freeform and the target area are required to reach a sufficient pattern contrast. This on the other hand contradicts requests from e.g. the 3D measurement, which desires a small as possible projection system and a minimum working distance, depending on the aspired application.

7.2 Reducing the pattern blurring of freeform projectors

7.2.1 Combining a zero-étendue single freeform and an imaging system

To tackle the above-mentioned problem, it is sensible to combine the benefits of a pattern generating single freeform surface and an imaging projection optics with a telecentric object

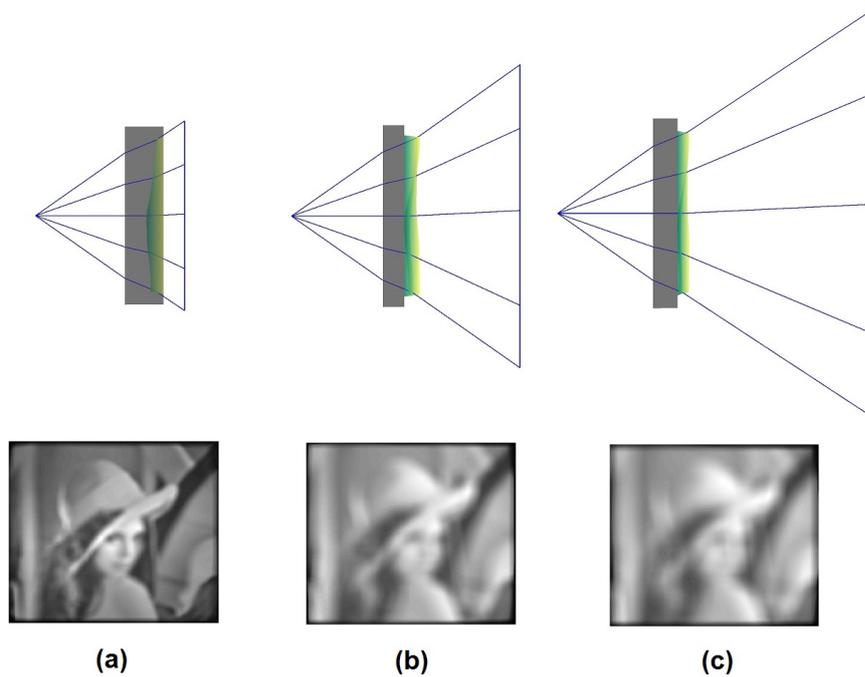


Figure 7.3: Setups and corresponding patterns for identical distances between the extended light source and the freeform surface, and different distances between the freeform and the target. The freeform surfaces are calculated under a point light source assumption. A light source extension of $3 \text{ mm} \times 3 \text{ mm}$ is assumed with an entrance surface distance of 15 mm and freeform surface distance of approximately 20 mm . The target planes are at (a) 25 mm , (b) 37.5 mm and (c) 50 mm . As expected, the fine structures and contrast of the patterns reduce with an increasing distance.

space, whereby the energy throughput of the system is fixed by the aperture of the pattern generating freeform. Whereas the generation of the exact predefined irradiance for extended light sources requires an infinite number of degrees of freedom or an infinite number of freeform surfaces, respectively, the additional degrees of freedom due to the imaging system enables the reduction of the degree of blurring.

The freeform surface, which is calculated for a zero-étendue light source, is used to generate the required illumination pattern on an intermediate target plane $z = z_T^{\text{interm}}$ as close as possible to the freeform itself. Hereby, the minimal achievable distance depends on the given emittance and irradiance and can be determined by the variation of the integration constant of the single freeform, which is defined relative to the target plane position, in the zero-étendue design process [Fig. 5.5 (a)]. The object structure on $z = z_T^{\text{interm}}$, defined by the single freeform and the extended light source, is then imaged by the projection optic onto the target plane $z = z_T$ to realize the required working distance [Fig. 7.4]. Hereby, the object cone angle of the telecentric imaging system is defined by the largest ray angle with the normal of $z = z_T^{\text{interm}}$. For an ideal imaging system the pattern and the *relative* blurring extension is therefore identical on $z = z_T^{\text{interm}}$ and $z = z_T$ [section 2.2.1]. For a large contrast, the relevant distances of the pattern generating part should thereby be chosen in a way that the blurring Δ_T is smaller than the feature size of the pattern. Considering the étendue conservation, this might require an extension of the lateral size

of the freeform to minimize the angular range of ray directions on the freeform surface [section 2.1.3]. For the 3D measurement example [Fig. 7.1], the feature size can be estimated by the minimal distance between the dark and bright spots.

For the minimization of the distance between the freeform and the intermediate target plane $z = z_T^{interm}$, the benefits of the developed design model from section 3 can be applied. This can be done e.g. by pre - or post shaping of the emittance and input wavefront with prescribed surfaces to minimize the necessary energy redistribution by the freeform surfaces or their curvature, respectively.

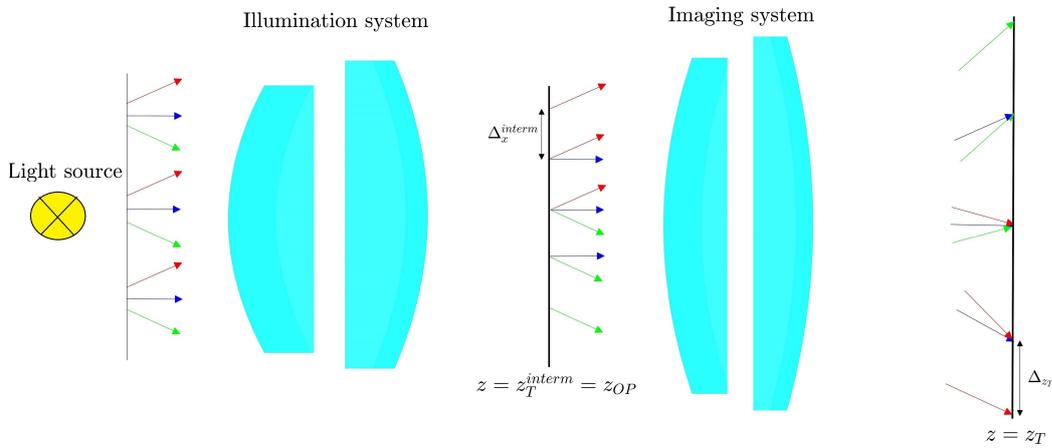


Figure 7.4: The freeform illumination system generates the required pattern on the intermediate plane $z = z_T^{interm}$, which corresponds to the object plane $z = z_{OP}$ of the imaging system. For this example geometry, the source is represented by a collimated beam with a residual divergence angle. Due to the source extension, different field angles induce different shifts Δ_x^{interm} compared to the ideal pattern for a residual divergence angle of 0 deg. An ideal imaging system preserves the pattern blurring of the illumination system. [88]

Until now, the scaling properties of the blurring effect for a single freeform projector due to the light source extension were discussed and the benefits of combining a single freeform projector with an imaging system were argued. To minimize the blurring of the pattern it was thereby suggested to place the intermediate illumination plane close to the freeform surface and then image from the intermediate plane to the target plane.

In the previous section, the blurred illumination pattern was interpreted as a superposition of pinhole contributions of the extended source from which the scaling properties of the blurring extension Δ_T can be seen directly. From another perspective, the extended light source can also be considered as a decomposition into an infinite number of point light sources. The blurred pattern is then the result of an overlay of every point source contribution in the target plane [Fig. 7.5]. This interpretation will be crucial for the arguments given below.

At first, a freeform surface $z_{FF}(\mathbf{x})$, calculated for a PLS placed at $(x, y) = (0, 0)$ and a specified irradiance $I_T(\mathbf{x})$ [Fig. 7.5 (a)] is considered. A shifted PLS at $(x, y) = (\Delta_{PLS,x}, 0)$ will then generate a pattern in the target plane, which is shifted approximately by Δ_x^{interm} and distorted

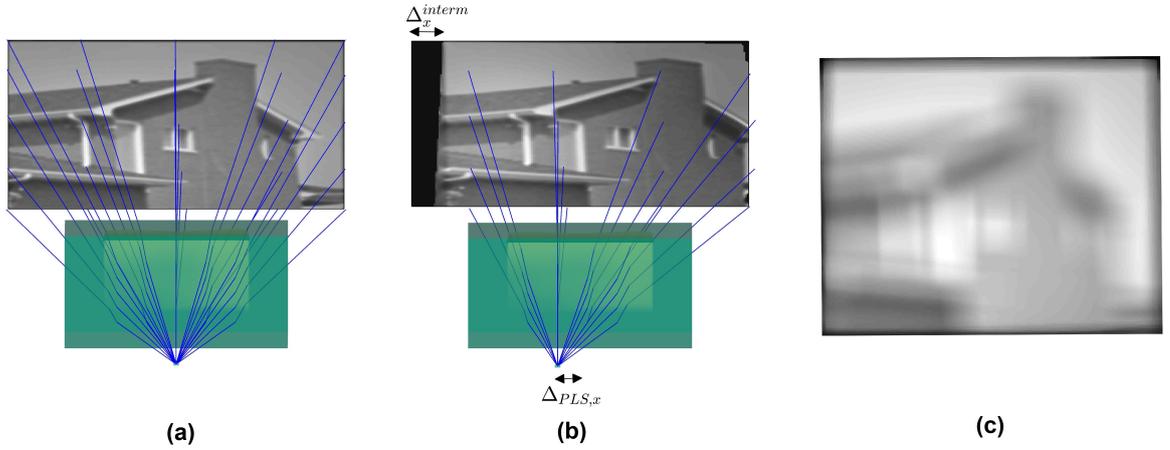


Figure 7.5: (a) Illumination pattern for a PLS placed at $(x, y) = (0, 0)$, for which the freeform surface was calculated. (b) A shift of the PLS by $\Delta_{PLS,x}$ induces a shift of the illumination pattern in the target plane by Δ_x^{interm} . Additionally, there is a locally varying distortion, compared to the ideal pattern. (c) Total illumination pattern by a superposition of all PLS contributions of the light source with an extension of $\Delta_{PLS,x} \times \Delta_{PLS,x}$ and 45 deg maximum half angle. [88]

[Fig. 7.5 (b)]. The shift Δ_x^{interm} of the irradiance is a result of the shift of the input ray direction vector field as a whole. The distortion of the pattern on the other hand is due to the fact that the freeform is calculated for the specific input ray directions of the PLS placed at $(x, y) = (0, 0)$, but illuminated by the locally incorrect ray directions of the shifted PLS.

Consequently, in the following, the pattern blurring will be interpreted as a composition of shift- and distortion contribution. The total blurring of the pattern is then a composition of both contributions [Fig. 7.5 (c)].

The shift- and distortion contribution concept can also be applied to an extended source, which is collimated by a corresponding optic. Due to the light source extension this results in a collimated beam consisting of a superposition planar wavefronts with a residual ray divergence angle α . It is assumed that the freeform is calculated for the central planar wavefront with $\alpha = 0$ and the target plane $z = z_T^{interm}$ [Fig 7.6 and Fig 7.7 (a)]. The blurring is then caused by the residual ray divergence angle α of the collimated extended light source, which results in a shifted and distorted irradiance pattern for every angle between 0 deg and α [Fig 7.7 (a)]. The distortion is hereby again due to the locally incorrect input ray directions at every freeform surface point, while the shift is due to the rotation of the input vector field by angles between 0 degree and α . The magnitude of the shift contribution is hereby determined by the source extension/divergence angle and the distance between the freeform and the target plane, and can be estimated directly from the law of refraction/reflection.

From a mathematical point of view this interpretation can be motivated from the single FID description (3.10). Assuming that the freeform surface $z_{FF}(\mathbf{x})$ was calculated for the input vector field $\hat{\mathbf{s}}_1(\mathbf{x}) = (0, 0, 1)$, a rotation of the input vector field in a paraxial approximation is described

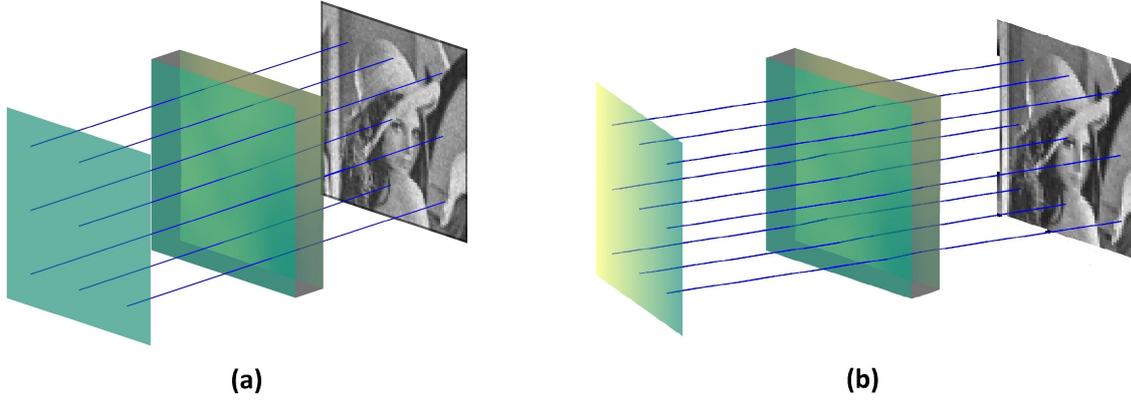


Figure 7.6: A single freeform surface is calculated for a perfectly collimated beam with $\alpha = 0$ deg and a prescribed irradiance pattern. (a) Irradiance pattern for $\alpha = 0$ deg. (b) Illuminating the freeform surface with a tilted planar wavefront under $\alpha = 20$ deg leads to shifted and distorted pattern in the target plane.

by $\hat{\mathbf{s}}_{1,\alpha}(\mathbf{x}) \sim (\alpha, 0, 1)$ and leads to a perturbation $\mathbf{u}(\mathbf{x}) \rightarrow \mathbf{u}_\alpha(\mathbf{x})$ of the mapping³:

$$\begin{aligned} \mathbf{u}_\alpha(\mathbf{x}) - \mathbf{x} &= \frac{[z_T - z_{FF,S}(\mathbf{x})]}{(\hat{\mathbf{s}}_{2,\alpha})_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_{2,\alpha})_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_{2,\alpha})_y(\mathbf{x}) \end{pmatrix} + \frac{[z_{FF,S}(\mathbf{x}) - z_0]}{(\hat{\mathbf{s}}_{1,\alpha})_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_{1,\alpha})_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_{1,\alpha})_y(\mathbf{x}) \end{pmatrix} \\ &\sim \frac{[z_T - z_{FF,S}(\mathbf{x})]}{(\hat{\mathbf{s}}_{2,\alpha})_z(\mathbf{x})} \begin{pmatrix} \alpha \\ 0 \end{pmatrix} + \dots + [z_{FF,S}(\mathbf{x}) - z_0] \begin{pmatrix} \alpha \\ 0 \end{pmatrix}. \end{aligned} \quad (7.1)$$

Thus, the nonvanishing x -component of the perturbed vector field $\hat{\mathbf{s}}_{1,\alpha}(\mathbf{x})$ introduces a term, which is proportional to $[z_T - z_{FF,S}(\mathbf{x})]\alpha/(\hat{\mathbf{s}}_{2,\alpha})_z(\mathbf{x})$. The proportionality to $[z_T - z_{FF,S}(\mathbf{x})]\alpha$ is interpreted as the shift contribution to the blurring, while the introduced local perturbations of $\mathbf{u}(\mathbf{x})$ with \mathbf{x} correspond to the distortion contribution of every plane wave to $I_T(\mathbf{x}) = \sum_\alpha I_{T,\alpha}(\mathbf{x})$. To eliminate the blurring, the imaging system, which images from the intermediate or object plane $z = z_T^{interm} = z_{OP}$, respectively, to $z = z_T$ should therefore compensate both effects. As explained in the following, both, the shift and the distortion contribution can be minimized simultaneously to a certain degree.

To do so, a collimated input beam for which the single freeform is calculated and a perfectly corrected imaging system with a telecentric object space is assumed [Fig. 7.4]. The required object cone angle of the imaging system thereby depends on the residual divergence angle of the collimated extended light source and should be chosen large enough to collect all ray directions at every point of $z = z_T^{interm} = z_{OP}$. In a paraxial approximation, a specific angle α then leads approximately to an averaged shift contribution of [Fig. 7.7 (a)]

$$\langle \Delta_x^{interm} \rangle \approx \alpha \langle z_T^{interm} - z_{FF}(\mathbf{x}) \rangle \quad (7.2)$$

compared to a single freeform system without a projection lens ($z_T^{interm} \rightarrow z_T$). The averaged

³The asymptotically equivalent symbol “ \sim ” refers to $\alpha \rightarrow 0$ throughout this chapter.

distance $\langle z_T^{interm} - z_{FF}(\mathbf{x}) \rangle$ is hereby measured along the z -axis for every value \mathbf{x} and the average of the spatial variant Δ_x^{interm} refers to the x - y -plane.

Considering the single freeform in Fig. 7.7 (a), which shows a projection of the design geometry in the y - z -plane, the shift contribution can be minimized by keeping the intermediate plane position $z = z_T^{interm}$ and placing the object plane $z = z_{OP}$ closer to the freeform surface. The displacement of $z = z_T^{interm}$ and $z = z_{OP}$ on the other hand will lead to a degradation of the illumination pattern quality since the single freeform is calculated for a specific plane $z = z_T^{interm}$ and, despite the compensation of the shift Δ_x^{interm} , the pattern quality on $z = z_T$ will not improve. As explained in the next section, this motivates the introduction of a second freeform surface for an additional control of the outgoing wavefront to decouple the pattern quality from the position of the plane $z = z_T^{interm}$.

