



## OPEN ACCESS

## EDITED BY

Tong Yang,  
Beijing Institute of Technology, China

## REVIEWED BY

Huazhong Xiang,  
University of Shanghai for Science and  
Technology, China  
Haisong Tang,  
Beijing Institute of Technology, China

## \*CORRESPONDENCE

J. H. M. ten Thije Boonkkamp,  
✉ j.h.m.tenthijeboonkkamp@tue.nl

RECEIVED 28 October 2024

ACCEPTED 24 January 2025

PUBLISHED 09 May 2025

## CITATION

ten Thije Boonkkamp JHM, Mitra K,  
Anthonissen MJH, Kusch L, Braam PA and  
IJzerman WL (2025) Inverse freeform design  
in non-imaging optics: Hamilton's theory of  
geometrical optics, optimal transport, and  
least-squares solvers.  
*Front. Phys.* 13:1518660.  
doi: 10.3389/fphy.2025.1518660

## COPYRIGHT

© 2025 Ten Thije Boonkkamp, Mitra,  
Anthonissen, Kusch, Braam and IJzerman.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Inverse freeform design in non-imaging optics: Hamilton's theory of geometrical optics, optimal transport, and least-squares solvers

J. H. M. ten Thije Boonkkamp<sup>1\*</sup>, K. Mitra<sup>1</sup>, M. J. H. Anthonissen<sup>1</sup>,  
L. Kusch<sup>1</sup>, P. A. Braam<sup>1</sup> and W. L. IJzerman<sup>1,2</sup>

<sup>1</sup>Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherlands, <sup>2</sup>Signify Research, Eindhoven, Netherlands

We present a systematic derivation of three mathematical models of increasing complexity for optical design, based on Hamilton's characteristic functions and conservation of luminous flux, and briefly explain the connection with the mathematical theory of optimal transport. We outline several iterative least-squares solvers for our models and demonstrate their performance for a few challenging problems.

## KEYWORDS

geometrical optics, Hamilton's characteristic functions, optimal transport, optical map, conservation of luminous flux, inverse methods, least-squares methods

## 1 Introduction

Nowadays, traditional incandescent light bulbs have largely been replaced by LED lamps, mainly because LEDs are highly energy efficient and last for long duration. However, to create a desired light output with LEDs, optical systems are required. Therefore, in the illumination optics industry, a lot of research is carried out on the design of optical systems for a myriad of applications.

The industry standard in optical design is (quasi-)Monte Carlo ray tracing. These are forward methods, which compute the target light distribution for a given source, typically LED, and a *given* optical system, tracing a large collection of rays from the source to target. Convergence of ray tracing is slow and requires millions or, maybe, even billions of rays to accurately compute the target distribution; see ([1], Chapter3). Moreover, to realize the *desired* target distribution, these methods have to be embedded in an iterative procedure, adjusting system parameters. More specifically to update these parameters, sophisticated design methods combine an iterative gradient-based optimization method, to minimize a suitably chosen merit function, with differentiable ray tracing to compute the required gradient. These methods are efficient and broadly applicable; see e.g., [31, 32]

A promising alternative are inverse methods. These methods compute the shape/location of the optical surface(s) in one go, given a source distribution and a *desired* target distribution. The surfaces are referred to as freeform since they have no imposed symmetries. Mathematical models for inverse methods are based on the principles of geometrical optics and conservation of luminous flux. Our models contain the following: a geometrical equation describing the shape/location of the optical surface(s), an (implicit)

relation between the location of an optical surface and the optical map, which expresses target coordinates as a function of source coordinates, a matrix equation for the optical map, a conservation law of luminous flux, and a constraint on the luminous flux. The geometrical equation can be derived evaluating one of Hamilton's characteristic functions and has multiple solutions. The optical map satisfies the luminous flux balance, which is a Jacobian equation, and a matrix equation, which has to be supplemented with the unconventional transport boundary condition. Our models are restricted to zero-étendue sources; more specifically, we consider only planar sources emitting a parallel beam of light or point sources.

In mathematics, the abstract theory of optimal transport is the correct framework to formulate optical design problems. The subject of optimal transport theory is to compute a transport plan (mapping) that transforms a given density function into another one, minimizing the transportation cost and subject to a(n) (integral) conservation law; see [2]. One of the governing equations is precisely the geometrical equation, containing a so-called cost function in the right hand side. The transportation cost is then a cost functional (an integral of the cost function). The analogy with optical design is obvious, the two density functions correspond to the source and target distributions, the transport plan to the optical map, the cost functional can be associated with the optical path length, and the conservation law applies to the luminous flux.

Based on the analogy with optimal transport theory, we can distinguish three mathematical models of increasing complexity. First, in the simplest model, the geometrical equation contains a quadratic cost function, and the Jacobian equation reduces to the standard Monge–Ampère (MA) equation, which is a second-order, non-linear, elliptic partial differential equation (PDE) [3], defining the location of an optical surface. In the next model, the cost function is no longer quadratic and the Jacobian equation corresponds to a generalized Monge–Ampère (GMA) equation. Finally, the most complicated model does not allow a description in terms of a cost function anymore; instead, a so-called generating function is required. The Jacobian equation now corresponds to a generated Jacobian (GJ) equation.

The procedure to derive our mathematical models consists of the following steps:

1. By applying principles of geometrical optics, derive the geometrical equation.
2. Formulate the luminous flux balance, which is a Jacobian equation for the optical map.
3. Based on optimal transport theory, select a uniquely defined solution from the geometrical equation.
4. Derive the necessary (zero-gradient) condition for this solution to define a unique relation between the location of an optical surface and the optical map.
5. Derive from the zero-gradient condition the matrix equation for the Jacobi matrix of the optical map.
6. Combine the luminous flux balance with the matrix equation to derive a constraint on the luminous flux.

There are many numerical methods available for inverse freeform optical design; for a detailed survey, see [4], Chapter 5. Without trying to be complete, a possible categorization is as follows: first, numerical methods for (G)MA equations governing the shape/location of an optical surface; second, optimization

methods; third, geometrical construction methods; and fourth, ray mapping methods. In the first category are methods that directly solve the second-order PDE of MA type by discretization and subsequent iteration for the resulting algebraic system. Examples of these are reported in papers by Froese [5], using a wide-stencil finite difference scheme, Benamou et al. [6], using finite differences, Brix et al. [35], using B-spline collocation, and Kawecki et al. [7], using finite elements, to just name a few examples. Methods in the second category are based on optimal transport theory and determine the optical surface(s) by minimizing the cost functional subject to conservation of luminous flux, see, e.g., [33, 34]. Next, based on the work of Caffarelli and Oliker [8], geometrical construction methods for optical surfaces with paraboloids have been investigated by Kitagawa et al. [9] and Mérigot and Thibert [10] for the GMA equation, and by Galouët et al. [11] for the GJ equation. Finally, in ray mapping methods, the computation of the optical surface(s) is split into two parts, i.e., first, an estimate of the optical map is computed from an optimal transport problem, and subsequently, the optical surface is computed employing the law of reflection/refraction; see e.g., [36–38].

Our methods are of the first category. The publication that inspired us to develop our numerical solvers is by Caboussat et al. [12]. In this paper, the authors present a least-squares method to solve the Dirichlet problem for the standard MA equation. Prins adjusted their method to include the transport boundary condition; see [13]. In a series of papers, Yadav and Romijn extended the method for the GMA equation, corresponding to a non-quadratic cost function; see [14], [15, 16]. Romijn further extended our least-squares methods to deal with the GJ equation; see [17, 18]. Specifically, for the cost function models, we have developed a two-stage algorithm, i.e., we first compute the optical map and subsequently the shape/location of the optical surface(s). For both stages, we apply a least-squares method. For the generating function model, the optical map and the optical surface(s) have to be computed simultaneously, also in a least-squares fashion.

In this paper, we present a systematic derivation of all models and outline numerical simulation methods. The content is then as follows. First, in Section 2, we introduce Hamilton's characteristic functions needed to derive the geometrical equations. In Section 3, we give an outline of optimal transport theory, needed to select a uniquely defined solution of the geometrical equation. A generic conservation law for luminous flux is presented in Section 4. Next, in Section 5, we apply the results of the previous sections to present a hierarchy of models based on three example systems. A concise description of our numerical methods is presented in Section 6. To demonstrate the performance of our methods, we present a few challenging examples in Section 7. Concluding remarks are given in Section 8, and the paper is concluded with a nomenclature list.

## 2 Hamilton's characteristic functions

In this section, we introduce Hamilton's characteristic functions which are needed for the derivation of the geometrical equations in Section 5; see, e.g., [19], pp. 94–107; [20], Section 4.1 for a more

detailed account. Our starting point is the eikonal equation given by

$$|\nabla\varphi| = n, \tag{1}$$

where  $\varphi = \varphi(\underline{\mathbf{r}})$  is the phase of an electromagnetic wave and  $n = n(\underline{\mathbf{r}})$  is the refractive index field as functions of the (three-dimensional) position vector (or coordinates)  $\underline{\mathbf{r}} \in \mathbb{R}^3$  of a typical point. A surface  $\varphi(\underline{\mathbf{r}}) = \text{Const}$  is a wavefront. We denote three-dimensional vectors in bold font with an underbar ( $\underline{\quad}$ ), to distinguish them from two-dimensional vectors, which are only written in bold font. Equation 1 describes free propagation of light waves and no longer holds when a light wave is reflected or refracted at an optical surface. The eikonal Equation 1 is a first-order non-linear PDE, which has the characteristic strip equations

$$\frac{d\underline{\mathbf{r}}}{ds} = \frac{1}{n}\nabla\varphi, \quad \frac{d\varphi}{ds} = n, \quad \frac{d}{ds}(\nabla\varphi) = \nabla n. \tag{2}$$

See, e.g., Kevorkian [21]. The curves  $C:\underline{\mathbf{r}} = \underline{\mathbf{r}}(s)$ , parameterized by the arc length  $s$ , are the characteristics of (1). The system of ordinary differential equations in (2) determines the location of the characteristics and the solution along the characteristics, from which the solution of (1) can be reconstructed. From the first equation in (2), we conclude that the characteristics are orthogonal trajectories to the wavefronts, i.e., characteristics are light rays; see, e.g., [20], p. 114.

Hamilton's characteristic functions are related to the optical path length of a ray, which we, therefore, introduce next. We assume that  $n(\underline{\mathbf{r}})$  is piecewise continuous, with jumps at lens surfaces; consequently, rays are continuous and consist of piecewise differentiable curve segments; cf. Equation 2. In this setting, we define the optical path length (OPL) between the points  $P_1$  (position vector  $\underline{\mathbf{r}} = \underline{\mathbf{r}}_1$ ) and  $P_2$  (position vector  $\underline{\mathbf{r}} = \underline{\mathbf{r}}_2$ ), connected by a light ray along the curve  $C$ , as

$$[P_1, P_2] = d_{\text{opt}}(\underline{\mathbf{r}}_1, \underline{\mathbf{r}}_2) := \int_C n(\underline{\mathbf{r}}(s)) ds. \tag{3}$$

A consequence of Fermat's principle is that the optical path length  $[P_1, P_2]$  is stationary with respect to variations in the curve  $C$ ; see, e.g., [20], p. 740.

Consider an optical system consisting of a source, located in a plane  $z = z_s$  (source plane), emitting light which via a sequence of reflectors and/or lenses arrives at a plane  $z = z_t$  (target plane). The  $z$ -axis is referred to as the optical axis. In this paper, we use the subscripts  $s$  and  $t$  for variables/vectors to denote their value at the source and target, respectively. Let  $\underline{\hat{\mathbf{v}}} = (v_1 \ v_2 \ v_3)^T = d\underline{\mathbf{r}}/ds$  denote the unit tangent vector of a ray; thus,  $\nabla\varphi = n\underline{\hat{\mathbf{v}}}$ ; cf. Equation 2. Unit vectors are denoted with a hat ( $\hat{\quad}$ ). Let  $Q_s$  (position vector  $\underline{\mathbf{r}} = \underline{\mathbf{r}}_s$ ) and  $Q_t$  (position vector  $\underline{\mathbf{r}} = \underline{\mathbf{r}}_t$ ) be two points on the source and target plane, respectively; then, from Equation 3, we obtain

$$[Q_s, Q_t] = d_{\text{opt}}(\underline{\mathbf{r}}_s, \underline{\mathbf{r}}_t) = \int_C n\underline{\hat{\mathbf{v}}} \cdot \underline{\hat{\mathbf{v}}} ds = \int_C \nabla\varphi \cdot d\underline{\mathbf{r}} = \varphi(\underline{\mathbf{r}}_t) - \varphi(\underline{\mathbf{r}}_s). \tag{4}$$

Note that reversing the direction of integration yields the symmetry property of  $d_{\text{opt}}$ , i.e.,

$$\begin{aligned} d_{\text{opt}}(\underline{\mathbf{r}}_t, \underline{\mathbf{r}}_s) &= \int_{-C} (-n\underline{\hat{\mathbf{v}}}) \cdot (-\underline{\hat{\mathbf{v}}}) ds \\ &= \int_{-C} -\nabla\varphi \cdot d\underline{\mathbf{r}} \\ &= -\varphi(\underline{\mathbf{r}}_s) + \varphi(\underline{\mathbf{r}}_t) \\ &= d_{\text{opt}}(\underline{\mathbf{r}}_s, \underline{\mathbf{r}}_t) \end{aligned} \tag{5}$$

as one would expect from the definition in Equation 3. Observe that  $d_{\text{opt}}$  is a metric since  $d_{\text{opt}}(\underline{\mathbf{r}}_s, \underline{\mathbf{r}}_t) \geq 0$ , with equality only if  $\underline{\mathbf{r}}_s = \underline{\mathbf{r}}_t$ , and it satisfies the triangle inequality due to Fermat's principle. The symmetry of  $d_{\text{opt}}$  is stated in Equation 5.

For ease of presentation, we use the shorthand notation  $\underline{\hat{\mathbf{v}}}_s = \underline{\hat{\mathbf{s}}}$  and  $\underline{\hat{\mathbf{v}}}_t = \underline{\hat{\mathbf{t}}}$ . Straightforward differentiation of the expression for  $d_{\text{opt}}(\underline{\mathbf{r}}_s, \underline{\mathbf{r}}_t)$  in Equation 4 gives

$$\nabla_s d_{\text{opt}}(\underline{\mathbf{r}}_s, \underline{\mathbf{r}}_t) = -\nabla\varphi(\underline{\mathbf{r}}_s) = -n_s \underline{\hat{\mathbf{s}}}, \quad \nabla_t d_{\text{opt}}(\underline{\mathbf{r}}_s, \underline{\mathbf{r}}_t) = \nabla\varphi(\underline{\mathbf{r}}_t) = n_t \underline{\hat{\mathbf{t}}}, \tag{6}$$

where  $\nabla_s$  denotes the gradient with respect to the source coordinates  $\underline{\mathbf{r}}_s$  and likewise for  $\nabla_t$ . From these relations, we readily conclude that

$$|\nabla_s d_{\text{opt}}| = n_s, \quad |\nabla_t d_{\text{opt}}| = n_t,$$

i.e.,  $d_{\text{opt}}$  satisfies the eikonal equation in both source and target coordinates.

