A Monge-Ampère-solver for free-form reflector design

SIAM J. SCI. COMPUT.
Vol. 36, No. 3, pp. B640–B660

A MONGE–AMPÈRE-SOLVER FOR FREE-FORM REFLECTOR DESIGN*

C. R. PRINS†, J. H. M. TEN THIJE BOONKKAMP†, J. VAN ROOSMALEN†, W. L. IJZERMAN‡, AND T. W. TUKKER§

Abstract. In this article we present a method for the design of fully free-form reflectors for illumination systems. We derive an elliptic partial differential equation of the Monge–Ampère type for the surface of a reflector that converts an arbitrary parallel beam of light into a desired intensity output pattern. The differential equation has an unusual boundary condition known as the transport boundary condition. We find a convex or concave solution to the equation using a state of the art numerical method. The method uses a nonstandard discretization based on the diagonalization of the Hessian. The discretized system is solved using standard Newton iteration. The method was tested for a circular beam with uniform intensity, a street light, and a uniform beam that is transformed into a famous Dutch painting. The reflectors were verified using commercial ray tracing software.

Key words. illumination optics, nonimaging optics, Monge–Ampère equation, optimal mass transport, nonlinear partial differential equation, reflector design, convex solution, wide stencil

AMS subject classifications. 65Z05, 35J96, 65N06

DOI. 10.1137/130938876

1. Introduction.

Optical design for illumination applications is a difficult and time-consuming job. The design is in large part done by trial and error. A reflector for a car headlamp, for example, has a complicated shape consisting of many facets, designed to fulfill the complicated output pattern regulations that exist for headlamps. Tools exist to evaluate the performance of a *given* reflector, but there are no good tools for the *inverse problem*: finding a reflector surface that gives the required light output. The common design procedure is to start with an initial reflector, evaluate the output intensity, modify the angles and positions of the reflector facets to improve the output intensity, and repeat this until a satisfactory design has been reached. This is a slow procedure, usually done by hand. The design process would significantly speed up if we could calculate the reflector surface directly.

A differential equation for the reflector surface can be derived using the law of reflection and conservation of luminous flux. The equation describes the geometrical deformation of flux density by the surface of the reflector. This equation is of Monge–Ampère-type and has an unusual boundary condition known as the transport boundary condition [20, 33, 36]. The Monge–Ampère equation with this boundary condition is known as the second boundary value problem.

Several attempts have been made to solve this boundary value problem. Oliker and Kochengin [24, 25, 26] developed a provably convergent algorithm to design a reflector for a point light source based on an intersection or union of a set of ellipsoids or paraboloids. Unfortunately, the method is very slow and not feasible for even

*Submitted to the journal's Computational Methods in Science and Engineering section September 27, 2013; accepted for publication (in revised form) March 17, 2014; published electronically June 26, 2014.

http://www.siam.org/journals/sisc/36-3/93887.html

†CASA, Eindhoven University of Technology, PO Box 513 5600 MB Eindhoven, The Netherlands (c.r.prins@tue.nl, j.h.m.tenthijeboonkkamp@tue.nl, jvroosmalen@hotmail.com).

‡Philips Lighting and CASA, Eindhoven University of Technology, PO Box 513 5600 MB Eindhoven, The Netherlands (wilbert.ijzerman@philips.com).

§Philips Research, Eindhoven, The Netherlands (teus.tukker@philips.com).

moderate requirements of detail in the target illumination. This method has been further developed by Fournier, Cassarly, and Rolland [12, 13]. Several attempts at a finite difference approach have been made as well, for example, by Norris and Westcott [28], and more recently by a group from the Zejiang University in China [31, 32]. Both solve the boundary value problem using standard finite differences. The boundary condition is only implemented for specific shapes of the target distribution, for example, circular or elliptical. These methods have given numerical solutions in several cases but are likely to fail in general cases. Specificaly, they are unable to solve the equation if the contrast within the target distribution is large.

The second boundary value problem for the Monge–Ampère equation is closely related to the problem of optimal mass transport (OMT): given a pile of sand and a hole, find a plan to transport the sand into the hole while minimizing the transportation cost. It can be shown that, when the cost of transporting a unit mass of sand is chosen to be the Euclidian distance of transportation (known as the quadratic cost function), this problem is equivalent to finding the convex solution to the second boundary value problem of a Monge–Ampère equation [35]. In the case of reflector design, we have luminous flux instead of sand, and the convex solution is the reflector surface.

A popular method to solve the OMT problem with quadratic cost function was developed by Benamou and Brenier [3]. They add a time dimension to the problem and look for a continuous transformation by a velocity field of the source density into the target density, satisfying the continuity equation. The squared velocity multiplied by the density integrated over space and time is minimized, and it is shown that this space-time minimization problem can be solved as a saddle-point problem. The method was further developed by Tannenbaum, Angenent, Haker, Haber, and Rehman [1, 21]. These methods find an optimal transport map from which a solution to the Monge–Ampère equation can be derived. Unfortunately, the methods are inefficient or no convergence proof is given. Also of interest is the work of Glimm and Oliker [19]. They show that the problem of designing a reflector for a point light source is equivalent to the OMT problem with a different cost function. However, it is not clear if it is possible to find cost functions for general optical design problems.

In this paper, we introduce a numerical method for free-form reflector design for a light source consisting of a parallel beam of light. In contrast to earlier methods, it is reliable, able to handle large contrasts, and can be used for arbitrary shapes of the target distribution. In addition, the method is sufficiently fast. We derive the Monge–Ampère equation and the corresponding boundary condition. We calculate a convex solution to this equation using a state of the art numerical method by Benamou, Froese, and Oberman [4, 5, 14, 15, 16, 17, 29, 30]. We present the discretization of the equation, based on a nonstandard wide stencil, and the discretization of the boundary condition. The discretized system is solved using Newton iteration. A convergence proof exists for this numerical method, and it is documented thoroughly in the papers of Benamou, Froese, and Oberman. However, in the case of a nonconvex target distribution, the Monge–Ampère solver does not work. We developed a numerical Legendre–Fenchel transform to circumvent this problem. Finally, we show how to use the same algorithm to find a concave reflector as well. We design a few example reflectors and verify the result by Monte Carlo ray tracing using the software package LightTools [27]. The designed reflectors reproduce the required target distributions in great detail.

The contents of this paper are as follows. In section 2 we give an introduction in illumination optics and derive the required differential equation and the boundary

condition describing the reflector surface. In section 3 we describe the numerical method used to solve this differential equation. Section 4 describes the performance of the algorithm and the verification by LightTools. Finally, we formulate concluding remarks in section 5.

2. Mathematical model.

2.1. Derivation of the Monge–Ampère equation. Our goal is to find the shape of the surface which describes a reflector that reflects the light from a source to a desired target intensity, as illustrated in Figure 1. We assume that the source is a subset of a plane parallel to and below the x-y-plane and emits light in the direction $\hat{\boldsymbol{s}}_1 = (0,0,1)^T$ with emittance $M(x,y)$ [lm/m^2], where lumen is the unit for the radiated power of the light source, corrected for the sensitivity of the eye. Let $\mathcal{S} = \overline{\operatorname{supp} M}$ be the closure of the support of $M(x,y)$ and let $\mathcal{R}_{\mathcal{S}}$ be the smallest enclosing rectangle of $\mathcal{S}$ with the edges parallel to the axes. The reflector surface is described by $z = u(x,y), (x,y) \in \mathcal{R}_{\mathcal{S}}$. We assume the size of the reflector and its distance to the origin of the coordinate system is negligible compared to the distance the reflected rays travel. Therefore, after reflection, we consider the reflector a point at the origin. In optical design this is known as the far-field approximation and is similar to the far-field approximation used for Fraunhofer diffraction [22]. After reflection, we measure the luminous flux per steradian (sr), known as luminous intensity, which should coincide with a required target intensity distribution $G(\theta,\phi)$ [lm/sr]. The target intensity distribution is given in spherical coordinates, where θ denotes the angle of the light with respect to the positive z-axis (inclination), and ϕ denotes the angle of the projection of the direction of the ray in the plane $z = 0$ with respect to the x-axis (azimuth). We assume all light of the source ends up at the target, which means that the integral of the source emittance over $\mathcal{S}$ equals the integral of the target intensity over the unit sphere S^2.

