Inverse reflector design for a point source and far-field target

(Preprint submitted to SIAM Journal on Scientific Computing (SISC))

by

L.B. Romijn, J.H.M. ten Thije Boonkkamp, W.L. IJzerman

Centre for Analysis, Scientific computing and Applications
Department of Mathematics and Computer Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven, The Netherlands
ISSN: 0926-4507
INVERSE REFLECTOR DESIGN FOR A POINT SOURCE AND FAR-FIELD TARGET∗

L. B. ROMIJN†, J. H. M. TEN THIJE BOONKKAMP†, AND W. L. IJZERMAN‡

Abstract. We present a method for the design of a single freeform reflector that converts the light distribution of a point source to a desired light distribution in the far field. Using the geometrical-optics law of reflection and requiring energy conservation, this optical design problem can be represented by a generalized Monge-Ampère equation for the shape of the reflector with transport boundary condition. We use a generalized least-squares algorithm that can handle a non-quadratic cost function in the corresponding optimal transport problem. The algorithm first computes the optical map and subsequently constructs the optical surface. We introduce a novel approach to handle the transport boundary condition and show with numerical examples that it leads to an improved solution. We also demonstrate that the algorithm can generate reflector surfaces for a number of complicated examples.

Key words. geometrical optics, optimal transport, non-quadratic cost function, generalized Monge-Ampère equation, least-squares

AMS subject classifications. 35J66, 35J96, 49K20, 65K10, 65N99

1. Introduction. A central problem in illumination optics is to determine an optical system that transforms a given light distribution on a source domain into a required distribution on a target domain. As an example of an application, an LED lamp contains one or multiple reflectors, lenses, diffusers and/or absorbers. These surfaces need to be designed in such a way that they redistribute the light from the LED into the required light output pattern of the lamp. Broadly, the methods for optical system design can be categorized as either forward or inverse methods.

Forward methods compute the target distribution from a known source distribution and optical system, most commonly using Monte-Carlo ray tracing techniques. The design of the optical system can be improved by making modifications to the optical elements and subsequently evaluating the output target distribution via ray tracing. Drawbacks of forward methods are that ray tracing can be slow if high precision is required and that the approach to create an improved design is often based on trial and error [12]. Filosa et al. [7, 8] recently developed a new ray tracing method based on the phase space representation of the source and target domains, which improves the accuracy and reduces computation time of the classical approach.

Inverse methods directly compute the optical system converting the light from the source into the specified output. One approach for the inverse design of freeform (i.e., non-axially symmetric) optical surfaces uses the principles of geometrical optics and conservation of energy to derive a partial differential equation for the location of the optical surface. With the laws of reflection and/or refraction of geometrical optics it is possible to construct an optical mapping that connects coordinates on the source and target domains. Substituting the mapping into the relation for energy conservation

∗Submitted to the journal’s Computational Methods in Science and Engineering section December 12, 2018.

†CASA, Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands (l.b.romijn@tue.nl, j.h.m.tenthijeboonkkamp@tue.nl).

‡Signify Research, High Tech Campus 7, 5656 AE Eindhoven, The Netherlands (wilbert.ijzerman@signify.com).
leads to a fully nonlinear second order elliptic partial differential equation, which is a generalized Monge-Ampère equation [23, p. 282].

Generalized Monge-Ampère equations frequently arise in the field of optimal transport [22, 23, 1]. The optimal transport problem is concerned with finding an optimal transport plan which minimizes a transport cost functional – i.e., an integral of a cost function weighted with the source distribution – under the constraint of energy conservation. More specifically, in optical design problems we seek an optimal transport plan or a mapping, such that all the light from the source is transmitted to the target and traverses the least optical path length. Each optical system corresponds to a different Monge-Ampère equation and associated cost function. For optical systems with quadratic cost functions, the optical surface can be described by a Monge-Ampère equation in its usual definition: a fully nonlinear second order partial differential equation which is linear in the determinant of the Hessian matrix [19, 28].

Currently, nearly all existing algorithms only work for quadratic cost functions, for which existence and uniqueness of the solution can also be proven [10, 1, 20, 28]. A typical example of an optical system that does not fit in this framework is a reflector that transforms the energy distribution of a point source into a required far-field intensity. This system has a logarithmic cost function in the optimal transport problem and the corresponding partial differential equation is nonlinear in the Hessian. Generalized Monge-Ampère equations with conditions for the existence, uniqueness and smoothness of a solution to reflector-type problems with a point source were derived in [18, 24, 25, 11].

Currently, there exist only a few numerical algorithms for generalized Monge-Ampère equations, which involve mostly quadratic cost functions. Here we mention a few numerical algorithms for the point source specifically, which do not assume any symmetry of the optical surface. Oliker [16] solves the problem using the method of supporting ellipsoids for a discrete target distribution and introduces an iterative optimization algorithm in [14]. Ries et al. [21] derive a set of partial differential equations for a point source using curvatures of wave fronts but do not present details of a numerical solution. Glimm and Oliker [13] develop a heuristic constructive algorithm for the dual Monge-Kantorovich problem using a linear programming approach. Fournier et al. [9] extend Oliker’s method of supporting ellipsoids [14] and construct 3D reflectors that produce continuous illuminance distributions using Monte-Carlo ray tracing. Canavesi et al. [5] replace the Monte-Carlo ray tracing in this algorithm by implementing a flux estimation method which calculates the intersection points between triplets of ellipsoids. Wu et al. [26] derive the Monge-Ampère equation for a lens surface and solve the equations using standard finite differences and Newton iteration. Finally, Brix et al. [3, 4] derive the Monge-Ampère equation for a point source with a near-field target and use a collocation method with a tensor-product B-spline basis to calculate reflectors and lenses capable of producing a detailed image on a near-field projection screen.

In this paper, we present a numerical algorithm for a point source with corresponding logarithmic cost function in the optimal transport problem. We show that an optical mapping can be derived both via the law of reflection and via the cost function, using stereographic coordinates to represent the source and target domains. The corresponding generalized Monge-Ampère equation can be formulated in terms of the cost function and solved numerically using a generalization of the least-squares approach inspired by the methods presented in [19, 20, 29]. The method works by first computing the optical map in an iterative procedure which minimizes the defect in the energy balance. It also imposes a transport boundary condition by minimizing
the deviation of the boundary of the optical map to the boundary of the target. Upon convergence of the iterative procedure, the location of the optical surface is calculated from the mapping also in a least-squares sense.

This paper is structured as follows. In Section 2 we present the derivation of the Monge-Ampère equation for the freeform reflector with corresponding logarithmic cost function. In Section 3 the generalized least-squares method is described. In Section 3.1 we present an improvement to the method described in [19, 20] to impose the transport boundary condition and show that it leads to improved convergence at corners of the target domain. In Section 4 we apply the algorithm to a few test cases. Notably, we challenge the algorithm to compute the reflector surface which transforms light of a point source into a complicated image on a screen in the far field and verify the result by a ray tracing method. Finally, we make some concluding remarks in Section 5.

2. Mathematical formulation. In this section we derive the partial differential equation which describes the shape of the reflector that transforms a beam of light originating from a point source into a specified output intensity distribution in the far field.