A ray is completely determined by its position and direction coordinates. Position coordinates  $\underline{\mathbf{q}} \in \mathbb{R}^2$  are defined as the projection on a plane  $z = \text{Const}$  of the position vector  $\underline{\mathbf{r}}$  of a typical point on the ray. Position coordinates in the source and target planes are denoted by  $\underline{\mathbf{q}}_s$  and  $\underline{\mathbf{q}}_t$ , respectively. Thus,  $Q_s$ , with position vector  $\underline{\mathbf{r}}_s = (\underline{\mathbf{q}}_s, z_s)$ , will denote the point at which a light ray is emitted from the source plane, and  $Q_t$ , with the position vector  $\underline{\mathbf{r}}_t = (\underline{\mathbf{q}}_t, z_t)$ , as the point at which it arrives at the target plane. Next, selecting the first two components from the equation  $n \frac{d\underline{\mathbf{r}}}{ds} = n\underline{\hat{\mathbf{v}}} = \nabla\varphi$ , we introduce the momentum vector  $\underline{\mathbf{p}} \in \mathbb{R}^2$  as

$$\underline{\mathbf{p}} := n \frac{d\underline{\mathbf{q}}}{ds} = n \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial\varphi}{\partial q_1} \\ \frac{\partial\varphi}{\partial q_2} \end{pmatrix} =: \frac{\partial\varphi}{\partial \underline{\mathbf{q}}}. \tag{7}$$

Thus, the momentum vector is the projection of  $\nabla\varphi(\underline{\mathbf{x}})$  on a plane  $z = \text{Const}$  and determines the direction of a light ray. Momentum vectors at the source and target are denoted with  $\underline{\mathbf{p}}_s$  and  $\underline{\mathbf{p}}_t$ , respectively. The variables  $\underline{\mathbf{q}} = \underline{\mathbf{q}}(z)$  and  $\underline{\mathbf{p}} = \underline{\mathbf{p}}(z)$  are referred to as phase-space coordinates and together constitute phase space.

In the remainder of this paper, we assume that  $n(\underline{\mathbf{r}})$  is piecewise constant, with possible discontinuities at optical interfaces, and consequently, a light ray consists of piecewise straight line segments. Below, we define four characteristic functions which relate the optical path length to phase-space coordinates at the source and target.

## Point characteristic

As a function of the position coordinates  $\underline{\mathbf{q}}_s$  and  $\underline{\mathbf{q}}_t$ , the point characteristic  $V$  is the optical path length of the ray connecting the points  $Q_s$  and  $Q_t$ , i.e.,

$$V(\underline{\mathbf{q}}_s, \underline{\mathbf{q}}_t) := d_{\text{opt}}((\underline{\mathbf{q}}_s, z_s), (\underline{\mathbf{q}}_t, z_t)) = [Q_s, Q_t]. \tag{8}$$

From Equation 4, we conclude that  $V(\underline{\mathbf{q}}_s, \underline{\mathbf{q}}_t) = \varphi(\underline{\mathbf{q}}_t, z_t) - \varphi(\underline{\mathbf{q}}_s, z_s)$ . Combining this relation with the equations in Equation 6, selecting the first two components in both equations, we find for the directional derivatives  $\frac{\partial V}{\partial \underline{\mathbf{q}}_s}$  and  $\frac{\partial V}{\partial \underline{\mathbf{q}}_t}$  the following expressions:

$$\frac{\partial V}{\partial \underline{\mathbf{q}}_s} = -n_s \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = -\underline{\mathbf{p}}_s, \quad \frac{\partial V}{\partial \underline{\mathbf{q}}_t} = n_t \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \underline{\mathbf{p}}_t; \tag{9}$$

cf. Equation 7. Typically, Equations 8, 9 are used to derive a geometrical equation for optical systems described in terms of the position coordinates  $\mathbf{q}_s$  and  $\mathbf{q}_t$ . An example is the parallel-to-near-field reflector, discussed in Section 5.3.

### Mixed characteristic of the first kind

Let a light ray connect the points  $Q_s(\mathbf{q}_s, z_s)$  and  $Q_t(\mathbf{q}_t, z_t)$  with momentum  $\mathbf{p}_s$  and  $\mathbf{p}_t$ , respectively; then, the mixed characteristic of the first kind  $W$  is defined as

$$W(\mathbf{q}_s, \mathbf{p}_t) := V(\mathbf{q}_s, \mathbf{q}_t) - \mathbf{q}_t \cdot \mathbf{p}_t. \tag{10}$$

Straightforward differentiation yields  $\partial W / \partial \mathbf{p}_s = \mathbf{0}$  and  $\partial W / \partial \mathbf{q}_t = \partial V / \partial \mathbf{q}_t - \mathbf{p}_t = \mathbf{0}$ ; cf. Equation 9, second relation. Consequently,  $W = W(\mathbf{q}_s, \mathbf{p}_t)$  indeed. Moreover, we readily verify that

$$\frac{\partial W}{\partial \mathbf{q}_s} = -\mathbf{p}_s, \quad \frac{\partial W}{\partial \mathbf{p}_t} = -\mathbf{q}_t.$$

$W$  can be interpreted as a modification of  $[Q_s, Q_t]$ , which we show as follows. Assume that  $z_t > z_s$  and define the point  $P_t$  as the intersection of the ray segment arriving at  $Q_t$ , parameterized by  $\mathbf{r}(\lambda) = \mathbf{r}_t + \lambda \hat{\mathbf{t}}$ , and the plane passing through the origin  $O_t(0, z_t)$  of the target plane (position vector  $\mathbf{o}_t$ ) perpendicular to this ray, given by the normal equation  $(\mathbf{r} - \mathbf{o}_t) \cdot \hat{\mathbf{t}} = 0$ ; see Figure 1. The intersection point is given by  $\lambda = \lambda(P_t) = -(\mathbf{r}_t - \mathbf{o}_t) \cdot \hat{\mathbf{t}}$ . Since the third component of  $\mathbf{r}_t - \mathbf{o}_t$  vanishes, this can be rewritten as  $\lambda(P_t) = -\mathbf{q}_t \cdot \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = -\mathbf{q}_t \cdot \mathbf{p}_t / n_t$ ; cf. Equation 7 or Equation 9. For the  $z$ -component of  $P_t$ , we have  $\delta z_t := z(P_t) - z_t = \lambda(P_t)t_3$  with  $t_3 > 0$  since the ray is directed toward the target plane ( $z_t > z_s$ ). From this, we conclude that  $\lambda(P_t) > 0$  if  $\delta z_t > 0$ , i.e.,  $P_t$  is located behind the target plane as seen from the source, and  $\lambda(P_t) < 0$  otherwise. Obviously, the distance  $d(Q_t, P_t) = \pm \lambda(P_t)$  since  $\hat{\mathbf{t}}$  is a unit vector. If  $\delta z_t > 0$ , then  $\lambda(P_t) > 0$ , implying that  $d(Q_t, P_t) = \lambda(P_t)$  and  $\mathbf{q}_t \cdot \mathbf{p}_t = -n_t d(Q_t, P_t) = -[Q_t, P_t]$ ; cf. Equation 3. Analogously, if  $\delta z_t < 0$ , then  $\mathbf{q}_t \cdot \mathbf{p}_t = [Q_t, P_t]$ . Finally, we can easily verify that  $\sigma_t := \text{sgn}(\mathbf{q}_t \cdot \mathbf{p}_t) = -\text{sgn}(\delta z_t)$ ; thus,  $W$  can be interpreted as

$$W(\mathbf{q}_s, \mathbf{p}_t) = [Q_s, Q_t] - \sigma_t [Q_t, P_t].$$

The characteristic function  $W$  is convenient to describe optical systems in terms of  $\mathbf{q}_s$  and  $\mathbf{p}_t$ . An example is the parallel-to-far-field reflector, as described in Section 5.1.

### Mixed characteristic of the second kind

Under the same setting, the mixed characteristic of the second kind  $W^*$  is defined as

$$W^*(\mathbf{p}_s, \mathbf{q}_t) := V(\mathbf{q}_s, \mathbf{q}_t) + \mathbf{q}_s \cdot \mathbf{p}_s.$$

In this case, using the first relation in Equation 9, we readily verify that  $\partial W^* / \partial \mathbf{q}_s = \partial V / \partial \mathbf{q}_s + \mathbf{p}_s = \mathbf{0}$  and  $\partial W^* / \partial \mathbf{p}_t = \mathbf{0}$ , so  $W^*$  is a function of  $\mathbf{p}_s$  and  $\mathbf{q}_t$  indeed. Furthermore, straightforward differentiation yields

$$\frac{\partial W^*}{\partial \mathbf{p}_s} = \mathbf{q}_s, \quad \frac{\partial W^*}{\partial \mathbf{q}_t} = \mathbf{p}_t.$$

In addition,  $W^*$  can be interpreted as a modified OPL. We assume again that  $z_t > z_s$  and introduce the point  $P_s$  as the intersection of the ray segment emitted from  $Q_s$ , parameterized by  $\mathbf{r}(\lambda) = \mathbf{r}_s + \lambda \hat{\mathbf{s}}$ , and the plane passing through the origin  $O_s(0, z_s)$  of the source plane (position vector  $\mathbf{o}_s$ ) perpendicular to this ray, given by  $(\mathbf{r} - \mathbf{o}_s) \cdot \hat{\mathbf{s}} = 0$ ; see Figure 1. Analogous to the previous derivation, we obtain  $\lambda = \lambda(P_s) = -(\mathbf{r}_s - \mathbf{o}_s) \cdot \hat{\mathbf{s}} = -\mathbf{q}_s \cdot \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = -\mathbf{q}_s \cdot \mathbf{p}_s / n_s$ . The  $z$ -component of  $P_s$  is given by  $z(P_s) = z_s + \lambda(P_s)s_3$  with  $s_3 > 0$ . Thus, if  $\delta z_s := z(P_s) - z_s > 0$ , then  $\lambda(P_s) > 0$ ; otherwise,  $\lambda(P_s) < 0$ . Following the same line of reasoning as before, we find that  $\mathbf{q}_s \cdot \mathbf{p}_s = \sigma_s [Q_s, P_s]$  with  $\sigma_s = \text{sgn}(\mathbf{q}_s \cdot \mathbf{p}_s)$ . Therefore, the second mixed characteristic can be interpreted as

$$W^*(\mathbf{p}_s, \mathbf{q}_t) = [Q_s, Q_t] + \sigma_s [Q_s, P_s].$$

Note that  $W^*(-\mathbf{p}_t, \mathbf{q}_s) = W(\mathbf{q}_s, \mathbf{p}_t)$ , i.e., the mixed characteristic of the second kind of the ray in reversed direction equals the mixed characteristic of the first kind of the original ray. The function  $W^*$  can be used to characterize optical systems in terms of  $\mathbf{p}_s$  and  $\mathbf{q}_t$ . An example is the point-to-near-field reflector.

### Angular characteristic

The angular characteristic  $T$  is another modification of  $[Q_s, Q_t]$  and is a function of the momenta  $\mathbf{p}_s$  and  $\mathbf{p}_t$ . Its definition and interpretation (for  $z_t > z_s$ ) are given by

$$T(\mathbf{p}_s, \mathbf{p}_t) := V(\mathbf{q}_s, \mathbf{q}_t) + \mathbf{q}_s \cdot \mathbf{p}_s - \mathbf{q}_t \cdot \mathbf{p}_t = [Q_s, Q_t] + \sigma_s [Q_s, P_s] - \sigma_t [Q_t, P_t]. \tag{11}$$

From Equations 9, 11, we readily verify that  $\partial T / \partial \mathbf{q}_s = \partial V / \partial \mathbf{q}_s + \mathbf{p}_s = \mathbf{0}$  and  $\partial T / \partial \mathbf{q}_t = \partial V / \partial \mathbf{q}_t - \mathbf{p}_t = \mathbf{0}$ , implying that  $T = T(\mathbf{p}_s, \mathbf{p}_t)$ . Moreover, it is evident that

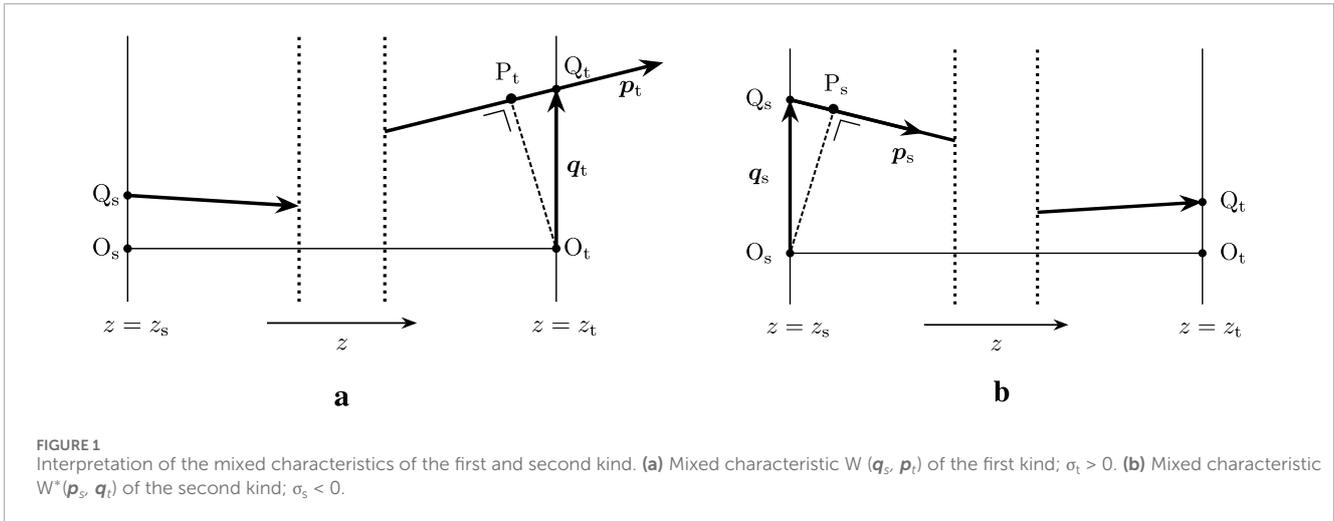
$$\frac{\partial T}{\partial \mathbf{p}_s} = \mathbf{q}_s, \quad \frac{\partial T}{\partial \mathbf{p}_t} = -\mathbf{q}_t.$$

Note that  $T(-\mathbf{p}_t, -\mathbf{p}_s) = T(\mathbf{p}_s, \mathbf{p}_t)$ , i.e., the angular characteristic is invariant if the direction of the ray is reversed. This function is used to describe optical systems in terms of  $\mathbf{p}_s$  and  $\mathbf{p}_t$ , such as the point-to-far-field lens; see Section 5.2.

## 3 Geometrical description of optical systems: an optimal transport point of view

To derive a geometrical description of an optical system, i.e., an algebraic relation defining the location and shape of the optical surface(s), we have to evaluate one of the four Hamilton's characteristic functions. However, this geometrical equation does not have a unique solution. In this section, we discuss how to select a uniquely defined solution from this equation. Subsequently, we derive an equation for the optical map.

First, we introduce some notation. We denote with  $\mathcal{S}$  and  $\mathcal{T}$  the (physical) source and target domain, respectively. The source and target domain are parameterized by  $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^2$  and  $\mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^2$ , respectively, i.e.,  $\mathcal{X}$  and  $\mathcal{Y}$  are the parameter domains for the source



and target. We assume that  $\mathcal{X}$  and  $\mathcal{Y}$  are compact, i.e., bounded and closed, and consequently, any continuous function defined on one of these domains assumes a minimum and a maximum. The key in the following discussion is the optical map  $\mathbf{m}:\mathcal{X} \rightarrow \mathcal{Y}$ , defining how a point on the source domain (specified by  $\mathbf{x}$ ) is connected via a light ray with a point on the target domain (specified by  $\mathbf{y} = \mathbf{m}(\mathbf{x})$ ). Given source and target domains, the optical map has to satisfy the condition  $\mathcal{Y} = \mathbf{m}(\mathcal{X})$ , i.e.,  $\mathcal{Y}$  is the image of  $\mathcal{X}$  under the mapping  $\mathbf{m}$ .