We derive a differential equation for the reflector surface using the law of reflection and conservation of luminous flux. Within the infinitesimal rectangle $(x, x + \mathrm{d}x) \times (y, y + \mathrm{d}y)$, the source emits a luminous flux $\mathrm{d}\Phi = M(x,y)\,\mathrm{d}x\,\mathrm{d}y$ [lm]. We denote the surface derivatives by $p = u_x$ and $q = u_y$. The unit surface normal of the reflector, directed towards the light source, is given by

$$\hat{\boldsymbol{n}} = \boldsymbol{n}/|\boldsymbol{n}|, \qquad \boldsymbol{n} = \begin{pmatrix} p \\ q \\ -1 \end{pmatrix}. \tag{2.1}$$

Throughout this paper we use the convention that a hat denotes a unit vector, and $|\cdot|$ denotes the 2-norm. The direction of the reflected ray $\hat{\boldsymbol{s}}_2$ is given by the law of

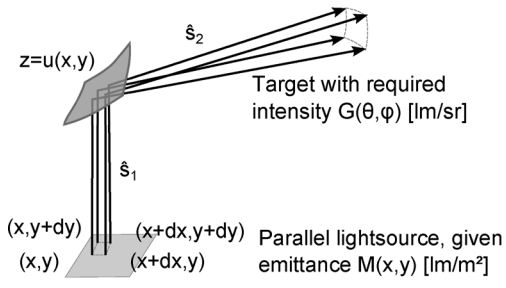


FIG. 1. *Transformation of a small rectangle of the source after reflection.*

reflection [18]:

$$\hat{\boldsymbol{s}}_2 = \hat{\boldsymbol{s}}_1 - 2(\hat{\boldsymbol{s}}_1 \cdot \hat{\boldsymbol{n}})\hat{\boldsymbol{n}}. \tag{2.2}$$

The vector $\hat{\boldsymbol{s}}_2$ is a unit vector, which is verified by

$$\hat{\boldsymbol{s}}_2 \cdot \hat{\boldsymbol{s}}_2 = \hat{\boldsymbol{s}}_1 \cdot \hat{\boldsymbol{s}}_1 - 4(\hat{\boldsymbol{s}}_1 \cdot \hat{\boldsymbol{n}})\,(\hat{\boldsymbol{s}}_1 \cdot \hat{\boldsymbol{n}}) + 4(\hat{\boldsymbol{s}}_1 \cdot \hat{\boldsymbol{n}})^2\,(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}) = 1. \tag{2.3}$$

The direction of the reflected ray depends on the coordinates x and y: $\hat{\boldsymbol{s}}_2 = \hat{\boldsymbol{s}}_2(x, y)$. The flux $\mathrm{d}\Phi$ is reflected into an infinitesimal parallelogram on S^2 spanned by $\partial\hat{\boldsymbol{s}}_2/\partial x\ \mathrm{d}x$ and $\partial\hat{\boldsymbol{s}}_2/\partial y\ \mathrm{d}y$. The size of this parallelogram is calculated using the cross product. Conservation of luminous flux now gives us the following relation:

$$\left|\frac{\partial\hat{\boldsymbol{s}}_2}{\partial x} \times \frac{\partial\hat{\boldsymbol{s}}_2}{\partial y}\right| G(\theta, \phi) = M(x, y). \tag{2.4}$$

To calculate the cross product, we use that vectors with constant 2-norm, such as $\hat{\boldsymbol{s}}_2$, are orthogonal to their derivatives. This can be seen by taking the derivative of the square of the 2-norm of such a vector $\boldsymbol{a}$,

$$0 = \frac{\partial\,(\boldsymbol{a} \cdot \boldsymbol{a})}{\partial x} = 2\,\boldsymbol{a} \cdot \frac{\partial \boldsymbol{a}}{\partial x},$$

and likewise for y. Therefore $\hat{\boldsymbol{s}}_2$ is orthogonal to its derivatives with respect to x and y. As a result, the cross product of the derivatives of $\hat{\boldsymbol{s}}_2$ is parallel to $\hat{\boldsymbol{s}}_2$. The norm of the cross product is calculated by projection on the unit vector $\hat{\boldsymbol{n}}$:

$$\left|\frac{\partial\hat{\boldsymbol{s}}_2}{\partial x} \times \frac{\partial\hat{\boldsymbol{s}}_2}{\partial y}\right| = \frac{\left|\left(\frac{\partial\hat{\boldsymbol{s}}_2}{\partial x} \times \frac{\partial\hat{\boldsymbol{s}}_2}{\partial y}\right) \cdot \hat{\boldsymbol{n}}\right|}{|\hat{\boldsymbol{s}}_2 \cdot \hat{\boldsymbol{n}}|}. \tag{2.5}$$

Using the law of reflection (2.2) and the relation $\hat{\boldsymbol{s}}_1 \cdot \hat{\boldsymbol{n}} = -1/|\boldsymbol{n}|$, we find

$$\hat{\boldsymbol{s}}_2 = \hat{\boldsymbol{s}}_1 + \frac{2}{|\boldsymbol{n}|^2}\boldsymbol{n}, \tag{2.6}$$

which is expressed in $\boldsymbol{n}$ for simplicity of the following calculations. We find for the derivatives

$$\frac{\partial\hat{\boldsymbol{s}}_2}{\partial x} = \frac{\partial}{\partial x}\left(\frac{2}{|\boldsymbol{n}|^2}\right)\boldsymbol{n} + \frac{2}{|\boldsymbol{n}|^2}\frac{\partial\boldsymbol{n}}{\partial x}, \tag{2.7a}$$

$$\frac{\partial\hat{\boldsymbol{s}}_2}{\partial y} = \frac{\partial}{\partial y}\left(\frac{2}{|\boldsymbol{n}|^2}\right)\boldsymbol{n} + \frac{2}{|\boldsymbol{n}|^2}\frac{\partial\boldsymbol{n}}{\partial y}. \tag{2.7b}$$

The cross product of these two expressions contains four terms, but three terms vanish in the numerator of (2.5): the cross product of $\boldsymbol{n}$ with itself is $\boldsymbol{0}$, and a cross product involving $\boldsymbol{n}$ is orthogonal to $\hat{\boldsymbol{n}}$, so we find for the numerator

$$\left(\frac{\partial\hat{\boldsymbol{s}}_2}{\partial x} \times \frac{\partial\hat{\boldsymbol{s}}_2}{\partial y}\right) \cdot \hat{\boldsymbol{n}} = \frac{4}{|\boldsymbol{n}|^5}\left(\frac{\partial\boldsymbol{n}}{\partial x} \times \frac{\partial\boldsymbol{n}}{\partial y}\right) \cdot \boldsymbol{n}. \tag{2.8}$$

The cross product of the surface normals results in the determinant of the Hessian matrix multiplied with $\hat{\boldsymbol{e}}_z$,

$$\frac{\partial\boldsymbol{n}}{\partial x} \times \frac{\partial\boldsymbol{n}}{\partial y} = \left(u_{xx}u_{yy} - u_{xy}^2\right)\hat{\boldsymbol{e}}_z = \det\left(D^2u\right)\hat{\boldsymbol{e}}_z, \tag{2.9}$$

so (2.8) simplifies to

$$\left(\frac{\partial \hat{\boldsymbol{s}}_2}{\partial x} \times \frac{\partial \hat{\boldsymbol{s}}_2}{\partial y}\right) \cdot \hat{\boldsymbol{n}} = -\frac{4}{|\boldsymbol{n}|^5}\left(u_{xx}u_{yy} - u_{xy}^2\right). \tag{2.10}$$

For the denominator in (2.5) we find

$$\hat{\boldsymbol{s}}_2 \cdot \hat{\boldsymbol{n}} = \hat{\boldsymbol{s}}_1 \cdot \hat{\boldsymbol{n}} - 2(\hat{\boldsymbol{s}}_1 \cdot \hat{\boldsymbol{n}})(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}) = \frac{1}{|\boldsymbol{n}|}. \tag{2.11}$$

