Single freeform surface design for prescribed input wavefront and target irradiance

Christoph Bösel$^{1,2}$ and Herbert Gross$^{1,2}$

$^1$Friedrich-Schiller-Universität Jena, Institute of Applied Physics, Albert-Einstein-Str. 15, 07745 Jena, Germany
$^2$Fraunhofer-Institut für Angewandte Optik und Feinmechanik, Albert-Einstein-Str. 7, 07745 Jena, Germany

In beam shaping applications, the minimization of the number of necessary optical elements for the beam shaping process can benefit the compactness of the optical system and reduce its cost. The single freeform surface design for input wavefronts, which are neither planar nor spherical, is therefore of interest. In this work, the design of single freeform surfaces for a given zero-étendue source and complex target irradiances is investigated. Hence, not only collimated input beams or point sources are assumed. Instead, a predefined input ray direction vector field and irradiance distribution on a source plane, which has to be redistributed by a single freeform surface to give the predefined target irradiance, is considered. To solve this design problem, a partial differential equation (PDE) or PDE system, respectively, for the unknown surface and its corresponding ray mapping is derived from energy conservation and the ray-tracing equations. In contrast to former PDE formulations of the single freeform design problem, the derived PDE of Monge-Ampère type is formulated for general zero-étendue sources in cartesian coordinates. The PDE system is discretized with finite differences and the resulting nonlinear equation system solved by a root-finding algorithm. The basis of the efficient solution of the PDE system builds the introduction of an initial iterate construction approach for a given input direction vector field, which uses optimal mass transport with a quadratic cost function. After a detailed description of the numerical algorithm, the efficiency of the design method is demonstrated by applying it to several design examples. This includes the redistribution of a collimated input beam beyond the paraxial approximation, the shaping of point source radiation and the shaping of an astigmatic input wavefront into a complex target irradiance distribution.

PACS numbers: 42.15.-i, 42.15.Eq.

I. INTRODUCTION

The development of numerical design algorithms for the calculation of freeform surfaces for irradiance control in the geometrical optics limit is of great interest in current research. The task of these methods is thereby the calculation of freeform surfaces, which redistribute a given light source into a required target irradiance pattern. In particular the calculation of continuous freeform surfaces, which is considered here, is of importance for manufacturability and the reduction of possible diffraction effects. Especially the irradiance control with continuous single freeform surfaces for input beams with planar or nonspherical wavefronts lacks of detailed discussion in the literature. Such design algorithms can potentially reduce the number of necessary optical elements or increase the efficiency in practical applications and are investigated in this work.

In principle, the freeform design for shaping of ideal sources can be divided into three major groups. The first group consists of design approaches using Oliker’s supporting quadrics method (SQM) [1–4], which is arguably the theoretically best justified, most reliable freeform design method and usable in complex optical systems [1]. As it was shown by Oliker [2], the geometric properties of quadric surfaces can be used effectively for the freeform design process. The shaping of e.g. a point source with a freeform mirror can be done by placing the point source in the common focal point of a number of ellipsoids. By changing the focal parameters of each ellipsoid appropriately, a freeform mirror, which gives the required irradiance distribution on the target area, can be designed by a unification of the ellipsoids. Due to the high computational effort, necessary for high resolution target irradiances and large number of quadrics, the major difficulty, considering the SQM, is the development of efficient algorithms.

The second group of design approaches consists of methods, which are modelling the freeform design problem by a highly nonlinear partial differential equation (PDE) of Monge-Ampère type for point sources [5–7] and collimated beams [8–9]. The main problem is thereby the development of numerical algorithms to solve the PDE. An advantage of the PDE methods over the SQM is the lower computational effort, enabling them to handle high resolution target irradiances. The major drawback of the PDE methods is the lack of theoretical justification and reliability as pointed out by Oliker [4].

The third group are ray-mapping or source-target mapping methods, respectively [10–13]. These are based on the calculation of a ray mapping between the source and the target irradiance distribution an the subsequent construction of the freeform surface from the mapping. The major difficulty of these methods is the finding of an integrable mapping, enabling the design of a freeform surface, which is on the one hand continuous and on the other hand maps the source onto the target distribution. Some of the ray-mapping methods [11–13] explore a mapping...
calculated from optimal mass transport (OMT) theory with a quadratic cost function. This special mapping solves the collimated beam shaping problem in a paraxial approximation \[13\], but is in general not integrable. As we will demonstrate in this work, the quadratic cost function mapping is nevertheless of great usefulness for the freeform design process for more general optical configurations.

In this work, we want to generalize the design approach for collimated beam shaping with continuous single freeform surfaces presented in Ref. \[13\] to general input wavefronts of ideal sources, whereby, contrary to former PDE formulations, parallel input beams or point sources are not assumed. Instead, we derive a system of PDE’s and a Monge-Ampère equation, respectively, for a predefined input direction vector field and present a detailed numerical algorithm to solve the corresponding problem. As indicated in the previous paragraph, the basis of the numerical algorithm builds thereby a ray mapping, derived from OMT with a quadratic cost function, which will serve as an initial iterate for solving the derived PDE system. As we will demonstrate explicitly, the design algorithm extends the work of Ref. \[13\] firstly to collimated beam shaping in the nonparaxial regime, secondly to the irradiance control of point sources, and thirdly, to other input wavefront shapes.

To introduce the design algorithm, this work is structured as follows. In section II and III, a PDE system is derived from energy conservation and the ray-tracing equations for single freeform lens and mirror surfaces. For the numerical solution of the PDE system, a method for the calculation of an appropriate initial iterate is introduced in section IV. In section V the numerical algorithm for determining the desired freeform surface is presented and in section VI several design examples, illustrating the effectiveness and usefulness of the design algorithm, are discussed.