We consider optical systems having one or two freeform optical surfaces, i.e., surfaces without any symmetry. Consider an arbitrary ray connecting the source and target, for which  $\mathbf{y} = \mathbf{m}(\mathbf{x})$ . Evaluation of the appropriate Hamilton's characteristic function for such a ray leads to either one of the two algebraic equations:

$$u_1(\mathbf{x}) + u_2(\mathbf{y}) = c(\mathbf{x}, \mathbf{y}), \tag{12a}$$

$$u_2(\mathbf{y}) = H(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})), \tag{12b}$$

which we refer to as geometrical equations, where  $u_1(\mathbf{x})$  defines the location of (one of) the optical surface(s) and  $u_2(\mathbf{y})$  is either an auxiliary variable or defines the location of the other surface. We like to emphasize that in Equation 12, source and target coordinates are related via  $\mathbf{y} = \mathbf{m}(\mathbf{x})$ . In Section 5, we will derive geometrical equations for several optical systems. Note that in Equation 12a, source and target coordinates are separated in the left-hand side. This equation occurs in the theory of optimal transport, where  $u_1(\mathbf{x})$  and  $u_2(\mathbf{y})$  are referred to as Kantorovich potentials and  $c(\mathbf{x}, \mathbf{y})$  as the cost function; see, e.g., [2] for a rigorous mathematical account. In Equation 12b, source and target coordinates are no longer separated. It will turn out that for fixed  $\mathbf{x}$  and  $\mathbf{y}$ , the function  $H(\mathbf{x}, \mathbf{y}, \cdot)$  has a unique inverse  $H^{-1}(\mathbf{x}, \mathbf{y}, \cdot) = G(\mathbf{x}, \mathbf{y}, \cdot)$ , referred to as the generating function; see [22]. Thus, the following holds:

$$\forall \mathbf{x} \in \mathcal{X} \quad \forall \mathbf{y} \in \mathcal{Y} \quad (u_2(\mathbf{y}) = H(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})) \iff u_1(\mathbf{x}) = G(\mathbf{x}, \mathbf{y}, u_2(\mathbf{y}))).$$

Obviously, Equation 12a is a special case of Equation 12b if we choose  $H(\mathbf{x}, \mathbf{y}, w) = c(\mathbf{x}, \mathbf{y}) - w$ . Nevertheless, for the sake of completeness, we discuss both cases separately.

### c-convex analysis

Equation 12a has multiple solution pairs  $(u_1(\mathbf{x}), u_2(\mathbf{y}))$ . From the theory of optimal transport, we know that a possible solution of Equation 12a is given by

$$\forall \mathbf{x} \in \mathcal{X} \quad u_1(\mathbf{x}) = \max_{\mathbf{y} \in \mathcal{Y}} (c(\mathbf{x}, \mathbf{y}) - u_2(\mathbf{y})), \tag{13a}$$

$$\forall \mathbf{y} \in \mathcal{Y} \quad u_2(\mathbf{y}) = \max_{\mathbf{x} \in \mathcal{X}} (c(\mathbf{x}, \mathbf{y}) - u_1(\mathbf{x})), \tag{13b}$$

which implicitly defines the optical map as an alternative to a geometrical optics derivation. To be more precise, for any  $\mathbf{x} \in \mathcal{X}$ , its image  $\mathbf{y}^* = \mathbf{m}(\mathbf{x}) \in \mathcal{Y}$  is the maximizer in Equation 13a. In addition, for any  $\mathbf{y} \in \mathcal{Y}$ , we have that  $\mathbf{x}^* \in \mathcal{X}$  such that  $\mathbf{y} = \mathbf{m}(\mathbf{x}^*)$  is the maximizer in Equation 13b, which is possible since  $\mathcal{Y} = \mathbf{m}(\mathcal{X})$ . Under certain conditions, to be specified shortly, the maxima in (13) are unique, which we henceforth assume. The solution in (13) is referred to as a *c*-convex pair; see [23]. The proof of the equivalence of (Equation 13a) and (Equation 13b) proceeds as follows. Suppose, Equation 12a and the expression for  $u_1(\mathbf{x})$  in Equation 13a hold, then we have to derive the expression for  $u_2(\mathbf{y})$  in Equation 13b. From Equation 13a we conclude

$$\forall \mathbf{x} \in \mathcal{X} \quad \forall \mathbf{y} \in \mathcal{Y} \quad u_1(\mathbf{x}) \geq c(\mathbf{x}, \mathbf{y}) - u_2(\mathbf{y}),$$

or equivalently, swapping  $u_1(\mathbf{x})$  and  $u_2(\mathbf{y})$ ,

$$\forall \mathbf{x} \in \mathcal{X} \quad \forall \mathbf{y} \in \mathcal{Y} \quad u_2(\mathbf{y}) \geq c(\mathbf{x}, \mathbf{y}) - u_1(\mathbf{x}).$$

Since the latter inequality holds for all  $\mathbf{x} \in \mathcal{X}$ , we can take the maximum over all  $\mathbf{x} \in \mathcal{X}$  to obtain

$$\forall \mathbf{y} \in \mathcal{Y} \quad u_2(\mathbf{y}) \geq \max_{\mathbf{x} \in \mathcal{X}} (c(\mathbf{x}, \mathbf{y}) - u_1(\mathbf{x})).$$

Recall that  $\mathcal{Y} = \mathbf{m}(\mathcal{X})$ . Let  $\mathbf{y} \in \mathcal{Y}$ , then  $\mathbf{y} = \mathbf{m}(\tilde{\mathbf{x}})$  for some  $\tilde{\mathbf{x}} \in \mathcal{X}$ . From Equation 12a, we conclude that

$$u_2(\mathbf{y}) = c(\tilde{\mathbf{x}}, \mathbf{y}) - u_1(\tilde{\mathbf{x}}) \leq \max_{\mathbf{x} \in \mathcal{X}} (c(\mathbf{x}, \mathbf{y}) - u_1(\mathbf{x})).$$

Combining both inequalities, we obtain the expression for  $u_2(\mathbf{y})$  in Equation 13b and conclude that  $\tilde{\mathbf{x}} = \mathbf{x}^*$  is the (unique)

maximizer of  $c(\mathbf{x}, \mathbf{y}) - u_1(\mathbf{x})$ . Thus, the expression for  $u_1(\mathbf{x})$  in Equation 13a combined with Equation 12a implies the expression for  $u_2(\mathbf{y})$  in Equation 13b. Conversely, given the expression for  $u_2(\mathbf{y})$  in Equation 13b and Equation 12a, we can derive in a similar manner the expression for  $u_1(\mathbf{x})$  defined in Equation 13a.

Alternatively, Equation 12a has a  $c$ -concave solution pair, for which the maxima in Equation 13 are replaced by minima, defining another mapping  $\mathbf{m}$ . In either case, a necessary condition is that  $(\mathbf{x}, \mathbf{m}(\mathbf{x}))$  is a stationary point of  $c(\mathbf{x}, \mathbf{y}) - u_1(\mathbf{x})$  for all  $\mathbf{x} \in \mathcal{X}$ , i.e.,

$$\nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{m}(\mathbf{x})) - \nabla u_1(\mathbf{x}) = \mathbf{0}, \tag{14}$$

where  $\nabla_{\mathbf{x}} c$  is the gradient of  $c$  with respect to the source coordinates  $\mathbf{x}$ . A sufficient condition for the  $c$ -convex pair (max/max solution) in (Equation 13) is that  $D_{\mathbf{xx}} c(\mathbf{x}, \mathbf{m}(\mathbf{x})) - D^2 u_1(\mathbf{x})$  be symmetric negative definite (SND) for all  $\mathbf{x} \in \mathcal{X}$ , where  $D_{\mathbf{xx}} c$  and  $D^2 u_1$  are the Hessian matrices of  $c$  (with respect to  $\mathbf{x}$ ) and  $u_1$ , respectively. The function  $c(\cdot, \mathbf{y}) - u_1$  is then concave, guaranteeing a unique maximum in Equation 13b. Alternatively, for a  $c$ -concave pair (min/min solution), the sufficient condition is that  $D_{\mathbf{xx}} c(\mathbf{x}, \mathbf{m}(\mathbf{x})) - D^2 u_1(\mathbf{x})$  be symmetric positive definite (SPD) for all  $\mathbf{x} \in \mathcal{X}$ . Then,  $c(\cdot, \mathbf{y}) - u_1$  is convex and has a unique minimum.

According to the implicit function theorem, the optical map  $\mathbf{y} = \mathbf{m}(\mathbf{x})$  is well-defined by Equation 14, provided the mixed derivative matrix  $D_{\mathbf{xy}} c = \left( \frac{\partial^2 c}{\partial x_i \partial y_j} \right)$  is regular in  $(\mathbf{x}, \mathbf{m}(\mathbf{x}))$  for all  $\mathbf{x} \in \mathcal{X}$ , the so-called twist condition. However, for many optical systems, the actual computation of  $\mathbf{y} = \mathbf{m}(\mathbf{x})$  from Equation 14 is quite complicated, if not impossible. Therefore, to derive an equation for the optical map, we differentiate the zero-gradient condition in Equation 14 to obtain the equation

$$\begin{aligned} \mathbf{C}(\mathbf{x}, \mathbf{m}(\mathbf{x})) \mathbf{D}\mathbf{m} &= \mathbf{P}(\mathbf{x}, \mathbf{m}(\mathbf{x})), \\ \mathbf{C}(\mathbf{x}, \mathbf{y}) &= D_{\mathbf{xy}} c(\mathbf{x}, \mathbf{y}), \\ \mathbf{P}(\mathbf{x}, \mathbf{y}) &= D^2 u_1(\mathbf{x}) - D_{\mathbf{xx}} c(\mathbf{x}, \mathbf{y}), \end{aligned} \tag{15}$$

where  $\mathbf{D}\mathbf{m} = \left( \frac{\partial m_i}{\partial x_j} \right)$  is the Jacobi matrix of the optical map. We refer to Equation 15 as the matrix-Jacobi equation. Note that the matrix  $\mathbf{P}$  is either SPD, for a  $c$ -convex, or SND for a  $c$ -concave solution pair.

Equation 15 has to be supplemented with the so-called transport boundary condition

$$\mathbf{m}(\partial\mathcal{X}) = \partial\mathcal{Y}, \tag{16}$$

stating that the boundary of the source domain  $\partial\mathcal{X}$  is mapped to the boundary of the target domain  $\partial\mathcal{Y}$ . This condition is a consequence of the edge-ray principle of geometrical optics [24] and guarantees that all light emitted from the source arrives at the target.

For some basic optical systems  $c(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ , implying that  $\mathbf{C} = \mathbf{I}$  and  $D_{\mathbf{xx}} c = \mathbf{O}$ ; see Section 5.1. The  $c$ -convex solution pair in (13a) reduces to a conjugate pair of Legendre–Fenchel transforms, see e.g., [4], for which the necessary condition (14) reduces to  $\mathbf{m}(\mathbf{x}) - \nabla u_1(\mathbf{x}) = \mathbf{0}$ . The sufficient condition for the  $c$ -convex solution pair is that  $D^2 u_1(\mathbf{x})$  is SPD for all  $\mathbf{x} \in \mathcal{X}$ , i.e.,  $u_1(\mathbf{x})$  is a convex function. Likewise, for the  $c$ -concave solution pair,  $D^2 u_1(\mathbf{x})$  has to be SND for all  $\mathbf{x} \in \mathcal{X}$ , and  $u_1(\mathbf{x})$  is then concave. In both cases,  $c(\cdot, \mathbf{y}) - u_1$  has a unique extremum.

In our numerical algorithm, the matrix-Jacobi (Equation 15), subject to the transport boundary condition (16), is employed to compute the optical map, and subsequently,  $u_1(\mathbf{x})$  is reconstructed from Equation 14. If needed,  $u_2(\mathbf{y})$  is computed from Equation 12a. From  $u_1(\mathbf{x})$  and possibly  $u_2(\mathbf{y})$ , the shape of the optical surface(s) is computed. A concise account of the numerical algorithm is given in Section 6.

## H-convex analysis

Next, we consider the geometrical Equation 12b, which can always be formulated such that  $H_w(\mathbf{x}, \mathbf{y}, \cdot) > 0$  for all  $\mathbf{x} \in \mathcal{X}$  and  $\mathbf{y} \in \mathcal{Y}$ , implying that  $H(\mathbf{x}, \mathbf{y}, \cdot)$  has an inverse. Also, Equation 12b has multiple solution pairs  $(u_1(\mathbf{x}), u_2(\mathbf{y}))$ . In analogy with (Equation 13), a possible solution reads

$$\forall \mathbf{x} \in \mathcal{X} \quad u_1(\mathbf{x}) = \max_{\mathbf{y} \in \mathcal{Y}} G(\mathbf{x}, \mathbf{y}, u_2(\mathbf{y})), \tag{17a}$$

$$\forall \mathbf{y} \in \mathcal{Y} \quad u_2(\mathbf{y}) = \min_{\mathbf{x} \in \mathcal{X}} H(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})), \tag{17b}$$

implicitly defining the optical map as follows:  $\mathbf{y}^* = \mathbf{m}(\mathbf{x}) \in \mathcal{Y}$  is the maximizer in Equation 17a and  $\mathbf{x}^* \in \mathcal{X}$  such that  $\mathbf{y} = \mathbf{m}(\mathbf{x}^*)$  is the minimizer in Equation 17b. Provided a condition on  $H$  holds, which we specify shortly, the extrema in Equation 17 are unique, which we henceforth assume. The functions  $u_1(\mathbf{x})$  and  $u_2(\mathbf{y})$  are referred to as a  $G$ -convex  $H$ -concave solution pair; see [4]. We prove that (Equation 17a) together with (Equation 12b) implies (Equation 17b); the implication in the reverse order proceeds in a similar way. Assume that the expression for  $u_1(\mathbf{x})$  in (Equation 17a) holds, then

$$\forall \mathbf{x} \in \mathcal{X} \quad \forall \mathbf{y} \in \mathcal{Y} \quad u_1(\mathbf{x}) \geq G(\mathbf{x}, \mathbf{y}, u_2(\mathbf{y})).$$

Applying the inverse  $H(\mathbf{x}, \mathbf{y}, \cdot) = G^{-1}(\mathbf{x}, \mathbf{y}, \cdot)$  and using that  $H_w(\mathbf{x}, \mathbf{y}, \cdot) > 0$ , we obtain

$$\forall \mathbf{x} \in \mathcal{X} \quad \forall \mathbf{y} \in \mathcal{Y} \quad H(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})) \geq u_2(\mathbf{y}).$$

Since this inequality holds for all  $\mathbf{x} \in \mathcal{X}$ , we can take the minimum and find

$$\forall \mathbf{y} \in \mathcal{Y} \quad u_2(\mathbf{y}) \leq \min_{\mathbf{x} \in \mathcal{X}} H(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})).$$

Let  $\mathbf{y} = \mathbf{m}(\tilde{\mathbf{x}}) \in \mathcal{Y}$  for some  $\tilde{\mathbf{x}} \in \mathcal{X}$ . Then, by virtue of Equation 12b.

$$u_2(\mathbf{y}) = H(\tilde{\mathbf{x}}, \mathbf{y}, u_1(\tilde{\mathbf{x}})) \geq \min_{\mathbf{x} \in \mathcal{X}} H(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})).$$

Combining both inequalities, we arrive at the expression for  $u_2(\mathbf{y})$  in Equation 17b with  $\tilde{\mathbf{x}} = \mathbf{x}^*$  the unique minimizer. Conversely, given the expression for  $u_2(\mathbf{y})$  in Equation 17b, we can derive in a similar manner the expression for  $u_1(\mathbf{x})$  in Equation 17a.