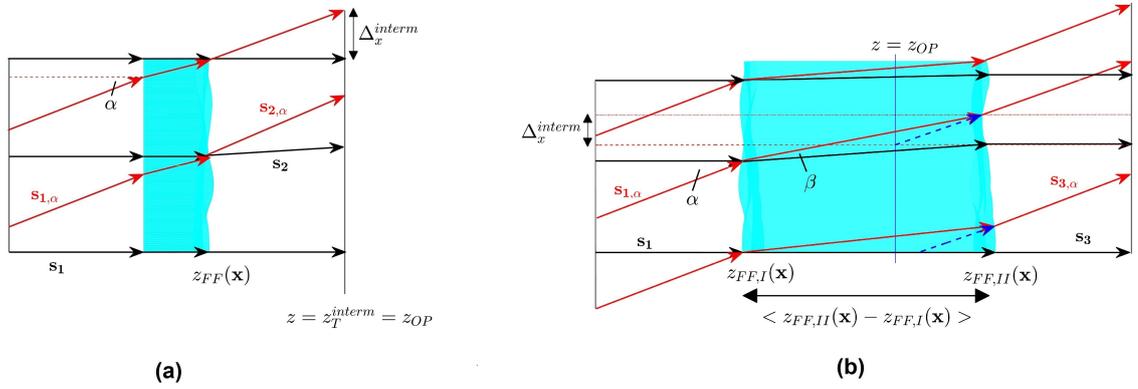


Figure 7.7: Difference between single freeform system for irradiance control and double freeform system for irradiance and phase control in a paraxial approximation. (a) Single freeform for collimated input beam with a residual ray divergence α . The freeform generates a target pattern on the intermediate plane $z = z_T^{interm}$, which corresponds to the object plane of the subsequent imaging system. To minimize the blurring or Δ_x^{interm} , respectively, the freeform should be calculated so that $z = z_T^{interm}$ is as close to the freeform as possible. A mismatch of $z = z_T^{interm}$ and the object plane will lead to a strong distortion of the pattern on $z = z_T$. (b) Double freeform illumination system for collimated beam with plane output wavefront. To minimize Δ_x^{interm} , the object plane of the imaging system should be approximately placed at the plane $z = z_{OP}$. The nonplanar output wavefronts thereby lead to a distortion of the generated pattern depending on the propagation distance between $z_{FF,II}(\mathbf{x})$ and $z = z_{OP}$. The double freeform system should be calculated so that the averaged freeform-freeform distance $\langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle$ is as small as possible. [88]

7.2.2 Combining a zero-étendue double freeform and an imaging system

In the following, instead of using a single freeform lens for irradiance control, a double freeform lens for irradiance and phase control will be applied. By requiring a collimated output wavefront, this then allows a variable placement of the object plane $z = z_{OP}$ relative to the freeform system without changing the irradiance pattern for $\alpha = 0$.

When applying a collimated beam shaper, the interpretation of the blurring as a composition of a shift and distortion contribution can be motivated mathematically analogous to the single FID as in Eq. (7.1) from section 7.2.1 from Eq. (3.10). It is hereby assumed that the freeform

surfaces $z_{FF,I}(\mathbf{x})$ and $z_{FF,II}(\mathbf{x})$ are calculated for the input and output direction vector fields $\hat{\mathbf{s}}_1(\mathbf{x}) = (0, 0, 1)$ and $\hat{\mathbf{s}}_3(\mathbf{x}) = (0, 0, 1)$ [Fig. 5.5 (b)]. In that case, z_T in Eq. (3.10) needs to be replaced by the second freeform surface $z_{FF,II}(\mathbf{x})$ at $\mathbf{u}(\mathbf{x})$ so that the irradiance is considered on the second freeform surface. Again, the perturbed vector field $\hat{\mathbf{s}}_{1,\alpha}(\mathbf{x}) \sim (\alpha, 0, 1)$ in Eq. (3.10) leads to a perturbation $\mathbf{u}(\mathbf{x}) \rightarrow \mathbf{u}_\alpha(\mathbf{x})$ with an additional term proportional to $n[z_{FF,II}(\mathbf{u}_\alpha(\mathbf{x})) - z_{FF,I}(\mathbf{x})]\alpha/(\hat{\mathbf{s}}_{2,\alpha})_z(\mathbf{x})$. Furthermore, it leads to a perturbation of the predefined outgoing ray directions $\hat{\mathbf{s}}_3(\mathbf{x}) = (0, 0, 1)$ after the second freeform, which can be calculated from the ray tracing equations. Thus, the distortion contribution in this case is additionally caused by the perturbed ray direction vector field $\hat{\mathbf{s}}_{3,\alpha}(\mathbf{u}_\alpha(\mathbf{x})) \sim (\alpha, 0, 1)$, which depends on the local surface normals at $z_{FF,II}(\mathbf{u}_\alpha(\mathbf{x}))$ and deviates from a planar outgoing wavefront for $\alpha \neq 0$. This leads to a variation of the irradiance during propagation between $z = z_{FF,II}(\mathbf{x})$ and $z = z_{OP}$. The surface normals at $z_{FF,II}(\mathbf{u}_\alpha(\mathbf{x}))$ and sufficiently large α can thereby result locally in total internal reflection, which causes a reduced energy throughput.

According to Fig. 7.7 (b), the object plane position for minimal shift contribution relative to the first freeform $z_{FF,I}(\mathbf{x})$ can be estimated roughly in a paraxial approximation by a simple geometrical calculation and the law of refraction, which gives

$$z_{OP} \approx \frac{n_1}{n_2} \langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle \quad (7.3)$$

with the refractive indices $n_1 < n_2$ and the averaged distance between both freeform surfaces $\langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle$, which is measured along the z -axis. It was hereby assumed that for different input angles α the outgoing wavefronts are collimated under the same angle and that the vectors $\hat{\mathbf{s}}_3(\mathbf{u}(\mathbf{x}))$ and $\hat{\mathbf{s}}_{3,\alpha}(\mathbf{u}_\alpha(\mathbf{x}))$ propagate in the same plane, which is only approximately true and dependent on the freeform surface shape. In particular, if the freeform surfaces are calculated for a short freeform-freeform distance, which is defined through the integration constants [Fig. 5.5 (b)], the freeforms might show strong local curvature and the object structure deviates significantly from a plane, as discussed below.

According to the argument given above, for larger (virtual) propagation distances of the non-planar outgoing wavefronts for $\alpha \neq 0$ between the exit freeform $z_{FF,II}(\mathbf{x})$ and $z = z_{OP}$, the pattern quality will decrease. Considering the calculation of the double freeform surfaces for different integration constants [Fig. 5.5 (b)], the shift contribution to the blurring in a paraxial regime remains nearly constant on $z = z_{OP}$ for different freeform-freeform distances, whereas the distortion contribution changes. Thus, to maximize the pattern quality of the projection system, the freeform surfaces should be calculated so that the freeform-freeform distance is as small as possible, since the object structure converges to $z_{FF,II}(\mathbf{x})$ for smaller distances. The minimal achievable freeform-freeform distance will be limited by effects like total internal reflection due to the specific emittance and irradiance, which define the local energy redistribution and accordingly the curvature of the freeform lenses. This is illustrated in Fig. 7.8, in which two double freeform surfaces were calculated for the same irradiance ‘‘IAP’’ and different emittances. Com-

pared to the Gaussian emittance [Fig. 7.8 (b)], the uniform emittance [Fig. 7.8 (a)] requires a stronger focusing of energy from the corners of the emittance towards the letters and therefore causes a larger local surface curvature. As a result, the achievable minimum distance between both freeform surfaces for the Gaussian emittance ($\langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle \approx 0.25$ arb. unit) is significantly smaller than for the uniform emittance ($\langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle \approx 0.325$ arb. unit), which potentially leads to a smaller pattern blurring.

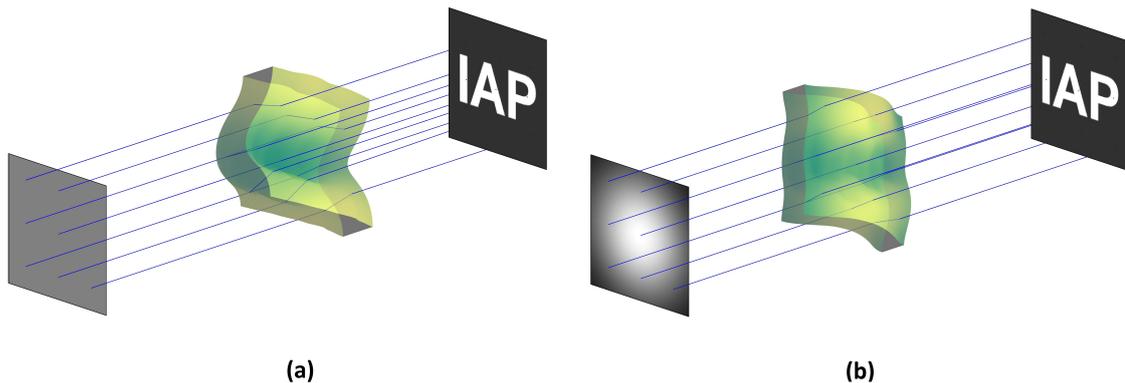


Figure 7.8: Illustration of the influence of the emittance on the achievable minimum double freeform distance for a collimated beam shaper. (a) Uniform emittance is mapped onto “IAP”. (b) Gaussian emittance is mapped onto “IAP”.

Also, as demonstrated in section 7.3, above a certain distance $\langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle$ the benefit of minimizing the shift contribution will be negligible compared to the distortion contribution and the double freeform concept will perform worse than the single freeform concept from section 7.2.1.

Furthermore, as pointed out above, the assumption of a planar object of the imaging system as in Fig. 7.7 (b) might not be the best solution for a homogeneous compensation of the shift contribution for a thin double freeform lens. This is due to the required energy redistribution, which leads to a strong curvature and large longitudinal extension of the freeform surfaces in the z -direction (see Fig. 7.8 (a)) so that the distance between $z = z_{OP}$ and the second freeform surface might vary strongly. In Fig. 7.9, this is illustrated for a collimated beam shaper consisting of a plano-freeform and a freeform-plano lens. The collimated beam shaper was calculated for the irradiance distribution “Elaine”, which is used for the design examples in section 7.3 [Fig. 7.10 (a)]. Hereby, a thin double freeform is calculated and the surface curvature is taken into account for the estimation of the object structure. Thus, the object $z = z_{OP}(\mathbf{x})$ should be shaped approximately like the second freeform surface [Fig. 7.9 (a)] and placed between both freeform surfaces [Fig. 7.9 (b)] to improve the pattern quality in the image plane. Due to the telecentricity of the imaging system, an appropriate position of the object can be determined straightforward by a simple shift along the z -axis towards the first freeform [Fig. 7.9 (a) and (b)] and comparing the irradiances. In Fig. 7.9 (d) and (e), the irradiance distributions on the corresponding object structures are shown for a divergence angle of $a = \pm 5$ deg. These

irradiances were computed with the geometric image analysis tool of Zemax OpticStudio 17. Since the distortion contribution to the blurring increases with distance between the object and the second freeform, the specification of an *exact* object shape of least blurring is nontrivial. In case of a finite freeform-freeform distance there will always be a local trade-off between the distortion- and the shift-contribution to the pattern blurring. It is therefore more reasonable to assume a *volume*, in which the object of the imaging system should be placed to realize a minimal blurring, than a *surface*. For the design examples in section 7.3, a planar object [Fig. 7.9 (c)] of the imaging system will be assumed for simplicity.

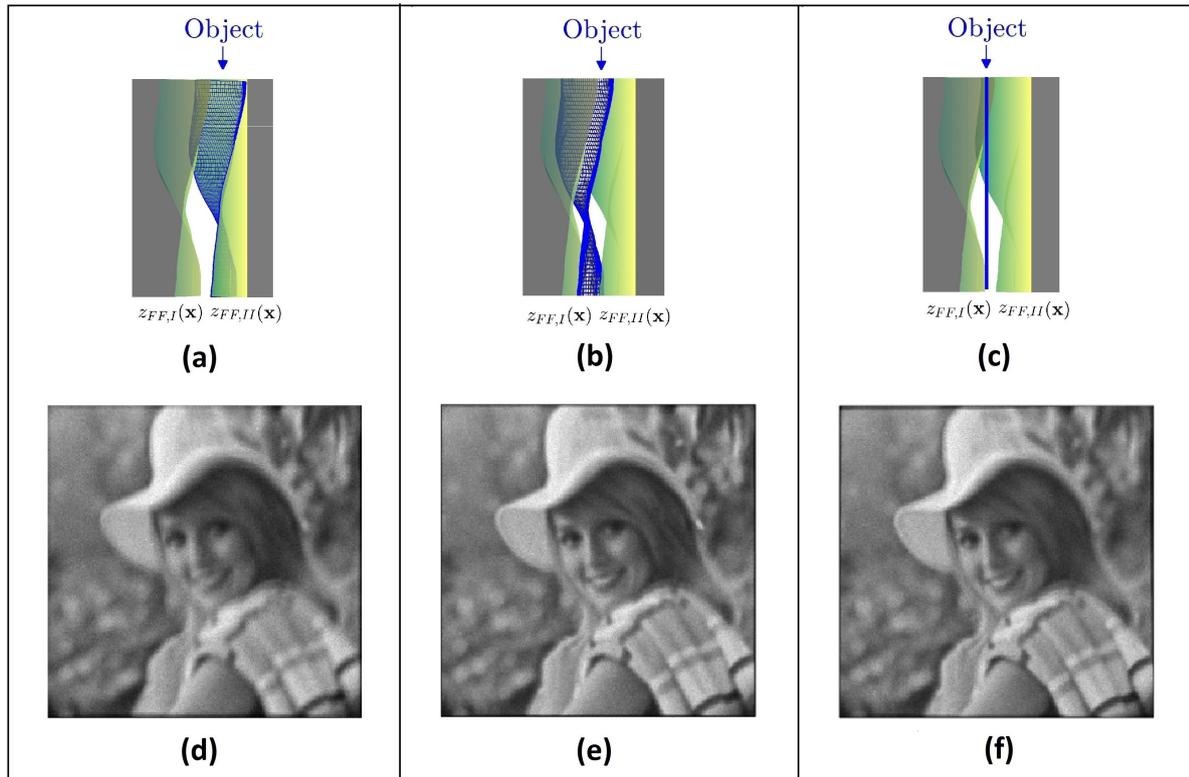


Figure 7.9: (a) The object of the imaging system is placed on the second freeform surface. (b) Approximate shape of object structure for double freeform surface with two plano-freeform lenses. Due to the small distance between both freeform surfaces the curvature varies strongly along the z -direction. For a homogenous reduction of the shift contribution to the blurring, the object of the imaging system needs to differ from a plane. For close freeform-freeform distances, the corresponding object builds a complex three dimensional structure, which is similar shaped as the second freeform surface and located between both freeform surfaces. Since there is always a balance between the shift- and distortion contribution to the blurring, an exact object shape of least blurring is difficult to specify. (c) Approximation of the object structure by a plane. (d), (e), (f): Irradiance distributions on the corresponding object structures for a residual divergence angle of ± 5 deg.

Hence, it can be concluded that, besides numerical accuracy, the following factors are limiting the theoretically achievable blurring reduction with the presented double freeform design concept: the nonparaxial ray divergence angle α and curvature of the freeform lenses, which leads to a variation of the position of $z = z_{OP}$ for different angles; the deviations from the ideal irradiance and planar output wavefront at the second freeform surface for divergence angles $\alpha \neq 0$; the finite

freeform-freeform distance; and aberrations caused by the imaging system. Additionally, there is an inevitable energy loss due to the finite freeform lens thickness since rays for divergence angles $\alpha \neq 0$ will partly be scattered or not reach the second freeform.

From the considerations above, the blurring extension Δ_T can be estimated in a paraxial approximation for each design concept. For a single freeform projector and a combination of a single freeform and an imaging system it follows [Fig. 7.7 (a)]

$$\frac{\Delta_T}{2} \approx \langle \Delta_x^{interm} \rangle \approx \alpha \langle z_T^{interm} - z_{FF}(\mathbf{x}) \rangle. \quad (7.4)$$

By applying the law of refraction, a thin double freeform lens and an imaging system leads in a paraxial regime roughly to a blurring of [Fig. 7.7 (b)]

$$\frac{\Delta_T}{2} \approx \langle \Delta_x^{interm} \rangle \approx \alpha \frac{n_1}{n_2} \langle z_{FF,II}(\mathbf{u}(\mathbf{x})) - z_{FF,I}(\mathbf{x}) \rangle \approx \alpha \frac{n_1}{n_2} \langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle \quad (7.5)$$

with $n_1 < n_2$.⁴ Since $(z_T^{interm} - z_{FF}(\mathbf{x}))$ in Eq. (7.4) varies strongly for small freeform-target plane distances, the blurring effect will also vary strongly over the irradiance pattern for the single freeform and the imaging system. Due to the constant OPL on the other hand $z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x})$ in Eq. (7.5) remains nearly constant for different \mathbf{x} (see Eq. (4.16)), which homogenizes the blurring for the double freeform concept.

Furthermore, regarding the design of the imaging system, it is important to note that the introduction of the second freeform has the advantage of a smaller required object cone angle of the imaging system due to the defined outgoing ray directions of the double freeform. According to the étendue conservation, the object cone angle can be estimated to be approximately the input ray divergence angle α . The angle α together with the object plane position then fixes the required object size, which for thin double freeform lenses will correspond approximately to the aperture size of the freeform lens.

The arguments, which led to the presented design concept, and Eq. (7.5) represent an important difference to the outcome of Ref. [26, 27] as they illustrate that the potential minimal blurring Δ_T is determined by the design of the zero-étendue double freeform for irradiance and phase control. The design constraints of the imaging system like the object structure, object size, object position, and object cone angle follow then directly from the design of the double freeform and the light source. This is important since the double freeform should not simply be designed to emulate slide projector geometries, but to effectively minimize the blurring, which is the main objective of the freeform projection with extended light sources and further discussed in section 7.5.