II. GEOMETRY

In the following, the problem of mapping an incoming beam, defined by the ray direction vector field

\[
\hat{s}_1 = \left(\begin{array}{c}
\hat{s}_{1x} \\
\hat{s}_{1y} \\
\hat{s}_{1z}
\end{array}\right),
\]

and the irradiance distribution \(I_S(x, y)\) on the plane \(z = z_0\), onto a predefined output irradiance distribution \(I_T(x, y)\) on the target plane \(z = z_T\) is considered (see Fig. 1).

The goal is thereby the calculation of a continuous reflecting or refracting surface, which enables the mapping of the source onto the target distribution. In a geometrical optics framework, this problem is fully governed by two equations. This is on the one hand the Jacobian equation

\[
\det(\nabla I(x, y))I_T(u(x, y)) = I_S(x, y),
\]

FIG. 1. The input \(I_S(x, y)\) on the plane \(z = z_0\), the output irradiance \(I_T(x, y)\) on the plane \(z = z_T\) and the input directions \(\hat{s}_1\) are given. The input vector field is redirected by the freeform surface \(z(x, y)\) to the target points \(u(x, y)\) according to the law of refraction/reflection to give the desired output distribution. The coordinates \((x_S, y_S)\) represent the intersection points of the input vectors with \(z(x, y)\).

with the surface unit normal vector field \(\hat{n}\). Equation (3) relates the input coordinates \((x, y)\) and the target coordinates \(u(x, y)\) by the unknown surface \(z(x, y)\). Unit vector fields are hereby denoted by the “hat”. To determine the unknown surface, we want to derive a PDE system for the mapping \(u(x, y)\) and the surface \(z(x, y)\). This can be done by expressing the considered geometry, according to Fig. 1 through the vector fields

\[
\hat{s}_1 = \left(\begin{array}{c}
x_S - x \\
y_S - y \\
z(x_S, y_S) - z_0
\end{array}\right), \quad \hat{s}_3 = \left(\begin{array}{c}
u_x(x, y) - x \\
u_y(x, y) - y \\
z_T - z_0
\end{array}\right),
\]

and by calculating the intersection points \(u(x, y)\) with the target plane \(z = z_T\) from the relation

\[
\hat{n} = \left(\begin{array}{c}
-\partial_{x_S}z(x_S, y_S) \\
-\partial_{y_S}z(x_S, y_S) \\
1
\end{array}\right),
\]

\[
s_2 = n\hat{s}_1 + \left\{-n \cdot \hat{n} \cdot \hat{s}_1 + \sqrt{1 - n^2(\hat{n} \cdot \hat{s}_1)^2}\right\}\hat{n},
\]

\[
s_2 = \hat{s}_1 - 2(\hat{n} \cdot \hat{s}_1)\hat{n}
\]
\[ s_1 + \lambda \cdot \hat{s}_2 = s_3 \quad \Rightarrow \quad \lambda = \frac{z_T - z(x_s, y_s)}{\langle \hat{s}_2 \rangle_z}. \] (5)

Hereby \( x_s, y_s \) and \( z(x_s, y_s) \) are the intersection coordinates of the input vector field \( s_1 \) with the surface \( z(x, y) \). From Eq. (5) we then directly get

\[
\begin{align*}
  u_x(x, y) &= \lambda \cdot \langle \hat{s}_2 \rangle_x + \langle s_1 \rangle_x, \\
  u_y(x, y) &= \lambda \cdot \langle \hat{s}_2 \rangle_y + \langle s_1 \rangle_y,
\end{align*}
\] (6)

which define a relation between the surface \( z(x, y) \), the surface gradient and the target coordinates. For the case of collimated input beams, analogous equations were derived by Wu et al. [8] and Oliker [16]. Together with Eq. (2), Eq. (6) builds a PDE system for the unknown functions \( u_x(x, y), u_y(x, y) \) and \( z(x, y) \). According to Eq. (1), the remaining difficulty is thereby the dependency of Eq. (6) on the surface intersection points \((x, y_s)\), which are determined by the target \( z(x, y) \) and therefore unknown. A similar problem arising in their work on the reconstruction of an optical surface from experimental ray data [17], was solved by Rubinstein and Wolansky using the geometrical relation

\[
\begin{align*}
  \frac{x_s - x}{z(x_s, y_s) - z_0} &= \frac{\langle \hat{s}_1 \rangle_x}{\langle \hat{s}_1 \rangle_z} \Rightarrow x_s = \frac{\langle \hat{s}_1 \rangle_x}{\langle \hat{s}_1 \rangle_z} [z(x_s, y_s) - z_0] + x, \\
  \frac{y_s - y}{z(x_s, y_s) - z_0} &= \frac{\langle \hat{s}_1 \rangle_y}{\langle \hat{s}_1 \rangle_z} \Rightarrow y_s = \frac{\langle \hat{s}_1 \rangle_y}{\langle \hat{s}_1 \rangle_z} [z(x_s, y_s) - z_0] + y,
\end{align*}
\] (7)

between the given ray directions defined in Eq. (1) and the vector field \( \hat{s}_1 \) in Eq. (1). By applying this coordinate transformation to the coordinates \( x_s, y_s \) and the gradient \( \nabla z(x_s, y_s) = (\partial_x z(x_s, y_s), \partial_y z(x_s, y_s)) \) occurring in Eq. (4), Eq. (6) can be expressed by the initial coordinates \((x, y)\) on the plane \( z = z_0 \). After using the chain rule it follows

\[
\begin{align*}
  \partial_x z(x_s, y_s) &= \frac{\langle \partial_y z \rangle_y [\partial_z z(x_s, y_s)] - \langle \partial_z y \rangle_y [\partial_y z(x_s, y_s)]}{\langle \partial_z x \rangle_x \cdot \langle \partial_y y \rangle_y - \langle \partial_z y \rangle_y \cdot \langle \partial_y x \rangle_x}, \\
  \partial_y z(x_s, y_s) &= \frac{-\langle \partial_x y \rangle_y [\partial_z z(x_s, y_s)] + \langle \partial_x z \rangle_x [\partial_y z(x_s, y_s)]}{\langle \partial_z x \rangle_x \cdot \langle \partial_y y \rangle_y - \langle \partial_z y \rangle_y \cdot \langle \partial_y x \rangle_x}.
\end{align*}
\] (8)