Equation 12b also has a  $G$ -concave  $H$ -convex (min/max) solution pair. The necessary condition for both solutions is that  $(\mathbf{x}, \mathbf{m}(\mathbf{x}))$  is a stationary point of  $\tilde{H}(\mathbf{x}, \mathbf{y}) = H(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x}))$ , i.e.,

$$\nabla_{\mathbf{x}} \tilde{H}(\mathbf{x}, \mathbf{m}(\mathbf{x})) = \nabla_{\mathbf{x}} H(\mathbf{x}, \mathbf{m}(\mathbf{x}), u_1(\mathbf{x})) + H_w(\mathbf{x}, \mathbf{m}(\mathbf{x}), u_1(\mathbf{x})) \nabla u_1(\mathbf{x}) = \mathbf{0}. \tag{18}$$

A sufficient condition for the max/min pair is that  $D_{\mathbf{xx}} \tilde{H}(\mathbf{x}, \mathbf{m}(\mathbf{x}))$  be SPD for all  $\mathbf{x} \in \mathcal{X}$ . Alternatively, for the min/max pair, the Hessian

matrix  $D_{xx}\tilde{H}(\mathbf{x}, \mathbf{m}(\mathbf{x}))$  should be SND for all  $\mathbf{x} \in \mathcal{X}$ . In both cases, the solution pair is unique.

Equation 18 implicitly defines the optical map  $\mathbf{y} = \mathbf{m}(\mathbf{x})$ , provided the mixed derivative matrix  $D_{xy}\tilde{H}(\mathbf{x}, \mathbf{y})$  is regular in  $(\mathbf{x}, \mathbf{m}(\mathbf{x}))$  for all  $\mathbf{x} \in \mathcal{X}$ , but the actual computation of the optical map from Equation 18 is virtually impossible for most optical systems. So, we differentiate the zero-gradient condition in Equation 18 and recover the matrix-Jacobi equation in Equation 15; however, with matrices  $\mathbf{C}$  and  $\mathbf{P}$  given by

$$\mathbf{C}(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})) = D_{xy}\tilde{H}(\mathbf{x}, \mathbf{y}), \quad \mathbf{P}(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})) = -D_{xx}\tilde{H}(\mathbf{x}, \mathbf{y}). \quad (19)$$

Notice, there is one difficulty, i.e., the function  $\tilde{H}(\mathbf{x}, \mathbf{y})$  depends on  $u_1(\mathbf{x})$ , and consequently, also the matrices  $\mathbf{C}$  and  $\mathbf{P}$  do. Therefore, the computation of the optical map and the function  $u_1(\mathbf{x})$  can no longer be decoupled; see Section 6 for more details. Finally, also, in this case, the transport boundary condition (16) applies.

## 4 Conservation of luminous flux

In the previous section, we presented equations describing the geometry of an optical system, more specifically (Equation 12) for the shape/location of the optical surface(s), the zero-gradient conditions (Equations 14, 18), and the matrix-Jacobi equation for the optical map, defined in Equations 15, 19. To close the model, we have to formulate the conservation law of luminous flux.

We first introduce stereographic projections of a unit vector  $\hat{\mathbf{v}} \in S^2$ , needed in some of the flux balances. A unit vector  $\hat{\mathbf{v}}$  can be represented by a point P or Q on the unit sphere; see Figure 2. There are two stereographic projections of  $\hat{\mathbf{v}}$ , i.e., the projection from the north pole and from the south pole. To compute the stereographic projection from the north pole N, we have to compute the intersection of the line through N and P, given by

$$\mathbf{r}(\lambda) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} v_1 \\ v_2 \\ v_3 - 1 \end{pmatrix},$$

with the equator plane  $z = 0$ . In this way, we obtain the stereographic projection  $\mathbf{z} = \mathbf{S}_N(\hat{\mathbf{v}})$  given by

$$\mathbf{S}_N(\hat{\mathbf{v}}) = \frac{1}{1 - v_3} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad (20a)$$

which is defined for  $v_3 \neq 1$ . The latter condition means that P and N should not coincide. Using the relations  $|\hat{\mathbf{v}}|^2 = 1$  and  $v_3 \neq 1$ , we can compute the inverse  $\hat{\mathbf{v}} = \mathbf{S}_N^{-1}(\mathbf{z})$  and find

$$\mathbf{S}_N^{-1}(\mathbf{z}) = \frac{1}{1 + |\mathbf{z}|^2} \begin{pmatrix} 2z_1 \\ 2z_2 \\ -1 + |\mathbf{z}|^2 \end{pmatrix}. \quad (20b)$$

In a similar manner, we can determine the stereographic projection from the south pole and its inverse. These are given by

$$\mathbf{z} = \mathbf{S}_S(\hat{\mathbf{v}}) = \frac{1}{1 + v_3} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad \hat{\mathbf{v}} = \mathbf{S}_S^{-1}(\mathbf{z}) = \frac{1}{1 + |\mathbf{z}|^2} \begin{pmatrix} 2z_1 \\ 2z_2 \\ 1 - |\mathbf{z}|^2 \end{pmatrix} \quad (21)$$

and are defined for  $v_3 \neq -1$ . Both projections are parameterizations of  $S^2$ , and in either case, an area element  $dS(\mathbf{z})$  on  $S^2$  generated by  $\hat{\mathbf{v}}$  is given by

$$dS(\mathbf{z}) = \left| \frac{\partial \hat{\mathbf{v}}}{\partial z_1} \times \frac{\partial \hat{\mathbf{v}}}{\partial z_2} \right| d\mathbf{z} = \frac{4}{(1 + |\mathbf{z}|^2)^2} d\mathbf{z} =: J(\hat{\mathbf{v}}(\mathbf{z})) d\mathbf{z}. \quad (22)$$

The choice for the stereographic projection depends on the direction of  $\hat{\mathbf{v}}$ . If  $v_3 > 0$ , a suitable choice is the projection from the south pole; since then, the domain for  $\mathbf{z}$  is included in the interior of the disc  $|\mathbf{z}| \leq 1$ . The projection from the north pole gives a domain outside this disc; see Figure 2. Likewise, if  $v_3 < 0$ , the projection from the north pole is the proper choice.

Assuming there is no energy lost, the most generic form of the luminous flux balance reads

$$\int_{\mathcal{A}} f(\mathbf{x}) dS(\mathbf{x}) = \int_{\mathbf{m}(\mathcal{A})} g(\mathbf{y}) dS(\mathbf{y}), \quad (23)$$

for every subset  $\mathcal{A} \subset \mathcal{X}$  and image set  $\mathbf{m}(\mathcal{A}) \subset \mathcal{Y}$ , where  $f(\mathbf{x})$  and  $g(\mathbf{y})$  are the flux densities, to be specified shortly, at the source and target, respectively. Here,  $dS(\mathbf{x})$  denotes an area element on the (curved) surface  $\mathcal{S}$  describing the source and parameterized by  $\mathbf{x}$ . Analogously,  $dS(\mathbf{y})$  is an area element on the target surface  $\mathcal{T}$ , parameterized by  $\mathbf{y}$ . Equation 23 should hold for any subsets  $\mathcal{A} \subset \mathcal{X}$  and  $\mathbf{m}(\mathcal{A}) \subset \mathcal{Y}$ , so also for  $\mathcal{A} = \mathcal{X}$  and  $\mathbf{m}(\mathcal{A}) = \mathcal{Y}$ , giving the global luminous flux balance. This implies that an inverse design problem can only have a solution if  $f$  and  $g$  satisfy this global flux balance.

The precise form of the flux balance depends on the source and target. For the source, we consider two options: first, a planar source located in  $z = 0$  and emitting a parallel beam of light with  $\hat{\mathbf{s}} = \hat{\mathbf{e}}_z$ , and second, a point source located in the origin and emitting a conical beam of light with  $\hat{\mathbf{s}} = \hat{\mathbf{e}}_r$ . Both are zero-étendue (English: zero-extent) sources, meaning that the emitted beam has either a spatial or angular extent, but in the embedding four-dimensional phase space, it has zero volume; see Chaves [25]. For the planar source  $\mathcal{S} = \{\mathbf{r}(\mathbf{x}) = (\mathbf{x}, 0) | \mathbf{x} \in \mathcal{X} \subset \mathbb{R}^2\}$ ,  $\mathbf{x} = \mathbf{q}_s$  are spatial coordinates (Cartesian/polar),  $dS(\mathbf{x}) = d\mathbf{x}$ , the area element in the source plane, and  $f(\mathbf{x}) = M(\mathbf{r}(\mathbf{x}))$ , the emittance, i.e., the luminous flux per unit area emitted by the source. The point source has no spatial extent but emits light rays in an angular domain, specified by  $\hat{\mathbf{s}} \in \mathcal{S} \subset S^2$ . As parameterization, we choose a stereographic projection, either from the north or south pole, the area element  $dS(\mathbf{x}) = J(\hat{\mathbf{s}}(\mathbf{x}))d\mathbf{x}$ , cf. Equation 22, and  $f(\mathbf{x}) = I_s(\hat{\mathbf{s}}(\mathbf{x}))$ , the intensity of the source, i.e., the emitted luminous flux per unit solid angle. We can express both flux densities in the form  $f = \tilde{f}_0 \mathbf{P}_s^{-1}$ , where  $\tilde{f} = M$  and  $\mathbf{P}_s: (\mathbf{x}, 0) \mapsto \mathbf{x}$ , the trivial projection from  $\mathcal{S}$  to  $\mathcal{X}$ , for the planar source and  $\tilde{f} = I_s$  and  $\mathbf{P}_s = \mathbf{S}_N$  or  $\mathbf{P}_s = \mathbf{S}_S$  for the point source.

For the target, we consider the following cases; first, a near-field target located on a curved surface  $\mathcal{T} = \{\mathbf{r}(\mathbf{y}) = (\mathbf{y}, v(\mathbf{y})) | \mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^2\}$ , and second, a far-field target. For the near-field target,  $\mathbf{y} = \mathbf{q}_t$  are spatial coordinates in a reference plane  $z = z_t$ , the area element

$$dS(\mathbf{y}) = \left| \frac{\partial \mathbf{r}}{\partial y_1} \times \frac{\partial \mathbf{r}}{\partial y_2} \right| d\mathbf{y} = \sqrt{|\nabla v(\mathbf{y})|^2 + 1} d\mathbf{y},$$

and  $g(\mathbf{y}) = E(\mathbf{r}(\mathbf{y}))$ , the illuminance, i.e., the luminous flux per unit area incident on the target surface. The domain of a far-field target is an angular domain, determined by the transmitted rays with  $\hat{\mathbf{t}} \in \mathcal{T} \subset S^2$ . Therefore, we choose for  $\mathbf{y}$  one of the two stereographic

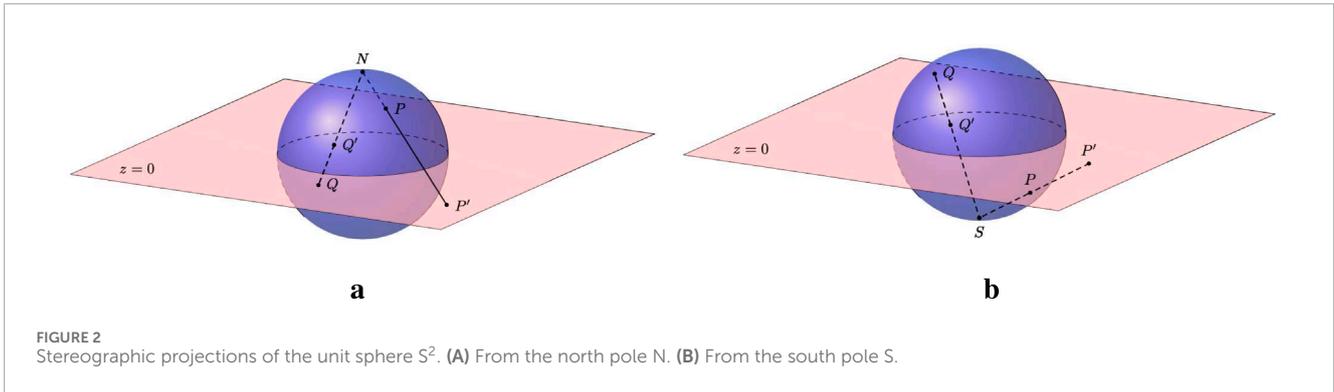


FIGURE 2 Stereographic projections of the unit sphere  $S^2$ . (A) From the north pole N. (B) From the south pole S.

projections, for which  $dS(\mathbf{y}) = J(\hat{\mathbf{t}}(\mathbf{y}))d\mathbf{y}$ ; cf. Equation 22. However,  $g(\mathbf{y}) = I_t(\hat{\mathbf{r}}(\mathbf{y}))$ , the intensity in the target domain as a function of  $\hat{\mathbf{r}}$ , the unit direction vector directed straight from source to target domain, skipping all optical surfaces, rather than a function of  $\hat{\mathbf{t}}$ . In the far-field approximation, the distance from the source to target is large compared to the size of the optical system, specifically  $|\underline{\mathbf{r}} - d\hat{\mathbf{t}}|/|\underline{\mathbf{r}}| \ll 1$ , where  $d$  is the distance from the optical surface to target, measured along a reflected ray; see, e.g., [4], pp. 59–60. The optical system is considered a point contracted at the origin, so we simply consider  $g(\mathbf{y}) = I_t(\hat{\mathbf{t}}(\mathbf{y}))$ . In addition, for the target, both flux densities can be unified in the expression  $g = \tilde{g} \circ \mathbf{P}_t^{-1}$ , where for the near field  $\tilde{g} = E$  and  $\mathbf{P}_t: (\mathbf{y}, \nu(\mathbf{y})) \mapsto \mathbf{y}$  is the orthogonal projection from  $\mathcal{T}$  to  $\mathcal{Y}$ , and where for the far-field target,  $\tilde{g} = I_t$  and either  $\mathbf{P}_t = \mathbf{S}_N$  or  $\mathbf{P}_t = \mathbf{S}_S$ .

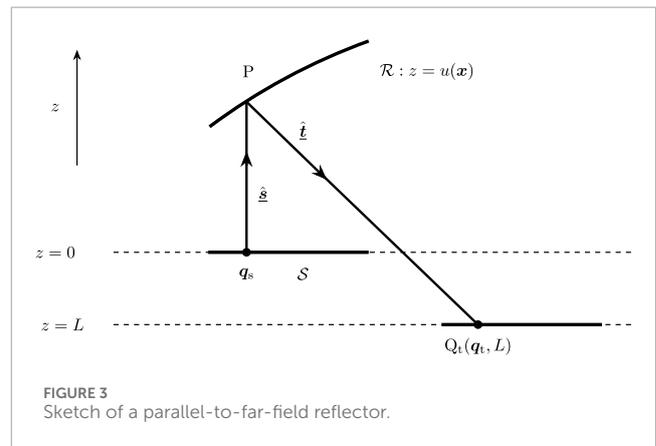


FIGURE 3 Sketch of a parallel-to-far-field reflector.

### 5 Reflector and lens equations: a hierarchy of mathematical models

We present a hierarchy of mathematical models for optical systems with a zero-étendue source, based on three example systems, i.e., a parallel-to-far-field reflector, point-to-far-field lens, and parallel-to-near-field reflector. These example systems are selected based on the mathematical model (quadratic cost function, non-quadratic cost function, and generating function) but are otherwise arbitrary. Other choices like a parallel-to-far-field-lens (quadratic cost function) or point-to-near-field reflector (generating function) would be equally appropriate. In this section, we combine the geometrical equations with the conservation law for luminous flux.