2.1. Geometrical formulation of the optical map. The geometry of the optical system transforming light from a point source into a far-field intensity is shown schematically in Figure 1a. We assume that the light source is a point located at the origin $O$ of the Cartesian coordinate system with $(x, y, z) \in \mathbb{R}^3$. In spherical coordinates the source emits light radially outward in the direction $\hat{s} = \hat{e}_r$. We use hats to denote unit vectors. The reflector surface is described by the parametrization $r(\phi, \theta) = u(\phi, \theta) \hat{e}_r$, where $u(\phi, \theta) > 0$ is the radial parameter that describes the location of the reflector surface, $0 \leq \phi \leq \pi$ is the zenith and $0 \leq \theta < 2\pi$ is the azimuth in the spherical coordinate system. The source has a given emittance $f(\phi, \theta)$ [lm/sr]. The target in the far field has a specified output distribution $g(\psi, \chi)$ [lm/sr], with respect to a different set of spherical coordinates $(\psi, \chi)$, with zenith $0 \leq \psi \leq \pi$ and azimuth $0 \leq \chi < 2\pi$, taking the origin to be the reflector surface approximated as a point in space (far-field approximation).

2.1.1. The geometrical optics approach. The mapping $m$ can be determined by tracing a typical ray through the optical system. We consider an incident ray propagating in the direction $\hat{s} = \hat{e}_r$, which intercepts the reflector $R$ and reflects off in direction $\hat{t}$. The unit surface normal of the parametrized reflector surface $r(\phi, \theta) = u(\phi, \theta) \hat{e}_r$, directed towards the point source, is given by

$$\hat{n} = \left| \frac{\partial r}{\partial \phi} \times \frac{\partial r}{\partial \theta} \right| = -\hat{e}_r + \nabla_r u \sqrt{1 + |\nabla_r u|^2},$$

with

$$\nabla_r u = \frac{1}{u} \frac{\partial u}{\partial \phi} \hat{e}_\phi + \frac{1}{u \sin(\phi)} \frac{\partial u}{\partial \theta} \hat{e}_\theta,$$

which is the gradient of $u$ restricted to the surface $r = $ constant. Using the vectorial law of reflection $\hat{t} = \hat{s} - 2 (\hat{s} \cdot \hat{n}) \hat{n}$, we obtain the direction $\hat{t}$ of the reflected ray

$$\hat{t} = \hat{e}_r + \frac{2}{1 + |\nabla_r u|^2} (-\hat{e}_r + \nabla_r u).$$
Point source with emittance \( f(\phi, \theta) \) [lm/sr]

Target with required intensity \( g(\psi, \chi) \) [lm/sr]

Fig. 1. Single reflector converting the emittance \( f(\phi, \theta) \) of a point source into a far-field target intensity \( g(\psi, \chi) \) and a diagram for the coordinate transformations of the incoming and outgoing rays.

Since the vectors \( \hat{s} = (s_1, s_2, s_3)^T \) and \( \hat{t} = (t_1, t_2, t_3)^T \) in Cartesian coordinates are both defined on the unit sphere \( S^2 \), we can express the first two components in terms of the third component. For this reason it is convenient to perform coordinate transformations from spherical to stereographic. Figure 1b schematically illustrates possible coordinate transformations for the source and target domains. We define

\[
\begin{align*}
\mathbf{x}(\hat{s}) &= \left( \begin{array}{c} x_1 \\ x_2 \end{array} \right) = \frac{1}{1 + s_3} \left( \begin{array}{c} s_1 \\ s_2 \end{array} \right) = \frac{1}{1 + \cos(\phi)} \left( \begin{array}{c} \sin(\phi) \cos(\theta) \\ \sin(\phi) \sin(\theta) \end{array} \right), \\
\mathbf{y}(\hat{t}) &= \left( \begin{array}{c} y_1 \\ y_2 \end{array} \right) = \frac{1}{1 - t_3} \left( \begin{array}{c} t_1 \\ t_2 \end{array} \right) = \frac{1}{1 - \cos(\psi)} \left( \begin{array}{c} \sin(\psi) \cos(\chi) \\ \sin(\psi) \sin(\chi) \end{array} \right),
\end{align*}
\]

with corresponding inverse projections

\[
\begin{align*}
\hat{s}(\mathbf{x}) &= \hat{e}_r = \frac{1}{1 + |x|^2} \left( \begin{array}{c} 2x_1 \\ 2x_2 \\ 1 - |x|^2 \end{array} \right), \\
\hat{t}(\mathbf{y}) &= \frac{1}{1 + |y|^2} \left( \begin{array}{c} 2y_1 \\ 2y_2 \\ -1 + |y|^2 \end{array} \right).
\end{align*}
\]

We represent the incoming rays \( \hat{s} \) by using a stereographic projection from the south pole \((0, 0, -1)\) onto the plane \( z = 0 \). The stereographic projection in (2.3) is undefined at the south pole, and we consider \( s_3 \neq -1 \) and \( 0 \leq \phi < \pi \). On the other hand, for the outgoing rays we use a stereographic projection from the north pole \((0, 0, 1)\) with the reflector surface as origin in the far-field approximation. The stereographic projection in (2.4) is undefined at the north pole, and we consider \( t_3 \neq 1 \) and \( 0 < \psi \leq \pi \). The reason for using the south pole for the incoming rays is that if we consider the point source to emit a conical beam of rays in the upward direction, as in Figure 1, we obtain a bounded source domain in stereographic coordinates which we can easily discretize. Likewise, we choose the north pole for the outgoing rays to ensure that the stereographic projection is always defined, assuming the reflector surface does not reflect rays upwards parallel to the \( z \)-axis.
We can derive the mapping \( y = m(x) \) from Equation (2.2) and outline the broad steps below. First, let \( u_1 = \log(u) \), so that

\[
\nabla_r u = \nabla_r (e^{u_1}) = \frac{\partial u_1}{\partial \phi} \hat{e}_\phi + \frac{1}{\sin(\phi)} \frac{\partial u_1}{\partial \theta} \hat{e}_\theta.
\]

Assuming \( u_1 = u_1(x) \) and applying the chain rule to the partial derivatives of \( u_1 \) gives

\[
\nabla_r u = \left( \frac{\partial u_1}{\partial x_1} \frac{\partial x_1}{\partial \phi} + \frac{\partial u_1}{\partial x_2} \frac{\partial x_2}{\partial \phi} \right) \hat{e}_\phi + \frac{1}{\sin(\phi)} \left( \frac{\partial u_1}{\partial x_1} \frac{\partial x_1}{\partial \theta} + \frac{\partial u_1}{\partial x_2} \frac{\partial x_2}{\partial \theta} \right) \hat{e}_\theta.
\]

Using Equation (2.3) we can rewrite the partial derivatives with respect to \( \hat{x} \) and applying (2.5) to \( \hat{r} \) as follows:

\[
\frac{\partial x_1}{\partial \phi} = \frac{\cos(\theta)}{1 + \cos(\phi)} = \frac{s_1}{(1 + s_3) \sqrt{s_1^2 + s_2^2}}, \quad \frac{\partial x_1}{\partial \theta} = \frac{\sin(\phi) \sin(\theta)}{1 + \cos(\phi)} = \frac{s_2}{1 + s_3},
\]

\[
\frac{\partial x_2}{\partial \phi} = \frac{\sin(\phi)}{1 + \cos(\phi)} = \frac{s_2}{(1 + s_3) \sqrt{s_1^2 + s_2^2}}, \quad \frac{\partial x_2}{\partial \theta} = \frac{\sin(\phi) \cos(\theta)}{1 + \cos(\phi)} = \frac{s_1}{1 + s_3}.
\]