The remaining derivatives of \( x_s \) and \( y_s \) can be evaluated with Eq. (4). With these transformations and the definition

\[ z_s(x, y) \equiv z(x_s, y_s), \] (9)

Eq. (6) is fully expressed in terms of the \( x, y \)-coordinates. Hence, the Jacobian equation (2) and Eq. (6) build a system of PDEs for the unknown functions \( u(x, y) \) and \( z(x, y) \). After solving the equation system for \( z_s(x, y) \), the coordinates \((x_s, y_s)\) can then be calculated from Eq. (7), which gives the surface \( z(x, y) \) on the scattered \( x_s-y_s \) grid.

### III. PDE SYSTEM AND MONGE-AMPE`RE EQUATION

In the previous section it was shown that the single freeform design problem for zero-´ etendue sources is governed by the PDE system

\[
[(\partial_x u_x) (\partial_y u_y) - (\partial_x u_y) (\partial_y u_x)] I_T(u_x, u_y) = I_S(x, y),
\]

\[
u_x(x, y) - x = [z_T - z_S(x, y)] \langle \hat{s}_2 \rangle_x \langle \hat{s}_1 \rangle_z + [z_S(x, y) - z_0] \langle \hat{s}_1 \rangle_x \langle \hat{s}_2 \rangle_z,
\]

\[
u_y(x, y) - y = [z_T - z_S(x, y)] \langle \hat{s}_2 \rangle_y \langle \hat{s}_1 \rangle_z + [z_S(x, y) - z_0] \langle \hat{s}_1 \rangle_y \langle \hat{s}_2 \rangle_z,
\] (10)

whereby \( \hat{s}_1 \) is the predefined input ray direction vector field and \( \hat{s}_2 \) is defined according to section III in terms of \( z_s(x, y), \partial_x z_s(x, y) \) and \( \partial_y z_s(x, y) \).

For a given source and target distribution \( I_S(x, y) \) and \( I_T(x, y) \), and a given input ray direction vector field \( \hat{s}_1 \) the equation system in Eq. (10) has to be solved simultaneously for the surface \( z_s(x, y) \) and the ray mapping \( u(x, y) \).

Equation (10) can be further simplified by plugging \( u_x(x, y) \) and \( u_y(x, y) \) into the Jacobian equation to obtain a nonlinear second order PDE of Monge-Amp`ere type for the function \( z_s(x, y) \):

\[
\begin{align*}
  \partial_{xx} z_s + 2B \partial_{xy} z_s + C \partial_{yy} z_s + \left[ (\partial_{xx} z_S)(\partial_{yy} z_S) - (\partial_{xy} z_S)^2 \right] &= D.
\end{align*}
\] (11)

Hereby, \( A, B, C, D \) are functions of \( x, y, z_S(x, y), \partial_x z_S(x, y) \) and \( \partial_y z_S(x, y) \), which can be calculated straightforward.

Compared to the Eq. system (10), the Eq. (11) has numerically the advantage of a lesser number of design variables. Despite this advantage, the numerical method presented in the following will nevertheless be based on the description in Eq. (10). In contrast to the Monge-Amp`ere equation in Eq. (11), we will therefore omit the discretization of second order derivatives. In addition to that, the boundary conditions can be directly controlled through the mapping \( u(x, y) \) by requiring that boundary points of the support of \( I_S(x, y) \) are mapped onto boundary points of the support of \( I_T(x, y) \). This is contrary to the Monge-Amp`ere equation for which the same goal has to be accomplished through the surface gradients \( \nabla z_s(x, y) \) at the boundary of the surface, which leads to the dependency of the boundary conditions on the input field \( \hat{s}_1 \). Furthermore, the properties of the Jacobian equation and the mapping equations in Eq. (10) can be controlled separately.

### IV. INITIAL ITERATE

The construction of freeform surfaces for complex target irradiances from nonlinear PDEs, like Eqs. (10) or (11) is usually done by discretizing the PDEs and
solving the resulting nonlinear equation system by an iterative process. A high resolution target irradiance distribution and therefore large number of unknowns requires in general an appropriate initial iterate to ensure (fast) convergence. Especially for discretization schemes, which are not provably convergent, those initial iterates are of great importance. For the published PDE methods for single freeform surface design an initial iterate producing a uniform illumination on the target plane or other simple surface shapes are usually chosen.

Here, we present a different method to construct an initial iterate for a given input direction vector field $\mathbf{s}_1$. It is based on the OMT with the quadratic cost function and inspired by Ref. [17]. Following the argumentation from Ref. [17], an unknown surface can be experimentally reconstructed from the known input ray directions and measured target plane intersection points of the rays send out from $(x, y)$. Since the target points are not given for the freeform design problem considered here, the approach from Ref. [17] can not be directly applied. Nevertheless, as we will demonstrate in the following, the approach is of great usefulness to generate an initial iterate for the freeform surface design.