#### 5.1 Parallel-to-far-field reflector

We consider a light source located in the plane  $z = 0$  emitting a parallel beam of light, for which  $\hat{\mathbf{s}} = \hat{\mathbf{e}}_z$ . Light rays strike the reflector  $\mathcal{R}$  defined by  $z = u(\mathbf{x})$ ,  $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^2$ , and are reflected off in the direction  $\hat{\mathbf{t}} \in \mathcal{T} \subset S^2$ . The reflected rays intersect an auxiliary plane  $z = L$ , which we introduce to be able to determine an optical distance; see Figure 3. More precisely, to derive a geometrical equation for the reflector surface, we determine from Equation 10 the mixed characteristic of the first kind  $W(\mathbf{q}_s, \mathbf{p}_t)$  for an arbitrary ray with  $\mathbf{q}_s = \mathbf{x}$  and  $\mathbf{p}_t = (t_1 \ t_2)^T$ , connecting the source with the plane  $z = L$ . The

point characteristic  $V(\mathbf{q}_s, \mathbf{q}_t)$  involved is given by

$$V(\mathbf{q}_s, \mathbf{q}_t) = [Q_s, Q_t] = u(\mathbf{x}) + d, \quad d = \sqrt{|\mathbf{q}_t - \mathbf{x}|^2 + (L - u(\mathbf{x}))^2}, \tag{24}$$

where  $d = d(P, Q_t)$  is the distance between the points  $P(\mathbf{x}, u(\mathbf{x}))$  and  $Q_t(\mathbf{q}_t, L)$ , which are the points where the reflected ray hits the reflector and target plane, respectively, loosely referred to as intersection points. For the direction vector  $\hat{\mathbf{t}}$  of the reflected ray, we have

$$\mathbf{p}_t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \frac{1}{d}(\mathbf{q}_t - \mathbf{x}), \quad t_3 = \frac{1}{d}(L - u(\mathbf{x})). \tag{25}$$

Moreover,  $\partial W / \partial \mathbf{q}_s = -\mathbf{p}_s = \mathbf{0}$  since  $\hat{\mathbf{s}}$  is perpendicular to the source plane, implying that  $W = W(\mathbf{p}_t)$ . Elaborating the expression for  $W$ , using the relations in Equations 24, 25, we find

$$\begin{aligned} W(\mathbf{p}_t) &= u(\mathbf{x}) + d - \mathbf{q}_t \cdot \mathbf{p}_t \\ &= u(\mathbf{x}) + \frac{1}{d}(|\mathbf{q}_t - \mathbf{x}|^2 + (L - u(\mathbf{x}))^2) - \mathbf{q}_t \cdot (\mathbf{q}_t - \mathbf{x}) \\ &= u(\mathbf{x}) + \frac{1}{d}(-\mathbf{x} \cdot (\mathbf{q}_t - \mathbf{x}) + dt_3(L - u(\mathbf{x}))) \\ &= u(\mathbf{x})(1 - t_3) - \mathbf{x} \cdot \mathbf{p}_t + Lt_3. \end{aligned}$$

To separate the variables  $\mathbf{x}$  (source) and  $\mathbf{p}_t$  (target), we bring all terms that solely depend on  $\hat{\mathbf{t}}$  to the left-hand side, and we obtain

$$\frac{W(\mathbf{p}_t) - Lt_3}{1 - t_3} = u(\mathbf{x}) - \frac{\mathbf{x} \cdot \mathbf{p}_t}{1 - t_3}. \tag{26}$$

Note that  $t_3 \neq 1$  since rays cannot pass straight through the reflector. Using the expressions for  $d$  and  $t_3$ , we can show that

$$\frac{\partial W}{\partial L} = \frac{\partial d}{\partial L} - \frac{\partial \mathbf{q}_t}{\partial L} \cdot \mathbf{p}_t, \quad \frac{\partial d}{\partial L} = \mathbf{p}_t \cdot \frac{\partial \mathbf{q}_t}{\partial L} + t_3,$$

from which we readily conclude that the left-hand side in Equation 26 is independent of  $L$ , which makes sense since a far-field target domain should not depend on the choice of a particular plane  $z = L$ . As stated in Section 4, the proper choice to parameterize the target  $\mathcal{T}$  is the stereographic projection from the north pole; thus,  $\mathbf{y} = \mathbf{p}_t/(1 - t_3)$  and  $\mathcal{Y} = \mathbb{S}_N(\mathcal{T})$ ; cf. Equation 20. Substituting the expression for  $\mathbf{y}$  in Equation 26, we arrive at the geometrical equation

$$u_1(\mathbf{x}) + u_2(\mathbf{y}) = c(\mathbf{x}, \mathbf{y}), \quad c(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}, \quad (27)$$

where  $u_1(\mathbf{x}) = u(\mathbf{x})$  and  $u_2(\mathbf{y}) = (Lt_3 - W(\mathbf{p}_t))/(1 - t_3)$ .

Referring to Section 3, a possible solution of Equation 27 is the max/max solution in Equation 13, which is a conjugate pair of Legendre–Fenchel transforms since  $c(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ . The necessary condition in Equation 14 and the matrix equation for the optical map in Equation 15 reduce to

$$\mathbf{y} - \nabla u_1(\mathbf{x}) = \mathbf{0}, \quad (28a)$$

$$D\mathbf{m} = \mathbf{P}(\mathbf{x}), \quad \mathbf{P}(\mathbf{x}) = D^2 u_1(\mathbf{x}). \quad (28b)$$

Recall that Equation 28b has to be supplemented with the transport boundary condition (Equation 16).

To close the model, we elaborate the flux balance in Equation 23. Substituting  $f(\mathbf{x}) = M(\underline{\mathbf{r}}(\mathbf{x}))$ ,  $dS(\mathbf{x}) = d\mathbf{x}$ ,  $g(\mathbf{y}) = I_t(\underline{\hat{\mathbf{t}}}(\mathbf{y}))$ , and  $dS(\mathbf{y}) = J(\underline{\hat{\mathbf{t}}}(\mathbf{y}))d\mathbf{y}$ , we obtain

$$\int_{\mathcal{A}} M(\underline{\mathbf{r}}(\mathbf{x})) d\mathbf{x} = \int_{m(\mathcal{A})} I_t(\underline{\hat{\mathbf{t}}}(\mathbf{y})) J(\underline{\hat{\mathbf{t}}}(\mathbf{y})) d\mathbf{y},$$

for arbitrary  $\mathcal{A} \subset \mathcal{X}$  and  $m(\mathcal{A}) \subset \mathcal{Y}$ . Substituting  $\mathbf{y} = m(\mathbf{x})$  and the expression for  $J(\underline{\hat{\mathbf{t}}}(\mathbf{y}))$ , replacing  $\underline{\hat{\mathbf{y}}}(z)$  by  $\underline{\hat{\mathbf{t}}}(\mathbf{y})$  in Equation 22, we can rewrite the integral in the right-hand side and find

$$\int_{\mathcal{A}} M(\underline{\mathbf{r}}(\mathbf{x})) d\mathbf{x} = \int_{\mathcal{A}} I_t(\underline{\hat{\mathbf{t}}}(m(\mathbf{x}))) \frac{4}{(1 + |m(\mathbf{x})|^2)^2} \det(D\mathbf{m}) d\mathbf{x}, \quad (29)$$

where we used that  $\det(D\mathbf{m}) > 0$ , which is correct because  $D\mathbf{m}$  is either SPD or SND. Since Equation 29 should hold for arbitrary  $\mathcal{A} \subset \mathcal{X}$ , we obtain the differential form

$$\det(D\mathbf{m}) = \frac{1}{4} (1 + |m(\mathbf{x})|^2)^2 \frac{M(\underline{\mathbf{r}}(\mathbf{x}))}{I_t(\underline{\hat{\mathbf{t}}}(m(\mathbf{x})))} =: F(\mathbf{x}, m(\mathbf{x})),$$

referred to as the Jacobian equation. Substituting  $\mathbf{m} = \nabla u_1$ , this equation reduces to the standard MA equation  $\det(D^2 u_1) = F(\mathbf{x}, \nabla u_1(\mathbf{x}))$ ; see, e.g., Gutiérrez [3].

## 5.2 Point-to-far-field lens

We consider a point source located in the origin  $O_s$  (position vector given by  $\mathbf{q}_s = 0$  and  $z_s = 0$ ), emitting upward a conical beam of light with the direction vector  $\underline{\hat{\mathbf{s}}} = \underline{\hat{\mathbf{e}}}_r$ . Light rays hit

a lens of refractive index  $n$ ; see Figure 4. The first (entrance) surface is spherical with center  $O_s$  and radius  $R$ ; hence, rays are not refracted there; however, the second (exit) surface is freeform and given by the parameterization  $\underline{\mathbf{r}}(\underline{\hat{\mathbf{s}}}) = u(\underline{\hat{\mathbf{s}}})\underline{\hat{\mathbf{s}}}$ . Rays are refracted at the exit surface and arrive at an auxiliary plane  $z = L > 0$ . To derive a geometrical equation for the freeform surface, we evaluate the angular characteristic  $T(\mathbf{p}_s, \mathbf{p}_t)$  defined in Equation 11 for an arbitrary ray connecting the source with the plane  $z = L$ . The point characteristic  $V(\mathbf{q}_s, \mathbf{q}_t)$  needed to evaluate  $T(\mathbf{p}_s, \mathbf{p}_t)$  is given by

$$V(\mathbf{q}_s, \mathbf{q}_t) = [O_s, Q_t] = R + n(u(\underline{\hat{\mathbf{s}}}) - R) + d \quad (30)$$

$$d = \sqrt{|\mathbf{q}_t - u(\underline{\hat{\mathbf{s}}})\mathbf{p}_s|^2 + (L - u(\underline{\hat{\mathbf{s}}})s_3)^2}$$

where  $d = d(P, Q_t)$  is the distance between the points  $P(u(\underline{\hat{\mathbf{s}}})\mathbf{p}_s, u(\underline{\hat{\mathbf{s}}})s_3)$  and  $Q_t(\mathbf{q}_t, L)$ , which are the intersection points of the refracted ray with exit surface and target screen, respectively. The direction vector  $\underline{\hat{\mathbf{t}}}$  of the refracted ray is given by

$$\mathbf{p}_t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \frac{1}{d}(\mathbf{q}_t - u(\underline{\hat{\mathbf{s}}})\mathbf{p}_s), \quad t_3 = \frac{1}{d}(L - u(\underline{\hat{\mathbf{s}}})s_3). \quad (31)$$

Note that  $\partial T/\partial \mathbf{p}_s = \mathbf{q}_s = 0$ , implying that  $T = T(\mathbf{p}_t)$ . Evaluating the expression for  $T(\mathbf{p}_t)$  in Equation 11 using the relations in Equations 30, 31, we obtain

$$\begin{aligned} T(\mathbf{p}_t) &= R + n(u(\underline{\hat{\mathbf{s}}}) - R) + d - \mathbf{q}_t \cdot \mathbf{p}_t \\ &= (1 - n)R + nu(\underline{\hat{\mathbf{s}}}) + \frac{1}{d}(|\mathbf{q}_t - u(\underline{\hat{\mathbf{s}}})\mathbf{p}_s|^2 + (L - u(\underline{\hat{\mathbf{s}}})s_3)^2) \\ &\quad - \mathbf{q}_t \cdot (\mathbf{q}_t - u(\underline{\hat{\mathbf{s}}})\mathbf{p}_s) \\ &= (1 - n)R + nu(\underline{\hat{\mathbf{s}}}) + \frac{1}{d}(-u(\underline{\hat{\mathbf{s}}})\mathbf{p}_s \cdot (\mathbf{q}_t - u(\underline{\hat{\mathbf{s}}})\mathbf{p}_s) \\ &\quad + dt_3(L - u(\underline{\hat{\mathbf{s}}})s_3)) \\ &= (1 - n)R + nu(\underline{\hat{\mathbf{s}}}) - u(\underline{\hat{\mathbf{s}}})\mathbf{p}_s \cdot \mathbf{p}_t + t_3(L - u(\underline{\hat{\mathbf{s}}})s_3) \\ &= (1 - n)R + (n - \underline{\hat{\mathbf{s}}} \cdot \underline{\hat{\mathbf{t}}})u(\underline{\hat{\mathbf{s}}}) + Lt_3. \end{aligned} \quad (32)$$

To derive the last expression for  $T$  in Equation 32, we substituted the relation  $\underline{\hat{\mathbf{s}}} \cdot \underline{\hat{\mathbf{t}}} = \mathbf{p}_s \cdot \mathbf{p}_t + s_3 t_3$ . Next, we move all terms that solely depend on  $\underline{\hat{\mathbf{t}}}$  and the constant  $(1 - n)R$  to the left-hand side and find

$$T(\mathbf{p}_t) - Lt_3 + (n - 1)R = (n - \underline{\hat{\mathbf{s}}} \cdot \underline{\hat{\mathbf{t}}})u(\underline{\hat{\mathbf{s}}}). \quad (33)$$

Differentiating the first expression in Equation 32 with respect to  $L$  and combining the expression for  $d$  with Equation 31, we obtain

$$\frac{\partial T}{\partial L} = \frac{\partial d}{\partial L} - \frac{\partial \mathbf{q}_t}{\partial L} \cdot \mathbf{p}_t, \quad \frac{\partial d}{\partial L} = \mathbf{p}_t \cdot \frac{\partial \mathbf{q}_t}{\partial L} + t_3,$$

from which we readily see that the left-hand side in Equation 33 is independent of  $L$ , as anticipated.

To separate source and target coordinates, analogous to Equation 27, we have to take the logarithm of the left- and right-hand sides of Equation 33. Note that  $n - \underline{\hat{\mathbf{s}}} \cdot \underline{\hat{\mathbf{t}}} > 0$  and  $u(\underline{\hat{\mathbf{s}}}) > 0$ ; consequently, the left-hand side is positive as well, and we can take the logarithms. In addition, we substitute the stereographic projections (from the south pole)  $\mathbf{x}$  of  $\underline{\hat{\mathbf{s}}}$  and  $\mathbf{y}$  of  $\underline{\hat{\mathbf{t}}}$  since both  $\underline{\hat{\mathbf{s}}}$  and  $\underline{\hat{\mathbf{t}}}$  are directed upward; cf. Equation 21. In this way, we obtain

$$u_1(\mathbf{x}) + u_2(\mathbf{y}) = c(\mathbf{x}, \mathbf{y}), \quad (34a)$$

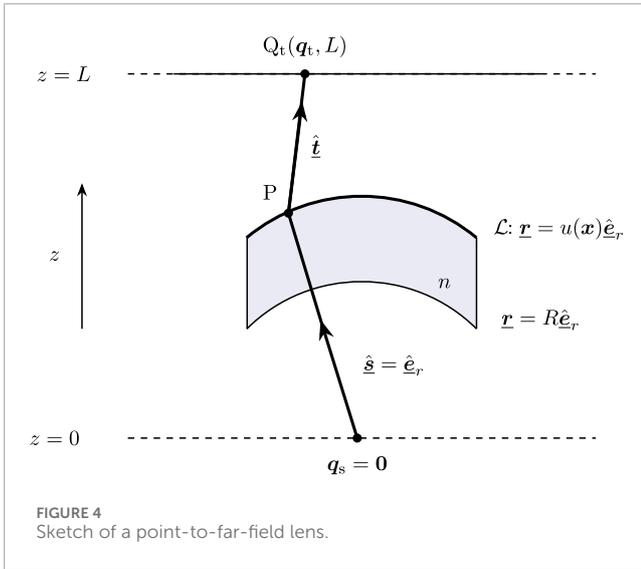


FIGURE 4 Sketch of a point-to-far-field lens.

where the variables  $u_1(\mathbf{x})$  and  $u_2(\mathbf{y})$  and the cost function  $c(\mathbf{x}, \mathbf{y})$  are defined by

$$u_1(\mathbf{x}) = \log u(\underline{\hat{\mathbf{s}}}(\mathbf{x})), \tag{34b}$$

$$u_2(\mathbf{y}) = -\log(T(\underline{\hat{\mathbf{p}}}_t(\mathbf{y})) - Lt_3(\mathbf{y}) + (n-1)R), \tag{34c}$$

$$c(\mathbf{x}, \mathbf{y}) = -\log(n - \underline{\hat{\mathbf{s}}}(\mathbf{x}) \cdot \underline{\hat{\mathbf{i}}}(\mathbf{y})). \tag{34d}$$

Formulated in terms of stereographic coordinates, the cost function reads

$$c(\mathbf{x}, \mathbf{y}) = -\log\left(n - 1 + \frac{2|\mathbf{x} - \mathbf{y}|^2}{(1 + |\mathbf{x}|^2)(1 + |\mathbf{y}|^2)}\right),$$

which is no longer quadratic. A possible solution of Equation 34a is the  $c$ -convex pair in Equation 13, for which the necessary condition (Equation 14) holds. For the optical map, matrix Equation 15 accompanied with the transport boundary condition (Equation 16) is given.