Second, we can express the basis vectors \( \hat{e}_\phi \) and \( \hat{e}_\theta \) in \( \hat{x} \)-coordinates (\( \hat{e}_r = \hat{s} \), see (2.5)) as follows:

\[
\hat{e}_\phi = \begin{pmatrix} \cos(\phi) \cos(\theta) \\ \cos(\phi) \sin(\theta) \\ -\sin(\phi) \end{pmatrix} = \frac{1}{\sqrt{s_1^2 + s_2^2}} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix},
\]

\[
\hat{e}_\theta = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \end{pmatrix} = \frac{1}{\sqrt{s_1^2 + s_2^2}} \begin{pmatrix} s_2 \\ s_1 \end{pmatrix} - 2 \cos(\phi) \sin(\phi) \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\]

Third, substituting \( \sin(\phi) = \sqrt{s_1^2 + s_2^2} = 2 |x|/(1 + |x|^2) \) and (2.8) into (2.7) and using the inverse projection in (2.5), we get an expression for \( \nabla_r u \) in terms of \( x \) and the partial derivatives of \( u_1 \) with respect to \( x \). Fourth, substituting this expression into (2.2) and applying (2.5) to \( \hat{e}_r = \hat{s} \), we find the outgoing ray \( \hat{t}(x) \) expressed in \( \hat{x} \)-coordinates. Finally, transforming \( \hat{t}(x) \) to stereographic coordinates \( y \) using (2.4), we derive \( y = m(x) \) as

\[
y = \frac{||\nabla u_1|^2 - 4 (1 + |x|^2)^{-1}|}{(\nabla u_1 \cdot |x|)^2 + 4 (1 + |x|^2)^{-1} + \nabla u_1 \cdot x},
\]

where \( \nabla u_1 \) is the gradient of \( u_1 \) with respect to \( x \) and

\[
A(x) = 2I - 2\begin{pmatrix} x_1^2 - x_2^2 \\ 2 x_1 x_2 \\ -x_1^2 + x_2^2 \end{pmatrix} = 2I - 2\rho^2 \begin{pmatrix} \cos(2\xi) \\ \sin(2\xi) \end{pmatrix} = \begin{pmatrix} \cos(2\xi) \\ \sin(2\xi) \end{pmatrix},
\]

letting \( x_1 = \rho \cos(\xi) \) and \( x_2 = \rho \sin(\xi) \) in polar stereographic coordinates, \( \rho \geq 0 \) and \( 0 \leq \xi < 2\pi \), and \( I \) is the \( 2 \times 2 \) identity matrix. We define our source domain \( S \) as the supporting domain of \( \hat{f}(x) = f(\phi(x), \theta(x)) \), and our target domain \( T \) as the supporting domain of \( \hat{g}(y) = g(\psi(y), \chi(y)) \). We refer to \( m : S \rightarrow T \) as the optical map \( y = m(x) \) from the source set of stereographic coordinates \( S \) to the target set of stereographic coordinates \( T \).
2.1.2. The cost function approach. The cost function of the optical system can be derived using Fermat’s principle, which states that a ray connecting the source $O$ with a point on the reflected ray has stationary optical path length $L(s, t)$. Let us consider an incident ray propagating from the source $O$ in the direction $\hat{s}$, which intercepts the reflector $R$ and reflects off in the direction $\hat{t}$, as shown in Figure 2. Let $P$ be the intersection point between the incident ray and the reflector and $Q$ the intersection point between the reflected ray and the wave front perpendicular to it and going through $O$. Then $L(s, t) = u(s) + d(P, Q)$, where we introduce $u(s)$ to denote $u(\phi, \theta)$ and $d(P, Q)$ is the distance between $P$ and $Q$. We see that $d(P, Q)$ is the projection of $u(s)$ on $\hat{t}$ and $d(P, Q) = -(\hat{s} \cdot \hat{t}) u(s)$, $\hat{s} \cdot \hat{t} < 0$, and $L(s, t) = u(s)(1 - \hat{s} \cdot \hat{t})$.

Next, we consider the tangent parabola of the reflector at $P$ for the reflected direction $\hat{t}$ as the locus of points in the plane of incidence that are equidistant from the focal point $O$, which is the point source, and directrix with focal parameter $\rho(\hat{t})$. By construction, the optical path length $L(s, t)$ is equal to the focal parameter $\rho(\hat{t})$ of the parabolic reflector and $\rho(\hat{t}) = u(s) (1 - \hat{s} \cdot \hat{t})$. Defining the new functions $\tilde{u}_1(s) = \log u(s)$ and $\tilde{u}_2(\hat{t}) = \log(1/\rho(\hat{t}))$ results in the relation

$$\tilde{u}_1(s) + \tilde{u}_2(\hat{t}) = \frac{1}{\rho(\hat{t})} = c(s, \hat{t}).$$

The function $\tilde{c}(s, \hat{t})$ is known as the cost function in optimal transport. The derivation described above closely resembles the method used by Oliker [17], who constructs the reflector surface as an envelope of paraboloids. An alternative method to derive the cost function is based on Hamilton’s characteristic functions, measuring the optical path length between specified source and target planes [15].

Transforming the cost function in (2.10) to the stereographic coordinates in Equation (2.3) and (2.4), defining

$$u_1(x) = \tilde{u}_1(s) - \log(1 + |x|^2), \quad u_2(y) = \tilde{u}_2(\hat{t}) + \log\left(\frac{2}{1 + |y|^2}\right),$$

we arrive at the relation

$$u_1(x) + u_2(y) = -\log (1 - 2 x \cdot y + |x|^2 |y|^2) = c(x, y).$$
In summary, we have derived a relation of the form \( u_1(x) + u_2(y) = c(x, y) \) for the location of the optical surface \( u \), where \( u_1(x) = \log(u/(1 + |x|^2)) \) and \( c(x, y) \) is a non-quadratic cost function in optimal transport theory.

Equation (2.12) has many solutions for \( u_1(x) \), \( u_2(y) \), and consequently for \( u(x) \). We can find a solution by assuming that \( u_1 \) and \( u_2 \) are either c-convex or c-concave functions [23, p. 54]. The surfaces \( u_1 \) and \( u_2 \) are c-convex if

\[
\begin{align}
(2.13a) & \quad u_1(x) = \max_{y \in T} (c(x, y) - u_2(y)), & \forall x \in S \\
(2.13b) & \quad u_2(y) = \max_{x \in S} (c(x, y) - u_1(x)), & \forall y \in T,
\end{align}
\]

which we call the maximum solution, or c-concave if

\[
\begin{align}
(2.14a) & \quad u_1(x) = \min_{y \in T} (c(x, y) - u_2(y)), & \forall x \in S \\
(2.14b) & \quad u_2(y) = \min_{x \in S} (c(x, y) - u_1(x)), & \forall y \in T,
\end{align}
\]

which we call the minimum solution. For a continuously differentiable function \( c \in C^1(S \times T) \), the c-convex/concave functions \( u_1, u_2 \) are Lipschitz continuous [27, p. 60], and the mapping \( y = m(x) \) is implicitly given by the critical point of Equation (2.13b) or (2.14b), i.e.,

\[

\nabla u_1(x) = \nabla_x c(x, m(x)),
\]

where \( \nabla_x c \) is the gradient of \( c \) with respect to \( x \), under the condition that the Jacobi matrix \( C \), defined by

\[
\begin{align}
(2.16a) & \quad C = C(x, m(x)) = D_{xy}c = \begin{pmatrix}
\frac{\partial^2 c}{\partial x_1 \partial y_1} & \frac{\partial^2 c}{\partial x_1 \partial y_2} \\
\frac{\partial^2 c}{\partial x_2 \partial y_1} & \frac{\partial^2 c}{\partial x_2 \partial y_2}
\end{pmatrix} \\
& \quad = \frac{4}{N(x,y)^2} (-x + |x|^2 y y T + \frac{2}{N(x,y)} (I - 2 x y T)
\end{align}
\]

is invertible, with

\[
(2.16b) \quad N(x,y) = 1 - 2 x \cdot y + |x|^2 |y|^2 = |x - y|^2 + (1 - |x|^2) (1 - |y|^2).
\]

In fact, we can verify by substituting the expression for \( c \) in (2.12) into (2.15) and solving for \( y \), that this implicit mapping is identical to the mapping derived in Equation (2.9) via the law of reflection.