### A. Initial mapping

For the single freeform surface design process a frequently used mapping $\mathbf{u}(x, y)$ is calculated from OMT by considering Eq. (2) together with the quadratic cost function condition

\[
d(I_s, I_T)^2 = \inf_{\mathbf{u} \in M} \int |\mathbf{u}(x) - \mathbf{x}|^2 I_s(x) \, dx, \quad (12)
\]

which we will refer to in the following by $\mathbf{u}^\infty(x, y)$. $M$ in Eq. (12) denotes the set of energy preserving mappings. This is arguably the most studied OMT problem in numerical mathematics and provides us therefore with efficient algorithms for the mapping calculation. As it was demonstrated in several publications, the corresponding mapping $\mathbf{u}^\infty(x, y)$ in general will not generate a continuous surface $z(x, y)$ fulfilling the required design goal of mapping the given input $I_s(x, y)$ onto the desired output irradiance $I_T(x, y)$. However, due to its properties, the mapping $\mathbf{u}^\infty(x, y)$ appears to be a good candidate for creating an appropriate initial iterate. According to Eq. (12), the quadratic cost function can in principle be thought of as describing a ray propagation in a homogeneous medium [18]. Since a continuous surface can be interpreted as a perturbation of this free propagation, it can be expected that the mapping $\mathbf{u}^\infty(x, y)$ serves as a reasonable initial iterate for solving Eq. (10). From a numerical point of view it is also helpful that $\mathbf{u}^\infty(x, y)$ already fulfills the Jacobian equation in Eq. (10), which therefore basically constrains all numerical corrections to $\mathbf{u}^\infty(x, y)$ to the space of energy preserving mappings. Hence, we will use the mapping defined by Eqs. (2) and (12) as an initial mapping for solving Eq. (10). As in the previous publications [13, 19], we will make use of the algorithm by Sulman et al. [20] for the examples in section VI for demonstrational purposes.

### B. Initial surface

After its calculation, the mapping $\mathbf{u}^\infty(x, y)$ can be used to construct the initial surface $z_S^\infty(x, y)$ as demonstrated in the following. The resulting initial surface construction equations are thereby a generalization of the surface construction equations used in Ref. [13] to noncollimated beams.

The derivation can be done by using the law of refraction/reflection

\[
\frac{n(x_S, y_S)}{(n(x_S, y_S))^\perp} = \frac{n_1 \mathbf{s}_1 - n_2 \mathbf{s}_2}{n_1 \mathbf{s}_1 - n_2 \mathbf{s}_2} = \nabla_S(z - z_S(x, y)) \quad (13)
\]

and applying again the coordinate transformation defined in Eq. (7) to invert the gradient by using

\[
\begin{align*}
\partial_{x_S} z_S^\infty(x, y) &= -\frac{n_1(\mathbf{s}_1)_x - n_2(\mathbf{s}_2)_x}{n_1(\mathbf{s}_1)_x - n_2(\mathbf{s}_2)_x} = n_x(x, y), \\
\partial_{y_S} z_S^\infty(x, y) &= -\frac{n_1(\mathbf{s}_1)_y - n_2(\mathbf{s}_2)_y}{n_1(\mathbf{s}_1)_x - n_2(\mathbf{s}_2)_x} = n_y(x, y). \\
\end{align*} \quad (14)
\]

After using the chain rule, this gives the differential equations [17]

\[
\begin{align*}
\partial_{x_S} z_S^\infty(x, y) &= \frac{\partial_x z_S^\infty(x, y)}{1 - n_x(x, y)(\mathbf{s}_1)_x + n_y(x, y)(\mathbf{s}_1)_y} \\
\partial_{y_S} z_S^\infty(x, y) &= \frac{\partial_y z_S^\infty(x, y)}{1 - n_x(x, y)(\mathbf{s}_1)_x + n_y(x, y)(\mathbf{s}_1)_y}.
\end{align*} \quad (15)
\]

for the initial surface $z_S^\infty(x, y)$. For a given input direction vector field, defined by Eq. (1), and given target points $\mathbf{u}^\infty(x, y)$, Eq. (15) can be integrated along an arbitrary path on the support of $I_s(x, y)$ to generate the initial surface $z_S^\infty(x, y)$.

Together with the initial mapping $\mathbf{u}^\infty(x, y)$, the surface $z_S^\infty(x, y)$ builds an initial iterate for solving the PDE system (10).

We want to remark that the initial surface construction approach, presented in this section can also be used for the Monge-Ampère equation in Eq. (11). To implement the boundary conditions in that case, the surface boundary normal vectors defined by Eq. (15) have to be chosen in a way that the boundary of $I_s(x, y)$ is mapped onto the boundary of $I_T(x, y)$. 
V. NUMERICS

A. Discretization

After the calculation of the initial iterate, the PDE system (10) has to be solved. Hence, we have to discretize and solve the resulting nonlinear equation system by a root-finding algorithm. To do that, we follow Ref. [19] by introducing the corrections \( \Delta u(x, y) \) and \( \Delta z_S(x, y) \) to \( u^\infty(x, y) \) and \( z^\infty(x, y) \):

\[
\begin{align*}
  u(x, y) &= u^\infty(x, y) + \Delta u(x, y), \\
  z_S(x, y) &= z_S^\infty(x, y) + \Delta z_S(x, y)
\end{align*}
\]