Substituting (2.5), (2.10), and (2.11) in (2.4) we find

$$\frac{M(x,y)}{G(\theta,\phi)} = \frac{4\left|u_{xx}u_{yy} - u_{xy}^2\right|}{(p^2+q^2+1)^2}. \tag{2.12}$$

Equation (2.12) contains the variables θ and ϕ, so we need the dependence of (θ,ϕ) on $\nabla u = (p,q)$. After reflection, light has the direction $\hat{\boldsymbol{s}}_2$. We have

$$\hat{\boldsymbol{s}}_2 = \frac{1}{p^2+q^2+1}\begin{pmatrix} 2p \\ 2q \\ p^2+q^2-1 \end{pmatrix} = \begin{pmatrix} \sin(\theta)\,\cos(\phi) \\ \sin(\theta)\,\sin(\phi) \\ \cos(\theta) \end{pmatrix}. \tag{2.13}$$

From this we find

$$\theta(p,q) = \arccos((\hat{\boldsymbol{s}}_2)_z) = \arccos\left(1 - \frac{2}{p^2+q^2+1}\right), \tag{2.14a}$$

$$\phi(p,q) = \tan^{-1}\left((\hat{\boldsymbol{s}}_2)_x, (\hat{\boldsymbol{s}}_2)_y\right) = \tan^{-1}(p,q), \tag{2.14b}$$

where we define the inverse function $\tan^{-1}(p,q)$ as

$$\tan^{-1}(p,q) = \begin{cases} \arctan(q/p), & p,q \geq 0, \\ \arctan(q/p) + \pi, & p < 0, \\ \arctan(q/p) + 2\pi, & p \geq 0, q < 0. \end{cases}$$

Note that the range of the arctan is $(-\pi/2, \pi/2)$, so that $\tan^{-1}(p,q)$ has range $[0, 2\pi)$. Now we define the new target intensity function that depends on p and q instead of θ and ϕ:

$$\widetilde{G}(p,q) = G\left(\theta(p,q), \phi(p,q)\right). \tag{2.15}$$

This function comes with the supporting domain

$$\mathcal{T} = \overline{\text{supp}\left(\widetilde{G}\right)}. \tag{2.16}$$

We restrict ourselves to convex solutions u, because this results in an equation with boundary conditions that has a unique solution; see subsection 2.2. In 3.3 we show how to find a concave solution from a convex solution. When u is convex, we have $\det(D^2u) \geq 0$ and we can drop the modulus in (2.12). Finally, we have the following differential equation:

$$\det\left(D^2u\right) = \frac{M(x,y)\left(p^2+q^2+1\right)^2}{4\,\widetilde{G}(p,q)}, \qquad (x,y) \in \mathcal{R}_{\mathcal{S}}. \tag{2.17}$$

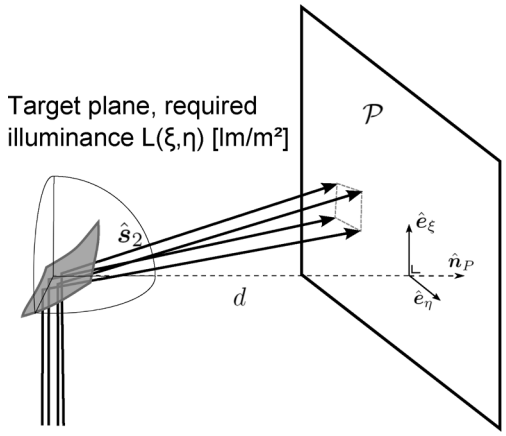


FIG. 2. *Conversion from a plane to the unit sphere.*

In many applications, we have a specified target illuminance $L(\xi,\eta)$ [lm/m^2] on a given plane $\mathcal{P}$ instead of a light intensity $G(\theta,\phi)$. In this case we may still use (2.17), but we need to derive an appropriate intensity function $\widetilde{G}(p,q)$. For this we convert $L(\xi,\eta)$ from illuminance to intensity, and write ξ and η as functions of p and q. Let $\boldsymbol{n}_P$ be the unique normal vector of the plane pointing at the origin. Let $d=|\boldsymbol{n}_P|$ and $\hat{\boldsymbol{n}}_P=-\frac{\boldsymbol{n}_P}{d}$. Let $\hat{\boldsymbol{e}}_\xi$, $\hat{\boldsymbol{e}}_\eta$ be an orthonormal basis of $\mathcal{P}$ as shown in Figure 2. We say that a point $\xi\hat{\boldsymbol{e}}_\xi+\eta\hat{\boldsymbol{e}}_\eta+d\hat{\boldsymbol{n}}_P$ has coordinates (ξ,η) on the plane. A ray with direction vector $\hat{\boldsymbol{s}}_2$ intersects the plane at (ξ,η) if

$$\hat{\boldsymbol{s}}_2=\frac{\xi\hat{\boldsymbol{e}}_\xi+\eta\hat{\boldsymbol{e}}_\eta+d\hat{\boldsymbol{n}}_P}{\sqrt{\xi^2+\eta^2+d^2}}. \tag{2.18}$$

Taking the inner product with $\hat{\boldsymbol{e}}_\xi$, $\hat{\boldsymbol{e}}_\eta$, and $\hat{\boldsymbol{n}}_P$, respectively, gives three equations, from which we derive that

$$\xi=d\frac{\hat{\boldsymbol{s}}_2\cdot\hat{\boldsymbol{e}}_\xi}{\hat{\boldsymbol{s}}_2\cdot\hat{\boldsymbol{n}}_P},\qquad \eta=d\frac{\hat{\boldsymbol{s}}_2\cdot\hat{\boldsymbol{e}}_\eta}{\hat{\boldsymbol{s}}_2\cdot\hat{\boldsymbol{n}}_P}. \tag{2.19}$$

By (2.13) we see that ξ and η are functions of p and q.

Next, we need to derive the luminous intensity $\widetilde{G}(p,q)$ [lm/sr] from the illuminance $L(\xi,\eta)$ [lm/m^2]. The luminous flux within a square $(\xi,\xi+\mathrm{d}\xi)\times(\eta,\eta+\mathrm{d}\eta)$ is given by $L(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta$. We need the luminous flux per unit area on the unit sphere. Therefore, we calculate the ratio between the surface of the infinitesimal square and its projection on the unit sphere. We introduce new spherical coordinates with ψ the inclination with respect to $\hat{\boldsymbol{n}}_P$ and χ the azimuth, according to

$$\psi=\arccos\left(\hat{\boldsymbol{s}}_2\cdot\hat{\boldsymbol{n}}_P\right), \tag{2.20a}$$

$$\chi=\tan^{-1}\left(\hat{\boldsymbol{s}}_2\cdot\hat{\boldsymbol{e}}_\xi,\hat{\boldsymbol{s}}_2\cdot\hat{\boldsymbol{e}}_\eta\right). \tag{2.20b}$$

A ray with angles ψ and χ in this coordinate system intersects the plane at a distance $d/\cos(\psi)$ from the origin. Therefore, the relation between the planar coordinates ξ, η, and the spherical coordinates ψ, χ is given by

$$\xi=d\tan(\psi)\cos(\chi), \tag{2.21a}$$

$$\eta=d\tan(\psi)\sin(\chi). \tag{2.21b}$$

An infinitesimal element on the plane has size $\mathrm{d}\xi\,\mathrm{d}\eta$, and an infinitesimal element on the unit sphere has size $\sin(\psi)\,\mathrm{d}\psi\,\mathrm{d}\chi$. By change of coordinates we find

$$\widetilde{G} = \left|\frac{1}{\sin(\psi)}\frac{\partial(\xi,\eta)}{\partial(\psi,\chi)}\right| L(\xi,\eta) = \left|\frac{d^2}{\cos(\psi)^3}\right| L(\xi,\eta). \tag{2.22}$$

Using the relation

$$\cos(\psi) = \frac{d}{\sqrt{d^2+\xi^2+\eta^2}}, \tag{2.23}$$

we find

$$\widetilde{G}(p,q) = \frac{(d^2+\xi^2+\eta^2)^{3/2}}{d} L(\xi,\eta). \tag{2.24}$$

This relation between illuminance and intensity is also derived in Born and Wolf [6, p. 196].