It is noteworthy that the replacement of the single freeform by the double freeform in this section can be motivated slightly different. This can be done, by observing that the single

⁴While Δ_T can be estimated more accurately, Eq. (7.5) offers a simple way to estimate the blurring from the integration constants of the double freeform surface and to initialize the design process described in section 7.4.

freeform blurring Δ_T varies strongly for small freeform-target distances [Fig. 7.2 and Eq. (7.4)] and that in this case the nonuniform output wavefront of a single freeform requires a telecentric projection optic with a significantly larger étendue than defined by the object size on z_T^{interm} and the residual divergence angle α . Both of these issues can be reduced significantly with a collimated beam shaper instead of a single freeform and by placing the object of the imaging system on the second freeform $z_{OP} = z_{FF,II}(\mathbf{x})$ [Fig. 7.9 (a)]. The constant optical path length [Eq. (4.16)] realizes thereby a homogeneous freeform-freeform distances and blurring [Eq. (7.5)], and the planar output wavefront enables a more symmetric output ray direction distribution for a residual divergence angle α .

By placing pinholes on the first freeform $z_{FF,I}(\mathbf{x})$ of Fig. 7.7 (b) analogous to Fig. 7.2 for the single freeform, it can then be seen that the blur Δ_T on the second freeform scales with the freeform-freeform distance [Eq. (7.5)], which thus needs to be minimized by appropriately chosen integration constants of the zero-étendue design process [Fig. 5.5 (b)]. Due to the approximately symmetric output wavefront distribution of the collimated beam shaper for a residual divergence angle α , it can be concluded that the shift contribution to the blurring can be minimized by shifting the object of the imaging system from $z_{FF,II}(\mathbf{x})$ towards $z_{FF,I}(\mathbf{x})$ [Fig. 7.9 (a) and (b)] until the shift and distortion contribution is balanced and a minimum blurring is reached. If necessary, the object structure can then be approximated by a simple shape like a plane [Fig. 7.9 (c)].

7.3 Design examples

For illustration purposes, the discussed design concepts are compared for concrete design examples with the design geometry shown in Fig. 7.10. Thereby a single freeform lens projector [Fig. 7.11 (a)] is compared to a single freeform lens with a projection lens [Fig. 7.11 (b)] and a double freeform lens with a projection lens [Fig. 7.11 (c)]. For a reasonable comparison of the illumination pattern blurring all projection systems have the same energy throughput and a similar lateral extension.

The design geometry [Fig. 7.10 (b)] corresponds to a typical 3D measurement setup. A collimated input beam with a uniform emittance and a ray divergence of ± 7 deg, and test image “Elaine” as the required irradiance distribution is chosen. The side length of the freeform surfaces is 1 arb. unit. All other lengths are measured relative to the side length. For both design concepts with an imaging lens, the distance between the single freeform and the intermediate target plane and the thickness of the double freeform, respectively, is chosen to be 0.2 arb. unit, which is fixed by the integration constants in the corresponding zero-étendue design processes [Fig. 5.5]. The data of the rotational symmetric imaging lens are summarized in Table 7.1.

The distance of the complete projection system to the target plane is 25 arb. unit and the side length of the square target area is 15 arb. unit. The results from the ray tracing simulation with 50 mio rays and a detector resolution of 250 pixels \times 250 pixels are shown in Fig. 7.11.

As expected, the blurring is drastically reduced compared to the single freeform design without

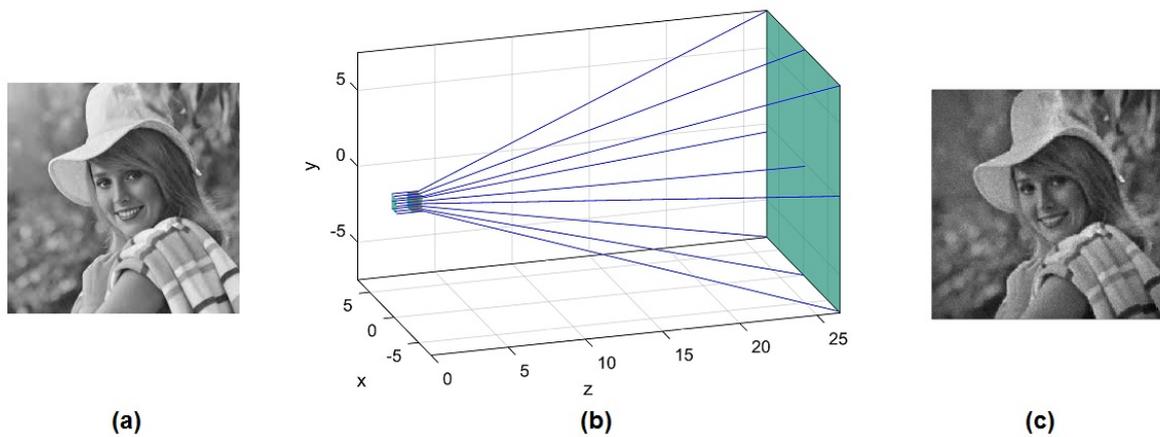


Figure 7.10: (a) Prescribed irradiance “Elaine”. (b) Design geometry. (c) Simulated irradiance pattern for an ideal zero-étendue light source. [88]

Table 7.1: Projection lens data

Object space	telecentric
Object size	1 arb. unit
Object cone angle	± 7 deg
Object position	-1 arb. unit
Radius 1 / Conic constant 1	1.313 arb. unit / -11.511
Distance	1.5 arb. unit
Radius 2 / Conic constant 2	-1.175 arb. unit / -0.905

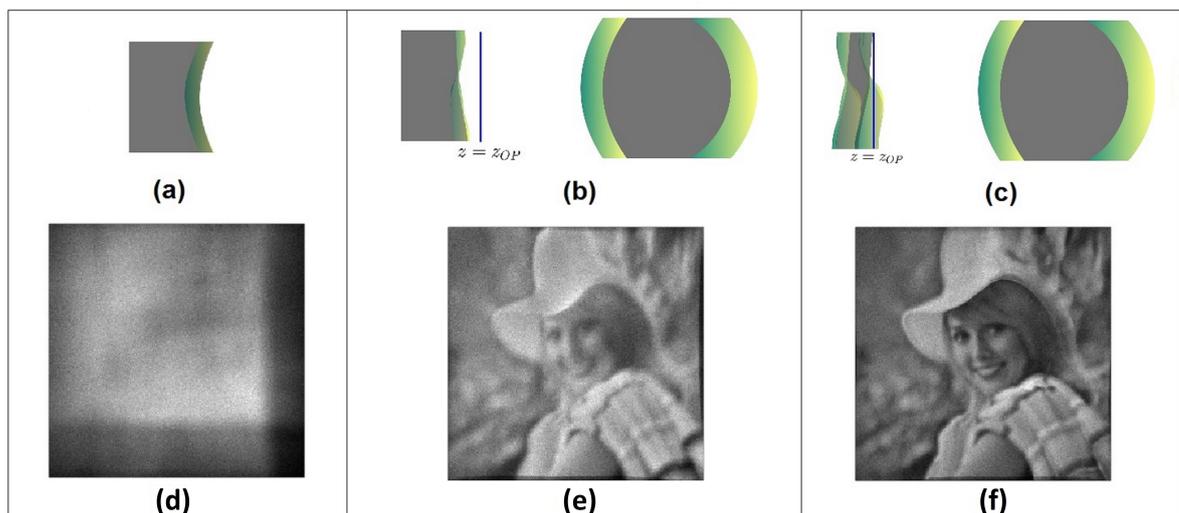


Figure 7.11: (a) Plano-freeform lens for direct generation of “Elaine”. The collimated input beam has a ray divergence angle of ± 7 deg. The plano-freeform is replaced by (b) a plano-freeform and imaging lens consisting of two aspheres and (c) double-freeform and the same imaging lens. The lower part shows the corresponding irradiance distributions at $z = z_T$, computed with the geometric image analysis tool of Zemax OpticStudio 17, for (d) the plano-freeform lens, (e) the plano-freeform with the imaging lens and (f) the double freeform with the imaging lens. [88]

the projection lens, which is in agreement with Eqs. (7.4) and (7.5) that give a relative blurring extension of $\Delta_T = 0.407$ (or 41 per cent) for the single freeform projector [Fig. 7.11 (d)] and $\Delta_T = 0.033$ (or 3 per cent) for the double freeform concept [Fig. 7.11 (f)].

For the designs with an imaging lens, especially the double freeform design shows the strongest blurring reduction. Both projection lens designs exhibit distortion towards the pattern boundary. This effect is due to the aberrations of the imaging lens and can be estimated by a comparison of the single freeform design pattern on the intermediate plane $z = z_T^{interm}$, which is aberration free, and the pattern on $z = z_T$.

The influence of the freeform-freeform distance in the double freeform concept is demonstrated in Fig. 7.12. In this examples the zero-étendue double freeforms are calculated for the irradiance “Elaine” with varying integration constants of both freeforms. The freeform-freeform distances correspond to 0.2 arb. unit [Fig. 7.12 (a)], to 0.4 arb. unit [Fig. 7.12 (b)] and to 0.8 arb. unit [Fig. 7.12 (c)]. The object of the imaging system in each case is placed at the position of minimal blurring, which due to the telecentricity of the object space can be determined by a simple shift along the z -axis.

From that it can be concluded that for a certain distance between both freeform surfaces, the application of a double freeform lens will have no advantage regarding the blurring reduction compared to a single freeform lens.

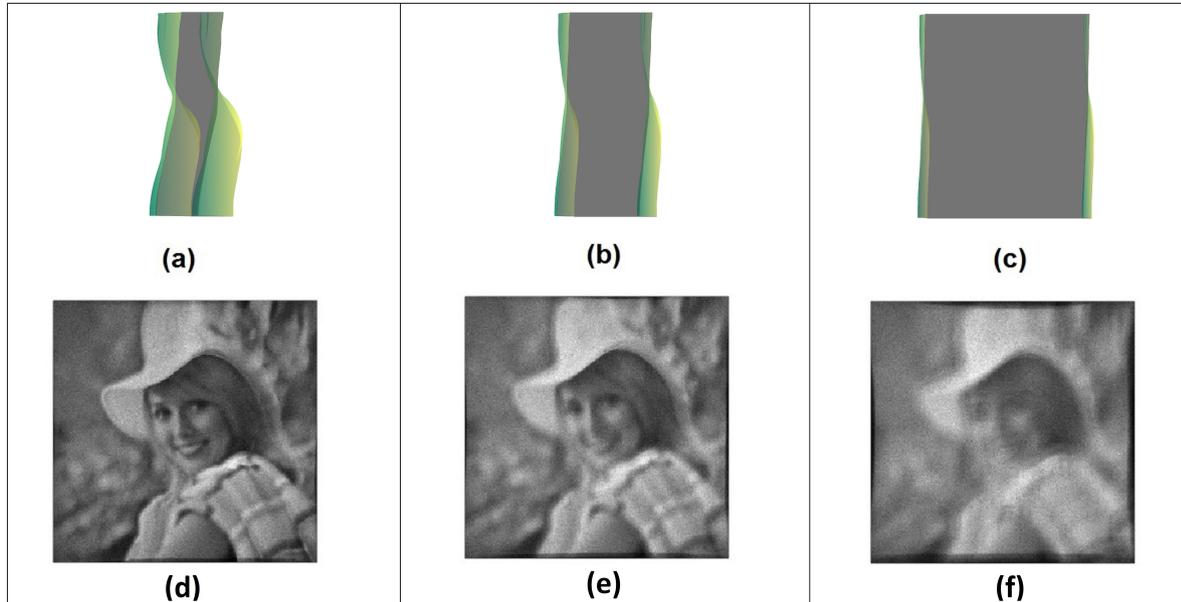


Figure 7.12: Influence on the double freeform lens thickness on the pattern blurring. The double freeform lenses are calculated for a collimated input beam with a uniform emittance and a ray divergence angle of ± 7 deg and the irradiance “Elaine”. The double freeform lenses are calculated for freeform-freeform distance of (a) 0.2 arb. unit, (b) 0.4 arb. unit and (c) 0.8 arb. unit, which is fixed through the integration constants in the design process. Simulated irradiance for double freeform distance of (d) 0.2 arb. unit, (e) 0.4 arb. unit and (f) 0.8 arb. unit. The object plane of the imaging systems in each case is placed so that a minimal blurring is reached. [88]

In Fig. 7.13, the design concept is applied to a more praxis relevant example, in which a Lambertian emitter of $3 \text{ mm} \times 3 \text{ mm}$ with a maximum half angle of 42 deg is used. For demonstrational purposes, the light source is collimated by a simple rotational symmetric aspherical lens, leading to a ray divergence of approximately $\pm 7 \text{ deg}$. To maximize the total energy throughput for larger input angles more advanced concepts like e.g. additional aspheres or CPC's can be used. The double freeform lens is defined on a square area of $15 \text{ mm} \times 15 \text{ mm}$ and the accordingly scaled projection lens of Table 7.1 is used.

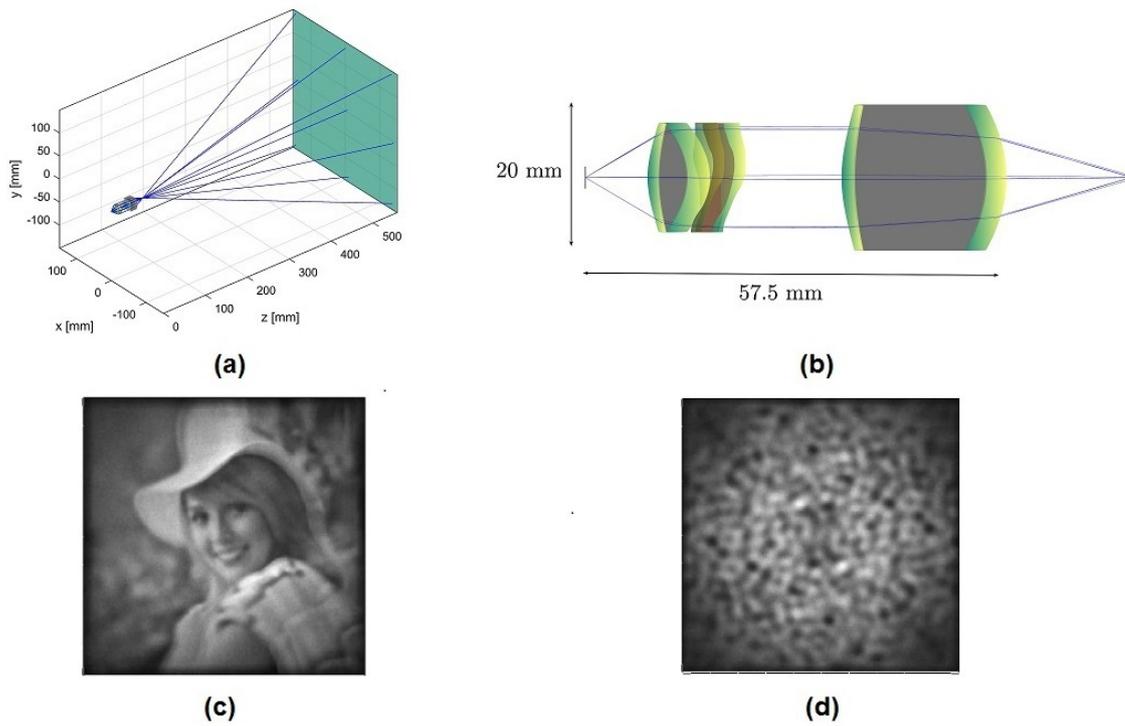


Figure 7.13: (a) Typical 3D measurement setup. (b) Projection system consisting of a collimating aspherical lens, a double freeform for irradiance and phase control and a projection lens. (c) Irradiance “Elaine” produced by a Lambertian emitter of $3 \text{ mm} \times 3 \text{ mm}$ and an opening angle of 42 deg . The divergence angle after the collimation optics is approximately $\pm 7 \text{ deg}$. (d) Typical speckle-like irradiance pattern for 3D measurement produced by the projection system with the corresponding double freeform lens. [88]

For minimization of energy losses (and maximization of the blurring), the double freeform lens was placed directly behind the collimator. As it can be seen, despite the large working distance of 500 mm and the compact projection system, the generated patterns show large contrast and minimal blurring. The pattern quality is thereby predominantly reduced due to inhomogeneous illumination towards the pattern boundary caused by the light source. These inhomogeneity's can in general be minimized by preshaping of the input light for a more uniform emittance e.g. by a Köhler integrator, which is also beneficial for further blurring minimization due to a smaller possible double freeform thickness. Also a feedback optimization [110] might be applied to minimize the irradiance inhomogeneity's, or a freeform array concepts [26, 27], which is helpful to reduce the longitudinal extension of the projection system. This is demonstrated in

the following.

In Fig. 7.14, the presented illumination concept is applied directly to an array of 11×11 identical channels with two coupled plano-freeform lenses and a projection lens⁵ for the generation of the test image “Elaine” and the speckle-like pattern at a distance of 500 mm between the projection system and the detector area of 300 mm \times 300 mm. Since for this simple design example, identical projection lenses, which are not designed to overlap exactly at the target, are applied, there is an additional blurring effect of the target pattern. This effect can be compensated by an individual design of each projection lens.