(16)

and redefine the initial iterates for the solving process to be \( \Delta u(x, y) = 0 \) and \( \Delta z(x, y) = 0 \). Accordingly, Eq. (10) becomes a PDE system for \( \Delta u(x, y) \) and \( \Delta z_S(x, y) \). To solve the PDE system, we discretize the functions \( \Delta u_x(x, y), \Delta u_y(x, y), \Delta z_S(x, y) \) and their derivatives in Eq. (10) by central finite differences

\[
\begin{align*}
  \partial_x (\Delta z_S) &\rightarrow \frac{1}{2\Delta x}[(\Delta z_S)_{i,j+1} - (\Delta z_S)_{i,j-1}], \\
  \partial_y (\Delta z_S) &\rightarrow \frac{1}{2\Delta y}[(\Delta z_S)_{i+1,j} - (\Delta z_S)_{i-1,j}]
\end{align*}
\]

(17)

for the inner points \( i; j = 2, ..., N - 1 \) and second order finite differences

\[
\begin{align*}
  \partial_x (\Delta z_S) &\rightarrow -\frac{1}{2\Delta x}[3(\Delta z_S)_{i,j+2} - 4(\Delta z_S)_{i,j+1} + (\Delta z_S)_{i,j}], \\
  \partial_y (\Delta z_S) &\rightarrow -\frac{1}{2\Delta y}[3(\Delta z_S)_{i+2,j} - 4(\Delta z_S)_{i+1,j} + (\Delta z_S)_{i,j}]
\end{align*}
\]

(18)

for the boundary points. Additionally, we apply boundary conditions by requiring that the boundary points of the source distribution \( I_S(x, y) \) are mapped onto the boundary points of the target distribution \( I_T(x, y) \). For the mapping of square supports, which is demonstrated in section VI, the boundary condition reads therefore:

\[
\begin{align*}
  \Delta u_x(-0.5, y) &= \Delta u_x(0.5, y) = 0, \quad y \in [-0.5, 0.5], \\
  \Delta u_y(x, -0.5) &= \Delta u_y(x, 0.5) = 0, \quad x \in [-0.5, 0.5].
\end{align*}
\]

(19)

After the discretization of Eq. (10), the resulting nonlinear Eq. system can then be solved by a root-find algorithm for the \( 3 \cdot N^2 \) unknown values \( (\Delta u_x)_{i,j}, (\Delta u_y)_{i,j} \) and \( (\Delta z_S)_{i,j} \). The solution will lead to a discretized surface \( z_S(x, y) \) from which, according to Eq. (7), the scattered grid points \( (x_S, y_S) \) can be calculated. As in Ref. [19], we use MATLAB’s \texttt{fsolve()} function with the trust-region-reflective solver for the solution of the nonlinear Eq. system (10).

B. Design algorithm

The design algorithm developed in the previous sections can be summarized as follows:

1. Define the input ray direction vector field \( \tilde{s}_1 \) and the input irradiance \( I_S(x, y) \) on \( z = z_0 \) and the output irradiance \( I_T(x, y) \) on \( z = z_T \).

2. Calculate the initial map \( u^\infty(x, y) \) between \( I_S(x, y) \) and \( I_T(x, y) \).

3. Integrate Eq. (15) along an arbitrary path on the support of \( I_S(x, y) \) to get the initial surface \( z_S^\infty(x, y) \).

4. Discretize the PDE system in Eq. (10) to get a nonlinear equation system.

5. Apply the boundary conditions and solve the nonlinear equation system with \( u^\infty(x, y) \) and \( z_S^\infty(x, y) \) as the initial iterate to get \( u(x, y) \) and \( z_S(x, y) \).

As mentioned above, for a general input vector field \( \tilde{s}_1 \), the presented design algorithm gives us the freeform surface \( z(x, y) \) on the scattered grid points \( (x_S, y_S) \). The quality of the surface \( z_S(x, y) \) is thereby already characterized by Eq. (10), which is equivalent to a ray-tracing calculation. For manufacturing processes it might still be useful to evaluate the surface quality in a ray-tracing simulation. We therefore interpolate the scattered data points \( z(x_S, y_S) \) onto a regular grid and import it to a ray-tracing software. The interpolation can be done by scattered data interpolation with e.g. MATLAB 2015b’s \texttt{griddata()} function. However, in our experience this function has two major drawbacks for the interpolation of freeform surfaces. This is its inability (for large data sets) to extrapolate outside of the convex hull of the grid points, which leads to inaccuracies at the boundary of the freeform surface in form of zig-zag patterns at the target irradiance boundary. Furthermore, the interpolation with \texttt{griddata()} leads to oscillations in the freeform surface, depending on the stepsize of the interpolation grid and the interpolation method.

To overcome this limitation, we therefore use the \texttt{RegularizeData3D()} function from the MATLAB file exchange website [21]. This function provides the possibility to choose how smooth the interpolated data is fitted to the scattered data. In our examples, we made good experience with a small smoothness factor for low noise data. Additionally, \texttt{RegularizeData3D()} extrapolates outside the convex hull of the scattered data grid, which may also lead to slight inaccuracies at the boundary of the freeform surface, which lower the energy efficiency, as we will point out in the example section VI.

Since the initial mapping \( u^\infty(x, y) \) defined by Eqs. (2) and (12) is in general not integrable, the integration of Eq. (15) will be path dependent and therefore lead to different initial surfaces \( z_S^\infty(x, y) \) for different integration
paths. Nevertheless, in our experience the path dependency has no significant influence on the suitability for using $z_{\infty}^\infty(x,y)$ as an initial iterate.