2.2. Implicit boundary condition and analogy with OMT. The gradient of u is regarded as a map $\nabla u : (x,y) \in \mathcal{S} \mapsto (p,q) \in \mathcal{T}$. We require that all light from the source is reflected toward the target. This gives us an implicit boundary condition

$$\nabla u(\mathcal{S}) = \mathcal{T}. \tag{2.25}$$

The Monge–Ampère equation (2.17) with implicit boundary condition (2.30) is closely related to OMT with a quadratic cost function. This problem is stated as follows. Let $f : \mathcal{X} \to \mathbb{R}^+$ and $g : \mathcal{Y} \to \mathbb{R}^+$ denote (mass) densities. The total mass in $\mathcal{X}$ is required to be equal to the total mass in $\mathcal{Y}$:

$$\int_{\mathcal{X}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathcal{Y}} g(p,q)\,\mathrm{d}p\,\mathrm{d}q. \tag{2.26}$$

Find a map $\boldsymbol{\gamma} : \mathcal{X} \to \mathcal{Y}$ such that

$$f(x,y) = g(\boldsymbol{\gamma}(x,y))\,|\det(D\boldsymbol{\gamma}(x,y))|\,, \tag{2.27}$$

that minimizes the transportation cost defined by

$$\mathcal{C}[\boldsymbol{\gamma}] = \int_{\mathcal{X}} \left|(x,y)^T - \boldsymbol{\gamma}(x,y)\right|^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{2.28}$$

An important theorem by Brenier [35, p. 66] states that there is a unique map $\boldsymbol{\gamma} : \mathcal{X} \to \mathcal{Y}$ that satisfies (2.27) and minimizes (2.28), and this map $\boldsymbol{\gamma}$ is the gradient of a convex function u. Substitution of $\boldsymbol{\gamma} = \nabla u$ in (2.27) yields the Monge–Ampère equation

$$\det\left(D^2 u(\boldsymbol{x})\right) = \frac{f(\boldsymbol{x})}{g(\nabla u(\boldsymbol{x}))},\ \boldsymbol{x} \in \mathcal{X}, \tag{2.29}$$

with implicit boundary condition

$$\nabla u(\mathcal{X}) = \mathcal{Y}. \tag{2.30}$$

Equation (2.29) with boundary condition (2.30) has a convex solution which is unique up to an additive constant [34, 35].

2.3. Explicit boundary condition. We replace the implicit boundary condition (2.30) by the following boundary condition, which is used in the numerical method:

$$\nabla u(\partial\mathcal{X}) = \partial\mathcal{Y}. \tag{2.31}$$

To show the equivalence, we first we show invertibility of the map ∇u.

LEMMA 2.1. *Let $u : \mathcal{X} \to \mathbb{R}$ be a strictly convex function that satisfies* (2.30)*. Then the map $\nabla u : \mathcal{X} \to \mathcal{Y}$ is bijective.*

Proof. Surjectivity of the map follows directly from (2.30). The map is injective if $\nabla u(\boldsymbol{a}) = \nabla u(\boldsymbol{b}) \Rightarrow \boldsymbol{a} = \boldsymbol{b}$. Suppose $\nabla u(\boldsymbol{a}) = \nabla u(\boldsymbol{b})$. A convex function has a graph that lies above its tangent plane [35, p. 53]. If $\boldsymbol{a} \neq \boldsymbol{b}$, we have because of the (strict) convexity of u

$$u(\boldsymbol{a}) > u(\boldsymbol{b}) + \nabla u(\boldsymbol{b}) \cdot (\boldsymbol{a} - \boldsymbol{b}),$$
$$u(\boldsymbol{b}) > u(\boldsymbol{a}) + \nabla u(\boldsymbol{a}) \cdot (\boldsymbol{b} - \boldsymbol{a}).$$

Adding these two inequalities we find

$$(\nabla u(\boldsymbol{b}) - \nabla u(\boldsymbol{a})) \cdot (\boldsymbol{b} - \boldsymbol{a}) > 0.$$

However, $\nabla u(\boldsymbol{b}) = \nabla u(\boldsymbol{a})$, so we have a contradiction. Hence $\boldsymbol{a} = \boldsymbol{b}$ and ∇u is injective, and thus bijective. ∎

Now we show that a convex solution to (2.29) subject to (2.30) maps the boundary of $\mathcal{X}$ to the boundary of $\mathcal{Y}$.

LEMMA 2.2. *Assume $u \in C^2(\mathcal{X})$ is a convex solution to* (2.29) *subject to* (2.30)*; then $\nabla u(\partial\mathcal{X}) = \partial\mathcal{Y}$.*

Proof. The map ∇u is continuous and thus maps open sets to open sets. Let $\mathrm{Int}(\mathcal{X}) = \bar{\mathcal{X}} \setminus \partial\mathcal{X}$ be the interior of $\mathcal{X}$. We find $\nabla u(\mathrm{Int}(\mathcal{X})) \subset \mathrm{Int}(\mathcal{Y})$ and $(\nabla u)^{-1}(\mathrm{Int}(\mathcal{Y})) \subset \mathrm{Int}(\mathcal{X})$. We find that

$$\mathrm{Int}(\mathcal{Y}) = \nabla u\left((\nabla u)^{-1}(\mathrm{Int}(\mathcal{Y}))\right) \subset \nabla u(\mathrm{Int}(\mathcal{X})) \subset \mathrm{Int}(\mathcal{Y})$$

and see that $\nabla u(\mathrm{Int}(\mathcal{X})) = \mathrm{Int}(\mathcal{Y})$. We conclude that $\nabla u(\partial\mathcal{X}) = \partial\mathcal{Y}$. ∎

The next lemma shows that (2.31) is sufficient to imply (2.30).

LEMMA 2.3. *Let $\mathcal{X}$ and $\mathcal{Y}$ be simply connected, and let u be a strictly convex solution to* (2.29) *and* (2.31)*; then u satisfies* (2.30)*.*

Proof. ∇u is continuous and bijective, and thus a homeomorphism. ∇u maps $\partial\mathcal{X}$ to $\partial\mathcal{Y}$. $\mathcal{X}$ is simply connected, and thus $\nabla u(\mathcal{X})$ is simply connected. Therefore, $\nabla u(\mathrm{Int}(\mathcal{X})) = \mathrm{Int}(\mathcal{Y})$. ∎

3. Numerical method. Recently Froese, Oberman, and Benamou [4, 5, 17] published several articles introducing a new discretization to find a convex solution to the Monge–Ampère equation (2.29) with boundary condition (2.31). We implemented a version of this method to solve our reflector equation (2.17). In this section we give an overview of the algorithm. For more details, we refer to the articles by Froese, Oberman, and Benamou. First we give the discretization of the Monge–Ampère equation (2.29) on the interior domain that enforces convexity. Subsequently, we give the discretization of the boundary condition (2.31). The discretization of the Monge–Ampère equation on the interior and the boundary gives a numerical scheme, which is known to converge [5]. This discrete system is solved using damped Newton iteration. The discretization requires the target domain to be convex. In many cases

our target domain is not convex, and in these cases we interchange source and target, and subsequently invert the mapping ∇u using the Legendre–Fenchel transform. This is described in more detail at the end of this section.