The channel array is illuminated by an intensity distribution, which corresponds to the VCSEL “VIX-850M-0000-OP06” with an emitting area of approximately 1.2 mm \times 1.2 mm, an opening angle of approximately ± 25 deg, a wavelength of 850 nm and a power of 4 W. This is then collimated by a rotational symmetric plano-asphere. The double freeform lenses of each channel are calculated for a uniform emittance and therefore identical, which is also the case for the projection lenses in the subsequent array. For computing the irradiances in a ray tracing simulation with 50 mio rays in the nonsequential mode of Zemax OpticStudio 17, the intensity of the VCSEL is modeled by “Source radial”, which is compared to the ray file of the light source. The corresponding emittance after the collimator is shown in Fig. 7.14 (b). From the 84.2 per cent of the light source energy, which are collected by the collimator and enter the freeform array, around 73.8 per cent [Fig. 7.14 (c)] or 75.0 per cent [Fig. 7.14 (d)] reach the detector area. The missing energy results from channel cross-talk, which leads to illuminated areas outside the detector and a reduced brightness at the pattern boundary. Thus, for a homogeneous brightness the channels need to be separated. The necessary spacing between the channels for avoiding channel cross talk is thereby determined by the residual divergence angle after the collimator and the distance between the double freeform lenses and the imaging lenses. While the separation between the channels allows a continuation of the freeform array onto a single substrate and consequently a simplified manufacturing, the energy throughput of the projection system is reduced. For instance, by introducing a spacing of 0.05 mm, compared to the sidelength of 0.4 mm of one double freeform, the pattern is nearly homogeneously illuminated but the energy throughput drops from 75.0 per cent to 63.0 per cent for the speckle-like pattern [Fig. 7.14 (e)]. A suitable collimation device, which redistributes the light of the source into the channels and/or a rearrangement of the channels can be applied to overcome the decrease in energy throughput. Also additional surfaces in the imaging system might be applied to decrease the distance between the double freeform lenses and the subsequent projection arrays. Furthermore, the patterns are slightly distorted at the boundary, which is due to the distortion introduced by the projection lenses. Since the grid distortion of given projection lenses is known, these effects can be compensated by an appropriate modification of the intermediate illumination pattern of the double freeform lens. This on the other hand has the disadvantage that the projection lens

⁵Compared to a single lens with two freeform surfaces, the thickness of the two plano-freeform lenses can be varied, which can benefit the mechanical stability. On the other hand an additional optical element is introduced in the design.

array can not be exchanged directly if a different working distance and/or detector area size is required.

It is important to note that for the statistical pattern the size of the projection system can be further reduced. This is due to the rather uniform redistribution of energy over the complete pattern, in contrast to the test image “Elaine” in which the energy needs to be focused more strongly into certain areas. The rather uniform redistribution of energy leads to reduced freeform surface curvatures required for the pattern generation. This leads to smaller possible distance between both freeform surfaces and to a better approximation of the intermediate object shape by a plane.

Due to the uniform emittance in each array channel, the array concept might also increase the blurring compared to a single channel setup with the same lateral extension. This can be understood directly by considering the design example in Fig. 7.8, which suffers from a uniform emittance.

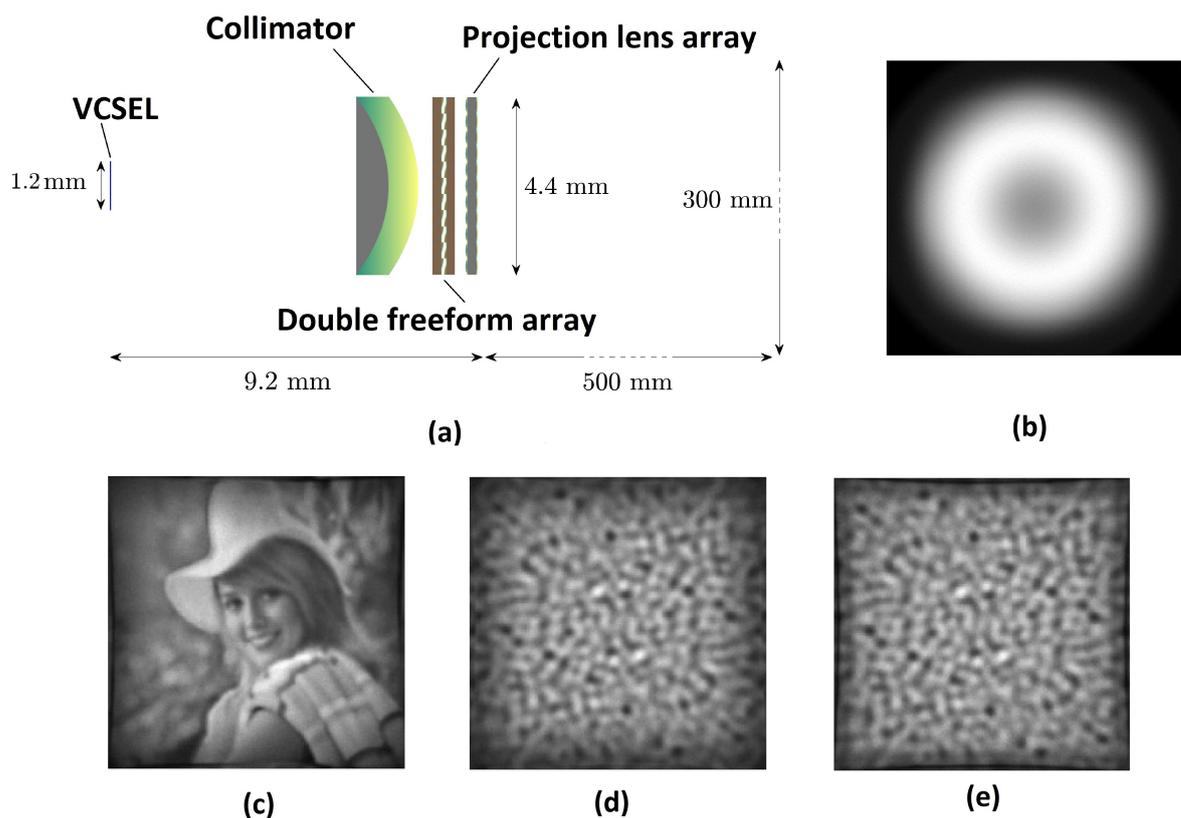


Figure 7.14: Application of design concept to freeform array for pattern generation with a real light source. The working distance is 500 mm and the target area 300 mm \times 300 mm. (a) Illumination system consisting of a light source, which intensity contribution corresponds to the VCSEL “VIX-850M-0000-OP06” with size of 1.2 mm \times 1.2 mm, an opening angle of ± 25 deg, a wavelength of 850 nm and a power of 4 W, a collimating plano-aspere, a 11 \times 11 array of identical double freeform lenses for collimated beam shaping and projection lenses. (b) Emittance after the collimator, simulated irradiances (c) “Elaine”, (d) speckle-like pattern and (e) speckle-like pattern with a spacing of 0.05 mm between the double freeform lenses. The spacing minimizes cross talk between the channels and leads to a more homogeneous brightness, but also to a decreased energy throughput.

7.4 Design procedure

To provide a better understanding and illustrate the simplicity of the proposed design concept, consisting of a zero-étendue double freeform for collimated beam shaping and an imaging system, the concrete design procedure [Fig. 7.15] will be explained.

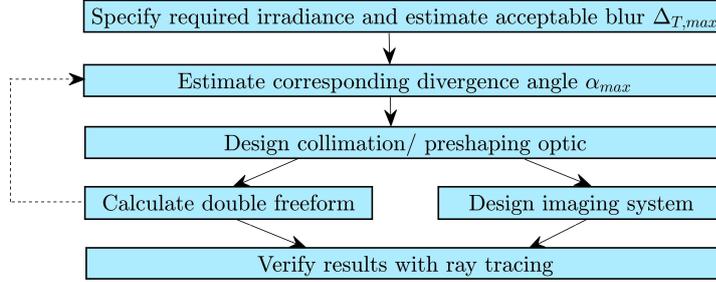


Figure 7.15: Design procedure

Assuming a reasonable application for the design concept and a given structured irradiance distribution, which needs to be projected, at first the *maximum* acceptable relative blurring extension $\Delta_{T,max}$ needs to be estimated. In case of a 3D measurement pattern [Fig. 7.1] this blurring extension is defined by the requirements for the measurement process like e.g. the minimum pattern contrast. A good initial estimate for $\Delta_{T,max}$ can for instance be determined through a convolution of the irradiance with blur kernels of different sizes. The maximum acceptable kernel size then defines $\Delta_{T,max}$. From $\Delta_{T,max}$ and Eq. (7.5) the *maximum* acceptable residual divergence angle α_{max} can be estimated. Hereby the minimum possible distance $\langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle$ depends on the emittance, which is defined by the light source and the collimation optic, and the prescribed irradiance. Hence, an initial guess needs to be provided. From experience, in many cases minimum values of $\langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle \approx 0.2$ arb. unit are achievable. According to the étendue conservation, the collimation optic of the given light source needs to be designed to achieve a residual divergence angle $\alpha \leq \alpha_{max}$. Consequently, the collimation optic defines the lateral extension of the subsequent double freeform lens.

This zero-étendue double freeform lens is then calculated [Fig. 5.5 (b)] for the given emittance and required irradiance, and a planar input and output wavefront so that a minimum distance between both freeform surfaces is reached, which is controllable by the integration constants of the freeform surfaces. The double freeform lens then defines the object position, the object cone angle of $\approx \alpha$ and the object size, which corresponds approximately to the aperture of the double freeform, of the subsequent imaging system with a telecentric object space. This imaging system is designed to achieve the required working distance and projection area, which finalizes the design that now can be improved iteratively. For instance, in case that effects like total internal reflection become significant, the design can be repeated for a smaller residual divergence angle α . Due to the telecentric object space, the position of the object of minimal

blurring can be determined by shifting the imaging system along the z -axis.

For a further minimization of the blurring, the collimation optic can be designed in a way that the resulting emittance enables a minimal energy redistribution by the double freeform lens so that $\langle z_{FF,II}(\mathbf{x}) - z_{FF,I}(\mathbf{x}) \rangle$ and thus Δ_T can be reduced (see Fig. 7.8). Due to the reduced curvature of the freeform surfaces this also leads to a better approximation of the ideal object shape of the imaging system by a planar surface.

7.5 Non-planar input and output wavefronts

In the previous part, a collimated input beam with a residual ray divergence for the double freeform concept was considered. The design concept consisting of a double freeform and an imaging system can also be applied to an extended light source represented by point light sources. Similar to the collimated design case, the source extension $\Delta_{PLS,x}$ causes a shift Δ_x^{interm} in the intermediate plane $z = z_T^{interm}$ or object plane $z = z_{OP}$ of the imaging system, for a single freeform [Fig. 7.16 (a)] and a double freeform [Fig. 7.16 (b)]. Hence, assuming a collimated outgoing beam, the shift Δ_x^{interm} can again be compensated by an appropriate placement of the object plane $z = z_{OP}$.

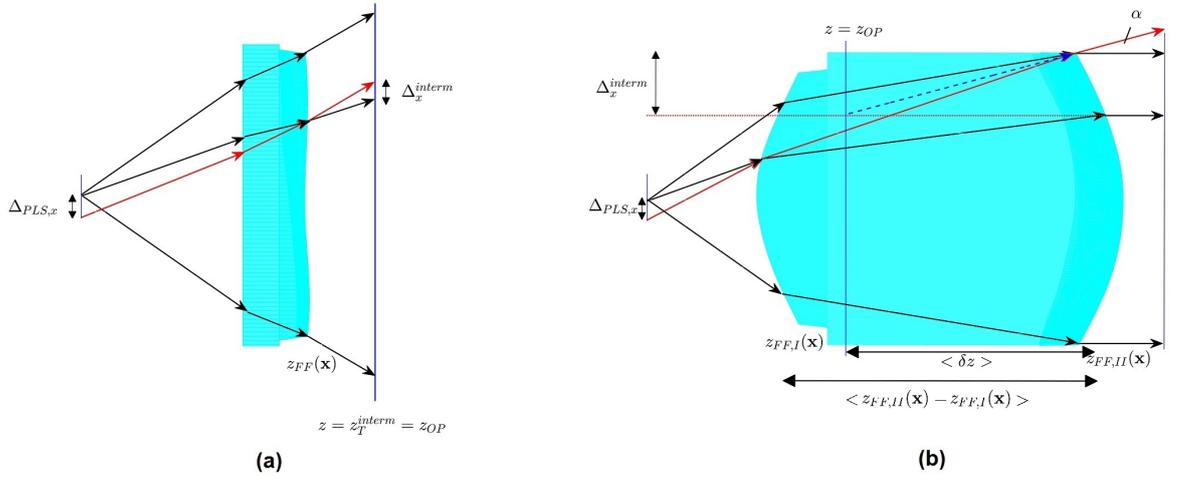


Figure 7.16: (a) Single freeform concept for a PLS. The shift $\Delta_{PLS,x}$ of the PLS causes a shift Δ_x^{interm} on $z = z_T^{interm}$. (b) Double freeform concept for PLS and collimated output beam. By an appropriate placement of the object plane $z = z_{OP}$, the shift Δ_x^{interm} can be compensated. [88]

Due to the stronger necessary ray bending for large opening angles of the PLS, this design geometry in general leads to larger double freeform lens thicknesses. Therefore, the propagation distance between the second freeform $z_{FF,II}(\mathbf{x})$ and $z = z_{OP}$ will be larger, which leads to an increased distortion contribution to the blurring. Furthermore, considering a double freeform, which was calculated for a PLS placed at $(0,0)$, and a PLS shifted by $\Delta_{PLS,x}$ [Fig. 7.17 (a)], the irradiance distribution at target plane $z = z_T$ will in general show the predefined pattern two times [Fig. 7.17 (b)] if the object plane of the imaging system is placed between both freeform

surfaces. This is due to the separate energy redistribution by the first and second freeform caused by a large Δ_x^{interm} [Fig. 7.16 (b)]. Consequently, the double freeform concept might perform worse than the single freeform concept. Thus, a preshaping with additional elements relaxes these critical conditions and might allow a significant blurring reduction. A limiting case is a collimation optic and leads to the design geometry discussed in the previous sections.

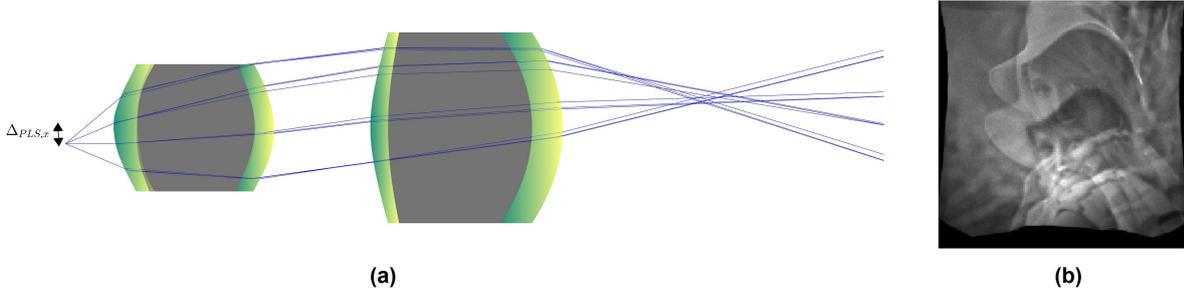


Figure 7.17: (a) Double freeform and imaging lens. The double freeform was calculated for a PLS place at $(0,0)$. The shift $\Delta_{PLS,x}$ of the PLS causes a (b) doubling of the irradiance pattern “Elaine” in the target plane. [88]

Another case, was considered by Michaelis et al. in Ref. [26, 27]. As explained in the introduction of this chapter, the authors try to emulate the functionality of a conventional slide projector by using a collimated incoming beam with a residual divergence angle and a double freeform surface to generate the required irradiance (object slide) and a spherical converging output wavefront (condensor lens) to image the irradiance by a projection lens to the target plane⁶. This setup, consisting of a double freeform and a projection lens, is applied to each channel in a freeform array concept to reduce irradiance inhomogeneities and minimize the longitudinal extension of the projection system.

According to the design geometry in Fig. 7.18, the distinction between a shift and a distortion contribution can also be applied here. In contrast to the collimated wavefront case discussed above, the converging output wavefront hereby causes a scaling of the irradiance with the position of the object plane $z = z_{OP}$. Hence, the shift Δ_x^{interm} has to be compensated by an appropriate $\Delta_{OP,x}$, which depends on the position of the focal point of the spherical wavefront, to minimize the shift blurring.

The advantage of the spherical wavefront compared to the planar wavefront is that the distortion contributions, caused by wavefront inhomogeneities due to the finite residual angles α will be less sensitive to the variation of the propagation distance between the second freeform $z_{FF,II}(\mathbf{x})$ and the $z = z_{OP}$. On the other hand, the increased lens thickness due to the wavefront redistribution in this case will be again a critical factor and will cause similar problems as discussed above like an increased shift contribution to the blurring and increased energy losses due to scattered

⁶Also for this configuration with a spherical output wavefront (as well as for every possible output wavefront of the double freeform) a projection lens with a telecentric object space can be applied. In this case, the disadvantage, compared to a double freeform lens for a planar output wavefront, is that the object cone (angle) does not match the angular distribution of rays.

significantly reduced. It is based on the combination of a thin zero-étendue double freeform lens for collimated beam shaping and an imaging system with a telecentric object space.

Furthermore, it was argued that the instead of emulating the functionality of the condensor lenses *and* the object slides of a typical slide projector array by a double freeforms for irradiance and phase control, as suggested in Ref. [26, 27], only the object slides should be replaced by the double freeforms to reach the stated objective of minimal blurring.

As opposed to other freeform illumination design methods in literature, the reduced blurring allows the design of comparably compact freeform illumination systems for the generation of patterns with high spatial frequency structures. Furthermore, the design concept works independent of the light source and the shape of its emission area and does not require a representation of the light source by individual wavefronts. Hence, every point of the emission area is considered equally by the projector, which results in a homogeneous reduction of the blurring over the target area and makes the design concept an universally applicable tool for FID with extended light sources.