By straightforward evaluation we note that the matrix \( C \) is a skew-symmetric matrix with equal diagonal elements, \( \det(C) > 0 \) (which holds if \( s \cdot t \neq 1 \) and the reflector changes the direction of the light rays, as we have assumed above), and

\[
(2.17) \quad \text{tr}(C) = \frac{4 (-1 + (m(x) - J m(x)) \cdot x)}{N(x,m(x))^2}
\]

where we introduce the symplectic matrix

\[
(2.18) \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},
\]

which acts upon a vector by rotating it over \( \pi/2 \) in the counterclockwise direction.
Noting that \(|(\mathbf{m}(x) + J \mathbf{m}(x))| = |(\mathbf{m}(x) - J \mathbf{m}(x))|\) and \((\mathbf{m}(x) + J \mathbf{m}(x)) \perp (\mathbf{m}(x) - J \mathbf{m}(x))\), we derive that
\[
\begin{align*}
\text{(2.19a)} & \quad (\mathbf{m}(x) + J \mathbf{m}(x)) \cdot x = |\mathbf{m}(x) + J \mathbf{m}(x)| |x| \cos(\beta), \\
\text{(2.19b)} & \quad (\mathbf{m}(x) - J \mathbf{m}(x)) \cdot x = |\mathbf{m}(x) - J \mathbf{m}(x)| |x| \sin(\beta),
\end{align*}
\]
where \(\beta\) is the angle between the vectors \(x\) and \(\mathbf{m}(x) + J \mathbf{m}(x)\). Hence, \(\text{tr}(C) \geq 0\) in (2.17) if both expressions in (2.19a) and (2.19b) are smaller than or equal to 1, or both are greater than or equal to 1.

A sufficient condition for a maximum/minimum solution requires
\[
\text{(2.20)} \quad D^2 u_1(x) - D_{xx} C(x, \mathbf{m}(x)) = P
\]
to be positive/negative semi-definite, respectively. Hence, for a c-convex pair we require \(\text{tr}(P) \geq 0\) and \(\det(P) \geq 0\). On the other hand, for a c-concave pair we need \(\text{tr}(P) \leq 0\) and \(\det(P) \geq 0\). Note that \(P\) is a symmetric matrix.

Differentiating Equation (2.15) again with respect to \(x\) gives
\[
\text{(2.21)} \quad D_{xx} C(x, \mathbf{m}(x)) + C \mathbf{m}(x) = D^2 u_1(x),
\]
where \(C \mathbf{m}(x)\) is the 2×2 Jacobi matrix of \(\mathbf{m}\) with respect to \(x\). Combining Equation (2.20) and (2.21) gives
\[
\text{(2.22)} \quad C \mathbf{m}(x) = P.
\]

2.2. Energy conservation. By transferring the light from source to target we require that all light from the source ends up at the target and energy is conserved, i.e.,
\[
\text{(2.23)} \quad \int_A f(\phi, \theta) \, dS(\phi, \theta) = \int_{\hat{t}(A)} g(\psi, \chi) \, dS(\psi, \chi),
\]
for an arbitrary set \(A \subset S^2\) and image set \(\hat{t}(A) \subset S^2\). If we substitute \(s = \hat{s}(x)\) and \(\hat{t} = \hat{t}(y)\) from Equation (2.5) we can write Equation (2.23) as
\[
\text{(2.24)} \quad \int_{x(A)} \hat{f}(x) \left| \frac{\partial \hat{s}}{\partial x_1} \right| \left| \frac{\partial \hat{s}}{\partial x_2} \right| \, dx = \int_{y(\hat{t}(A))} \hat{g}(y) \left| \frac{\partial \hat{t}}{\partial y_1} \right| \left| \frac{\partial \hat{t}}{\partial y_2} \right| \, dy,
\]
where \(\hat{f}(x) = f(\phi, \theta)\) and \(\hat{g}(y) = g(\psi, \chi)\). We can derive easily that
\[
\text{(2.25)} \quad \left| \frac{\partial \hat{s}}{\partial x_1} \right| \left| \frac{\partial \hat{s}}{\partial x_2} \right| = \frac{4}{(1 + |x|^2)^2}, \quad \left| \frac{\partial \hat{t}}{\partial y_1} \right| \left| \frac{\partial \hat{t}}{\partial y_2} \right| = \frac{4}{(1 + |y|^2)^2}.
\]
Substituting (2.25) and the mapping \(y = \mathbf{m}(x)\) into the energy conservation relation (2.24) gives
\[
\text{(2.26)} \quad \int_{x(A)} \hat{f}(x) \left( \frac{4}{(1 + |x|^2)^2} \right) \, dx = \int_{x(A)} \hat{g}(\mathbf{m}(x)) \left( \frac{4}{(1 + |\mathbf{m}(x)|^2)^2} \right) \det(\mathbf{m}(x)) \, dx.
\]
Using (2.22) we can rewrite Equation (2.26) to the generalized Monge-Ampère equation
\[
\text{(2.27)} \quad \det(\mathbf{m}(x)) = \frac{\hat{f}(x) (1 + |\mathbf{m}(x)|^2)^2}{\hat{g}(\mathbf{m}(x)) (1 + |\mathbf{m}(x)|^2)^2} = \frac{\det(P(x))}{\det(C(x, \mathbf{m}(x)))}.
\]
Our goal is to find a mapping \( y = m(x) : S \rightarrow T \) for a particular source domain \( S \) and target domain \( T \). We require global energy conservation (2.24) with \( x(A) = S \) and \( y(\hat{t}(\hat{A})) = T \) under any mapping \( y = m(x) : S \rightarrow T \). I.e., all light from the source is mapped to the target and the reflector functions as a perfect mirror. For this reason, we cannot choose any combination of \( S, T, \tilde{f}(x) \) and \( \tilde{g}(y) \). We eliminate this dependency by normalizing the source intensity \( \tilde{f}(x) \) by the total intensity over \( S \) (2.28a)
\[
\Gamma(S) = \int_S \tilde{f}(x) \frac{4}{(1 + |x|^2)^2} \, dx,
\]
and normalizing the target intensity \( \tilde{g}(y) \) by the total intensity over \( T \) (2.28b)
\[
\Gamma(T) = \int_T \tilde{g}(y) \frac{4}{(1 + |y|^2)^2} \, dy.
\]

Using (2.28) we rewrite the generalized Monge-Ampère equation in (2.27) to
\[
(2.29a) \quad \det(Dm(x)) = \frac{\det(P)}{\det(C)} \frac{\tilde{f}(x) (1 + |m(x)|^2)^2 \Gamma(T)}{\tilde{g}(m(x)) (1 + |x|^2)^2 \Gamma(S)} = \tilde{F}(x, m(x)),
\]

where we introduce \( \tilde{F}(x, m(x)) \) to equal the total right hand side. We define the corresponding boundary condition to (2.29a) as
\[
(2.29b) \quad m(\partial S) = \partial T,
\]

stating that all light from the boundary of the source \( S \) is mapped to the boundary of the target \( T \) [19, 20].