VI. EXAMPLES

In this section, we want to demonstrate basically three things. Firstly, the design algorithm presented in section V extends the work of Ref. [13] on collimated beam shaping with single freeform surfaces to the nonparaxial regime. Secondly, the design algorithm can be applied to the design of single freeform surfaces for shaping of point light sources, which is frequently discussed in literature. And thirdly, we want to demonstrate the applicability to a nonstandard design example by dealing with an input wavefront, which is neither planar nor spherical. This is done in the following three subsections. All the examples in these sections were calculated on an Intel Core i3 at $2 \times 2.4$Ghz with 16GB RAM. To make the ray-tracing results comparable, we used for each of the examples the target distribution “lake” (see Fig. 3(b)) with a resolution of $250 \times 250$ pixels. For the calculation of the initial surface by Eq. (25), we used MATLAB’s ode45 solver with tolerances of $10^{-10}$, which is in our experience necessary for a legitimate comparison of the initial and final surface in a ray-tracing simulation. For the root finding of Eq. (10) by fsolve() tolerances of $10^{-3}$ were used. For the examples with lenses refractive indices of $n = 1.5$ were used. The ray-tracing simulations for the following examples are done with $200 \cdot 10^6$ rays and detectors with $250 \times 250$ pixels by using a self-programmed MATLAB ray-tracing toolbox. Since the design examples are calculated in a geometrical optics framework, we will use arbitrary units as length units and omit the explicit declaration of lengths by “a.u.”.

A. Collimated beam shaping beyond the paraxial approximation

In Ref. [13] it was shown that the single freeform surface design problem for collimated input beams

$$s_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$  \hspace{1cm} (20)

can be solved by the quadratic cost function ray mapping, defined by Eqs. (2) and (12), in the paraxial regime. Here, we demonstrate that the presented algorithm from section V directly extends the results from Ref. [13] to nonparaxial calculations. Since the quadratic cost function mapping solves the single freeform surface design problem exactly for infinite distances between the freeform surface and the target plane[13], it can be argued analogously to Ref. [19] that the numerical convergence of the algorithm can theoretically be ensured by a stepwise reduction of the target plane distance $z_T$. However, in our experience, this stepwise reduction is in a practical calculation not necessary to ensure numerical convergence.

FIG. 2. (a) System layout. The incoming collimated Gaussian beam is reflected by the freeform mirror to give the required target distribution “lake” at the predefined position. (b) Freeform mirror.

From an application point of view, the extension to the nonparaxial regime is especially important for optical configurations in which the distance between the target plane and the freeform surface is small, the target area is strongly shifted or scaled, or if there are caustics between the target plane and the freeform surface. In the following, we therefore demonstrate the utility of the presented algorithm by designing a freeform mirror, which shifts and scales the target area compared to the position and size of the input beam. The corresponding system layout, considered in this section, can be seen in Fig. 2, showing a target area with a side length, which is 3 times as large as the input beam’s side length of 1 and shifted by 5. The distance between the freeform mirror...
and the target plane \( z_T \) is approximately 5.

![FIG. 3.](image)

(a) Normalized Gaussian input irradiance distribution \( I_S(x, y) \) and (b) normalized output irradiance distribution \( I_T(x, y) \) “lake”. The output irradiance in (b) has the side length of 1, but will be scaled and shifted for comparison with the ray tracing simulations according to the layouts of each example. \( I_S(x, y) \) and the input ray direction vector field in Eq. (20) at \( z = z_0 \) and \( I_T(x, y) \) at \( z = z_T \) build the input of the design algorithm presented in section V.

For the irradiance distribution, we choose the Gaussian distribution, which is represented in Fig. 3 (a) and the test image “lake” as the required target distribution (Fig. 3 (b)). According to section V, we first calculate the initial mapping between the input and output irradiances (320 sec), then the initial surface from Eq. (10) (208 sec) and finally solve the PDE system in Eq. (15) (68 sec). After the construction of the initial and final freeform mirror, we evaluate their quality by a ray-tracing simulation. The corresponding results and comparisons with the algorithm presented in Ref. [13], which is equal to the initial iterate of the root finding, are shown in Fig. 4. Additionally, the \( \text{rms} \) and the peak-to-valley of \( \Delta I_T(x, y) = I_T(x, y) - I_{T,RT}(x, y) \), and the correlation coefficient \( \text{corr}_{I_T} \) between the predefined irradiance \( I_T(x, y) \) and the irradiance from the ray tracing \( I_{T,RT}(x, y) \), and the energy efficiency \( \eta \) are given in Table I. The results thereby show clearly the improvement of the presented design algorithm compared to Ref. [13] and the mapping from the quadratic cost function [12], respectively.

**TABLE I. Comparison of \( \Delta I_T \) for example “lake” with a collimated input beam.**

| \( \Delta I_T \) | rms \( \Delta I_T \) | \( \text{pv} \Delta I_T \) | \( \text{corr} \Delta I_T \) | \( \eta \) |
|-----------------|-----------------|-----------------|-----------------|-------|
| \( z = z_0 \)   | \( 9.2288 \cdot 10^{-5} \)  | \( 2.3472 \cdot 10^{-5} \)  | \( 1.6158 \cdot 10^{-4} \)  | \( 3.6349 \cdot 10^{-5} \)  | \( 0.5750 \)  | \( 0.9607 \)  | \( 85.66\% \)  | \( 99.98\% \)  |

**B. Point source**

Besides the collimated beam shaping with a single freeform surface, we want to apply the design algorithm to the other standard configuration considered in literature, which is the mapping of a point source onto a required irradiance distribution on a target plane \( z_T \). In the case of single freeform lens design, the point source is assumed to be in the medium or the rays emitted by the point source first hit a spherical surface, respectively. The ray directions are therefore first unaltered and then redirected by the freeform surface to give the predefined target distribution \( I_T(x, y) \).