Referring to Section 4, in the far-field approximation, the luminous flux balance reads

$$\int_{\mathcal{A}} I_s(\underline{\hat{\mathbf{s}}}(\mathbf{x})) J(\underline{\hat{\mathbf{s}}}(\mathbf{x})) \, d\mathbf{x} = \int_{\mathbf{m}(\mathcal{A})} I_t(\underline{\hat{\mathbf{i}}}(\mathbf{y})) J(\underline{\hat{\mathbf{i}}}(\mathbf{y})) \, d\mathbf{y}, \tag{35}$$

for arbitrary  $\mathcal{A} \subset \mathcal{X}$  and image set  $\mathbf{m}(\mathcal{A}) \subset \mathcal{Y}$ . To derive the differential form, we substitute the expressions for  $J(\underline{\hat{\mathbf{s}}}(\mathbf{x}))$  and  $J(\underline{\hat{\mathbf{i}}}(\mathbf{y}))$ , according to Equation 22, and subsequently substitute  $\mathbf{y} = \mathbf{m}(\mathbf{x})$ . Assuming  $\det(D\mathbf{m}) > 0$ , we obtain

$$\det(D\mathbf{m}) = \left(\frac{1 + |\mathbf{m}(\mathbf{x})|^2}{1 + |\mathbf{x}|^2}\right)^2 \frac{I_s(\mathbf{x})}{I_t(\mathbf{m}(\mathbf{x}))} =: F(\mathbf{x}, \mathbf{m}(\mathbf{x})).$$

Combining this equation with the matrix equation in Equation 15, we obtain the GMA equation  $\det(D^2 u_1 - D_{xx} c) = \det(C(\mathbf{x}, \mathbf{m}(\mathbf{x})) F(\mathbf{x}, \mathbf{m}(\mathbf{x})))$ .

### 5.3 Parallel-to-near-field reflector

The last example concerns a planar source, located in  $z = 0$ , emitting a parallel beam of light rays in the direction  $\underline{\hat{\mathbf{s}}} = \underline{\hat{\mathbf{e}}}_z$ ,

consequently  $\underline{\mathbf{p}}_s = \mathbf{0}$ , which hits a reflector  $\mathcal{R}$  given by  $z = u(\mathbf{x})$ , with  $\mathbf{x} = \mathbf{q}_s$  spatial coordinates in the source domain, and are reflected off in the direction  $\underline{\hat{\mathbf{i}}}$  and lands at a target surface given by  $z = v(\mathbf{y})$ , with  $\mathbf{y} = \mathbf{q}_t$  spatial coordinates in the reference plane  $z = z_t = 0$ . Thus, source and target planes coincide; see Figure 5. We evaluate Hamilton's point characteristic and find

$$V(\mathbf{q}_s, \mathbf{q}_t) = [Q_s, Q_t] = u(\mathbf{x}) + d, \quad d = \sqrt{|\mathbf{y} - \mathbf{x}|^2 + (v(\mathbf{y}) - u(\mathbf{x}))^2},$$

where  $d = d(P, Q_t)$  is the distance between the intersection points  $P(\mathbf{x}, u(\mathbf{x}))$  and  $Q_t(\mathbf{y}, v(\mathbf{y}))$  of the reflected ray with the reflector and the target surface, respectively. Since  $\partial V / \partial \mathbf{q}_s = -\underline{\mathbf{p}}_s = \mathbf{0}$ , we find that  $V = V(\mathbf{q}_t)$ . In this case, it is no longer possible to separate the source and target coordinates  $\mathbf{x}$  and  $\mathbf{y}$  like in Equation 27 or (Equation 34a). Instead, we write

$$u_2(\mathbf{y}) = H(\mathbf{x}, \mathbf{y}, u_1(\mathbf{x})), \quad H(\mathbf{x}, \mathbf{y}, w) = w + \sqrt{|\mathbf{y} - \mathbf{x}|^2 + (v(\mathbf{y}) - w)^2}, \tag{36}$$

where  $u_1(\mathbf{x}) = u(\mathbf{x})$  and  $u_2(\mathbf{y}) = V(\mathbf{q}_t)$ . Straightforward differentiation gives  $H_w(\mathbf{x}, \mathbf{y}, \cdot) > 0$ , implying that for fixed  $\mathbf{x}, \mathbf{y}$ , the inverse  $G(\mathbf{x}, \mathbf{y}, \cdot) = H^{-1}(\mathbf{x}, \mathbf{y}, \cdot)$  exists. Therefore, we can explicitly determine  $u_1(\mathbf{x})$  from Equation 36, and we obtain

$$u_1(\mathbf{x}) = G(\mathbf{x}, \mathbf{y}, u_2(\mathbf{y})), \quad G(\mathbf{x}, \mathbf{y}, w) = \frac{1}{2} \left( w + v(\mathbf{y}) - \frac{|\mathbf{y} - \mathbf{x}|^2}{w - v(\mathbf{y})} \right). \tag{37}$$

A possible solution of Equations 36, 37 is the  $G$ -convex  $H$ -concave pair in Equation 17, for which the necessary condition (18) holds. The matrix equation for  $\mathbf{m}$  is given in Equation 15 with matrices  $\mathbf{C}$  and  $\mathbf{P}$  defined in Equation 19. Recall that these matrices explicitly depend on  $u_1(\mathbf{x})$ . Obviously, also the transport boundary condition (16) holds.

Making the proper choices for the flux densities and area elements, the flux balance (23) reduces to

$$\int_{\mathcal{A}} M(\underline{\mathbf{r}}(\mathbf{x})) \, d\mathbf{x} = \int_{\mathbf{m}(\mathcal{A})} E(\underline{\mathbf{r}}(\mathbf{y})) \sqrt{|\nabla v(\mathbf{y})|^2 + 1} \, d\mathbf{y}, \tag{38}$$

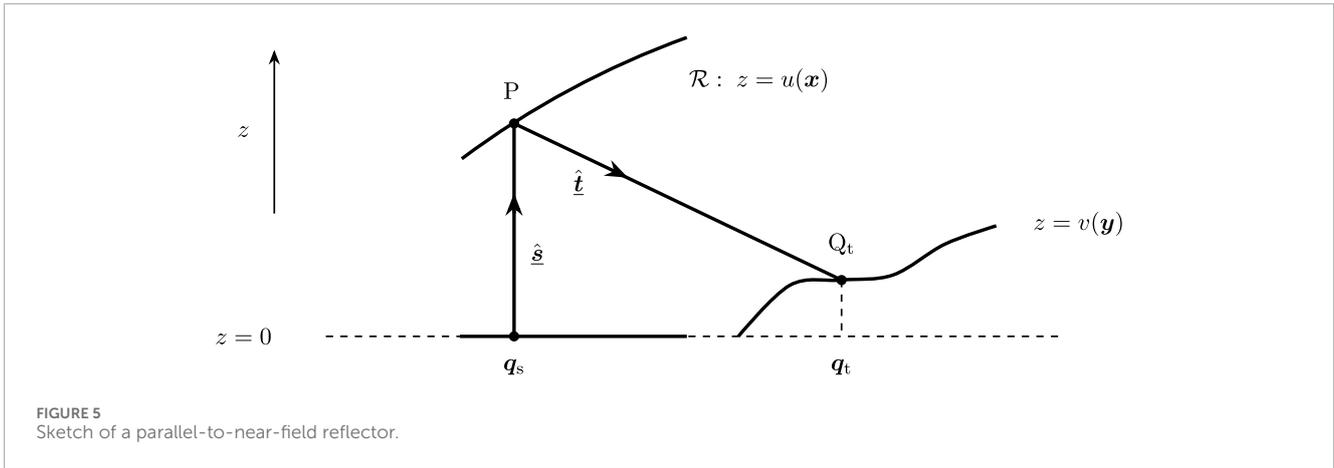
for an arbitrary set  $\mathcal{A} \subset \mathcal{X}$  and image set  $\mathbf{m}(\mathcal{A}) \subset \mathcal{Y}$ , where  $M(\underline{\mathbf{r}}(\mathbf{x}))$  and  $E(\underline{\mathbf{r}}(\mathbf{y}))$  denote the emittance and illuminance of the source and target, respectively. Substituting  $\mathbf{y} = \mathbf{m}(\mathbf{x})$ , assuming  $\det(D\mathbf{m}) > 0$ , we obtain

$$\det(D\mathbf{m}) = \frac{1}{\sqrt{|\nabla v(\mathbf{m}(\mathbf{x}))|^2 + 1}} \frac{M(\underline{\mathbf{r}}(\mathbf{x}))}{E(\underline{\mathbf{r}}(\mathbf{m}(\mathbf{x})))} =: F(\mathbf{x}, \mathbf{m}(\mathbf{x})).$$

Combining this equation with Equation 15 and the matrices  $\mathbf{C}$  and  $\mathbf{P}$  defined in Equation 19, we obtain the equation  $\det(D_{xx} \bar{H}) = \det(C(\mathbf{x}, \mathbf{m}(\mathbf{x}), u_1(\mathbf{x})) F(\mathbf{x}, \mathbf{m}(\mathbf{x})))$ , referred to as the GJ equation.

### 5.4 Summary of mathematical models

All mathematical models considered consist of a geometrical equation, a (zero-gradient) condition for a stationary point, a matrix equation for the optical map coupled to the transport boundary condition, and a luminous flux balance



$\det(D\mathbf{m}) = F(\mathbf{x}, \mathbf{m}(\mathbf{x}))$ . The conditions for the stationary point read

$$\text{cost function } \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{y}) - \nabla u_1(\mathbf{x}) = \mathbf{0}, \tag{39a}$$

$$\text{generating function } \nabla_{\mathbf{x}} \tilde{H}(\mathbf{x}, \mathbf{y}) = \mathbf{0}. \tag{39b}$$

Substituting the optical map  $\mathbf{y} = \mathbf{m}(\mathbf{x})$  and differentiating with respect to  $\mathbf{x}$  gives the matrix equation  $\mathbf{C}D\mathbf{m} = \mathbf{P}$ , where the matrices  $\mathbf{C}$  and  $\mathbf{P}$  are given by

$$\text{quadratic cost function } \mathbf{C} = \mathbf{I}, \quad \mathbf{P} = D^2 u_1, \tag{40a}$$

$$\text{general cost function } \mathbf{C} = D_{xy} c, \quad \mathbf{P} = D^2 u_1 - D_{xx} c, \tag{40b}$$

$$\text{generating function } \mathbf{C} = D_{xy} \tilde{H}, \quad \mathbf{P} = -D_{xx} \tilde{H}. \tag{40c}$$

Combining the matrix equation for  $\mathbf{m}$  with the flux balance gives

$$\det(\mathbf{C})F(\mathbf{x}, \mathbf{m}(\mathbf{x})) = \det(\mathbf{P}),$$

which we refer to as the luminous flux constraint. Note that for the cost-function model, the matrices  $\mathbf{C}$  and  $\mathbf{P}$  explicitly depend on  $\mathbf{x}$  and  $\mathbf{m}(\mathbf{x})$ ; on the other hand, for the generating-function model, both matrices in addition depend on  $u_1(\mathbf{x})$ . In [26], we have given a similar overview of mathematical models of 16 base optical systems.

In the next section, we outline numerical methods to compute the optical map  $\mathbf{y} = \mathbf{m}(\mathbf{x})$  and the auxiliary variable  $u_1(\mathbf{x})$ . The shape/location of the optical surface is then reconstructed from a simple algebraic equation, i.e.,  $u(\mathbf{x}) = u_1(\mathbf{x})$  for the parallel-to-far-field and parallel-to-near-field reflectors and  $u(\hat{\mathbf{s}}(\mathbf{x})) = e^{u_1(\mathbf{x})}$  for the parallel-to-far-field lens.

## 6 Iterative least-squares methods

In this section, we outline iterative least-squares methods for the models presented in the previous section, starting with the base scheme for the cost-function models and adding modifications for the generation-function model. A detailed account of these methods is presented in a series of theses—see [27]; [23]; [4]—and papers—see [13]; [28]; [14, 29]; [15–18]; [26].

## Base scheme for cost function

To compute the layout of the optical system, we apply a two-stage method, i.e., we first compute the optical map  $\mathbf{m}$  and subsequently, the auxiliary variable  $u_1$  and possibly  $u_2$ , from which we trivially can compute the shape/location of the optical surface(s). We employ a uniform rectangular grid covering the parameter space of the source domain.

We iteratively compute a symmetric matrix  $\mathbf{P}$  as approximation of the symmetric part of  $\mathbf{C}D\mathbf{m}$ , satisfying  $\det(\mathbf{P}) = \det(\mathbf{C}(\cdot, \mathbf{m}))F(\cdot, \mathbf{m})$ , and a vector field  $\mathbf{b}: \partial\mathcal{X} \rightarrow \partial\mathcal{Y}$ , from which we subsequently compute  $\mathbf{m}$ . Our solution strategy is then to minimize the functionals

$$J_1[\mathbf{m}, \mathbf{P}] = \frac{1}{2} \int_{\mathcal{X}} \|\mathbf{C}D\mathbf{m} - \mathbf{P}\|_F^2 \, dx, \tag{41a}$$

subject to  $\det(\mathbf{P}) = \det(\mathbf{C}(\cdot, \mathbf{m}))F(\cdot, \mathbf{m}),$

$$J_B[\mathbf{m}, \mathbf{b}] = \frac{1}{2} \oint_{\partial\mathcal{X}} |\mathbf{m} - \mathbf{b}|_2^2 \, ds, \tag{41b}$$

$$J[\mathbf{m}, \mathbf{P}, \mathbf{b}] = \alpha J_1[\mathbf{m}, \mathbf{P}] + (1 - \alpha) J_B[\mathbf{m}, \mathbf{b}], \quad (0 < \alpha < 1), \tag{41c}$$

where  $\|\cdot\|_F$  denotes the Frobenius norm. The minimization of  $J_B[\mathbf{m}, \mathbf{b}]$  is to enforce the transport boundary condition (16). Given an initial guess  $\mathbf{m}^0$ , the iteration scheme then reads

$$\mathbf{P}^{k+1} = \operatorname{argmin}_{\mathbf{P} \in \mathcal{P}(\mathbf{m}^k)} J_1[\mathbf{m}^k, \mathbf{P}], \tag{42a}$$

$$\mathbf{b}^{k+1} = \operatorname{argmin}_{\mathbf{b} \in \mathcal{B}} J_B[\mathbf{m}^k, \mathbf{b}], \tag{42b}$$

$$\mathbf{m}^{k+1} = \operatorname{argmin}_{\mathbf{m} \in \mathcal{M}} J[\mathbf{m}, \mathbf{P}^{k+1}, \mathbf{b}^{k+1}], \tag{42c}$$

where the corresponding function spaces are given by

$$\mathcal{P}(\mathbf{m}) = \{\mathbf{P} \in C^1(\mathcal{X})^{2 \times 2} | \mathbf{P}^T = \mathbf{P}, \det(\mathbf{P}) = \det(\mathbf{C}(\cdot, \mathbf{m}(\cdot)))F(\cdot, \mathbf{m}(\cdot))\}, \tag{43a}$$

$$\mathcal{B} = \{\mathbf{b} \in C(\partial\mathcal{X})^2 | \mathbf{b}(\mathbf{x}) \in \partial\mathcal{Y}\}, \tag{43b}$$

$$\mathcal{M} = C^2(\mathcal{X})^2. \tag{43c}$$

In common parlance,  $\mathcal{P}(\mathbf{m})$  is the space of  $2 \times 2$  matrix functions that are continuously differentiable on  $\mathcal{X}$ , are symmetric, and satisfy the luminous flux constraint.  $\mathcal{B}$  is the space of continuous vector functions that map the boundary  $\partial\mathcal{X}$  to the boundary  $\partial\mathcal{Y}$ .