The relative extension of the pattern blurring can be estimated roughly from Eq. (7.4) for single freeform projectors and from Eq. (7.5) for the developed design concept. For a design example with a working distance of 500 mm and a target area of 300 mm \times 300 mm a reduction of the relative blurring spot size from ≈ 41 per cent for a single freeform projector to ≈ 3 per cent for the new design concept was estimated. Furthermore the design concept was applied to the design of an array projector for a VCSEL, which led to the homogenization of the input distribution of the light source and a reduction of the longitudinal size of the projection system. Moreover, a systematic design procedure for the concept was proposed to compute projections systems for required irradiances with an acceptable blurring.

It is important to note that the functionality of the design concept does not necessarily require an energy and phase redistribution by a double freeform surface. According to the argumentes given above the double freeform can be replaced by other optical elements, which map a given input wavefront and emittance onto a required output wavefront and irradiance. If this optical element realizes collimated beam shaping and has two functional surfaces that implement this redistribution, the element should be calculated in a way that both surfaces are as close as possible to each other so that the blurring is minimized in the image plane of the telecentric imaging system.

To minimize the blurring extension further, a preshaping or a generalization of the presented design concept by a cascade of optical elements is conceivable. The preshaping thereby potentially reduces the required surface curvature of the freeforms for the energy redistribution, which minimizes the achievable freeform-freeform distance and improves the approximation of the object structure by a plane. An example for additional optical elements might be a repeated utilization of combinations of bi-telecentric imaging systems and thin-as-possible zero-étendue double freeform lenses for collimated beam shaping between the collimation optic and the object-space telecentric imaging system.

Further investigations of the presented design concept might consider the influence of aberrations

of the imaging system on the performance of the projection system, for instance in case of complex-shaped object structures defined by the double freeform.

8 Conclusion and outlook

Freeform surfaces are a valuable tool which allow the design of compact optical systems, while facilitating high specifications. In nonimaging optics, this means the possibility of realizing highly energy efficient illumination systems for complex illumination patterns without the necessity of absorbing structures. On the one side, the key challenge is therefore the development of mathematical models, efficient numerical algorithms and concepts for the design of freeform illumination systems and, on the other side, the practical implementation of these systems through modern manufacturing techniques. Especially a general formulation of the FID problem that is not restricted to certain geometrical assumptions, like systems with spherical and planar input and/or output wavefronts, can be beneficial for beam shaping applications. Possible benefits are, for instance, the simplification of systems for the shaping of laser diode radiation or the entering of geometrical regimes through the pre- and post shaping by predefined system entrance- and exit surfaces.

Consequently three major aims were addressed within the scope of this thesis: (a) the mathematical formulation of the single FID for irradiance control and the double FID for irradiance and phase control in optical systems with general zero-étendue light sources [Chapter 3], (b) the development of a numerical design strategy for the numerically efficient and stable solution of the mathematical models [Chapters 4 and 5] and (c) the development of a freeform illumination system design concept for the generation of illumination patterns by *compact* projection systems with extended light sources [Chapter 7].

The mathematical formulation of the FID problem was addressed in chapter 3 to overcome the above-mentioned restrictions of the PDE models in literature to spherical and planar wavefronts. At first, the single FID for a given emittance and irradiance with zero-étendue light sources was formulated in terms of PDE system for the ray-mapping components and the freeform surface. This PDE system can be reduced to a nonlinear, elliptic PDE of Monge-Ampère type for the surface only. The possibility to define general zero-étendue input wavefronts also allows the utilization of prescribed entrance surfaces, which was not described by previously published PDE models. Thus, the application of prescribed entrance surfaces, e.g. in case of single freeform lenses, can be used to overcome manufacturing limitations or geometrical constraints of the given design geometry. Furthermore, it was shown that the PDE system can be extended to predefined *exit* surfaces by applying additional geometrical constraints for the unknown surface intersection coordinates of every predefined exit surface.

Subsequently, the mathematical formulation of the FID was extended to double freeform surfaces

for irradiance and phase control, which overcomes the restriction of previous design models in literature to at least one planar wavefront or paraxial geometries. Due to the necessity of specifying not only the target points but also the outgoing ray directions of the illumination system, this design problem requires the introduction of additional degrees of freedom by a second freeform surface. This led to the PDE system of the form for one of the freeform surfaces and *projected* wavefront mapping coordinates, which is applicable to single lenses, double lenses, double mirrors and lens-mirror combinations. In contrast to the single FID, the PDE system thereby does not reduce to an MAE for the freeform only, if a nonplanar output wavefront is required. Furthermore, in section 3.3, the single FID and double FID was generalized to freeform surfaces in a system of predefined surfaces.

The development of the numerical design strategy to solve the new PDE models for general zero-étendue wavefronts was addressed in the chapters 4 and 5 and its evaluation was performed in chapter 6. Due to the nonlinearities, the discretization of the PDE models leads to systems of coupled nonlinear equations, which need to be solved simultaneously by nonlinear equation system solvers. Since the intention is a *stable* numerical solving process and the proof of convergence of numerical methods for solving discretized nonlinear PDEs is in general nontrivial, the focus was placed onto the development of an appropriate initial iterate for the mapping components and the surface.

In chapter 4, based on *numerical* evidence from literature, it was shown *theoretically* that the L_2 OMT can be utilized to calculate a ray-mapping, which is integrable for collimated beam shaping and infinite distances between the freeform and the target in single FID, or between both freeform lens surfaces in double FID. Furthermore, it was shown that the double freeform lens design problem with collimated beams can be formulated as the PDE system for the mapping components *only*.

In chapter 5, building on the shown connection of the L_2 OMT map to the FID, design strategies to solve the PDE models for general zero-étendue methods were proposed. In contrast to most other PDE methods in literature, it is not based on directly solving a discretized MAE. While solving an MAE with simple discretization schemes, e.g. standard FDs, leads to numerical instabilities for complex irradiance distributions with strong gradients and the development of convergent schemes is nontrivial, the solving of the equivalent PDE *systems* with *explicitely* stated energy conservation equation greatly benefits the numerical stability, if the L_2 OMT map is applied as an initial iterate. In that case, the energy conservation equation confines the solutions of the PDE systems to the space of energy conserving mappings close to the L_2 OMT map, while the solving of the PDE system leads to corrections to the initial map and surface by imposing the laws of refraction and reflection through both mapping equations.

To increase the numerical reliability of the root-finding of the PDE systems an initial surface construction approach was proposed. It is based on integrating ordinary differential equations, thereby incorporating information of the design geometry, wavefronts and the ray-mapping into the initial surface. Thus it is arguably more suitable than other initial surfaces proposed in literature so far.

Additionally, numerical procedures for ray aiming and noise free irradiance computation for extended light sources were proposed, which are based on solving the description of optical systems, introduced in section 3.3. These were applied to generate the input data for the design examples in chapter 6. In this chapter, several design examples were calculated to demonstrate the capability of the numerical algorithm to handle nonstandard predefined wavefronts and complex irradiance distributions and to evaluate its performance. This includes the single FID with two predefined entrance surfaces and the double FID of a double mirror system and a single lens system with two freeform surfaces for two nonplanar wavefronts (or planar but tilted relative to the target plane) beyond a paraxial regime, which had not been done in literature before. The numerical implementation shows the capability of calculating freeform surfaces for complex irradiances up to resolutions of $512 \text{ pixels} \times 512 \text{ pixels}$, limited by the memory of 16 GB RAM, within 111 min for single FID and 130 min for double FID on an Intel Core i3 at 2×2.4 Ghz.

In chapter 7, based on the application of the design methods from the previous chapters, an energy efficient freeform illumination system design concept for pattern generation with extended light sources was developed. The design concept enables the generation of complex patterns with high spatial frequency structures for extended light source, while allowing a projection system compactness and a pattern quality, that was arguably not achievable with previously published design methods and concepts for extended light sources. It is based on the combination of a *thin* double freeform lens for *collimated* beam shaping and a imaging system with a *telecentric* object space. Estimates for the pattern blurring are derived, which suggest that the blurring extension on the target plane is mainly determined by the distance between both freeform surfaces. Furthermore a systematic design procedure for the concept is proposed, which allows an efficient computation of corresponding freeform projection systems for a predefined acceptable blurring of the required pattern. Since the proposed concept allows the design of compact projection systems for pattern generation that cannot be realized by FID methods for extended light sources, it opposes the often quoted “five-times rule”, which claims that zero-étendue FID methods are only applicable if the distance between the extended light source and the illumination optics is at least five times as large as the largest dimension of the light source.

Of course there are numerous open questions, which motivate further investigations in multiple research areas. Regarding the PDE models for single and double FID, derived in chapter 3, there is a lack of mathematical analysis in literature. While publications so far have concentrated on single and double freeform system with spherical and planar wavefronts, a thorough investigation of the new PDE models, with regards to e.g. existence theorems for solutions, might benefit the understanding of the FID problem for general zero-étendue wavefronts. Also the convergence properties of the numerical design strategy as proposed in chapter 5 remains unclear. Despite the numerical evidence due to numerous design examples, which show that for a given initial L_2 OMT map the implemented algorithm converges to a root of the PDE system, the research field of numerical analysis requires convergence proofs [93]. Thus, the finding of a root for one design example does not guarantee the same for a different design example. This is an important,

but usually overlooked, aspect in the physical literature when numerical algorithms for FID are proposed.

A possible, relevant extension of the presented work for practical applications requires further investigations regarding more complicated boundaries or targets with singularities. As discussed this requires the implementation and application of adequate L_2 OMT algorithms, for instance as presented in Refs. [50, 93]. Since the proposed design strategies also worked for L_2 OMT concave target irradiances [Fig. 5.8] without modification of the implementation of the transport boundary conditions, the algorithms by [50, 93] might be directly applicable.

Another important aspect not discussed in this thesis is the shaping of coherent and partially coherent light sources. Since the applied design models and concepts are purely geometrical the calculated freeform surfaces do not incorporate diffraction effects in the beam shaping process, which diminishes the pattern quality [113]. While the geometrical freeform design methods have been applied for the generation of an initial phase for the calculation of diffractive optical elements for complicated irradiances [114–116], the influence of partially coherent light sources and the design of appropriate freeform surfaces got less attention.

Furthermore, manufacturing designed freeform surfaces deserves a more in-depth analysis and requires further improvements of corresponding manufacturing technologies. Especially the 3D printing technology [11, 117] seems to be a *promising candidate* for the mass production of freeform illumination system concepts like the one presented in chapter 7.

Bibliography

- [1] <https://ourfiniteworld.com/2012/03/12/world-energy-consumption-since-1820-in-charts/>
- [2] O.V. Losev, "Luminous carborundum detector and detection effect and oscillations with crystals, " The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **6**(39), 1024–1044 (1928).
- [3] https://en.wikipedia.org/wiki/Light-emitting_diode
- [4] <https://www.stouchlighting.com/blog/top-15-advantages-of-led-lighting>
- [5] R. Winston, J.C. Miñano, and P. Benítez, *Nonimaging optics*, (Elsevier Academic Press, 2005).
- [6] J. Chaves, *Introduction to nonimaging optics - second edition*, (CRC Press, 2016).
- [7] J. Koshel, *Illumination Engineering*, (Wiley, 2013).
- [8] M. Schaffer, M. Grosse, and R. Kowarschik, "High-speed pattern projection for three-dimensional shape measurement using laser speckles," *Appl. Opt.* **49**(18), 3622–3629 (2010).
- [9] A. Brahm, C. Bösel, J. Hartung, R. Jende, P. Boettner, C. Reinlein, H. Gross, P. Kuehmstedt, and G. Notni, "Active freeform mirror for optical 3D measurements", European Optical Society Biennial Meeting (EOSAM), Delft (2018).
- [10] F. Fang, X. Zhang, A. Weckenmann, G. Zhang, and C. Evans, "Manufacturing and measurement of freeform optics," *CIRP Annals* **62**(2), 823–846 (2013).
- [11] B. G. Assefa, T. Saastamoinen, J. Biskop, M. Kuittinen, J. Turunen, and J. saarinen, "3D printed plano-freeform optics for non-coherent discontinuous beam shaping," *Opt. Rev.* **25**(3), 456-462 (2018).
- [12] K.P. Thompson and J.P. Rolland, "Freeform Optical Surfaces: A Revolution in Imaging Optical Design," *Optics & Photonics News* **23**(6), 30–35 (2012).
- [13] K. Fuerschbach, J. P. Rolland, and K. P. Thompson, "Theory of aberration fields for general optical systems with freeform surfaces," *Opt. Express* **22**(22), 26585–26606 (2014).
- [14] J. Reimers, A. Bauer, K. P. Thompson, and J. P. Rolland, "Freeform spectrometer enabling increased compactness," *Light: Science and Applications* **6**, e17026 (2017).
- [15] M. Sieler, P. Schreiber, P. Dannberg, A. Bräuer, and A. Tünnermann, "Ultraslim fixed pattern projectors with inherent homogenization of illumination," *Appl. Opt.* **51** (1), 64–74 (2012).
- [16] P. Benítez, J. C. Miñano, J. Blen, R. Mohedano, J. Chaves, O. Dross, M. Hernández, and W. Falicoff, "Simultaneous multiple surface optical design method in three dimensions," *Opt. Eng.* **43**(7), 1489–1502 (2004).
- [17] O. Dross, R. Mohedano, P. Benítez, J. C. Miñano, J. Chaves, J. Blen, M. Hernández, and F. Muñoz, "Review of SMS design methods and real world applications," *Proc. SPIE* **5529**, 35–47 (2004).
- [18] Y. Ding, X. Liu, Z. R. Zheng, and P. F. Gu, "Freeform LED lens for uniform illumination," *Opt. Express* **16**(17), 12958–12966 (2008).
- [19] F. Fournier, *Freeform reflector design with extended sources* (Ph.D. Diss., University of Florida, 2010).
- [20] F. R. Fournier, W. J. Cassarly, and J. P. Rolland, "Fast freeform reflector generation using source-target maps," *Opt. Express* **18**(5), 5295–5304 (2010).
- [21] Y. Luo, Z. Feng, Y. Han, and H. Li, "Design of compact and smooth free-form optical system with uniform illuminance for LED source," *Opt. Express* **18**(9), 9055–9063 (2010).

-
- [22] M.A. Moiseev and L.L. Doskolovich “Design of refractive spline surface for generating required irradiance distribution with large angular dimension,” J. of Modern Optics volume **57** (7), 536–544 (2010).
- [23] J. Bortz and N. Shatz, “Relationships between the generalized functional method and other methods of nonimaging optical design,” Appl. Opt. **50**(10), 1488–1500 (2011)
- [24] R. Wester, G. Müller, A. Völl, M. Berens, J. Stollenwerk, and P. Loosen, “Designing optical free-form surfaces for extended sources,” Opt. Express **22**(S2), A552–A560 (2014).
- [25] S. Seroka, *Optimierung sequentieller optischer Systeme* (Ph.D. Diss., Universität Paderborn, 2014).
- [26] D. Michaelis, P. Schreiber, C.Li, A. Bräuer, and H.Gross, “Freeform Array Projection”, Proc. SPIE **9629**, 962909-1 – 962909-10 (2015).
- [27] D. Michaelis and P. Schreiber, “Projektionsvorrichtung und Verfahren zur Projektion mit optischen Freiformflächen ”, DE102015216985A1 (2015).
- [28] D. Ma, Z. Feng, and R. Liang, “Deconvolution method in designing freeform lens array for structured light illumination”, Appl. Opt. **54** (5), 1114-1117 (2015).
- [29] R. Wu and H. Hua, “Direct design of aspherical lenses for extended non-Lambertian sources in three-dimensional rotational geometry,” Opt. Express **24**(2), 1017–1030 (2016).
- [30] R. Wu, C.Y.Huang, X. Zhu, H. Cheng, and R. Liang, “Direct three-dimensional design of compact and ultra-efficient freeform lenses for extended light sources,” Optica **3**(8), 840-843 (2016).
- [31] A. Aksoylar, *Modelling and model-aware signal processing methods for enhancement of optical systems* (Ph.D. Diss., Boston University, 2016).
- [32] M. Brand, and A. Aksoylar, “Sharp images from freeform optics and extended light sources”, Frontiers in Optics 2016, OSA Technical Digest (online) (Optical Society of America, 2016), paper FW5H.2.
- [33] D. Rausch, M. Rommel, A. M. Herkommer, and T. Talpur, “Illumination design for extended sources based on phase space mapping,” Opt. Eng. **56**(6), 065103-1 – 065103-7(2017).
- [34] S. Sorgato, R. Mohedano, J. Chaves, M. Hernández, J. Blen, D. Grabovickic, P. Benítez, J. C. Miñano, H. Thienpont, and F. Duerr, “Compact illumination optic with three freeform surfaces for improved beam control,” Opt. express **25**(24), 29627–29641 (2017).
- [35] M. Brand, and D. A. Birch, “Freeform irradiance tailoring for light fields ”, Opt. Express **27**(12), A611—A619 (2019).
- [36] V.I. Oliker, “Mathematical aspects of design of beam shaping surfaces in geometrical optics,” in *Trends in Nonlinear Analysis*, M. Kirkilionis, S. Kromker, R. Rannacher, and F. Tomi, eds. (Springer-Verlag, 2003), pp. 193-222.
- [37] D. Michaelis, S. Kudaev, R. Steinkopf, A. Gebhardt, P. Schreiber, and A. Bräuer, “Cartesian oval representation of freeform optics in illumination systems,” Proc. SPIE **7059**, 705905-1 – 705905-6 (2008).
- [38] F.R. Fournier, W.J. Cassarly, and J.P. Rolland, “Fast freeform reflector generation using source-target maps,” Opt. Express **18**(5), 5295–5304 (2010).
- [39] D. Michaelis, P. Schreiber, and A. Bräuer, “Cartesian oval representation of freeform optics in illumination systems,” Opt. Lett. **36**(6), 918–920 (2011).
- [40] C. Canavesi, W.J. Cassarly, and J.P. Rolland, “Target flux estimation by calculating intersections between neighboring conic reflector patches,” Opt. Lett. **38**(23), 5012-5015 (2013).
- [41] V.I. Oliker, “Controlling light with freeform multifocal lens designed with supporting quadric method(SQM),” Opt. Express **25**(4), A58-A72 (2017).
- [42] V. Oliker, L. L. Doskolovich, and D. A. Bykov, “Beam shaping with a plano-freeform lens pair,” Opt. Express **26**(15), 19406–19419 (2018).
- [43] T. Glimm and V.I. Oliker, “Optical Design of Two-reflector Systems, the Monge-Kantorovich Mass Transfer Problem and Fermat’s Principle,” Indiana University Math. J. **53**(5), 1255–1277 (2004).