Since we want to apply the algorithm from section V, we first have to project the point source onto the input plane \( z = z_0 \), giving the distribution \( I_S(x, y) \). Subsequently, the mapping \( u^\infty(x, y) \) between the projected distribution \( I_S(x, y) \) on \( z = z_0 \) and the target distribution \( I_T(x, y) \) on \( z = z_T \) can be calculated. For the point source, we choose a Lambertian distribution with a maximum opening angle of 30 degrees, which refers to the angle between the \( z \)-axis and the corner points of the quadratic aperture and leads to the irradiance distribution \( I_S(x, y) \) presented in Fig. 3 (a).

The input vector field \( \hat{s}_1 \) on \( z = z_0 \) is given by

\[
\hat{s}_1 = \frac{1}{\sqrt{x^2 + y^2 + z_0^2}} \begin{pmatrix} x \\ y \\ z_0 \end{pmatrix}.
\]  

(21)

The placement of the plane \( z = z_0 \) relative to the point source does not matter. Depending on the value of \( z_0 \) only the mapping \( u^\infty(x, y) \) has to be scaled by a constant.
factor to create the predefined target distribution area size, which we choose to have a side length of 5. For the distance between the point source and the freeform surface we choose \( z(0,0) = 1 \), which enters the design process as an integration constant in Eq. (15), and for the distance between the point source and the target plane, we choose \( z_T = 10 \). The corresponding system layout can be seen in Fig. 4.

The calculation of the initial mapping \( u^\infty(x,y) \) took approximately 378 sec, the initial surface \( z_S^\infty(x,y) \) 183 sec and the final surface \( z_S(x,y) \) about 213 sec. After the calculation of the freeform surface \( z_S(x,y) \) by the presented design algorithm an additional complication, compared to the collimated beam shaping, arises since the freeform surface is calculated on the nonuniform \( x_S-y_S \)-grid. Fig. 5(b) shows every 4th point of the scattered grid points. The outermost points define the aperture of the freeform lens.

As described in section [V], the function \( z_S(x,y) = z(x_S,y_S) \) is interpolated and extrapolated onto a square, uniform grid by \( \text{RegularizeData3D()} \) to give \( z(x,y) \). For a fine boundary interpolation, we choose an interpolation grid of 500 \( \times \) 500 points, whereby the minimum and maximum \( x_S \) and \( y_S \) values correspond to the minimum and maximum \( x_S \) and \( y_S \) values (see Fig. 5(b)). The function values are then imported to a ray-tracing software. In Fig. 7 and Table II the results of the ray tracing are shown. From Fig. 7(a) and 7(c) it can be seen that the quadratic cost function in Eq. (12) provides a reasonable initial iterate for the root-finding process, as argued in section [IV]. Figure 7(b) and 7(c) and Table II show results, which are comparable to the example from the previous subsection, except for the slightly lowered efficiency \( \eta \) due to the interpolation of \( z_S(x,y) \) at the surface boundary.

From the viewpoint of manufacturing, it is usually preferred to have a plane instead of a spherical surface as the entrance surface of the lens. In the following section it is demonstrated that the design algorithm can be applied directly to these configurations.
FIG. 7. Target irradiance distribution $I_T(x, y)$ from the ray tracing simulation produced by the (a) initial surface and (b) optimized surface. Difference between the predefined distribution $I_T(x, y)$ and the ray tracing simulation $I_{T,RT}(x, y)$ for (c) the initial surface and (d) the optimized surface.

C. Astigmatic wavefront

In this section, we want to explore an example with a nonstandard input vector field $\hat{s}_1$. We therefore predefine a surface $z_{\text{pre}}(x, y) = \frac{1}{3}y^2 - \frac{2}{3}x^2$, deforming a collimated beam with a gaussian distribution (Fig. 9(a)) into a beam with an astigmatic wavefront. The corresponding system layout can be seen in Fig. 8.

As in the previous examples, we need the input vector field $\hat{s}_1$ and the irradiance distribution on an intermediate plane $z = z_0$. This plane is chosen to be $z_0 = 5$ and therefore after the refraction of the astigmatic beam by the plane entrance surface ($z = 3$ in Fig. 8(a)), and the aperture at $z = z_0$ is a square with a side length of 1. The vector field $\hat{s}_1$ and the source distribution $I_S(x, y)$ on $z = z_0$ can then be calculated numerically, e.g. from the law of refraction and the Jacobian equation (2). $I_S(x, y)$ is shown in Fig. 9(b) and serves as the input for the calculation of $u^\infty(x, y)$ (417 sec).

For the calculation of the initial surface with Eq. (15), we have to fix one point of the freeform surface relative to the plane $z = z_0$, which we choose to be $z(0, 0) = z_0 + 1$ and the side length of the target area is chosen to be 2. The calculation of $z^*_S(x, y)$ took 145 sec and the root finding of Eq. (10) about 215 sec. Since the input vector field is noncollimated, the surface $z(x, y)$ is again calculated on a scattered grid and has to be interpolated for the ray-tracing simulation. The results of the ray tracing are shown in Fig. 10 and Table III.

Also for this example, the results show that the initial map is appropriate for the root finding of Eq. (10) and that by the presented design algorithm, single freeform surfaces for complex target irradiances and nonplanar or
nonspherical input wavefronts can be calculated.

The quantitative differences between the ray-tracing results for the examples are thereby mainly due to interpolation and extrapolation effects. It is also a priori not clear, if for the particular predefined system parameters of each example an exact solution to Eq. \( I_0 \) exists. This might also lead to a degradation of the quality of the ray-tracing results, since an accurate root-finding of the respective nonlinear equation systems might not be possible. In contrast to the first two examples, additional errors are introduced in the third design example by the numerical calculation of \( I_S(x, y) \) and the input direction vector field.