3. Numerical method. We first compute the mapping \( m \) by using a generalized least-squares method [19, 20, 29]. The mapping \( m \) can be calculated efficiently by an iterative procedure that involves finding the numerical solution of a constrained minimization problem, imposing the transport boundary condition using skew projections on line segments, and computing the numerical solution of a linear elliptic boundary value problem. Upon convergence the location of the optical surface \( u \) is calculated from the mapping also in a least-squares sense.

To compute a c-convex solution for \( u_1 \), we can write the Monge-Ampère equation (2.29) as the matrix equation \( C(x, m(x))Dm(x) = P(x) \), where \( P(x) \) is a symmetric, positive definite (SPD) matrix satisfying \( \det(P(x)) = \tilde{F}(x, m(x)) \det(C(x, m(x))) \). We write \( m = m(x) \) and enforce the matrix equation by minimizing the functional
\[
(3.1) \quad J_I[m, P] = \frac{1}{2} \int_S \| C \, Dm - P \|^2 \, dx.
\]

under the constraint \( \det(P) = \tilde{F} \det(C) \). The norm used is the Frobenius norm. To impose the transport boundary condition we minimize the functional
\[
(3.2) \quad J_B[m, b] = \frac{1}{2} \oint_{\partial S} |m - b|^2 \, ds,
\]

where \( | \cdot | \) denotes the \( l_2 \)-norm. We combine the functionals \( J_I \) and \( J_B \) by a weighted average as
\[
(3.3) \quad J[m, P, b] = \alpha J_I(m, P) + (1 - \alpha) J_B[m, b],
\]
with $0 < \alpha < 1$.

Starting from an initial guess $m^0$ and cost function matrix $C(\cdot, m^0)$ we perform the iteration:

\begin{align}
(3.4a) & \quad b^{n+1} = \arg\min_{b \in B} J_B[m^n, b], \\
(3.4b) & \quad P^{n+1} = \arg\min_{P \in \mathcal{P}(m^n)} J_I[m^n, P], \\
(3.4c) & \quad m^{n+1} = \arg\min_{m \in M} J[m, P^{n+1}, b^{n+1}],
\end{align}

where the minimization steps are performed over the spaces

\begin{align}
(3.5a) & \quad \mathcal{B} = \{ b \in C^1(\partial S)^2 \mid b(x) \in \partial T \}, \\
(3.5b) & \quad \mathcal{P}(m) = \{ P \in C(\mathcal{S})^{2 \times 2} \mid P \text{ SPD/SND}, \det(P) = \tilde{F}(\cdot, m) \det(C(\cdot, m)) \}, \\
(3.5c) & \quad \mathcal{M} = C^2(S)^2.
\end{align}

After each iteration we update the matrix $C(\cdot, m^n)$. As initial guess $m^0$ we map the smallest bounding box enclosing $S$ to the smallest bounding box enclosing $T$. Without loss of generality we assume the bounding box of the source $S$ has rectangular shape $[a_{\min}, a_{\max}] \times [b_{\min}, b_{\max}]$ and the bounding box of the target $T$ has rectangular shape $[c_{\min}, c_{\max}] \times [d_{\min}, d_{\max}]$. In order to find a $c$-convex $u_1$, we specify the initial guess $m^0 = (m_1^0, m_2^0)^T$ as

\begin{align}
(3.6a) & \quad m_1^0 = \frac{x_1 - a_{\min}}{a_{\max} - a_{\min}} c_{\max} + \frac{a_{\max} - x_1}{a_{\max} - a_{\min}} c_{\min}, \\
(3.6b) & \quad m_2^0 = \frac{x_2 - b_{\min}}{b_{\max} - b_{\min}} d_{\max} + \frac{b_{\max} - x_2}{b_{\max} - b_{\min}} d_{\min}.
\end{align}

The corresponding Jacobi matrix $Dm^0$ is symmetric positive definite.

If we would like to find a $c$-concave $u_1$ instead, we choose a slightly different initial guess $m^0 = (m_1^0, m_2^0)^T$ given by

\begin{align}
(3.7a) & \quad m_1^0 = \frac{x_1 - a_{\min}}{a_{\max} - a_{\min}} c_{\min} + \frac{a_{\max} - x_1}{a_{\max} - a_{\min}} c_{\max}, \\
(3.7b) & \quad m_2^0 = \frac{x_2 - b_{\min}}{b_{\max} - b_{\min}} d_{\min} + \frac{b_{\max} - x_2}{b_{\max} - b_{\min}} d_{\max}.
\end{align}

The corresponding Jacobi matrix $Dm^0$ is symmetric negative definite.

Using the initial mapping in (3.6) or (3.7), we can derive that $P = C(\cdot, m^0)Dm^0$ is positive or negative semi-definite, respectively, if $\text{tr}(C(\cdot, m^0)) \geq 0$. This holds since, in either case, $Dm^0$ is a diagonal matrix and the diagonal elements of $C$ are equal. Moreover, substituting $m^0$ into (2.17) we find that $\text{tr}(C(\cdot, m^0)) \geq 0$ is satisfied.

We remark that for a $c$-convex $u_1$ and initial guess (3.6) we need to choose a source domain such that $|x| \neq 1$ for all $x \in S$ in order to ensure that $N(\cdot, m) \neq 0$ in (2.16b) and $C(\cdot, m^0)$ is invertible.

We discretize the source $S$ using a standard rectangular $N_1 \times N_2$ grid for some $N_1, N_2 \in \mathbb{N}$ and introduce $x_{ij} = (x_{1,i}, x_{2,j})$ with

\begin{align}
(3.8a) & \quad x_{1,i} = a_{\min} + (i - 1) h_1, \quad h_1 = \frac{a_{\max} - a_{\min}}{N_1 - 1}, \quad i = 1, \ldots, N_1, \\
(3.8b) & \quad x_{2,j} = b_{\min} + (j - 1) h_2, \quad h_2 = \frac{b_{\max} - b_{\min}}{N_2 - 1}, \quad j = 1, \ldots, N_2.
\end{align}
After setting the initial guess \( m^0 \) we perform the minimization steps in (3.4) and subsequently update \( C \) in every iteration. The minimization steps (3.4a), (3.4b), and (3.4c) are explained in detail in Section 3.1, 3.2, 3.3, respectively. Finally, we compute the location of the reflector surface \( u \) as described in Section 3.4.

### 3.1. Minimization procedure for \( b \)

We assume \( m = m^n \) is given and we need to minimize \( J_{B}(m, b) \) over \( b \in B \). The minimization can be performed point-wise because the integrand does not depend on derivatives of \( b \). We drop the indices \( n \) and \( n + 1 \) for ease of notation. We denote \( m_{ij} = m(x_{ij}) \), \( b_{ij} = b(x_{ij}) \) and perform the minimization

\[
\min_{b_{ij} \in \mathbb{R}} \frac{1}{2} |m_{ij} - b_{ij}|^2.
\]

We discretize the boundary of \( \mathcal{T} \) using points \( z_k \in \partial \mathcal{T} \), \( k = 1, 2, ..., N_b \) with increasing index clockwise along the boundary, and we define \( z_{N_b+1} = z_1 \). We connect adjacent points by line segments \((z_k, z_{k+1})\) and determine the “closest” line segment to each \( m_{ij} \).

First, we calculate the outward normals \( n_k \) associated with each boundary point \( z_k \) as

\[
n_k = \frac{1}{2} \left( \frac{1}{|z_{k+1} - z_k|} J(z_{k+1} - z_k) + \frac{1}{|z_k - z_{k-1}|} J(z_k - z_{k-1}) \right),
\]

where \( J \) is defined as in (2.18) and \( n_k \) is the normal pointing in the average direction of the normals to the two adjacent line segments \( z_{k+1} - z_k \) and \( z_k - z_{k-1} \). It bisects the angle between the adjacent segments.