- [44] J. Rubinstein and G. Wolansky, “Intensity control with a free-form lens,” *J. Opt. Soc. Am. A* **24**(2), 463–469 (2007).
- [45] C.R. Prins, J.H.M. ten Thije Boonkamp, J. Van Roosmalen, W.L. IJzerman, and T.W. Tukker, “A Least-Squares Method for Optimal Transport Using the Monge–Ampère Equation,” *SIAM J. Sci. Comput.* **37**(6), B937–B961 (2014).
- [46] A. Bäuerle, A. Bruneton, P. Loosen, and R. Wester, “Algorithm for irradiance tailoring using multiple freeform optical surfaces,” *Opt. Express* **20**(13), 14477–14485 (2012).
- [47] A. Bruneton, A. Bäuerle, P. Loosen, and R. Wester, “High resolution irradiance tailoring using multiple freeform surfaces,” *Opt. Express* **21**(9), 10563–10571 (2013).
- [48] Z. Feng, L. Huang, G. Jin, and M. Gong, “Beam shaping system design using double freeform optical surfaces,” *Opt. Express* **21**(12), 14728–14735 (2013).
- [49] Z. Feng, L. Huang, G. Jin, and M. Gong, “Designing double freeform optical surfaces for controlling both irradiance and wavefront,” *Opt. Express* **21**(23), 28693–28701 (2013).
- [50] Y. Schwartzburg, R. Testuz, A. Tagliasacchi, and M. Pauly, “High-contrast Computational Caustic Design,” *ACM Trans. Graph.* **33**(11), 74:2–74:11 (2014).
- [51] D. Ma, Z. Feng, and R. Liang, “Freeform illumination lens design using composite ray mapping,” *Appl. Opt.* **54** (3), 498–503 (2015).
- [52] D. Ma, Z. Feng, and R. Liang, “Tailoring freeform illumination optics in a double-pole coordinate system,” *Appl. Opt.* **54**(9), 2395–2399 (2015).
- [53] Z. Feng, B.D. Froese, C. Huang, D. Ma, and R. Liang, “Creating unconventional geometric beams with large depth of field using double freeform-surface optics,” *Appl. Opt.* **54**(20), 6277–6281 (2015).
- [54] Z. Feng, B.D. Froese, R. Liang, D. Cheng, and Y. Wang, “Simplified freeform optics design for complicated laser beam shaping,” *Appl. Opt.* **56**(33), 9308–9314 (2017).
- [55] L.L. Doskolovich, A.A. Mingazov, D.A. Bykov, E.S. Andreev, and E. A. Bezus, “Variational approach to calculation of light field eikonal function for illuminating a prescribed region,” *Opt. Express* **25**(22), 26378–26392 (2017).
- [56] L.L. Doskolovich, D.A. Bykov, E.S. Andreev, E. A. Bezus, and V. Oliker, “Designing double freeform surfaces for collimated beam shaping with optimal mass transportation and linear assignment problems,” *Opt. Express* **26**(19), 24602–24613 (2018).
- [57] K. Desnijder, P. Hanselaer, and Y. Meuret, “Flexible design method for freeform lenses with an arbitrary lens contour,” *Opt. Letters* **42**(24), 5238–5241 (2017).
- [58] K. Desnijder, P. Hanselaer, and Y. Meuret, “Ray mapping method for off-axis and non-paraxial freeform illumination lens design,” *Opt. Letters* **44**(4), 771–774 (2019).
- [59] N. G. Boldyrev, “About calculation of asymmetrical specular reflectors,” *Svetotekhnika* **7**, 7–8, (1932).
- [60] V.D. Komissarov and N. G. Boldyrev, “The foundations of calculating specular prismatic fittings,” *Trudy VEI* **43**, 6–61 (1941).
- [61] V. Galindo, “Design of dual-reflector antennas with arbitrary phase and amplitude distributions,” *IEEE Transactions on Antennas and Propagation* **12**(4), 403–408 (1964).
- [62] J.S. Schruben, “Formulation of a Reflector-Design Problem for a Lighting Fixture,” *J. Opt. Soc. Am.* **62**(12), 1498–1501 (1972).
- [63] B.S. Westcott, and P. Norris, “Reflector synthesis for generalized far-fields,” *J. Phys. A: Math. Gen.*, **8**(4), 521–532 (1977).
- [64] F. Brickell, L. Marder, and B.S. Westcott, “The geometrical optics design of reflectors using complex coordinates,” *J. Phys. A: Math. Gen.*, **10**(2), 245–260 (1977).
- [65] P. Norris, and B.S. Westcott, “Computation of reflector surfaces for bivariate beamshaping in the elliptic case,” *J. Phys. A: Math. Gen.*, **9**(12), 2159–2169 (1976).

-
- [66] B.S. Westcott, F.A. Stevens, and F. Brickell, "GO synthesis of offset dual reflectors," *IEE Proc.* **128**(1), 11–18 (1981).
- [67] V. Galindo-Israel, W.A. Imbriale, and R. Mittra, "On the theory of the synthesis of single and dual offset shaped reflector antennas," *IEEE Transactions on Antennas and Propagation* **35**(8), 887–896 (1987).
- [68] V.I. Oliker and E. Newman. "On the energy conservation equation in the reflector mapping problem," *Appl. Math. Lett.*, **6**, 91–95 (1993).
- [69] A.L. Karakhanyan, "On the regularity of weak solutions to refractor problem," *Armenian J. of Math.* **2**(1), 28–37 (2009).
- [70] R. Wu, L. Xu, P. Liu, Y. Zhang, Z. Zheng, H. Li, and X. Liu, "Freeform illumination design: a nonlinear boundary problem for the elliptic Monge-Ampère equation," *Opt. Lett.* **38**(2), 229–231 (2013).
- [71] R. Wu, P. Liu, Y. Zhang, Z. Zheng, H. Li, and X. Liu, "A mathematical model of the single freeform surface design for collimated beam shaping," *Opt. Express* **21**(18), 20974–20989 (2013).
- [72] V. Oliker, "Differential equations for design of a freeform single lens with prescribed irradiance properties," *Opt. Eng.* **53**(3), 031302 (2014).
- [73] C.R. Prins, J.H.M. ten Thije Boonkkamp, J. Van Roosmalen, W.L. IJzerman, and T.W. Tukker, "A Monge-Ampère-Solver for free-form reflector design," *SIAM J. Sci. Comput.* **36**(3), B640–B660 (2014).
- [74] S. Chang, R. Wu, A. Li, and Z. Zheng, "Design beam shapers with double freeform surfaces to form a desired wavefront with prescribed illumination pattern by solving a Monge-Ampère type equation," *J. Opt.* **18**(12), 125602 (2016).
- [75] C. E. Gutiérrez and A. Sabra, "Freeform Lens Design for Scattering Data with General Radiant Fields," *Archive for rational mechanics and analysis* **228** (2), 341–399 (2017).
- [76] H. Ries and J. Muschaweck, "Tailored freeform optical surfaces," *J. Opt. Soc. Am. A* **19**(3), 590–595 (2002).
- [77] R. Wu, Y. Zhang, M.M. Sulman, Z. Zheng, P. Benítez, and J.C. Miñano, "Initial design with L2 Monge-Kantorovich theory for the Monge-Ampère equation method in freeform surface illumination design," *Opt. Express* **22**(13), 16161–16177 (2014).
- [78] Y. Zhang, R. Wu, P. Liu, Z. Zheng, H. Li, and X. Liu, "Double freeform surfaces design for laser beam shaping with Monge-Ampère equation method," *Opt. Comm.* **331**, 297–305 (2014).
- [79] K. Brix, Y. Hafizogullari, and A. Platen, "Designing illumination lenses and mirrors by the numerical solution of Monge-Ampère equations," *J. Opt. Soc. Am. A* **32**(11), 2227–2236 (2015).
- [80] R. Wu, S. Chang, Z. Zheng, L. Zhao, and X. Liu, "Formulating the design of two freeform lens surfaces for point-like light sources," *Opt. Lett.* **43** (7), 1619–1622 (2018).
- [81] P. Moon and D. E. Spencer, "Theory of the photic field," *J. of the Franklin Institute* **255**(1), 33–50 (1953).
- [82] C. Bösel and H. Gross, "Ray mapping approach for the efficient design of continuous freeform surfaces," *Opt. Express* **24**(12), 14271–14282 (2016).
- [83] C. Bösel and H. Gross, "Ray mapping approach in double freeform surface design for collimated beam shaping," *Proc. SPIE* **9950**, 995004 (2016).
- [84] C. Bösel, N.G. Worku and H. Gross, "Ray mapping approach in double freeform surface design for collimated beam shaping beyond the paraxial approximation," *Appl. Opt.* **56**(13), 3679–3688 (2017).
- [85] C. Bösel and H. Gross, "Single freeform surface design for prescribed input wavefront and target irradiance," *J. Opt. Soc. Am. A* **34**(9), 1490–1499 (2017).
- [86] C. Bösel and H. Gross, "Double freeform illumination design for prescribed wavefronts and irradiances," *J. Opt. Soc. Am. A* **35**(2), 236–243 (2018).
- [87] C. Bösel, J. Hartung, and H. Gross, "Irradiance and Phase control with two freeform surfaces using partial differential equations," *Proc. of SPIE Vol. 10693*, 106930C-1 (2018).
- [88] C. Bösel, N.G. Worku and H. Gross, "Compact freeform illumination system design for pattern generation with extended sources," *Appl. Opt.* **58**(10), 2713–2724 (2019).

- [89] C. Bösel, “Verfahren und Vorrichtung zur Projektion von Mustern,” Patent application 10 2019 001 922.9 (2019).
- [90] G. Berendt, and E. Weimar-Woods, *Mathematik für Physiker Band 1*, (VCH, 1990).
- [91] S. Haker, L. Zhu, A. Tannenbaum and S. Angenent, “Optimal mass transport for registration and warping,” *Int. J. Comp. Vis.* **60**(3), 225–240 (2004).
- [92] M.M. Sulman, J.F. Williams, and R.D. Russel, “An efficient approach for the numerical solution of Monge-Ampère equation,” *Appl. Numer. Math.* **61**(3), 298-307 (2011).
- [93] B. Froese, *Numerical methods for the elliptic Monge-Ampère equation and optimal transport* (Ph.D. Diss., Simon Fraser University, 2012).
- [94] J. Rubinstein, G. Wolansky, “Reconstruction of Optical Surfaces from Ray Data,” *Optical Review* **8**(4), 281-283 (2001).
- [95] J. Rubinstein, G. Wolansky, “A variational principle in optics,” *J. Opt. Soc. Am. A* **21**(11), 2164-2172 (2004).
- [96] M. Beier, J. Hartung, T. Peschel, C. Damm, A. Gebhardt, S. Scheiding, D. Stumpf, U. D. Zeitner, S. Risse, R. Eberhardt, and A. Tünnermann, “Development, fabrication, and testing of an anamorphic imaging snap-together freeform telescope,” *Appl. Opt.* **54**(12), 3530-3542 (2015).
- [97] J. Stock, A. Broemel, J. Hartung, D. Ochse, and H. Gross, “Description and reimplemention of real freeform surfaces,” *Appl. Opt.* **56** (3), 391-396 (2017).
- [98] A. Bruneton, A. Bäuerle, R. Wester, J. Stollenwerk, and P. Loosen, “Limitations of the ray mapping approach in freeform optics design,” *Opt. Lett.* **38**(11), 1945–1947 (2013).
- [99] Z. Feng, B.D. Froese, and R. Liang, “Freeform illumination optics construction following an optimal transport map,” *Appl. Opt.* **55**(16), 4301–4306 (2016).
- [100] T.F. Coleman, and Y. Li, “On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds,” *Mathematical Programming* **67** (2), 189–224, (1994).
- [101] T.F. Coleman, and Y. Li, “An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds,” *SIAM Journal on Optimization* **6**, 418–445, (1996).
- [102] J.R. Dormand, and P.J. Prince, “A family of embedded Runge-Kutta formulae,” *J. Comp. Appl. Math.*, **6**, 19–26, (1980).
- [103] L.F., Shampine and M.W. Reichelt, “The MATLAB ODE Suite,” *SIAM Journal on Scientific Computing*, **18**, 1–22, (1997).
- [104] G. Barles, and P.E. Souganidis, “Convergence of approximation schemes for fully nonlinear second order equations,” *Asymptotic Analysis*, **4**, 271–283, (1991).
- [105] <https://de.mathworks.com/matlabcentral/fileexchange/46223-regularizedata3d>
- [106] R. T. Rockafellar, “Characterization of the subdifferentials of convex functions,” *Pacific J. Math.* **17**(3), 497–510, (1966).
- [107] G. H. Spencer, and M. V. R. K. Murty, “General ray-tracing procedure,” *J. Opt. Soc. Am.* **52**(6), 672–678 (1962).
- [108] <https://de.mathworks.com/help/matlab/ref/ode45.html>
- [109] <https://de.mathworks.com/help/optim/ug/fsolve.html>
- [110] R. Wu, Z. Feng, Z. Zheng, R. Liang, P. Benítez, J. C. Miñano, F. Duerr, “Design of Freeform Illumination Optics”, *Laser Photonics Rev.* **12** (7), 1700310 (2018).
- [111] S. Zwick, S. Heist, R. Steinkopf, S. Huber, S. Krause, C. Bräuer- Burchardt, P. Kühmstedt, and G. Notni, “3D phase-shifting fringe projection system on the basis of a tailored free-form mirror,” *Appl. Opt.* **52**(14), 3134–3146 (2013).
- [112] S. Zwick, P. Kühmstedt, and G. Notni, “Phase-shifting fringe projection system using freeform optics,” *Proc. SPIE* **8169**, 81690W-1 – 81690W-9 (2011).

-
- [113] S. Zwick, R. Feßler, J. Jegorov, and G. Notni, “Resolution limitations for tailored picture-generating freeform surfaces,” *Opt. Express* **22**(4), 3642–3653 (2012).
- [114] D. Infante-Gómez, H.P. Herzig, “Micro-optical freeform elements for beam-shaping,” *Proc. SPIE* **9630**, 96300N-1 – 96300N-5 (2015).
- [115] D. Infante-Gómez, H.P. Herzig, “Design, simulation, and quality evaluation of micro-optical freeform beam shapers at different illumination conditions,” *Appl. Opt.* **55**(29), 8340–8346 (2016).
- [116] Zexin Feng, Brittany D. Froese, and Rongguang Liang, “Composite method for precise freeform optical beam shaping,” *Appl. Opt.* **54**(31), 9364–9369 (2015).
- [117] B.G. Assefa, T. Saastamoinen, M. Pekkarinen, J. Biskop, M. Kuittinen, J. Turunen, and J. Saarinen, “Design and characterization of 3D-printed freeform lenses for random illuminations,” *Proc. SPIE* **10554**, 105541J-1 – 105541J-8 (2018).
- [118] D. Kuzmin, *A Guide to Numerical Methods for Transport Equations* (University Erlangen-Nuremberg, 2010).

A Additional content

A.1 Local energy conservation for nonplanar target surface

In section 2.2.1 the emittance and irradiance were defined on two parallel planes $z = z_0$ and $z = z_T$. In case that a general target surface $z = z_T(\mathbf{x})$ is considered, Eq. (2.19) can still be applied to describe the local energy redistribution. Assuming the final ray coordinates $\mathbf{u}(\mathbf{x})$ and irradiance $I_T(\mathbf{x})$ on $z = z_T(\mathbf{x})$, Eq. (2.19) refers to an irradiance $I_T^{proj}(\mathbf{x})$. This irradiance $I_T^{proj}(\mathbf{x})$ is the projection of $I_T(\mathbf{x})$ onto a plane $z = z_T^{proj}$, which is parallel to the input plane $z = z_0$. Thus, the infinitesimal quadrilateral on $z = z_T^{proj}$ with the area $A_{z_T^{proj}}^{proj}(\mathbf{u}(\mathbf{x})) = \det(\nabla \mathbf{u}(\mathbf{x})) dx dy$ [Fig. 2.4 (b)] needs to be locally projected onto $z = z_T(\mathbf{x})$ to correctly describe the energy redistribution. This is done by utilizing the normal vector field $\mathbf{n}_{z_T}(\mathbf{x}) = (-\partial_x z_T(\mathbf{x}), -\partial_y z_T(\mathbf{x}), 1)^T$ of $z_T(\mathbf{x})$ at $\mathbf{u}(\mathbf{x})$ to calculate the area $A_{z_T}(\mathbf{u}(\mathbf{x}))$ on $z = z_T(\mathbf{x})$ by

$$\frac{A_{z_T}^{proj}(\mathbf{u}(\mathbf{x}))}{A_{z_T}(\mathbf{u}(\mathbf{x}))} = \hat{\mathbf{n}}_{z_T}(\mathbf{u}(\mathbf{x})) \cdot \hat{\mathbf{n}}_{z_T^{proj}}(\mathbf{u}(\mathbf{x})) = (\hat{\mathbf{n}}_{z_T}(\mathbf{u}(\mathbf{x})))_z \Rightarrow A_{z_T}(\mathbf{u}(\mathbf{x})) = \frac{\det(\nabla \mathbf{u}(\mathbf{x})) dx dy}{(\hat{\mathbf{n}}_{z_T}(\mathbf{u}(\mathbf{x})))_z} \quad (\text{A.1})$$

Hence, it follows the modified Jacobian equation

$$I_S(\mathbf{x}) = \det(\nabla \mathbf{u}(\mathbf{x})) I_T^{proj}(\mathbf{u}(\mathbf{x})) = \det(\nabla \mathbf{u}(\mathbf{x})) I_T(\mathbf{u}(\mathbf{x}), z_T(\mathbf{u})) \sqrt{1 + [(\partial_x z_T(\mathbf{x}))|_{\mathbf{u}}]^2 + [\partial_y z_T(\mathbf{x})|_{\mathbf{u}}]^2}, \quad (\text{A.2})$$

which describes the energy redistribution between $z = z_0$ and $z = z_T(\mathbf{x})$. Consequently, Eq. (A.2) can be directly applied to the noise-free irradiance computation in optical systems [section 5.5.3] with a target surface at $z = z_T(\mathbf{x})$.