3.1. Discretization of the Monge–Ampère equation. The rectangular source domain $\mathcal{R}_{\mathcal{X}} = [a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2$ is discretized using a standard grid. For simplicity we take a square grid. We choose $M_x, M_y \in \mathbb{N}$ such that $h = (b_1 - a_1)/(M_x - 1) = (b_2 - a_2)/(M_y - 1)$; if necessary we modify the dimensions of the domain slightly such that we can use a square grid. Each grid point is given by

$$\boldsymbol{x}_{i,j} = \begin{pmatrix} x_i \\ y_j \end{pmatrix} = \begin{pmatrix} a_1 + h(i-1) \\ a_2 + h(j-1) \end{pmatrix} \tag{3.1}$$

for $i = 1, 2, \ldots, M_x$ and $j = 1, 2, \ldots, M_y$. We denote the approximate solution of the boundary value problem by

$$u_{i,j} \approx u(\boldsymbol{x}_{i,j}). \tag{3.2}$$

The solution of the Monge–Ampère equation (2.29) with boundary condition (2.31) without the convexity requirement may not be unique. For stability of the numerical algorithm it is essential that convexity is enforced [15]. Benamou, Froese, and Oberman introduced a numerical scheme for the interior that ensures convexity of the solution and is provably convergent. The scheme is based on a convergence result by Barles and Souganidis [2, 5], which states that a consistent, stable, and monotone approximation scheme converges to the unique viscosity solution of a denegerate elliptic equation (such as the convex solution to the Monge–Ampère equation). A monotone scheme is defined as follows. Let $F^{i,j}(\boldsymbol{u}) = 0$ be the finite difference equation at $\boldsymbol{x}_{i,j}$. Then the scheme is called monotone if the following inequalities are satisfied [29]:

$$\frac{\partial F^{i,j}}{\partial u_{i,j}} \geq 0, \qquad \frac{\partial F^{i,j}}{\partial u_{k,l}} \leq 0 \text{ for } (i,j) \neq (k,l). \tag{3.3}$$

A standard finite difference scheme for the Monge–Ampère equation is not monotone because, among other reasons, the discretization of the mixed second derivative u_{xy} is not monotone [30]. Also, a standard finite difference scheme does not enforce convexity.

Therefore, an alternative representation of $\det\left(D^2 u\right)$ is introduced. The representation is based on the observation that $\det\left(D^2 u\right) = \lambda_{\min}\lambda_{\max}$, where $\lambda_{\min} \leq \lambda_{\max}$ are the smallest and largest eigenvalue of the matrix $D^2 u$. These eigenvalues correspond to the normalized eigenvectors $\boldsymbol{v}_{\min}$ and $\boldsymbol{v}_{\max}$, which are orthogonal because $D^2 u$ is symmetric and real. Let $\mathcal{V}$ be the set of all orthonormal bases of $\mathbb{R}^2$. The determinant of the Hessian for any base $(\boldsymbol{v}_1, \boldsymbol{v}_2) \in \mathcal{V}$ is given by

$$\frac{\partial^2 u}{\partial v_1^2}\frac{\partial^2 u}{\partial v_2^2} - \left(\frac{\partial^2 u}{\partial v_1 \partial v_2}\right)^2 = \det\left([\boldsymbol{v}_1\, \boldsymbol{v}_2]^T \left(D^2 u\right) [\boldsymbol{v}_1\, \boldsymbol{v}_2]\right) = \det\left(D^2 u\right). \tag{3.4}$$

Here $\frac{\partial}{\partial v_i}$ denotes differentiation in the direction of $\boldsymbol{v}_i$. The determinant of the Hessian is thus independent of the direction of differentiation. On the other hand, we have by the Raleigh–Ritz quotient [23, p. 176]

$$\det\left(D^2 u\right) = \lambda_{\min}\lambda_{\max} = \left(\boldsymbol{v}_{\min}^T (D^2 u)\boldsymbol{v}_{\min}\right)\left(\boldsymbol{v}_{\max}^T (D^2 u)\boldsymbol{v}_{\max}\right) = \frac{\partial^2 u}{\partial v_{\min}^2}\frac{\partial^2 u}{\partial v_{\max}^2}. \tag{3.5}$$

Thus, when the basis is the set of eigenvectors, the mixed derivative vanishes, which is exactly what we want, because the mixed derivative does not have a monotone discretization. The square of the mixed derivative is nonnegative, and thus we can find the eigenvectors by minimizing the product of the second order derivatives over all directions:

$$\det\left(D^2u\right) = \min_{(\boldsymbol{v}_1,\boldsymbol{v}_2)\in\mathcal{V}} \frac{\partial^2 u}{\partial v_1^2}\,\frac{\partial^2 u}{\partial v_2^2}. \tag{3.6}$$

We use this to rewrite the Monge–Ampère equation (2.29) as follows:

$$\min_{(\boldsymbol{v}_1,\boldsymbol{v}_2)\in\mathcal{V}} \frac{\partial^2 u}{\partial v_1^2}\,\frac{\partial^2 u}{\partial v_2^2} = \frac{f(\boldsymbol{x})}{g(\nabla u(\boldsymbol{x}))}. \tag{3.7}$$

Subsequently, we modify this equation to admit only convex solutions u. The function u is convex if and only if both of the eigenvalues are positive. For that reason, we introduce the operators

$$(z)^+ = \max(z,0), \qquad (z)^- = \min(z,0). \tag{3.8}$$

We modify (3.7) to

$$\min_{(\boldsymbol{v}_1,\boldsymbol{v}_2)\in\mathcal{V}} \left\{ \left(\frac{\partial^2 u}{\partial v_1^2}\right)^+ \left(\frac{\partial^2 u}{\partial v_2^2}\right)^+ \right\} = \frac{f(\boldsymbol{x})}{g(\nabla u(\boldsymbol{x}))}. \tag{3.9}$$

If the right-hand side is positive, the new formulation only allows convex u. However, if the right-hand side is 0, any function that is not convex can be a solution to (3.9). To prevent this, we add the negative part of the second order derivatives:

$$\min_{(\boldsymbol{v}_1,\boldsymbol{v}_2)\in\mathcal{V}} \left\{ \left(\frac{\partial^2 u}{\partial v_1^2}\right)^+ \left(\frac{\partial^2 u}{\partial v_2^2}\right)^+ + \left(\frac{\partial^2 u}{\partial v_1^2}\right)^- + \left(\frac{\partial^2 u}{\partial v_2^2}\right)^- \right\} = \frac{f(\boldsymbol{x})}{g(\nabla u(\boldsymbol{x}))}. \tag{3.10}$$

Now, if D^2u has a negative eigenvalue, the left-hand side is negative and therefore not a solution to (3.10). If both eigenvalues are positive, the equation is equivalent to (3.7).

Equation (3.10) is discretized using a finite number of directions $V \subset \mathcal{V}$. The directional resolution is called $\Delta\theta$. We use the 17-point stencil shown in Figure 3. The second order derivatives in the different directions are discretized using standard central differences:

$$\mathcal{D}_{\boldsymbol{v}\boldsymbol{v}} u_{i,j} = \frac{1}{{h_{\boldsymbol{v}}}^2}\left(u(\boldsymbol{x}_{i,j} - \boldsymbol{v}h_{\boldsymbol{v}}) + u(\boldsymbol{x}_{i,j} + \boldsymbol{v}h_{\boldsymbol{v}}) - 2u(\boldsymbol{x}_{i,j})\right). \tag{3.11}$$

Here $h_{\boldsymbol{v}}$ is the distance to the neighboring points on the line $\boldsymbol{x} = \boldsymbol{x}_{i,j} + s\,\boldsymbol{v}$, $s \in \mathbb{R}$. This leads to a monotone discretization of (3.10).