The minimization of  $J_1[\mathbf{m}, \mathbf{P}]$  to compute  $\mathbf{P}$  can be performed point-wise and requires the solution of a constrained minimization problem.  $\mathbf{P}$  has to be either SPD or SND, which can be enforced by a constraint on  $\text{tr}(\mathbf{P})$ . The minimization of  $J_B[\mathbf{m}, \mathbf{b}]$  to compute  $\mathbf{b}$  is a piecewise projection of  $\mathbf{m}$  on  $\partial\mathcal{X}$ . For the minimization of  $J[\mathbf{m}, \mathbf{P}, \mathbf{b}]$ , we impose the first variation with respect to  $\mathbf{m}$  to vanish, i.e.,

$$\delta J[\mathbf{m}, \mathbf{P}, \mathbf{b}](\boldsymbol{\eta}) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (J[\mathbf{m} + \epsilon \boldsymbol{\eta}, \mathbf{P}, \mathbf{b}] - J[\mathbf{m}, \mathbf{P}, \mathbf{b}]) = 0, \quad (44)$$

for arbitrary  $\boldsymbol{\eta} \in C^2(\mathcal{X})^2$ . Evaluating this limit, applying Gauss's theorem and the fundamental lemma of calculus of variations ([30], p. 185), we derive the following coupled boundary value problem (BVP) for  $\mathbf{m}$ :

$$\nabla \cdot (\mathbf{C}^T \mathbf{C} \mathbf{D} \mathbf{m}) = \nabla \cdot (\mathbf{C}^T \mathbf{P}), \quad \mathbf{x} \in \mathcal{X}, \quad (45a)$$

$$(1 - \alpha) \mathbf{m} + \alpha (\mathbf{C}^T \mathbf{C} \mathbf{D} \mathbf{m}) \cdot \hat{\mathbf{n}} = (1 - \alpha) \mathbf{b} + \alpha \mathbf{C}^T \mathbf{P} \hat{\mathbf{n}}, \quad \mathbf{x} \in \partial\mathcal{X}, \quad (45b)$$

where  $\hat{\mathbf{n}}$  is the unit outward normal on  $\partial\mathcal{X}$ . The divergence of a  $2 \times 2$  matrix function  $\mathbf{A} = (\mathbf{a}_1 \ \mathbf{a}_2)$  is defined as  $\nabla \cdot \mathbf{A} = \begin{pmatrix} \nabla \cdot \mathbf{a}_1 \\ \nabla \cdot \mathbf{a}_2 \end{pmatrix}$ . For discretization of the BVP in (Equation 45), we employ the standard finite volume method ([4], pp. Appendix B).

Next, upon convergence of the iteration (42), we compute  $u_1$  from (Equation 39a) by minimizing the functional

$$I[u_1] = \frac{1}{2} \int_{\mathcal{X}} |\nabla_x c(\cdot, \mathbf{m}) - \nabla u_1|_2^2 dx.$$

Analogous to the derivation of (Equation 45), we set the first variation  $\delta I[u_1](v) = 0$  for arbitrary  $v \in C^2(\mathcal{X})$ , to obtain the Neumann problem

$$\nabla^2 u_1 = \nabla \cdot \nabla_x c(\cdot, \mathbf{m}), \quad \mathbf{x} \in \mathcal{X}, \quad (46a)$$

$$\nabla u_1 \cdot \hat{\mathbf{n}} = \nabla_x c(\cdot, \mathbf{m}) \cdot \hat{\mathbf{n}}, \quad \mathbf{x} \in \partial\mathcal{X}. \quad (46b)$$

We employ standard finite differences for the BVP in 46.  $u_1$  is determined up to an additive constant, which translates in an additive constant in the location of the parallel-to-far-field reflector and a multiplicative constant for the location of the point-to-far-field lens. To compute a unique solution from 46 we either fix the distance from the source to optical surface along a specific ray, or prescribe an average for  $u_1$ .

## Extension to generating function

The optical map  $\mathbf{m}$  and the variable  $u_1$  have to be computed simultaneously since the matrices  $\mathbf{C}$  and  $\mathbf{P}$  depend on both variables. Like in the previous case, the optical map satisfies the equation  $\mathbf{C} \mathbf{D} \mathbf{m} = \mathbf{P}$ , with the matrices  $\mathbf{C}$  and  $\mathbf{P}$  defined in Equation 40c, and the luminous flux balance  $\det(\mathbf{D} \mathbf{m}) = F$ . Consequently, the functional  $J_1[\mathbf{m}, \mathbf{P}]$  remains the same,

albeit with different matrices  $\mathbf{C}$  and  $\mathbf{P}$ . The variable  $u_1$  has to be computed from Equation 39b, for which we introduce the functional

$$\begin{aligned} I[u_1, \mathbf{m}] &= \frac{1}{2} \int_{\mathcal{X}} |\nabla_x \tilde{H}(\cdot, \mathbf{m})|_2^2 dx \\ &= \frac{1}{2} \int_{\mathcal{X}} |\nabla_x H(\cdot, \mathbf{m}, u_1) + H_w(\cdot, \mathbf{m}, u_1) \nabla u_1|_2^2 dx. \end{aligned} \quad (47)$$

To minimize the functional in Equation 47, we require  $\delta I[u_1, \mathbf{m}](v) = 0$  for arbitrary  $v \in C^2(\mathcal{X})$ , with  $\delta I[u_1, \mathbf{m}](v)$  the first variation in  $I$  with respect to  $u_1$ , defined analogously to Equation 44. This way we can derive the BVP

$$\nabla \cdot (H_w \nabla_x H + H_w^2 \nabla u_1) = \frac{1}{2} \frac{d}{d w} |\nabla_x H + H_w \nabla u_1|_2^2, \quad \mathbf{x} \in \mathcal{X}, \quad (48a)$$

$$H_w (\nabla_x H + H_w \nabla u_1) \cdot \hat{\mathbf{n}} = 0, \quad \mathbf{x} \in \partial\mathcal{X}. \quad (48b)$$

We employ the standard finite volume method for discretization. We can show that also the solution of the BVP in Equation 48 is determined up to an additive constant and compute a unique solution prescribing the average value of  $u_1$ ; see ([4], pp. 166–167) for more details. Finally, given an initial guess  $\mathbf{m}^0$  and  $u_1^0$ , the iteration scheme then reads

$$\mathbf{P}^{k+1} = \operatorname{argmin}_{\mathbf{P} \in \mathcal{P}(m^k)} J_1[\mathbf{m}^k, u_1^k, \mathbf{P}], \quad (49a)$$

$$\mathbf{b}^{k+1} = \operatorname{argmin}_{\mathbf{b} \in \mathcal{B}} J_B[\mathbf{m}^k, \mathbf{b}], \quad (49b)$$

$$\mathbf{m}^{k+1} = \operatorname{argmin}_{\mathbf{m} \in \mathcal{M}} J[\mathbf{m}, u_1^k, \mathbf{P}^{k+1}, \mathbf{b}^{k+1}], \quad (49c)$$

$$u_1^{k+1} = \operatorname{argmin}_{v \in \mathcal{U}} I[v, \mathbf{m}^{k+1}], \quad (49d)$$

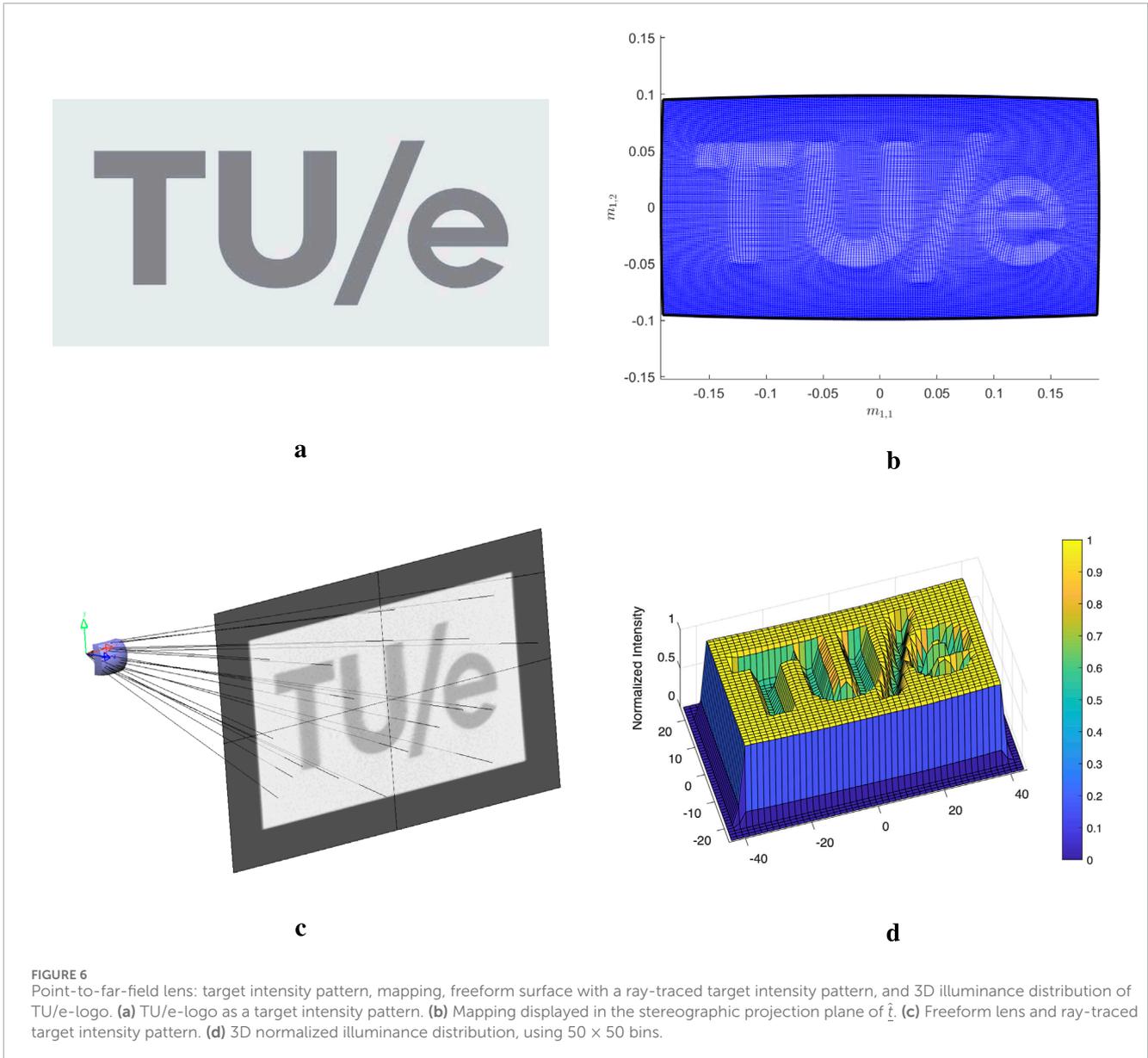
with  $\mathcal{U} = C^2(\mathcal{X})$ , where we have included  $u_1^k$  in the argument list of  $J_1$  and  $J$  to denote their implicit dependence on  $u_1$  via the matrix  $\mathbf{C}$ .

## 7 Numerical examples

We present two numerical examples, viz. the point-to-far-field lens discussed in Section 5.2 and the parallel-to-near-field reflector from Section 5.3.

### 7.1 Point-to-far-field lens

We consider a point source, located in the origin  $O_s$  of the source plane  $z = 0$ , emitting upward a beam of light with uniform intensity  $I_s(\hat{\mathbf{s}}(\mathbf{x}))$  on the circular stereographic source domain  $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^2 | |\mathbf{x}| < 0.2\}$ . In terms of spherical coordinates  $(\phi, \theta)$ , this corresponds to the domain  $0 \leq \phi \leq \arccos(12/13)$  and  $0 \leq \theta < 2\pi$  with  $\phi$  and  $\theta$  as the polar and azimuthal angle, respectively, i.e., the emitted conical light beam has an opening angle of  $\arccos(12/13) = 0.39$  rad. For the target, we choose the stereographic domain  $\mathcal{Y} = [-0.2, 0.2] \times [-0.1, 0.1]$  and intensity  $I_t(\hat{\mathbf{t}}(\mathbf{y}))$  corresponding to the gray-scale image of the TU/e logo given in Figure 6a. Both intensities are scaled such that the global luminous flux balance holds, i.e., Equation 35 for  $\mathcal{A} = \mathcal{X}$  and  $\mathbf{m}(\mathcal{A}) = \mathcal{Y}$ . Both stereographic coordinates are projections from the south pole; cf. Equation 21.



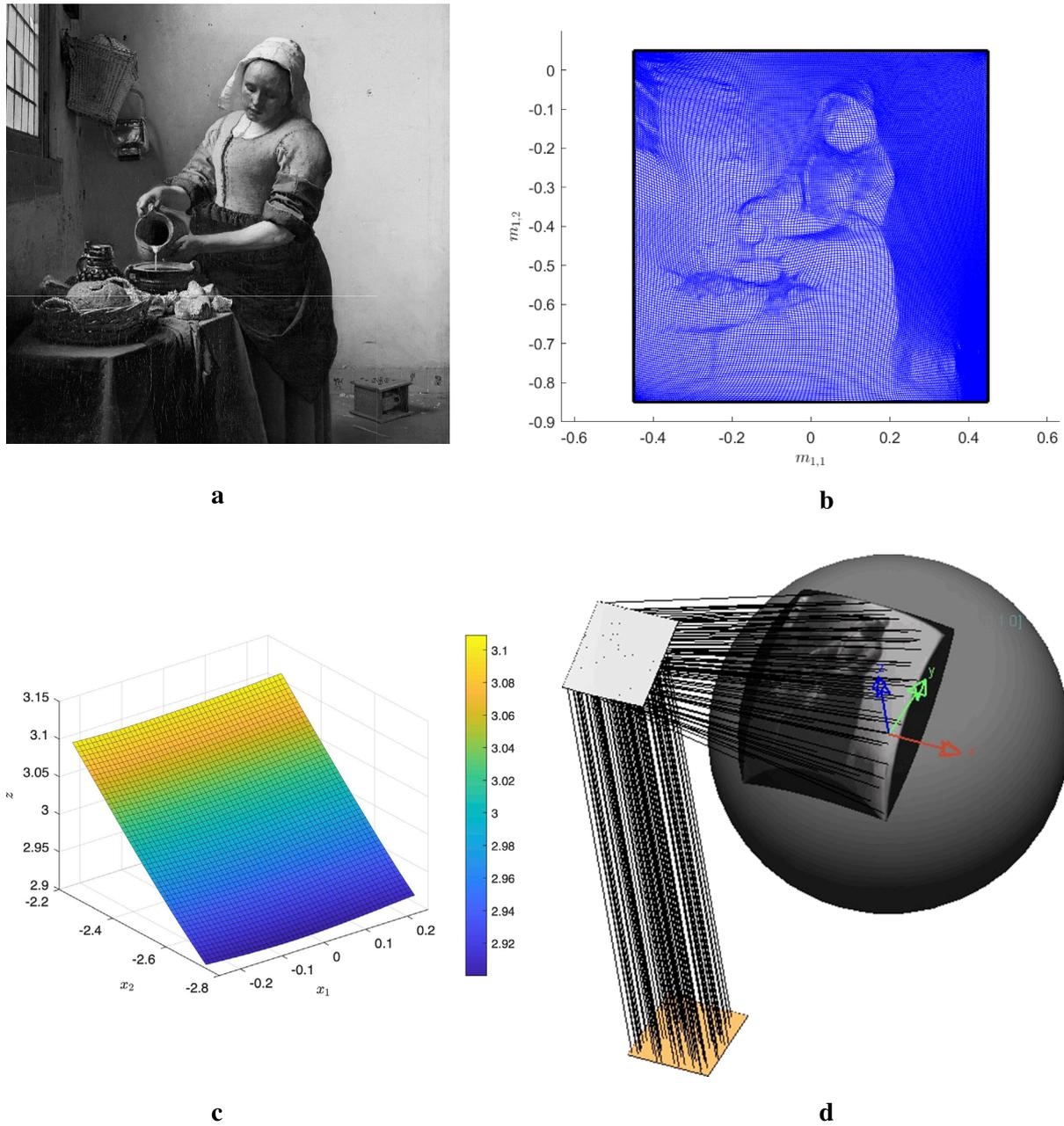
The refractive index of the lens  $n = 1.5$ , and we enforce a unique solution for the lens surface by setting  $u(\mathbf{x}_c) = 1$ , and therefore,  $u_1(\mathbf{x}_c) = 0$ , corresponding to the central light ray with stereographic source coordinate  $\mathbf{x}_c = 0$  and direction vector  $\hat{\mathbf{s}}(\mathbf{x}_c) = \hat{\mathbf{e}}_z$ ; cf. Equation 34b.