We define \( l_k : y = z_k + \lambda n_k \), \( k = 1, 2, ..., N_b \), as the points \( y \in \mathbb{R}^2 \) which are on the normal line through \( z_k \), and let \( \ell \) be the line through \( m_{ij} \), parallel to the segment \((z_k, z_{k+1})\). We let \( p_k \) and \( p_{k+1} \) be the intersection of \( \ell \) with \( l_k \) and \( l_{k+1} \), respectively, as shown in Figure 3.

The intersection points \( p_k \) and \( p_{k+1} \) of \( \ell \) with \( l_k \) and \( l_{k+1} \), respectively, can be written as

\[
\begin{align}
    p_k &= z_k + k_1 n_k, \\
p_{k+1} &= z_{k+1} + k_2 n_{k+1},
\end{align}
\]

with \( k_1 \) and \( k_2 \) constants. We can solve for \( k_1 \) and \( k_2 \) by setting the slope of \((m_{ij} - p_k)\) and \((m_{ij} - p_{k+1})\) equal to the slope of \((z_{k+1} - z_k)\), respectively, giving

\[
\begin{align}
k_1 &= \frac{\det(m_{ij} - z_k, z_{k+1} - z_k)}{\det(n_k, z_{k+1} - z_k)}, \\
k_2 &= \frac{\det(m_{ij} - z_{k+1}, z_{k+1} - z_k)}{\det(n_{k+1}, z_{k+1} - z_k)}.
\end{align}
\]

The normal lines \( l_k \) and \( l_{k+1} \) may cross as illustrated in Figure 3b. We determine whether this occurs by evaluating if both of the following two conditions hold for each line segment

\[
\begin{align}
d_k(p_{k+1}) d_k(z_{k+1}) &< 0, \\
d_{k+1}(p_k) d_{k+1}(z_k) &< 0,
\end{align}
\]
The first condition (3.12a) checks whether \( \mathbf{y} \) for which (3.12c) equals 0. Likewise, the line \( l_{k+1} \) through \( z_{k+1} \) and \( p_{k+1} \) are the points \( \mathbf{y} \) for which (3.12d) equals 0. The first condition (3.12a) checks whether \( p_{k+1} \) and \( z_{k+1} \) are located on opposite sides of the line segment \( p_k - z_k \). The second condition (3.12b) checks whether \( p_k \) and \( z_k \) are located on opposite sides of \( p_{k+1} - z_{k+1} \). Together they determine whether \( p_k - z_k \) crosses \( p_{k+1} - z_{k+1} \).

The projection \( m^k_{i,j} \) of \( m_{i,j} \) on the line segment \((z_k, z_{k+1})\) is given by

\[
\begin{align*}
(3.13a) & \quad m^k_{i,j} = z_k + t_k (z_{k+1} - z_k), \\
(3.13b) & \quad t_k = \frac{|m_{i,j} - \mathbf{p}^*|}{|p_{k+1} - p_k|}.
\end{align*}
\]

If (3.12) do not hold, i.e., the normal lines do not cross as in Figure 3a, we define \( \mathbf{p}^* = p_k \). If (3.12) is true, i.e., the normal lines cross as in Figure 3b, we use \( \mathbf{p}^* = p_{k+1} \).

For a given line segment \((z_k, z_{k+1})\), we check if \( m_{i,j} \) is located in between \( p_k \) and \( p_{k+1} \) by evaluating

\[
0 \leq (m_{i,j} - p_k) \cdot (p_{k+1} - p_k) \leq |p_{k+1} - p_k|^2.
\]

If this does not hold, we set the parameter \( t_k \) in (3.13b) to an arbitrarily large number.

The skew projection (3.13) ensures that the ratio between the distances \(|m_{i,j} - \mathbf{p}^*|\) and \(|p_{k+1} - p_k|\) is the same as the ratio between the distances \(|m^k_{i,j} - z_k|\) and \(|z_{k+1} - z_k|\).

Finally, we calculate the points \( \mathbf{b}_{ij} \) as

\[
(3.15) \quad \mathbf{b}_{ij} = \arg\min_{m^k_{i,j}} |m^k_{i,j} - m_{i,j}|^2.
\]

This procedure is repeated for all \( \mathbf{x}_{ij} \in \partial \mathbf{S} \). In Section 4.1 we compare this boundary method to a previous method which uses perpendicular projections on line segments [19, 20].

### 3.2. Minimization procedure for \( P \).

We assume \( \mathbf{m} = \mathbf{m}^n \) is fixed. The minimization of \( J_{ij}[\mathbf{m}, \mathbf{P}] \) can be performed point-wise because the integrand does not depend on derivatives of \( \mathbf{P} \). We minimize \( \| Q - \mathbf{P} \| \) for each grid point \( \mathbf{x}_{ij} \in \mathbf{S} \), where \( Q = \mathbf{CD} = (q_{ij}) \), and \( \mathbf{D} = (d_{ij}) \) is the central difference approximation of \( \mathbf{Dm} \). This gives rise to the minimization problem

\[
\begin{align*}
(3.16a) & \quad \text{minimize } H_S(p_{11}, p_{22}, p_{12}) = \frac{1}{2} \| Q_S - \mathbf{P} \|^2, \\
(3.16b) & \quad \text{subject to } \det(\mathbf{P}) = p_{11} p_{22} - p_{12}^2 = \tilde{F}(\cdot, \mathbf{m}) \det(\mathbf{C}(\cdot, \mathbf{m})),
\end{align*}
\]

where \( Q_S = \frac{1}{2}(Q + Q^T) \). We have replaced \( Q \) by its symmetric part \( Q_S \), since this gives the same minimizer \((p_{11}, p_{22}, p_{12})\). For a \( c \)-convex solution we impose the additional constraint

\[
(3.16c) \quad \text{tr}(\mathbf{P}) = p_{11} + p_{22} \geq 0,
\]
while for a c-concave solution we require

\[(3.16d) \quad \text{tr}(P) = p_{11} + p_{22} \leq 0.\]

Possible solutions are stationary points of the Lagrangian \( \Lambda = \Lambda(p_{11}, p_{22}, p_{12}; \mu) \) defined by

\[(3.17) \quad \Lambda(p_{11}, p_{22}, p_{12}; \mu) = \frac{1}{2} \| Q_S - P \|^2 + \mu \left( \det(P) - \tilde{F}(\cdot, m) \det(C(\cdot, m)) \right).\]

Setting all partial derivatives of \( \Lambda \) to 0 results in the algebraic system

\[(3.18a) \quad p_{11} + \lambda p_{22} = q_{11},\]
\[(3.18b) \quad \lambda p_{11} + p_{22} = q_{22},\]
\[(3.18c) \quad (1 - \lambda) p_{12} = \frac{1}{2} (q_{12} + q_{21}),\]
\[(3.18d) \quad p_{11} p_{22} - p_{12}^2 = \tilde{F}(\cdot, m) \det(C(\cdot, m)).\]

where \( \lambda = \mu / \det(C(\cdot, m)) \). We can always select the possible minimizers that satisfy the nonlinear constraints (3.16b), and the inequality constraint (3.16c) [29]. Details of the remaining analytical procedure to select the minimizers are presented in [19, 20, 29].