Adjacent to the boundary, some directions in the stencil are missing. The missing points are replaced by interpolated points on the boundary. Suppose, for example, we need to calculate the second order difference $\mathcal{D}_{\boldsymbol{v}\boldsymbol{v}} u_{i,2}$ for $\boldsymbol{v} = (1,2)^T/\sqrt{5}$. Using second order Taylor series around $u(x_i, y_2)$, we find the following approximation:

$$\mathcal{D}_{\boldsymbol{v}\boldsymbol{v}} u_{i,2} = \frac{2\,u_{i-\frac{1}{2},1} - 3\,u_{i,2} + u_{i+1,4}}{\frac{3}{4}\,h_{\boldsymbol{v}}^2}, \tag{3.12}$$

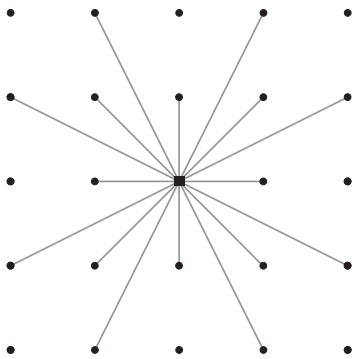


FIG. 3. *The* 17*-point stencil.*

where $u_{i-\frac{1}{2},1}$ is approximated using quadratic interpolation,

$$u_{i-1/2,1} \approx \frac{3}{8}u_{i-1,1} + \frac{6}{8}u_{i,1} - \frac{1}{8}u_{i+1,1}. \tag{3.13}$$

Substitution yields

$$\mathcal{D}_{\boldsymbol{vv}}u_{i,2} = \frac{3\,u_{i-1,1} + 6\,u_{i,1} - u_{i+1,1} + 4\,u_{i+1,4} - 12\,u_{i,2}}{3\,h_{\boldsymbol{v}}^2}. \tag{3.14}$$

This approximation is second order accurate, but monotonicity of the scheme is broken because of the minus sign before $u_{i+1,1}$. Alternatively, we could use linear interpolation, but as no convergence problems were encountered we chose the more accurate interpolation.

To ensure monotonicity, we introduce the following operators: $(\alpha)_\delta^+ = \max(\alpha, \delta)$ and $(\alpha)_\delta^- = \min(\alpha, \delta)$. The left-hand side of (3.10) is modified as follows:

$$\det{}_\delta^+\left(D^2u\right) = \min_{(\boldsymbol{v}_1,\boldsymbol{v}_2)\in\mathcal{V}} \left\{ \left(\frac{\partial^2 u}{\partial v_1^2}\right)_\delta^+ \left(\frac{\partial^2 u}{\partial v_2^2}\right)_\delta^+ + \left(\frac{\partial^2 u}{\partial v_1^2}\right)_\delta^- + \left(\frac{\partial^2 u}{\partial v_2^2}\right)_\delta^- \right\}. \tag{3.15}$$

The gradient in the right-hand side of (2.29) is approximated using a linear combination of the discretized derivatives for which the minimum was achieved in the discretization of the left-hand side. We denote this gradient by $\widetilde{\nabla}u_{i,j}$. Finally, we have the following discretization for the interior domain:

$$\begin{aligned} \mathrm{MA}^{h,\Delta\theta}[u]_{i,j} = -\frac{f(x_i,y_j)}{g\left(\widetilde{\nabla}u_{i,j}\right)} + \min_{(\boldsymbol{v}_1,\boldsymbol{v}_2)\in V} \Big\{ &\left(\mathcal{D}_{\boldsymbol{v}_1\boldsymbol{v}_1}u_{i,j}\right)_\delta^+ (\mathcal{D}_{\boldsymbol{v}_2\boldsymbol{v}_2}u_{i,j})_\delta^+ \\ &+ (\mathcal{D}_{\boldsymbol{v}_1\boldsymbol{v}_1}u_{i,j})_\delta^- + (\mathcal{D}_{\boldsymbol{v}_2\boldsymbol{v}_2}u_{i,j})_\delta^- \Big\}. \end{aligned} \tag{3.16}$$

For monotonicity we need $\frac{\partial \mathrm{MA}^{h,\Delta\theta}[u]_{i,j}}{\partial u_{k,l}} \le 0$ for all $(k,l) \neq (i,j)$. This is guaranteed when $\delta \ge \sqrt{5}hK/2$, where K is the Lipschitz constant in the variables p and q of $f(x,y)/g(p,q)$. This choice of δ is analogous to the choice made in [4, p. 12].

3.2. Discretization of the boundary. The boundary condition (2.31) is implemented by creating a nonlinear equation for each grid point at the boundary of $\mathcal{X}$. Let $H(\boldsymbol{y})$ be the signed distance to the boundary of $\mathcal{Y}$, defined to be negative

for values inside $\mathcal{Y}$; then $H(\boldsymbol{y}) = 0$ if and only if $\boldsymbol{y} \in \partial\mathcal{Y}$. A supporting line of a point $\boldsymbol{z} \in \partial\mathcal{Y}$ is a line that contains the point $\boldsymbol{z}$ but does not separate any two points of $\partial\mathcal{Y}$. Let $\mathcal{N}(\boldsymbol{z})$ be the set of normals of the lines supporting $\boldsymbol{z} \in \partial\mathcal{Y}$. The signed distance function is a convex function and can be written in terms of the supporting lines [7, p. 51] to $\mathcal{Y}$:

$$H(\boldsymbol{y}) = \max\left\{(\boldsymbol{y}-\boldsymbol{z})\cdot\hat{\boldsymbol{n}} \middle|\, \boldsymbol{z}\in\partial\mathcal{Y},\, \hat{\boldsymbol{n}}\in\mathcal{N}(\boldsymbol{z})\right\}. \tag{3.17}$$

This function can be rewritten as follows. Let $\mathcal{Z}(\hat{\boldsymbol{n}})$ be the set of points $\boldsymbol{z} \in \partial\mathcal{Y}$ that have a supporting line with normal $\hat{\boldsymbol{n}}$ and S^1 the unit circle. We find

$$\begin{aligned} H(\boldsymbol{y}) &= \max\left\{(\boldsymbol{y}-\boldsymbol{z})\cdot\hat{\boldsymbol{n}} \middle|\, \hat{\boldsymbol{n}}\in S^1,\, \boldsymbol{z}\in\mathcal{Z}(\hat{\boldsymbol{n}})\right\} \\ &= \max_{\hat{\boldsymbol{n}}\in S^1}\left\{\boldsymbol{y}\cdot\hat{\boldsymbol{n}} - H^*(\hat{\boldsymbol{n}})\right\} \end{aligned} \tag{3.18}$$

with $H^*(\hat{\boldsymbol{n}}) = \min_{\boldsymbol{z}\in\mathcal{Z}(\hat{\boldsymbol{n}})} \boldsymbol{z}\cdot\hat{\boldsymbol{n}}$. The last equality follows because $\boldsymbol{z}\cdot\hat{\boldsymbol{n}}$ is constant for $\boldsymbol{z}$ on a line with normal $\hat{\boldsymbol{n}}$. In the algorithm, the function $H^*(\hat{\boldsymbol{n}})$ is approximated by taking the minimum over a uniform discretization of S^1. We only need to calculate the function H^* once. In [5] it is shown that, because of the convexity of u, for any $\boldsymbol{x} \in \partial\mathcal{X}$ with unit outward normal $\hat{\boldsymbol{n}}_{\mathcal{X}}$, the maximum in (3.18) can be restricted to vectors $\hat{\boldsymbol{n}}_{\mathcal{Y}} \in S^1$ such that $\hat{\boldsymbol{n}}_{\mathcal{X}}\cdot\hat{\boldsymbol{n}}_{\mathcal{Y}} > 0$. This is used to discretize the boundary as follows: let

$$S^1_{\Delta\alpha} = \{\hat{\boldsymbol{n}} = (n_1,n_2) |\, n_1 = \cos(i\Delta\alpha), n_2 = \sin(i\Delta\alpha), i = 1,2,\dots 2\pi/\Delta\alpha\}. \tag{3.19}$$