For space discretization of the BVPs in (45) and (46), we cover the source domain with a uniform  $201 \times 201$  grid and evaluate 500 iterations of scheme (42)–(43), where the associated functionals are defined in Equation 41. We choose  $\alpha = 10^{-3}$ . The computational time, on a laptop with Intel Core i7 - 11800 H 2.30 GHz processor and a RAM of 16.0 GB, is 443.65 s (7.39 min). The optical mapping, computed as an image of the source grid, is shown in Figure 6b, and clearly exhibits the TU/e-logo. The computed freeform lens with a ray traced target intensity pattern is shown in Figure 6c. This intensity pattern is computed with the commercial software code LightTools using  $5 \times 10^6$  rays and clearly resembles the desired intensity. The corresponding (normalized) 3D

illuminance distribution, projected on the plane  $z = 100$ , is shown in Figure 6d.

## 7.2 Parallel-to-near-field reflector

We consider a reflector that converts a uniform source emittance  $M(\mathbf{x}, 0)$  into the illuminance  $E(\mathbf{r}(y))$  corresponding to the gray-scale image of the painting ‘the Milkmaid’ by Johannes Vermeer as shown in Figure 7a, projected on a spherical target surface. The source domain  $\mathcal{X} = [-0.25, 0.25] \times [-2.25, -2.75]$  and the target domain is located on the northern hemisphere of the unit sphere given by  $z = v(y) = \sqrt{1 - |y|^2}$  with  $\mathcal{Y} = [-0.42, 0.42] \times [-0.82, 0.02]$ . The parameters  $\mathbf{x}$  and  $\mathbf{y}$  are both Cartesian coordinates. The emittance and illuminance are scaled such that the global flux balance holds, i.e., Equation 38 for  $\mathcal{A} = \mathcal{X}$  and  $\mathbf{m}(\mathcal{A}) = \mathcal{Y}$ . We enforce a unique solution for the location of the reflector by setting  $u(\mathbf{x}_c) =$



**FIGURE 7** Parallel-to-near-field reflector: target illuminance pattern, mapping, sag map, and freeform surface with ray-traced target illuminance. **(a)** “Milkmaid” as target illuminance pattern. **(b)** Mapping displayed in the orthogonal projection plane of the target surface. **(c)** Sag map of the freeform reflector. **(d)** Free-form reflector with ray-traced target illuminance pattern on a sphere.

$u_1(\mathbf{x}_c) = 3$ , where  $\mathbf{x}_c = (0, -2.5)^T$  is the center point of the source domain  $\mathcal{X}$ .

To discretize the BVPs in Equations 45, 48, we cover the source domain  $\mathcal{X}$  with a uniform  $201 \times 201$  grid and 500 times apply the iteration scheme defined in Equations 49, 43. The associated functionals are defined in Equation 41, where we choose  $\alpha = 0.05$ , and Equation 47. On the same laptop/processor, the computational time is 496.63 s (8.28 min). The optical map as image of the

uniform rectangular source grid is shown in Figure 7b. Clearly, the girl shown in Figure 7a is recognizable in optical mapping as grid points come closer together in regions of high contrast, showing her contours. The sag map of the reflector is shown in Figure 7c.

To verify the reflector computed by the least-squares algorithm, we used LightTools. Implementing the reflector coordinates into this code and specifying the uniformly distributed rectangular source

along with the curved target surface, we obtained the reflector system illustrated in Figure 7d, where only 50 rays are shown. The resulting target intensity was obtained tracing  $5 \times 10^7$  rays. The image of the ‘Milkmaid’ is clearly visible, confirming the validity of the reflector computed by the least-squares algorithm.

## 8 Concluding remarks

We have presented a systematic and generic derivation of the governing equations for inverse optical design. Using Hamilton’s characteristic functions, we could derive a geometrical equation defining the shape/location of the optical surface(s). From this equation, applying concepts from optimal transport theory, we derived equations for the optical map. To close the model, we presented a generic conservation law of luminous flux. Combining all equations, we derived the luminous flux constraint. Subsequently, we elaborated the generic model in three different models of increasing complexity, on the basis of three example optical systems. We briefly outlined numerical least-squares methods for all models and demonstrated their performance for a few examples.

We intend to expand our research on inverse methods along the following lines. First, we will adjust our numerical methods to incorporate two target distributions; second, we will develop mathematical models and numerical methods for catadioptric systems, where light rays propagate via different paths from the source to target, and finally, and probably the most challenging, we intent to generalize our models and numerics to optical systems with finite sources, as opposed to zero-étendue sources.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

## Author contributions

JT: writing–original draft and writing–review and editing. KM: writing–original draft and writing–review and editing. MA:

writing–review and editing. LK: writing–review and editing. PB: writing–review and editing. WI: writing–review and editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was partially supported by the Dutch Research Council (*Dutch*: Nederlandse Organisatie voor Wetenschappelijk Onderzoek–NWO) through the (grants P15-36 and P21-20).

## Acknowledgments

The authors would like to thank R. Beltman from Signify for executing all LightTools simulations.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Filosa C. *Phase space ray tracing for illumination optics*. Eindhoven: Eindhoven University of Technology (2018). *PhD Thesis*.
- Santambrogio F. *Optimal transport for applied mathematicians*. Birkäuser, New York: Springer (2015).
- Gutiérrez C. *The Monge–Ampère equation*. 2nd ed. Switzerland: Birkhäuser: Springer International Publishing AG (2016).
- Romijn L. *Generated Jacobian equations in freeform optical design*. Eindhoven: Eindhoven University of Technology (2021). *PhD Thesis*.
- Froese B. A numerical method for the elliptic Monge–Ampère equation with transport boundary conditions. *SIAM J Scientific Comput* (2012) 34:A1432–59. doi:10.1137/110822372
- Benamou J, Froese B, Oberman A. Numerical solution of the optimal transportation problem using the Monge–Ampère equation. *J Comput Phys* (2014) 260:107–26. doi:10.1016/j.jcp.2013.12.015
- Kawecki E, Lakkis O, Pryer T. A finite element method for the Monge–Ampère equation with transport boundary conditions. *arXiv preprint arXiv:1807* (2018).
- Caffarelli L, Oliker V. Weak solutions of one inverse problem in geometric optics. *J Math Sci* (2008) 154:39–49. doi:10.1007/s10958-008-9152-x
- Kitagawa J, Mérigot Q, Thibert B. Convergence of a Newton algorithm for semi-discrete optimal transport. *J Eur Math Soc* (2019) 21:2603–51. doi:10.4171/jems/889
- Mérigot Q, Thibert B. Mirrors, lenses and Monge–Ampère equations. *Eur Math Soc Mag* (2021) 120:16–28. doi:10.4171/mag-21
- Gallouet A, Mérigot Q, Thibert B. A damped Newton algorithm for generated Jacobian equations. *Calculus of Variations and Partial Differential Equations* (2021) 61:49–59. doi:10.1007/s00526-021-02147-7
- Caboussat A, Glowinski R, Sorensen D. A least-squares method for the numerical solution of the Dirichlet problem for the elliptic Monge–Ampère equation in

dimension two. *ESAIM: Control, Optimisation and Calculus of Variations* (2013) 19:780–810. doi:10.1051/cocv/2012033

13. Prins C, Beltman R, ten Thije Boonkkamp J, IJzerman W, Tukker T. A least-squares method for optimal transport using the Monge-Ampère equation. *SIAM J Scientific Comput* (2015) 37:B937–61. doi:10.1137/140986414
14. Yadav N, ten Thije Boonkkamp J, IJzerman W. A Monge-Ampère problem with non-quadratic cost function to compute freeform lens surfaces. *J Sci Comput* (2019) 80:475–99. doi:10.1007/s10915-019-00948-9
15. Romijn L, ten Thije Boonkkamp J, IJzerman W. Freeform lens design for a point source and far-field target. *J Opt Soc Am A* (2019) 36:1926–39. doi:10.1364/josaa.36.001926
16. Romijn L, ten Thije Boonkkamp J, IJzerman W. Inverse reflector design for a point source and far-field target. *J Comput Phys* (2020) 408:109283. doi:10.1016/j.jcp.2020.109283
17. Romijn L, Anthonissen M, ten Thije Boonkkamp J, IJzerman W. The generating function approach for double freeform lens design. *J Opt Soc Am A* (2021) 38:356–68. doi:10.1364/OE.438920
18. Romijn L, ten Thije Boonkkamp J, Anthonissen M, IJzerman W. An iterative least-squares method for generated Jacobian equations in freeform optical design. *SIAM J Scientific Comput* (2021) 43:B298–B322. doi:10.1364/OE.438920
19. Luneburg R. *Mathematical theory of optics*. Berkeley and Los Angeles: University of California Press (1966).
20. Born M, Wolf E. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. 5th ed. Oxford: Pergamon Press (1975).
21. Kevorkian J. *Partial differential equations: analytical solution techniques*. Pacific Grove, California: Brooks/Cole (1990).
22. Guillen N. A primer on generated Jacobian equations: geometry, optics, economics. *Notices Am Math Soc* (2019) 66:1401–11. doi:10.1090/noti1956
23. Yadav N. *Monge-Ampère problems with non-quadratic cost function: application to freeform optics*. Eindhoven: Eindhoven University of Technology (2018). *PhD Thesis*.
24. Ries H, Rable A. Edge-ray principle of nonimaging optics. *J Opt Soc Am A* (2002) 19:590–5. doi:10.1364/OE.438920
25. Chaves J. *Introduction to nonimaging optics*. 2nd ed. Boca Raton, Florida: CRC Press (2016).
26. Anthonissen M, ten Thije Boonkkamp J, IJzerman W. Unified mathematical framework for a class of fundamental freeform optical systems. *Opt Express* (2021) 29:31650–64. doi:10.1364/OE.438920
27. Prins C. *Inverse methods for illumination optics*. Eindhoven: Eindhoven University of Technology (2014). *PhD Thesis*.
28. Beltman R, ten Thije Boonkkamp J, IJzerman W. A least-squares method for the inverse reflector problem in arbitrary orthogonal coordinates. *J Comput Phys* (2018) 367:347–73. doi:10.1016/j.jcp.2018.04.041
29. Yadav N, Romijn L, ten Thije Boonkkamp J, IJzerman W. A least-squares method for the design of two-reflector optical systems. *J Phys Photon* (2019) 1:034001. doi:10.1088/2515-7647/ab2db3
30. Courant R, Hilbert D. *Methods of mathematical physics*, 2. New York: John Wiley and Sons (1989). doi:10.1002/9783527617210
31. Wang C, Luo Y, Li H, Zang Z, Xu Y, Han Y, et al. Extended ray-mapping method based on differentiable ray-tracing for non-paraxial and off-axis freeform illumination lens design. *Optics Express* (2023) 31:30066–30078.
32. Heemels A, de Koning B, Möller M, Adam A. Optimizing freeform lenses for extended sources with algorithmic differentiable ray tracing and truncated hierarchical B-splines. *Optics Express* (2024) 32:9730–9746.
33. Doskolovich L, Bykov D, Andreev E, Bezus E, Olikier V. Designing double freeform surfaces for collimated beam shaping with optimal mass transportation and linear assignment problems. *Optics Express* (2018) 26:24602–24613.
34. Doskolovich L, Bykov D, Mingazov A, Bezus E. Optimal mass transportation and linear assignment problems in the design of freeform refractive optical elements generating far-field irradiance distributions. *Optics Express* (2019) 27:13083–13097.
35. Brix K, Hafizogullari Y, Platen A. Designing illumination lenses and mirrors by the numerical solution of Monge-Ampère equations. *J Opt Soc Am A* (2015) 32:2227–2236.
36. Feng Z, Froese B, Liang R. Freeform illumination optics construction following an optimal transport map. *Appl. Optics* (2016) 55:4301–4306.
37. Bösel C, Gross H. Single freeform surface design for prescribed input wavefront and target irradiance. *J Opt Soc Am* (2017) 34:1490–1499.
38. Bösel C, Gross H. Double freeform illumination design for prescribed wavefronts and irradiances. *J Opt Soc Am* (2017) 35:236–243.

## Nomenclature

### Scalar variables

$c(x, y)$	cost function
$d(\mathbf{P}, \mathbf{Q})$	distance between the points $\mathbf{P}$ and $\mathbf{Q}$
$d_{\text{opt}}(\underline{r}_s, \underline{r}_t)$	optical distance, i.e., optical path length $[\mathbf{P}_s, \mathbf{P}_t]$ as function of the position vectors $\underline{r}_s$ and $\underline{r}_t$
$E(\underline{r}(y))$	illuminance at a near-field target
$f(\mathbf{x})$	generic flux density of the source
$g(\mathbf{y})$	generic flux density of the target
$G(\mathbf{x}, \mathbf{y}, \mathbf{w})$	generating function
$H(\mathbf{x}, \mathbf{y}, \mathbf{w})$	inverse of generating function
$I_s(\underline{\hat{s}}(\mathbf{x}))$	intensity of the point source
$I_t(\underline{\hat{t}}(\mathbf{y}))$	intensity of the far-field target
$J(\underline{\hat{v}}(z))$	factor in the area element $dS(z)$ on $S^2$ generated by $\underline{\hat{v}}$ and parameterized by $z$
$M(\underline{r}(\mathbf{x}))$	emittance of the planar source
$n(\underline{r})$	refractive index field
$s$	arc length along a curve $C$ representing a ray
$T(\underline{p}_s, \underline{p}_t)$	angular characteristic (function)
$u_1(\mathbf{x})$	Kantorovich potential for source domain
$u_2(\mathbf{y})$	Kantorovich potential for target domain
$V(\underline{q}_s, \underline{q}_t)$	point characteristic (function)
$W(\underline{q}_s, \underline{p}_t)$	mixed characteristic (function) of the first kind
$W^*(\underline{p}_s, \underline{q}_t)$	mixed characteristic (function) of the second kind
$z$	coordinate along the optical axis
$\varphi(\underline{r})$	phase of an electromagnetic wave
$[\mathbf{P}_s, \mathbf{P}_t]$	optical path length between the points $\mathbf{P}_s$ and $\mathbf{P}_t$

### Vectors

$\underline{p}(z)$	2D momentum vector of a ray
$\underline{q}(z)$	2D position vector of a ray
$\underline{r}$	3D position vector of an arbitrary point
$\underline{\hat{s}}$	3D unit direction/tangent vector of a ray at the source plane
$\underline{\hat{t}}$	3D unit direction/tangent vector of a ray at the target plane
$\underline{\hat{v}}$	3D unit direction/tangent vector at an arbitrary point of a ray
$\mathbf{x}$	2D vector parameterizing the source domain
$\mathbf{y}$	2D vector parameterizing the target domain