FIG. 10. Target irradiance distribution \( I_T(x, y) \) from the ray tracing simulation produced by the (a) initial surface and (b) optimized surface. Difference between the predefined distribution \( I_T(x, y) \) and the ray tracing simulation \( I_{T,RT}(x, y) \) for (c) the initial surface and (d) the optimized surface.

TABLE III. Comparison of \( \Delta I_T \) for example “lake” with an astigmatic input wavefront

| \( z^x(x, y) \) | \( z(x, y) \) |
|----------------|----------------|
| \( \text{rms}_{\Delta I_T} \) | \( 3.2072 \times 10^{-6} \) |
| \( \text{pv}_{\Delta I_T} \) | \( 4.2670 \times 10^{-5} \) |
| \( \text{corr}_{\Delta I_T} \) | \( 0.9252 \) |
| \( \eta \) | \( 99.23\% \) |

VII. CONCLUSION

A single freeform surface design algorithm for prescribed input wavefronts and complex irradiance distributions was introduced. This was done by deriving the PDE system \( I_0 \) and Monge-Ampère equation \( I_1 \), which governs the single freeform design process for ideal sources in the geometrical optics limit, from the energy conservation and ray tracing equations in cartesian coordinates. After introducing a possible way to calculate an appropriate initial iterate in section [IV] the PDE system is then solved for the unknown functions \( u(x, y) \) and \( z(x, y) \) by a finite difference discretization and a root finding of the resulting nonlinear equation system. The design examples in section [VI] show that the presented design algorithm extends the results of Ref. [19] for collimated beam shaping to the nonparaxial regime and that the algorithm is applicable to point sources and more general input wavefronts. This might be helpful in applications in which one has to deal with the shaping of e.g. astigmatic wavefronts produced by laser diodes or simplify manufacturing processes. Also the relaxation of physical situations, e.g. for point sources with large opening angles, might be reached by a predefined entrance surface of a freeform lens.

Despite the presented results, there are still interesting problems left for future research and improvements, respectively. For example, a mathematical rigorous justification of the initial iterate construction method presented in section [IV] is of interest, since there is no guarantee that the proposed method will always give a suitable initial iterate, despite the arguments given in section [IV]. Also the numerical extension of the presented algorithm to more complicated boundaries of the irradiance distributions would help to increase the number of possible applications or energy efficiency, respectively. This requires the numerical implementation of the transport boundary conditions, which are realized for irradiance distributions with square boundary shapes by Eq. [19], for more complex boundary shapes and will be investigated in the future. As mentioned in section [III] a numerical investigation of the Monge-Ampère equation in Eq. [11] is of interest. Despite the listed additional complications, the reduced number of design variables can benefit the speed of the root-finding process. Another interesting future goal is the generalization of the presented design algorithm to wavefront and irradiance control with double freeform surfaces by combining it with the results presented in Ref. [19] and the investigation of the manufacturability of the calculated freeform surfaces.

FUNDING

We acknowledge support from the German Federal Ministry of Education and Research (BMBF) under contract number 031PT609X within the project KoSimO.

ACKNOWLEDGMENTS

The authors want to thank Norman Worku for providing and helping them with the MatLightTracer and Johannes Stock for comments on the manuscript.
[1] D. Michaelis, D. Schreiber, and A. Bräuuer, “Cartesian oval representation of freeform optics in illumination systems,” Opt. Lett. 36(6), 918–920 (2011).
[2] V.I. Oliker, “Mathematical aspects of design of beam shaping surfaces in geometrical optics,” in Trends in Nonlinear Analysis, M. Kirkilionis, S. Kromker, R. Ranacher, and F. Tomi, eds. (Springer-Verlag, 2003), pp. 193-222.
[3] 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).
[4] V.I. Oliker, “Controlling light with freeform multifocal lens designed with supporting quadric method(SQM),” Opt. Express 25(4), A58-A72 (2017)
[5] H. Ries and J. Muschaweck, “Tailored freeform optical surfaces,” J. Opt. Soc. Am. A 19(3), 590–595 (2002).
[6] 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).
[7] 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)
[8] 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).
[9] C.R. Prins, J.H.M. ten Thije Boonkamp, 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).
[10] V. Oliker, J. Rubinstein, and G. Wolansky, “Ray mapping and illumination control,” J. Photon. Energy 3(1), 035599 (2013).
[11] A. Brunton, A. Bäuerle, P. Loosen, and R. Wester, “High resolution irradiance tailoring using multiple freeform surfaces,” Opt. Express 21(9), 10563–10571 (2013).
[12] 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).
[13] C. Bösel and H. Gross, “Ray mapping approach for the efficient design of continuous freeform surfaces,” Opt. Express 24(12), 14271–14282 (2016).
[14] D. Ma, Z. Feng, and R. Liang, “Tailoring freeform illumination optics in a double-pole coordinate system,” Appl. Opt. 54(9), 2395-2399 (2015)
[15] C. Gannon and R. Liang, “Ray mapping with surface information for freeform illumination design,” Opt. Express 25(8), 9426-9434 (2017).
[16] V. Oliker, “Differential equations for design of a freeform single lens with prescribed irradiance properties,” Opt. Eng. 53(3), 031302 (2014).
[17] J. Rubinstein, G. Wolansky, “Reconstruction of Optical Surfaces from Ray Data,” Optical Review 8(4), 281-283 (2001).
[18] J. Rubinstein, G. Wolansky, “A variational principle in optics,” J. Opt. Soc. Am. A 21(11), 2164-2172 (2004).
[19] C. Bösel 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).
[20] 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).
[21] https://de.mathworks.com/matlabcentral/fileexchange/46223-
regularizedata3d