### 3.3. Minimization procedure for \( m \)

We minimize the combined functional \( J[m, P, b] \) over all \( m \in \mathcal{M} \). This step cannot be performed point-wise and we com-
We compute the first variation $\delta J[m, P, b](\eta)$ with respect to $m$ for $\eta \in M$, i.e.
\[
\delta J[m, P, b](\eta) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left( J[m + \epsilon \eta, P, b] - J[m, P, b] \right)
\]
\[
= \lim_{\epsilon \to 0} \left[ \frac{\alpha}{2} \int_S 2(CDm - P) : CD\eta + \epsilon |CD\eta|^2 \, dx \right.
\]
\[
+ \frac{1 - \alpha}{2} \int_{\partial S} 2(m - b) \cdot \eta + \epsilon |\eta|^2 \, ds \bigg] \quad \text{(3.19)}
\]
\[
= \alpha \int_S (CDm - P) : CD\eta \, dx + (1 - \alpha) \int_{\partial S} (m - b) \cdot \eta \, ds.
\]

The minimizer is given by $\delta J[m, P, b](\eta) = 0$ for all $\eta \in M$. Using Gauss’ law and the fundamental lemma of calculus of variations [6], we obtain the coupled elliptic boundary value problem
\[
\nabla \cdot (C^TC \, Dm) = \nabla \cdot (C^T \, P), \quad x \in S, \quad \text{(3.20a)}
\]
\[
(1 - \alpha)m + \alpha (C^TC \, \nabla m) \cdot \hat{n} = (1 - \alpha) b + \alpha C \cdot \hat{P}\hat{n}, \quad x \in \partial S. \quad \text{(3.20b)}
\]

We solve this system for $m$ using the standard central difference discretization.

### 3.4. Computation of $u$.\,\,\,\,\,Upon convergence of the iterative procedure for the mapping $m$, we can calculate the location of the optical surface from Equation (2.15).\,\,\,\,\,We compute the generalized least-squares solution by minimizing the functional
\[
I[u_1] = \frac{1}{2} \int_S |\nabla_x c(\cdot, m) - \nabla u_1|^2 \, dx. \quad \text{(3.21)}
\]

We cannot perform this step point-wise and analogous to the minimization procedure for $m$, we compute the first variation of $\delta I[u_1](v)$ with respect to $u_1$ for $v \in C^2(S)$ as
\[
\delta I[u_1](v) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left( I[u_1 + \epsilon v] - I[u_1] \right)
\]
\[
= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_S 2 \left( \nabla u_1 - \nabla_x c(\cdot, m) \right) \cdot \nabla v \, dx
\]
\[
+ \int_S \nabla u_1 \cdot \nabla_x c(\cdot, m) \cdot \nabla v \, dx. \quad \text{(3.22)}
\]

The minimizer is given by $\delta I[u_1](v) = 0$ for all $v \in C^2(S)$. Once more, using Gauss’ law and the fundamental lemma of calculus of variations [6], we obtain the boundary value problem
\[
\Delta u_1 = \nabla \cdot \nabla_x c(\cdot, m), \quad x \in S, \quad \text{(3.24a)}
\]
\[
\nabla u_1 \cdot \hat{n} = \nabla_x c(\cdot, m) \cdot \hat{n}, \quad x \in \partial S. \quad \text{(3.24b)}
\]

This is a Neumann problem which has a unique solution up to an additive constant, and consequently a corresponding finite difference matrix with incomplete rank. We calculate a unique least-squares solution by using the $QR$-decomposition of the finite difference matrix. The compatibility condition is satisfied for this Neumann problem up to discretization errors.

Finally, we compute the location of the optical surface $u$ as a function of $x_{ij}$ using $u(x) = (1 + |x|^2)^{u_1(x)}$, since $u_1 = \log(u/(1 + |x|^2))$. We transform our
stereographic source coordinates defined in (2.3) to Cartesian coordinates denoting \((x, y, z)^T = u(s) \hat{s}\) and plot the reflector surface using

\[
x = \frac{2u(x)}{1 + x_1^2 + x_2^2}, \quad y = \frac{2u(x)}{1 + x_1^2 + x_2^2}, \quad z = \frac{u(x)(1 - x_1^2 - x_2^2)}{1 + x_1^2 + x_2^2},
\]

(cf. (2.5), for all \(x_{ij} = (x_{1,i}, x_{2,j}) \in S\).

4. Numerical results. First, we compare the boundary method described in Section 3.1 to the method described in [19, 20]. Subsequently, we test our generalized least-squares method on a number of example problems. We compute the correct surface for a number of mappings corresponding to flat or spherical reflector surfaces. We can also compute the surface that transforms light into an image on a screen in the far field.

4.1. Comparison of boundary methods. We compare the minimization procedure for \(b\) described in Section 3.1 to the perpendicular projection previously presented in [20]. In our test problem, the source domain is given by the square \(S = [-0.5, 0.5]^2\) and the target domain by the parallelogram centred in the origin spanned by a vector \(w\) of length 1 parallel to the \(y_1\)-axis and a vector \(u\) of length 1 which forms an angle of \(\pi/6\) with \(w\), i.e., \(T = \{ y \in \mathbb{R}^2 \mid -1/2 + y_2 \sqrt{3} \leq y_1 \leq 1/2 + y_2 \sqrt{3}, -1/4 \leq y_2 \leq 1/4\}\). We have taken a uniform source distribution \(f(x) = 1\) and uniform target distribution \(g(y) = 1\), and in order to satisfy global energy conservation we calculate \(\Gamma(S) = 16/\sqrt{3} \cot^{-1}(\sqrt{3})\) and \(\Gamma(T) \approx 1.553\) using (2.28). \((\Gamma(T)\) is evaluated by using MATLAB’s inbuilt double numerical integration method). We solve the boundary value problem (2.29) for a 100 \times 100 grid and 4 corner boundary points. With 4 boundary points each side of the parallelogram is one segment \((z_k, z_{k+1})\). We use the initial guess \(m^0\) given in (3.7) for a c-concave \(u_1\), in order to ensure \(C(\cdot, m^0)\) is invertible. The convergence of \(J_1\) and \(J_B\) is shown in Figure 4c and the final mapping is shown in Figure 4a and Figure 4b. Using a skew projection results in a better and faster convergence in the corner points of the target domain. Figure 4d shows the influence of varying \(N_b\) on the final values of \(J_1\) and \(J_B\), after 2000 iterations. The orthogonal and skew projections become equivalent when the number of boundary points reaches approximately 200, due to the fact that the skew projection becomes more and more orthogonal as the lengths of the line segments decrease.

4.2. Exact solution: tilted flat surface. One way to test our algorithm is by pre-computing the target domains corresponding to a surface of which we can derive the correct mapping. For example, we consider a tilted flat reflector surface \(ax + by + cz = d\), with given constants \(a, b, c, d\), and \(c \neq 0, d > 0\). By substituting \(x = us_1, y = us_2, z = us_3\) and \(u = e^{\pi i / 2} (1 + |x|^2)^{1/2}\), and changing to stereographic coordinates using (2.5), we can derive an expression for the c-convex/c-concave solution \(u_1\)

\[
(4.1) \quad u_1(x) = \log(d) - \log\left(2a \cdot x + c (1 - |x|^2)\right), \quad a = \begin{pmatrix} a \\ b \end{pmatrix}.
\]

Note that we require our source domain \(S\) to lie within the interior of the ellipse \(c^2 |x - a_c^2| = |a|^2 + c^2\). Subsequently calculating \(\frac{\partial u}{\partial x_1}\) and \(\frac{\partial u}{\partial x_2}\), and substituting into Equation (2.9) we obtain the mapping as

\[
(4.2a) \quad y = m(x) = \frac{A(x)}{c^2 + 2c a \cdot x + |a|^2 |x|^2},
\]

where

\[
A(x) = \left(\frac{1 + |x|^2}{c^2 + 2c a \cdot x + |a|^2 |x|^2}\right)^{1/2}.
\]
Fig. 4. “Square to parallelogram” problem: the mapping on a $100 \times 100$ grid with $a = 0.2$ and $N_b = 4$. We have $\tilde{f}(x) = 1$ and $\tilde{g}(y) = 1$.