A.2 Single FID with general target surfaces

A.2.1 Generalized PDE model

The PDE models for single FID from chapter 3 can be directly generalized to target surfaces $z = z_T(\mathbf{x})$. For physically reasonable target surfaces the required irradiance is defined on $z = z_T(\mathbf{x})$. In this case the energy conservation equation in Eq. (3.10) or Eq. (3.30), respectively, is simply replaced by Eq. (A.2) and z_T in the ray mapping equations in Eq. (3.10) or Eq. (3.30) is replaced by $z_T(\mathbf{u})$.

The corresponding PDE systems, therefore describe the single FID for zero-étendue input wavefronts and target surfaces $z = z_T(\mathbf{x})$.

A.2.2 Numerics

For single FID with a general target surface $z = z_T(\mathbf{x})$, the workflow in Fig. 5.5 (a) is still applicable. In that case the initial mapping $\mathbf{u}^\infty(\mathbf{x})$ is calculated for the projected irradiance $I_T^{proj}(\mathbf{x})$, which can be calculated straightforwardly from $I_T(\mathbf{x})$ and $z = z_T(\mathbf{x})$.

The initial surface $z_S^\infty(\mathbf{x})$ is constructed from Eq. (5.15), in which the refracted vector $\mathbf{s}_2(\mathbf{x})$ field is defined through $z = z_T(\mathbf{x})$ instead of $z = z_T$: $\mathbf{s}_2(\mathbf{x}) = (\mathbf{u}^\infty(\mathbf{x}) - \mathbf{x}_S, z_T(\mathbf{u}^\infty(\mathbf{x})) - z_S^\infty(\mathbf{x}))^T$ with \mathbf{x}_S defined according to Eq. (3.7). The Runge-Kutta method can still be applied to integrate the corresponding differential equations.

The final surface is then obtained analogously to planar target surfaces by solving the nonlinear equation system that results from the discretization of the generalized PDE model from Appendix A.2.1 with the trust-region-reflective solver.

A.3 Double lens with two freeform surfaces

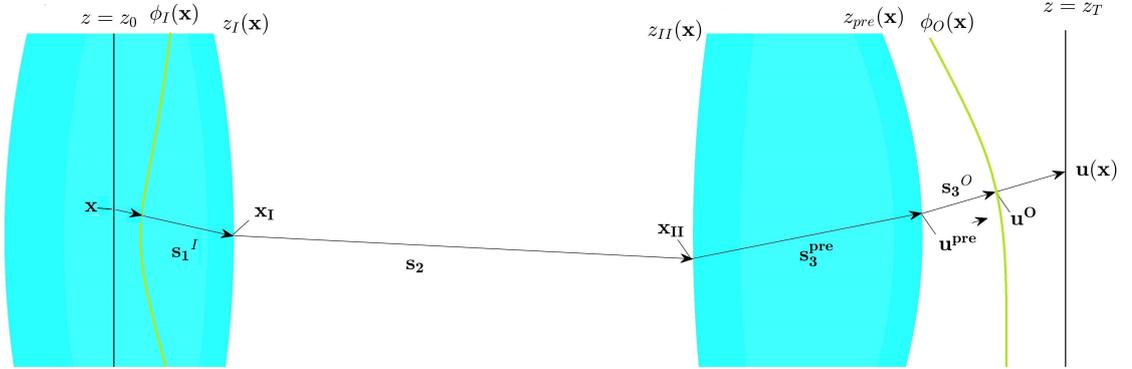


Figure A.1: Double lens system with two freeform surfaces $z_I(\mathbf{x})$ and $z_{II}(\mathbf{x})$. The ray intersection coordinates $\mathbf{u}^{pre}(\mathbf{x})$ of the predefined exit surface is defined by the projection of $\mathbf{u}(\mathbf{x})$ onto $z_{pre}(\mathbf{x})$ according to the output wavefront $\phi_O(\mathbf{x})$.

For double lens systems with two freeform surfaces and a predefined entrance surface of the first lens and exit surface $z_{pre}(\mathbf{x})$ of the second lens [Fig. A.1], Eq. (3.18) is replaced by

$$\begin{aligned} \mathbf{u}(\mathbf{x}) - \mathbf{x} = & \frac{z_T - z_{pre}(\mathbf{u}^{pre})}{(\hat{\mathbf{s}}_3)_z(\mathbf{u})} \begin{pmatrix} (\hat{\mathbf{s}}_3)_x(\mathbf{u}) \\ (\hat{\mathbf{s}}_3)_y(\mathbf{u}) \end{pmatrix} + \frac{z_{pre}(\mathbf{u}^{pre}) - z_{II,S}(\mathbf{x})}{(\hat{\mathbf{s}}_3^{pre})_z(\mathbf{u}^{pre})} \begin{pmatrix} (\hat{\mathbf{s}}_3^{pre})_x(\mathbf{u}^{pre}) \\ (\hat{\mathbf{s}}_3^{pre})_y(\mathbf{u}^{pre}) \end{pmatrix} \\ & + \frac{z_{II,S}(\mathbf{x}) - z_{I,S}(\mathbf{x})}{(\hat{\mathbf{s}}_2)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_2)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_2)_y(\mathbf{x}) \end{pmatrix} + \frac{z_{I,S}(\mathbf{x}) - z_0}{(\hat{\mathbf{s}}_1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1)_y(\mathbf{x}) \end{pmatrix}. \end{aligned} \quad (\text{A.3})$$

Hereby $\mathbf{u}^{pre}(\mathbf{x})$ is the projection of the mapping $\mathbf{u}(\mathbf{x})$ onto the exit surface $z_{pre}(\mathbf{x})$ according to the output wavefront $\phi_O(\mathbf{x})$.

The vector field $\mathbf{s}_3^{\text{pre}}(\mathbf{x})$ is defined between the freeform surface $z_{II,S}(\mathbf{x})$ and the exit surface $z_{\text{pre}}(\mathbf{x})$ and can be expressed directly by the raytracing equations through the given outgoing ray directions $\hat{\mathbf{s}}_3(\mathbf{x})$ and vice versa to get $\hat{\mathbf{s}}_3(\mathbf{u})$ in Eq. (A.3) as a function of $\mathbf{u}^{\text{pre}}(\mathbf{x})$. To eliminate $z_{II,S}(\mathbf{x})$ from Eq. (A.3), we again solve the constant OPL condition [Fig. A.1]

$$OPL = n_1|\mathbf{s}_1^I| + n_2|\mathbf{s}_2| + n_1|\mathbf{s}_3^{\text{pre}}| + n_2|\mathbf{s}_3^O| \quad (\text{A.4})$$

analytically to express $z_{II,S}(\mathbf{x})$ in terms of $z_{I,S}(\mathbf{x})$, $\phi_O(\mathbf{u}^O)$ and $z_{\text{pre}}(\mathbf{u}^{\text{pre}})$. By applying Eq. (3.21) in (A.3), the mapping $\mathbf{u}(\mathbf{x})$ is then projected onto the output wavefront. Similar to section 3.1.3, the resulting ray-mapping equation will depend on the unknown intermediate coordinate $\mathbf{u}^{\text{pre}}(\mathbf{x})$, which vary with the projected wavefront coordinates. In contrast to section 3.1.3, the outgoing ray directions $\hat{\mathbf{s}}_3(\mathbf{x})$ after the exit surface are predefined and the coordinates $\mathbf{u}^{\text{pre}}(\mathbf{x})$ can be fixed through the constraints

$$\mathbf{u}^{\text{pre}} = \frac{z_{\text{pre}}(\mathbf{u}^{\text{pre}}) - \phi_O(\mathbf{u}^O)}{(\hat{\mathbf{s}}_3)_z(\mathbf{u}^O)} \begin{pmatrix} (\hat{\mathbf{s}}_3)_x(\mathbf{u}^O) \\ (\hat{\mathbf{s}}_3)_y(\mathbf{u}^O) \end{pmatrix} + \mathbf{u}^O. \quad (\text{A.5})$$

This leads also to a PDE system like Eq. (3.22) for $\mathbf{u}^O(\mathbf{x})$ and $z_{I,S}(\mathbf{x})$ and the constraints Eq. (A.5) for the intermediate coordinates $\mathbf{u}^O(\mathbf{x})$.

A.4 Coupling of input wavefronts by freeform systems

The framework presented in section 3.3.2 can be directly applied to the description of irradiance control of N input wavefronts with N surfaces for irradiance control. In that case the PDE system in Eq. (3.30) has to be applied to every freeform surface and input wavefront *separately*, hence giving a system of $3 \cdot N$ coupled partial differential equations and $\sum_{i=1}^N (N - i)$ constraints according to Eq. (3.31) for the same numbers of unknown functions.

From the viewpoint of section 3.1, the k 's input wavefront is controlled by surface $z_k(\mathbf{x})$, which "considers" all surfaces $i \neq k$ as predefined surfaces in its sub-PDE system

$$\begin{aligned} \mathbf{u}^k(\mathbf{x}) &= \mathbf{f}(z_{FF,S}^k(\mathbf{x}), \partial_x z_{FF,S}^k(\mathbf{x}), \partial_y z_{FF,S}^k(\mathbf{x}), \mathbf{x}_{k+1}(\mathbf{x}), \dots, \mathbf{x}_N(\mathbf{x})) \\ \det(\nabla \mathbf{u}^k(\mathbf{x})) I_T^k(\mathbf{u}^k(\mathbf{x})) &= I_S^k(\mathbf{x}). \\ BC : \mathbf{u}^k(\partial \Omega_S^k) &= \Omega_T^k \end{aligned} \quad (\text{A.6})$$

with

$$\begin{pmatrix} \partial_{x_i} z_i(\mathbf{x}_i) \\ \partial_{y_i} z_i(\mathbf{x}_i) \end{pmatrix} = \frac{n_{i-1} \hat{\mathbf{s}}_{i-1}^k - n_i \hat{\mathbf{s}}_i^k}{(n_{i-1} \hat{\mathbf{s}}_{i-1}^k - n_i \hat{\mathbf{s}}_i^k)_z} \quad i = k + 1, \dots, N \quad (\text{A.7})$$

The superscript k hereby indicates that the PDE system and constraints refer to the freeform surface $i = k$.

In contrast to the single freeform design the other "predefined" surfaces are not fixed, since they are simultaneously described by a PDE system like (A.6). For the irradiance control with N surfaces and an extended source described by N input wavefronts $I_T^k(\mathbf{x}) \equiv I_T(\mathbf{x}), \forall k \in \{1, \dots, N\}$ is considered.

From a numerical perspective additional difficulties arise compared to the SFF Design. Firstly, all surfaces $i \neq k$ have to be interpolated during the simultaneous solving process of the sub-PDE systems due to the different ray paths of different input wavefronts through the system. Secondly, all surfaces $i \neq k$ have to be extrapolated, since the support of the freeform surfaces might not coincide with the ray-paths of different input wavefronts. To overcome this time consuming interpolation process a parameteric description of the surfaces might be advantageous. The numerical treatment of coupled input wavefront is not considered in this thesis.

A.4.1 Example: N=2

For better understanding, we state here the case for coupling two point light sources explicitly. In that case the total PDE system has two coupled sub-PDE systems, $k \in \{1, 2\}$, for two freeform surfaces $z_{FF,S}^1(\mathbf{x})$ and $z_{FF,S}^2(\mathbf{x})$. The first sub-PDE system treats $z_{FF,S}^2(\mathbf{x})$ as the predefined exit surface and the second sub-PDE system treats $z_{FF,S}^1(\mathbf{x})$ as the predefined entrance surface. For $k = 1$ we get

$$\mathbf{u}^1(\mathbf{x}) = \frac{z_T - z_{FF}^2(\mathbf{x}_2^1)}{(\hat{\mathbf{s}}_2^1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_2^1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_2^1)_y(\mathbf{x}) \end{pmatrix} + \frac{z_{FF}^2(\mathbf{x}_2^1) - z_{FF,S}^1(\mathbf{x})}{(\hat{\mathbf{s}}_1^1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1^1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1^1)_y(\mathbf{x}) \end{pmatrix} + \frac{z_{FF,S}^1(\mathbf{x}) - z_0}{(\hat{\mathbf{s}}_0^1)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_0^1)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_0^1)_y(\mathbf{x}) \end{pmatrix} + \mathbf{x}$$

$$\det(\nabla \mathbf{u}^1(\mathbf{x})) I_T^1(\mathbf{u}^1(\mathbf{x})) = I_S^1(\mathbf{x}) \quad (\text{A.8})$$

with the constraints for the surface interception points $\mathbf{x}_2^1(\mathbf{x})$

$$\begin{pmatrix} \partial_{x_2^1} z_{FF}^2(\mathbf{x}_2^1) \\ \partial_{y_2^1} z_{FF}^2(\mathbf{x}_2^1) \end{pmatrix} = \frac{n_1 \hat{\mathbf{s}}_1^1 - n_2 \hat{\mathbf{s}}_2^1}{(n_1 \hat{\mathbf{s}}_1^1 - n_2 \hat{\mathbf{s}}_2^1)_z} \quad (\text{A.9})$$

and for $k = 2$ we get

$$\mathbf{u}^2(\mathbf{x}) = \frac{z_T - z_{FF,S}^2(\mathbf{x})}{(\hat{\mathbf{s}}_2^2)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_2^2)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_2^2)_y(\mathbf{x}) \end{pmatrix} + \frac{z_{FF,S}^2(\mathbf{x}) - z_{FF}^1(\mathbf{x}_1^2)}{(\hat{\mathbf{s}}_1^2)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_1^2)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_1^2)_y(\mathbf{x}) \end{pmatrix} + \frac{z_{FF}^1(\mathbf{x}_1^2) - z_0}{(\hat{\mathbf{s}}_0^2)_z(\mathbf{x})} \begin{pmatrix} (\hat{\mathbf{s}}_0^2)_x(\mathbf{x}) \\ (\hat{\mathbf{s}}_0^2)_y(\mathbf{x}) \end{pmatrix} + \mathbf{x}$$

$$\det(\nabla \mathbf{u}^2(\mathbf{x})) I_T^2(\mathbf{u}^2(\mathbf{x})) = I_S^2(\mathbf{x}) \quad (\text{A.10})$$

and

$$\begin{pmatrix} \partial_{x_1^2} z_{FF}^1(\mathbf{x}_1^2) \\ \partial_{y_1^2} z_{FF}^1(\mathbf{x}_1^2) \end{pmatrix} = \frac{n_1 \hat{\mathbf{s}}_1^2 - n_2 \hat{\mathbf{s}}_2^2}{(n_1 \hat{\mathbf{s}}_1^2 - n_2 \hat{\mathbf{s}}_2^2)_z} \quad (\text{A.11})$$

We therefore have to solve the PDE system simultaneously for seven unknown functions: the

surfaces $z_{FF,S}^1(\mathbf{x})$ and $z_{FF,S}^2(\mathbf{x})$, the mapping function $\mathbf{u}^1(\mathbf{x})$ and $\mathbf{u}^2(\mathbf{x})$ and the surface interception coordinates of the second surface $\mathbf{x}_2^1(x)$.