The boundary condition is discretized using a second order upwind scheme

$$\begin{aligned} &H(\nabla u(\boldsymbol{x}_{i,j})) \\ &\quad\approx \max_{(n_1,n_2)^T\in S^1_{\Delta\alpha}} \Big\{ \max\{n_1,0\}\frac{3u_{i,j}-4u_{i-1,j}+u_{i-2,j}}{2h} \\ &\qquad\qquad + \min\{n_1,0\}\frac{-3u_{i,j}+4u_{i+1,j}-u_{i+2,j}}{2h} \\ &\qquad\qquad + \max\{n_2,0\}\frac{3u_{i,j}-4u_{i,j-1}+u_{i,j-2}}{2h} \\ &\qquad\qquad + \min\{n_2,0\}\frac{-3u_{i,j}+4u_{i,j+1}-u_{i,j+2}}{2h} \\ &\qquad\qquad - H^*(n_1,n_2) \,\Big|\, \boldsymbol{n}_{\mathcal{X}}(\boldsymbol{x}_{i,j})\cdot(n_1,n_2)^T > 0 \Big\} \\ &\quad=: H^{h,\Delta\alpha}[u]_{i,j}, \end{aligned} \tag{3.20}$$

where $\boldsymbol{n}_{\mathcal{X}}(\boldsymbol{x}_{i,j})$ is the normal of $\mathcal{X}$ at $\boldsymbol{x}_{i,j}$. The relation $\hat{\boldsymbol{n}}_{\mathcal{X}}\cdot\hat{\boldsymbol{n}}_{\mathcal{Y}} > 0$ guarantees that the scheme only relies on values of u inside the grid. Close to the corners, a first order scheme is used for some derivatives when necessary. We implemented the first order discretization as well. The combination $\hat{\boldsymbol{n}}_{\mathcal{X}}\cdot\hat{\boldsymbol{n}}_{\mathcal{Y}} > 0$ with a first order discretization gives a monotone scheme [5]. From experiments we found that a first order discretization causes a relatively large error at the boundary compared to the interior of the solution. As we did not encounter convergence problems in practice, we use the second order scheme in our numerical experiments.

Equation (2.29) with boundary condition (2.31) contains only first and second order derivatives of u, and therefore if $u(x,y)$ is a solution, then $u(x,y)+C$ is also a solution for any constant C. Therefore, we solve instead the equation

$$\det\left(D^2u(\boldsymbol{x})\right) = \frac{f(\boldsymbol{x})}{g(\nabla u(\boldsymbol{x}))} + w_{\text{anchor}}u(\boldsymbol{x}_0) \tag{3.21}$$

for some anchor point $\boldsymbol{x}_0 \in \mathcal{X}$ and some $w_{\text{anchor}} > 0$, which must be positive to ensure monotonicity of the scheme. In the numerical experiments we decided to use the grid point with indices

$$i = \left\lfloor \frac{M_x}{3} \right\rfloor,\ j = \left\lfloor \frac{M_y}{3} \right\rfloor, \tag{3.22}$$

which is an arbitrary choice. This new formulation has a unique solution with $u(\boldsymbol{x}_0) = 0$. Other values for $u(\boldsymbol{x}_0)$ would not allow a solution, as (2.29) only has a solution if the integrability condition (2.26) is satisfied. The equation $w_{\text{anchor}} u(\boldsymbol{x_0}) = 0$ is added to each equation in the discretized system. We chose the value $w_{\text{anchor}} = 4$, because this minimizes the condition number of the Jacobi-matrix of the nonlinear system.

The numerical scheme given by (3.16) and (3.20) and the anchor equation converges to the unique viscosity solution, shown in [5]. The system is solved using damped Newton. Denote by $\boldsymbol{N}(\boldsymbol{u}) = \boldsymbol{0}$ this system, and let μ be the damping factor with initial value $\mu = 1$. The next iterant is calculated by $\boldsymbol{u}^{n+1} = \boldsymbol{u}^n + \mu \boldsymbol{s}$, where $\boldsymbol{s}$ is the descent direction, calculated using the UMFpack implementation in MATLAB [8, 9, 10, 11]. When $|\boldsymbol{N}(\boldsymbol{u}^{n+1})|_1 \geq |\boldsymbol{N}(\boldsymbol{u}^n)|_1$, μ is divided by two, and this is repeated until $|\boldsymbol{N}(\boldsymbol{u}^{n+1})|_1 < |\boldsymbol{N}(\boldsymbol{u}^n)|_1$, so the norm always decreases. We start the iteration with an initial guess $\boldsymbol{u}^0$. As an initial guess, we use a convex quadratic function of which the gradient maps $\mathcal{S}$ to a bounding box $[c_1, d_1] \times [c_2, d_2]$ of $\mathcal{T}$:

$$u^0_{i,j} = \frac{1}{2}\frac{d_1 - c_1}{b_1 - a_1} x_i^2 + \frac{b_1 c_1 - a_1 d_1}{b_1 - a_1} x_i + \frac{1}{2}\frac{d_2 - c_2}{b_2 - a_2} y_j^2 + \frac{b_2 c_2 - a_2 d_2}{b_2 - a_2} y_j - u_r. \tag{3.23}$$

The value u_r is chosen such that $u(\boldsymbol{x}_0) = 0$.

3.3. Mapping inversion and concave solutions. The discretization used to solve the Monge–Ampère equation fails when the target domain is not convex. As we will see in the examples, even if the original target domain supp $G(\theta, \phi)$ or supp $L(\xi, \eta)$ is convex, the domain $\mathcal{T}$ may not be so. The source domain, on the other hand, which is the light emitting area, is typically convex. Therefore, we often switch the roles of source and target, and let $\mathcal{X} = \mathcal{T}$, $\mathcal{Y} = \mathcal{S}$, $f = 4\,\widetilde{G}(p, q)/(1 + p^2 + q^2)^2$, and $g = M$. Solving this yields a surface $u^*(p, q)$ defined on the smallest enclosing rectangle of $\mathcal{T}$. We find the reflector of the original problem by taking the Legendre–Fenchel transform of u^*:

$$u(x, y) = \max_{(p,q)\in\mathcal{T}} \{x\,p + y\,q - u^*(p, q)\}\,. \tag{3.24}$$

The relation between mapping inversion and the Legendre–Fenchel tranform is discussed in [35, p. 57]. Because the function $u^*(p, q)$ was calculated on a grid, we approximate $u(x_i, y_j)$ by taking a maximum of

$$L^{i,j}_{k,l} = x_i\,p_k + y_j\,q_l - u^*(p_k, q_l) \tag{3.25}$$

over the grid points (p_k, q_l). This maximum can be found by simply calculating all values of $L^{i,j}_{k,l}$ for each i, j, k, and l. This method is simple and robust, but slow for large grid sizes. A faster algorithm is constructed as follows. Because u^* is convex, $L^{i,j}_{k,l}$ is concave, and we can find the maximum using a grid search. We start at a grid point (k, l) and move to a neighboring point (k', l') if the value of $L^{i,j}_{k',l'}$ is larger than $L^{i,j}_{k,l}$. This is repeated until we cannot find such a neighboring point anymore. For the

next values of (i, j), the last maximum (k, l) is used as a starting point, because it is most likely very close to the next maximum. This makes the method very efficient. Unfortunately, the global maximum is not always found because the calculated u^* is not always perfectly convex if $f = 0$. This turned out not to be a problem in practice, because when $f = 0$, no light is involved for that part of the reflector.

Subsequently the maximum is refined using parabolic quasiinterpolation on the 3×3 subgrid centered around the point found by the grid search just described (or centered around its neighbor if the maximum is on the boundary of the grid). Let $(\tilde{k}, \tilde{l})$ be the center of the subgrid. We interpolate the function on the subgrid with a a polynomial of the form

$$P(p, q) = c_1 p^2 + c_2 pq + c_3 q^2 + c_4 p + c_5 q + c_6. \tag{3.26}$$

We perform a least squares fit with the coefficients of this polynomial. To achieve that, we construct the matrix $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} p_{\tilde{k}-1}^2 & p_{\tilde{k}-1} q_{\tilde{l}-1} & \cdots & 1 \\ p_{\tilde{k}}^2 & p_{\tilde{k}} q_{\tilde{l}-1} & \cdots & \\ & \ddots & & \\ p_{\tilde{k}+1}^2 & p_{\tilde{k}+1} q_{\tilde{l}+1} & \cdots & 1 \end{pmatrix} \tag{3.27}$$

and the vector $\boldsymbol{b}$

$$\boldsymbol{b} = \left(L^{i,j}_{\tilde{k}-1,\tilde{l}-1}, L^{i,j}_{\tilde{k},\tilde{l}-1}, \ldots, L^{i,j}_{\tilde{k}+1,\tilde{l}+1} \right)^T. \tag{3.28}$$