where

\[ A(x) = \begin{pmatrix} -a^2 + b^2 + c^2 & -2 a b \\ 2 a b & a^2 - b^2 + c^2 \end{pmatrix}. \]

We consider a square source domain $S = [-0.2, 0.2]^2$ and a $100 \times 100$ grid with $N_b = 1000$. We choose $a = 2$, $b = 1$, $c = 3$, and $d = 1$. Using Equation (4.2) we can compute the target domain $T$, the corresponding target boundary, and the correct mapping. We solve the boundary value problem (2.29) for a uniform source and target distribution, i.e., $f(x) = 1$ on $S$ and $g(y) = 1$ on $T$. In order to satisfy global energy conservation we require $\Gamma(S) = \Gamma(T)$. In fact, we know that $\Gamma(S) = \Gamma(T)$ is satisfied already by our choice of target domain $T$ corresponding to the mapping of a flat surface and choice of equal uniform intensities. We use the initial guess $m^0$ given in (3.6) for a $c$-convex $u_1$. Figure 5 shows the results after 1000 iterations. The converged mapping is displayed in Figure 5a and the reflector surface upon convergence in Figure 5b. The absolute difference between the reflector surface...
and the exact solution is shown in Figure 5c, which is of the order $10^{-5}$. The error convergence is given in Figure 5d. Choosing a value of $\alpha = 0.2$ ensures that $J_I$ and $J_B$ are close together. The value of $N_b$ does not influence the convergence if chosen large enough.

![Graphs](a) (b) (c) (d)

**Fig. 5.** "Tilted flat surface" problem: the mapping and reflector surface after 1000 iterations on a $100 \times 100$ grid with $N_b = 1000$. We have $\tilde{f}(x) = 1$ and $\tilde{g}(y) = 1$. The absolute difference between the computed reflector surface and exact solution is displayed in Figure 5c.

### 4.3. Picture on a projection screen in the far field.

As a final example we challenge our algorithm to compute a reflector surface that converts the light from a point source into a far-field target intensity distribution corresponding to a picture. The source domain in stereographic coordinates is given by the square $S = [-0.2, 0.2]^2$ and has a uniform light distribution $f(x) = 1$ and $\Gamma(S) = 8\sqrt{2/13}\cot^{-1}(\sqrt{26})$. The reflected rays are projected on a screen $P$ in the far field, parallel to the $xz$-plane. The required illumination $L(\xi, \eta)$ [lm/m²], with $(\xi, \eta)$ the Cartesian coordinates on the projection screen, is derived from the gray scale values of a picture of a koala, see Figure 6. The target distribution $\tilde{g}(y)$ is a deformation of the illuminance $L(\xi, \eta)$; the conversion from $L(\xi, \eta)$ to $\tilde{g}(y)$ is explained in detail in [19, 2].
The gray scale values of the picture prescribe the illuminance. However, the conversion from the coloured image to gray scale values creates black regions in the target distribution for which \( \tilde{g}(y) = 0 \). To avoid division by 0 in the right hand side of Equation (2.29a), we increase values of \( \tilde{g}(y) \) which are below a threshold of 5% of its maximum value to this threshold. We determine \( \Gamma(T) \) by dividing the region on the projection screen into quadrants and approximating the integral of \( L(\xi, \eta) \) over each quadrant using the 2D composite trapezoidal rule.

We use the least-squares algorithm to compute the optical map \( m \) and the reflector surface. We discretize the source domain using a 1000 \( \times \) 1000 grid and use \( \alpha = 0.5 \) to give the domain and boundary equal weights in the minimization procedure. We use the initial guess \( m^0 \) given in (3.6) to compute a \( c \)-convex \( u_1 \). The optical map, plotted using a coarsened version of the source grid, the reflector surface, and convergence results are shown in Figure 7. The error \( J_I \) in the interior converges to a larger value than the error \( J_B \) in the boundary, as the finite grid is unable to capture the finest details of the koala.

Subsequently, we validated the resulting reflector image using ray tracing. We traced approximately 3000 \( \times \) 3000 uniformly distributed rays from source to far field. The resulting target illuminance \( L(\xi, \eta) \) is plotted in Figure 7. The ray trace image closely resembles the original picture, showing details such as the hairs on the koala’s coat.

Fig. 6. “Square-to-koala” problem: the original image (left) and target distribution \( \tilde{g}(y) \) and target boundary (right).

5. Concluding remarks. In this paper, we presented a method to compute the shape of a single reflector which transforms the light distribution of a point source to a desired far-field distribution. First, we derived the optical mapping using both the geometrical-optics law of reflection and the logarithmic cost function in the optimal transport problem. We subsequently combined the optical mapping with energy conservation to derive the generalized Monge-Ampère equation. We developed a generalized least-squares method for the numerical solution of this nonlinear second order elliptic partial differential equation. The method first computes the optical map in an iterative procedure, and proceeds by using the converged mapping to calculate
the location of the optical surface. We introduced an improved method to deal with the transport boundary condition leading to better convergence at corner points of the target domain. We tested the algorithm for a couple of examples. For an exact solution, such as flat reflector surfaces, we can check the accuracy of the algorithm. We can also construct a reflector converting the light of a point source into a picture on a screen in the far field. Currently, there are no known algorithms for the point source which are capable of calculating reflector surfaces that produce this level of detail in the far field.

A natural extension is to apply this algorithm to the generalized Monge-Ampère equation corresponding to a lens surface. In fact, the formulation of the Monge-Ampère equation using the cost function in (2.12) and corresponding solution using c-convexity theory allows for an extension of the method to any optical system with a continuously differentiable cost function and a relation of the form $u_1(x) + u_2(y) = c(x, y)$, where $u_1$ and $u_2$ are some geometrical parameters describing the optical element(s). We aim to apply the method to real optical design problems, using a different variety of source and target distributions, e.g., target intensity profiles for street lighting purposes.
REFERENCES


<table>
<thead>
<tr>
<th>Number</th>
<th>Author(s)</th>
<th>Title</th>
<th>Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-04</td>
<td>J. de Graaf</td>
<td>Some Mathematical Comments on Thermostatics</td>
<td>August ‘18</td>
</tr>
<tr>
<td>18-05</td>
<td>B.S. van Lith, J.H.M. ten Thije, Boonkamp, W.L. IJzerman</td>
<td>Active flux schemes on moving meshes with applications to geometric optics</td>
<td>September ‘18</td>
</tr>
<tr>
<td>18-06</td>
<td>A.J. Vromans, A.A.F. van de Ven, A. Muntean</td>
<td>Periodic homogenization of a pseudo-parabolic equation via a spatial-temporal decomposition</td>
<td>October ‘18</td>
</tr>
<tr>
<td>18-07</td>
<td>T.G.J. Beelen, T. Verhoeff</td>
<td>Determining the essentially different partitions of all Japanese convex tangrams</td>
<td>October ‘18</td>
</tr>
<tr>
<td>18-08</td>
<td>L.B. Romijn, J.H.M. ten Thije, Boonkamp, W.L. IJzerman</td>
<td>Inverse reflector design for a point source and far-field target</td>
<td>December ‘18</td>
</tr>
</tbody>
</table>