A.5 Derivation of Equation (4.5)

In this section, we will derive Eq. (4.5). This is done by plugging the law of reflection/refraction (4.1) into the integrability condition (4.2) and using the vanishing curl of an incident field $\hat{\mathbf{s}}_1(\mathbf{x})$, representing a zero-étendue wavefront. We get

$$\mathbf{n} \cdot \left(\nabla \times \frac{\mathbf{s}_2}{|\mathbf{s}_2|} \right) = \mathbf{n} \cdot \left(\frac{1}{|\mathbf{s}_2|} \nabla \times \mathbf{s}_2 - \mathbf{s}_2 \times \nabla \frac{1}{|\mathbf{s}_2|} \right) = 0 \quad (\text{A.12})$$

and after applying

$$\nabla \frac{1}{|\mathbf{s}_2|} = -\frac{1}{2|\mathbf{s}_2|^3} \nabla(\mathbf{s}_2 \cdot \mathbf{s}_2) = -\frac{1}{|\mathbf{s}_2|^3} [\mathbf{s}_2 \times (\nabla \times \mathbf{s}_2) + (\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2] \quad (\text{A.13})$$

we obtain

$$\mathbf{s}_2 \times \nabla \frac{1}{|\mathbf{s}_2|} = -\frac{1}{|\mathbf{s}_2|^3} \left\{ \underbrace{\mathbf{s}_2 \times [\mathbf{s}_2 \times (\nabla \times \mathbf{s}_2)]}_{= [\mathbf{s}_2 \times (\nabla \times \mathbf{s}_2)] \mathbf{s}_2 - |\mathbf{s}_2|^2 (\nabla \times \mathbf{s}_2)} + \mathbf{s}_2 \times [(\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2] \right\}. \quad (\text{A.14})$$

Hence, Eq. (A.12) can be written as

$$\mathbf{n} \cdot \{ [\mathbf{s}_2 \cdot (\nabla \times \mathbf{s}_2)] \mathbf{s}_2 + \mathbf{s}_2 \times [(\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2] \} = 0. \quad (\text{A.15})$$

We now insert Eq. (4.1) and use $\hat{\mathbf{s}}_1 = (0, 0, 1)$ as well as $\hat{\mathbf{s}}_2 \cdot (\mathbf{s}_2 \times \dots) \equiv 0$ so that

$$(\mathbf{n} \cdot \mathbf{s}_2) [\mathbf{s}_2 \cdot (\nabla \times \mathbf{s}_2)] + n_1 \{ \mathbf{s}_2 \times [(\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2] \}_z = 0. \quad (\text{A.16})$$

From $\mathbf{s}_2 = \mathbf{s}_3 - \mathbf{s}_1$ we get

$$\mathbf{s}_2 (\nabla \times \mathbf{s}_1) = n_1 \frac{\{ \mathbf{s}_2 \times [(\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2] \}_z}{\mathbf{n} \cdot \mathbf{s}_2} + \mathbf{s}_2 (\nabla \times \mathbf{s}_3). \quad (\text{A.17})$$

With the definition $\mathbf{v} = (-(\mathbf{s}_2)_y, (\mathbf{s}_2)_x)$ and Eq. (4.3), the LHS of Eq. (A.17) can be written as

$$\mathbf{s}_2 (\nabla \times \mathbf{s}_1) = \begin{pmatrix} -(\mathbf{s}_2)_y \\ (\mathbf{s}_2)_x \end{pmatrix} \cdot \begin{pmatrix} \partial_x z(x, y) \\ \partial_y z(x, y) \end{pmatrix} = \mathbf{v} \nabla z(x, y) \quad (\text{A.18})$$

and for the RHS it follows

$$\begin{aligned} \{ \mathbf{s}_2 \times [(\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2] \}_z &= (\mathbf{s}_2)_x \cdot [(\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2]_y - (\mathbf{s}_2)_y \cdot [(\mathbf{s}_2 \cdot \nabla) \mathbf{s}_2]_x \\ &= \begin{pmatrix} -(\mathbf{s}_2)_y \\ (\mathbf{s}_2)_x \end{pmatrix} \cdot \left[(\mathbf{s}_2 \cdot \nabla) \begin{pmatrix} (\mathbf{s}_2)_x \\ (\mathbf{s}_2)_y \end{pmatrix} \right] \\ &= \mathbf{v} \cdot [(\mathbf{v}^\perp \cdot \nabla) \mathbf{v}^\perp] \end{aligned} \quad (\text{A.19})$$

and

$$\begin{aligned}\mathbf{s}_2 \cdot (\nabla \times \mathbf{s}_3) &= (z_T - z(\mathbf{x}))[\partial_x(\mathbf{s}_2)_y - \partial_y(\mathbf{s}_2)_x] \\ &= -(z_T - z(\mathbf{x}))\nabla \mathbf{v}.\end{aligned}\tag{A.20}$$

A.6 Integrating the linear advection equation (4.9)

Hereafter, we will show how to solve the linear advection equation (4.9) for a given ray mapping $\mathbf{u}(\mathbf{x})$ or velocity field $\mathbf{v}(\mathbf{x})$, respectively. Therefore, we need to calculate the surface $z(\mathbf{x})$ on the inflow part of the integration area $\Omega_S := \overline{\{(x, y) \in \mathbb{R}_2 \mid I_S(x, y) \neq 0\}}$ [118], which is defined by the energy conservation law. On this inflow part $\partial\Omega_{S-} := \{(x, y) \in \partial\Omega_S \mid \mathbf{v} \cdot \hat{\mathbf{r}} < 0\}$ of the boundary $\partial\Omega_S$ with the outward boundary normal $\hat{\mathbf{r}}$, the velocity field points into the integration area. Using this boundary conditions, the linear advection equation (4.9) has at most one solution [118].

To calculate the boundary conditions, we at first have to calculate the surface on the boundary of the integration area. This is done by observing that the boundary of a freeform surface determines only the *tangential* deflection of an incoming boundary ray. The deflection *normal* to the boundary is determined by the inner area of the freeform surface $z(\mathbf{x})$ (see Fig. (A.2))

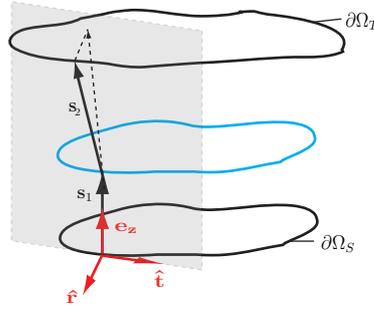


Figure A.2: The boundary $\partial\Omega_S$ of the area Ω_S is parameterized by s . The tangential vector $\hat{\mathbf{t}}$, the normal vector $\hat{\mathbf{r}}$, and the ray direction vector \mathbf{e}_z span a local coordinate system at every point of the boundary $\partial\Omega_S$. To determine the boundary values $z(s)$ of the freeform surface, only the tangential deflection of a boundary ray in every $\hat{\mathbf{t}}(s)$ - \mathbf{e}_z -plane is considered.

Hence, we consider a local coordinate system

$$\mathbf{t} = \frac{d}{ds} \begin{pmatrix} x(s) \\ y(s) \\ 0 \end{pmatrix}, \quad \mathbf{r} = \frac{d}{ds} \begin{pmatrix} y(s) \\ -x(s) \\ 0 \end{pmatrix}, \quad \mathbf{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\tag{A.21}$$

at every point of the boundary $\partial\Omega_S$, which is parametrized through the parameter s [Fig. (A.2)]. To derive a differential equation for the surface $z(s)$, we apply Eq. (4.6) in every $\hat{\mathbf{t}}(s)$ - \mathbf{e}_z -plane tangential of the surface boundary. From this, we obtain

$$\frac{\mathbf{n}}{(\mathbf{n})_z} \stackrel{!}{=} \begin{pmatrix} -\partial_l z(l) \\ 1 \end{pmatrix}, \quad l(s) := \int_0^s \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt, \quad (\text{A.22})$$

with the path length $l(s)$. Consequently this leads to the differential equation

$$\partial_s z(s) = -\frac{\mathbf{s}_2 \cdot \mathbf{t}}{(z_T - z(s)) - \frac{n_1}{n_2} \sqrt{(\mathbf{s}_2 \cdot \hat{\mathbf{t}})^2 + (z_T - z(s))^2}}, \quad (\text{A.23})$$

which asymptotically behaves as

$$\partial_s z(s) \stackrel{z_T \rightarrow \infty}{\sim} -\frac{\mathbf{s}_2 \cdot \mathbf{t}}{(z_T - z_P) \left(1 - \frac{n_1}{n_2}\right)} = \frac{v_x \partial_s y - v_y \partial_s x}{(z_T - z_P) \left(1 - \frac{n_1}{n_2}\right)} \quad (\text{A.24})$$

and similar to the paraxial approximation condition (4.8). Hereby the position of the freeform surface in space was fixed through the integration constant z_P .

These kind of ordinary differential equations can straightforwardly solved for $z(s)$ by standard techniques like the Runge-Kutta method [Eq. (A.23)] or by a simple numerical integration [Eq. (A.24)], which gives the required inflow boundary conditions. The linear advection equation (4.9) can then be solved by standard fluid dynamic approaches like finite volume methods.

Alternatively, we can also directly apply the differential equations (A.23) or (A.24). Since the boundary conditions are determined by the velocity itself, which is already known from the mapping calculation, we can also integrate these differential equations along an arbitrary path on Ω_S to obtain $z(\mathbf{x})$. A disadvantage is the error accumulation, due to the non-exact integrability of the map, along the integration path, which might lead to visible artefacts in the irradiance pattern of the calculated freeform surface.

We also want to note that Eq. (A.23) for double freeform surfaces can be derived analogously. This introduced the dependency of (A.23) on the second surface $z_{II}(\mathbf{u})$, which can be eliminated by applying the constant OPL condition [Eq. (4.16)]. Thus also in that case the ray-mapping can be integrated according to arguments given above to calculate the desired freeform surfaces.

A.7 Remarks on manufacturability of freeform surfaces

An important practical aspect, we did not discuss so far is the manufacturability of the calculated freeform surfaces in section 6. When considering certain geometrical parameters, emittances, irradiances and input and output vector fields of the design geometry, the roots of the PDE systems (3.11) and (3.22) are *fixed*. Even though PDE methods, compared to other design methods, automatically enforce a certain differentiability through the partial derivatives of the freeform surfaces, it does obviously not guarantee manufacturability. This will in general depend on the chosen fabrication method. Common methods are e.g. diamond turning, injection molding or 3D printing, which differ in their capabilities, regarding precision, manufacturing speed, etc.

For instance, in diamond turning, the freeform sag values and the azimuthal accelerations of the diamond turning machine are the important parameters with regard to the manufacturability.

Consequently, after the design of the freeform surfaces, the corresponding freeform data needs to be translated into a toolpath for the diamond turning machine.

The data from such a manufacturing analysis are shown in Fig. A.3. In that design example, a double freeform consisting of two plano-freeform surfaces was calculated.

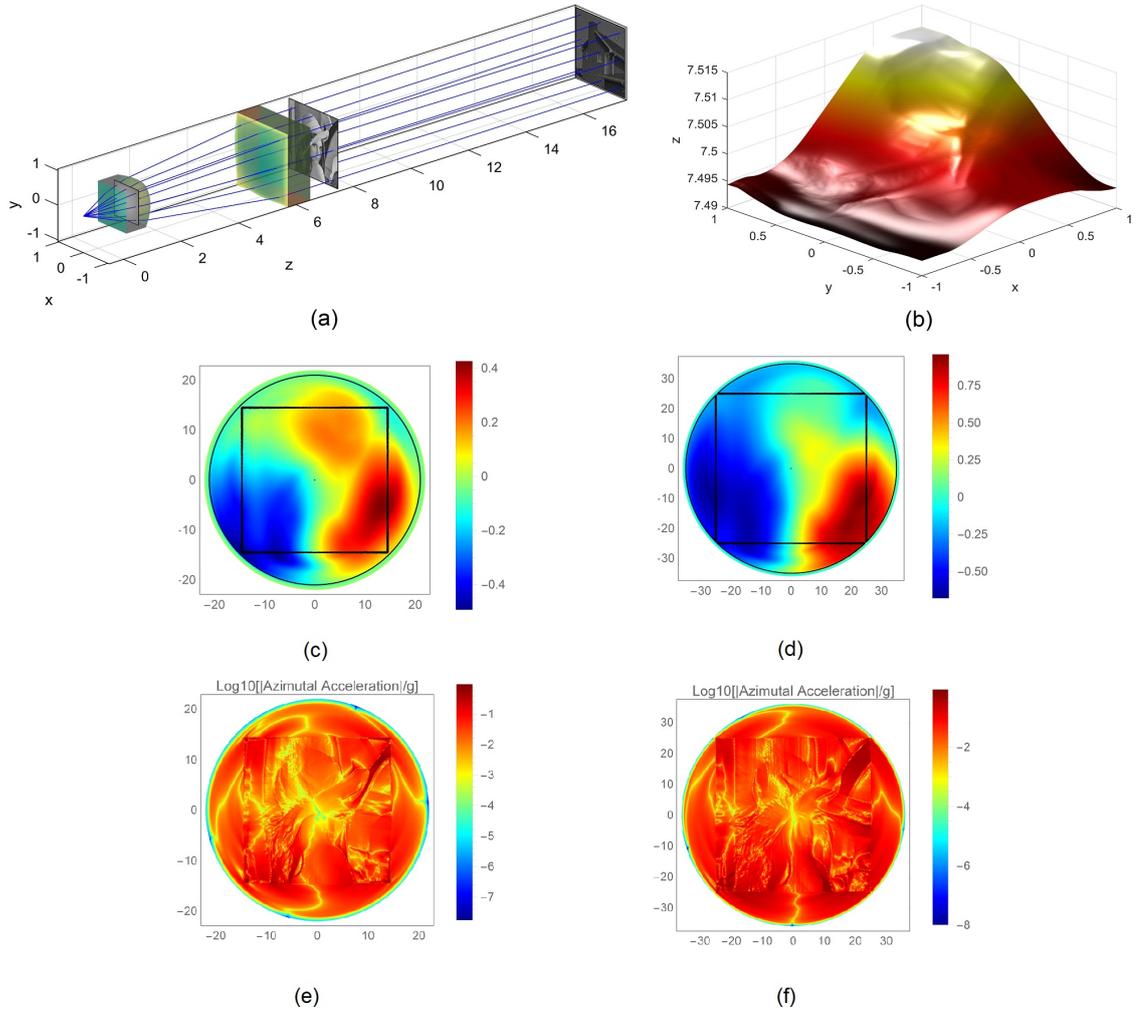


Figure A.3: (a) System geometry with two plano-freeform lenses for the redistribution of a PLS with a Lambertian intensity distribution into two irradiance distributions. (b) Output wavefront transforming “Lena” into “house”. Sag values of the (c) first freeform and (d) second freeform from the toolpath generation. Azimuthal acceleration of the diamond turning for the (e) first freeform and (f) second freeform. [87]

These freeform surfaces redistribute a PLS with a Lambertian Intensity into two irradiance distributions “Lena” and “house” by defining an appropriate output wavefront [Fig. A.3 (b)] (see Ref. [87]). Figures A.3 (c) and (d) show the freeform surfaces from the generated toolpath. The freeform surfaces have a side length 29 mm and 50 mm, respectively, and were extrapolated onto a circular area to minimize the manufacturing errors at the boundary due to the rotational toolpath of the diamond turning machine. Figures A.3 (e) and (f) show the azimuthal accelerations of the diamond turning. According to the manufacturing analysis, for the chosen system scale, the

freeform surfaces can be manufactured at 200 RPM conveniently with the fast tool servo.

B Publications

Journal articles

1. C. Bösel, and H. Gross, “*Ray mapping approach for the efficient design of continuous freeform surfaces*” in *Opt. Express* **24**, 14271 (2016).
2. C. Bösel, N.G. Worku, and H. Gross, “*Ray mapping approach in double freeform surface design for collimated beam shaping beyond the paraxial approximation*” in *Appl. Opt.* **56**, 3679 (2017).
3. C. Bösel, and H. Gross, “*Single freeform surface design for prescribed input wavefront and target irradiance*” in *J. Opt. Soc. Am. A* **34**, 1490 (2017).
4. C. Bösel and H. Gross, “*Double freeform illumination design for prescribed wavefronts and irradiances*” in *J. Opt. Soc. Am. A* **35**, 236 (2018).
5. C. Bösel and H. Gross, “*Compact freeform illumination system design for pattern generation with extended light sources*” in *Appl. Opt.* **58**(10), 2713 (2019).

Journal articles not directly related to this thesis

1. R. Filter, C. Bösel, G. Toscano, F. Lederer, and C. Rockstuhl, “*Nonlocal effects: relevance for the spontaneous emission rates of quantum emitters coupled to plasmonic structures*” in *Opt. Lett.* **39**, 6118 (2014).

International conferences

1. C. Bösel, and H. Gross, “*Ray mapping approach in double freeform surface design for collimated beam shaping*”, Oral presentation, SPIE Optics + Photonics, San Diego (2016).
2. C. Bösel, N.G. Worku, and H. Gross, “*Laser beam shaping using freeform surfaces*”, Oral presentation, DokDok, Suhl (2017).

3. C. Bösel, and H. Gross, “*Irradiance and Phase control with two freeform surfaces using partial differential equations*”, Oral presentation, SPIE Optical System Design, Frankfurt (2018).
4. A. Brahm, C. Bösel, J. Hartung, R. Jende, P. Boettner, C. Reinlein, H. Gross, P. Kuehstedt, G. Notni, “*Active freeform mirror for optical 3D measurements*”, European Optical Society Biennial Meeting (EOSAM), Delft (2018).
5. C. Bösel, J. Hartung, H. Gross, “*Ultra-compact freeform array projection for pattern generation with extended light sources*”, UK Optical Design Meeting, London (2019).
6. C. Munkelt, H. Speck, C. Bösel, C. Junger, S. Töpfer, G. Notni, “*Continuous low-latency 3D measurements using efficient freeform GOBO pattern projection and close-to-sensor image rectification*”, SPIE Defense+Commercial sensing, Anaheim (2020).

C Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich geholfen:

- Der MatLightTracer, welcher zur Durchführung der Raytracing Simulationen in Kapitel 6 verwendet wurde, ist von Norman G. Worku programmiert worden.
- Die Herstellungsanalysen der Freiformlinsen in Appendix A.7 wurden von Johannes Hartung vom Fraunhofer Institut für angewandte Optik und Feinmechanik durchgeführt und interpretiert.
- Das Rayfile der VCSEL in Kapitel 7 wurde von Henri Speck vom Fraunhofer Institut für angewandte Optik und Feinmechanik zur Verfügung gestellt.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für die Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Die geltende Prüfungsordnung der Physikalisch-Astronomischen Fakultät ist mir bekannt. Ich versichere ehrenwörtlich, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Jena, den 01.07.2019

Christoph Bösel