The linear least squares fit is the vector of coefficients $\boldsymbol{c} = (c_1, \ldots, c_6)$ such that

$$|\mathbf{A}\boldsymbol{c} - \boldsymbol{b}|_2^2 \tag{3.29}$$

is minimized. The minimizer is given by

$$\boldsymbol{c} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{b}. \tag{3.30}$$

Given the coefficients of the polynomial, we find the critical point of $P(p, q)$ by differentiation. The critical point occurs at

$$\bar{p} = \frac{2c_3 c_4 - c_2 c_5}{c_2^2 - 4c_1 c_3}, \tag{3.31a}$$

$$\bar{q} = \frac{2c_1 c_5 - c_2 c_4}{c_2^2 - 4c_1 c_3} \tag{3.31b}$$

and is given by

$$P_{\max} = \frac{c_3 c_4^2 - c_2 c_4 c_5 + c_1 c_5^2 + c_2^2 c_6 - 4c_1 c_3 c_6}{c_2^2 - 4c_1 c_3}. \tag{3.32}$$

Because u^* is convex, the critical point will be a maximum. We take $P_{\max}$ as an improved approximation of the maximum if the critical point $(\bar{p}, \bar{q})$ of the polynomial is within the rectangle $[p_{\tilde{k}-1}, p_{\tilde{k}+1}] \times [q_{\tilde{l}-1}, q_{\tilde{l}+1}]$. If it is not, then $u^*(p, q)$ may not be strictly convex locally, and $\max_{k,l}\{L^{i,j}_{k,l}\}$ is used.

For many applications, a concave reflector is preferred over a convex reflector. We could adapt the algorithm above to yield concave solutions, but it is easier to find concave solutions using the same Monge–Ampère solver. To find a concave solution to (2.29) and (2.30), we calculate the convex solution v for

$$\det\left(D^2 v(\boldsymbol{x})\right) = \frac{f(\boldsymbol{x})}{g(-\nabla v(\boldsymbol{x}))}, \tag{3.33a}$$

$$H(-\nabla v) = 0, \tag{3.33b}$$

and then $u = -v$ is the concave solution to (2.29) with $H(\nabla u) = 0$. This technique is applied to the numerical example of the street light.

4. Numerical results. We test the algorithm on three cases: a convex reflector for a circular beam with uniform intensity, a concave reflector for a lamppost with a uniform rectangular illuminance on the street, and a convex reflector that projects a painting on a wall. The reflectors are evaluated using the LightTools software package [27]. In this software package we create three-dimensional models of the light sources, the reflectors, and, if applicable, the target plane. The systems are simulated using ray tracing: rays are emitted randomly from the light source with a distribution proportional to the surface emittance. The software calculates the reflection at the reflector surface and subsequently the distribution of the rays over the target plane or the angular distribution by dividing the target in small rectangles (bins). The number of rays in each bin is counted and used to estimate the illuminance or intensity. We use one million rays in each simulation.

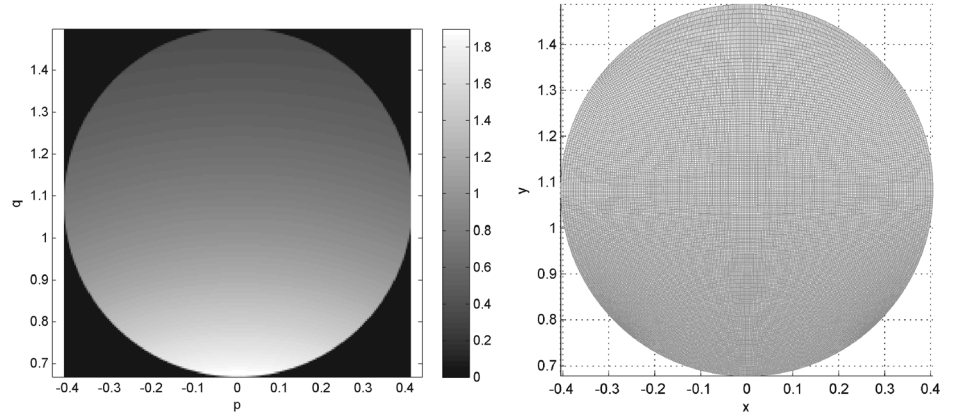
4.1. A circle with uniform intensity. The first test problem is the design of a convex reflector for constant output intensity. We have a square light source with constant emittance $M(x, y) = 1$. Let ρ be the angle with respect to the positive y-axis. We aim for a constant output intensity $G_0 > 0$ for $\rho < \frac{\pi}{8}$ and zero intensity for $\rho > \frac{\pi}{8}$. We choose $M_x = M_y = 200$. As the target is convex, we choose $f(x, y) = M(x, y) = 1$ and $g(p, q) = 4\,G_0/(p^2 + q^2 + 1)^2$. The value of G_0 is chosen such that (2.26) is satisfied. The converted target function can be seen in Figure 4(a). The Monge–Ampère equation is solved using the damped Newton iteration. After 14 steps the solver finds a solution with $|\boldsymbol{N}(\boldsymbol{u})|_1/M_x M_y < 10^{-10}$.

The gradient map of the reflector is shown in Figure 4(b). The reflector surface is shown in Figure 4(c). The surface was exported to LightTools and simulated using raytracing. The resulting intensity profile can be seen in Figure 4(d). The profile is uniform and circular as expected. It can be seen that the circle has radius $\pi/8$ indeed.

4.2. The street light. The second problem is the design of a reflector for street illumination. A sketch of the lamppost is shown in Figure 5(a). We illuminate the area $(1, 8) \times (-6, 6) \times \{-6\} \in \mathbb{R}^3$ with constant illuminance L_0, which is chosen later to satisfy (2.26). The light source is a circle with radius 0.01 around the origin and uniform emittance $M(x, y) = 1$. We choose a concave reflector in this setting because it bends rays more effectively away from the light source, thus allowing a more compact optical system.

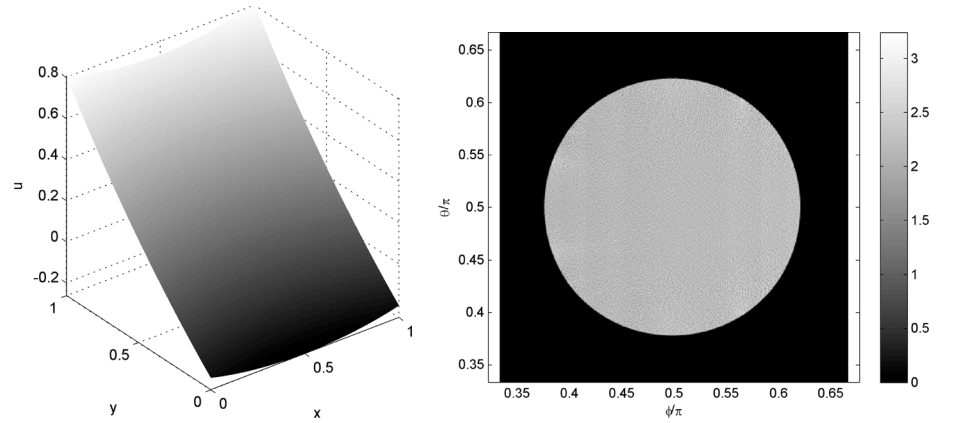
Using (2.19) with $\hat{\boldsymbol{n}}_P = (0, 0, 1)^T$, $\hat{\boldsymbol{e}}_\xi = (1, 0, 0)$, $\hat{\boldsymbol{e}}_\eta = (0, 1, 0)$, and $d = 6$. We find

$$\xi = \frac{2\,d\,p}{p^2 + q^2 - 1}, \quad \eta = \frac{2\,d\,q}{p^2 + q^2 - 1}. \tag{4.1}$$



(a) Converted target function $g(p, q)$

(b) The square grid on (x, y) after deformation by the gradient map $(p(x, y), q(x, y))$ of the reflector



(c) Reflector surface

(d) Intensity in the simulation

FIG. 4. *The target function, reflector, and simulation results of the reflector for a square light source and a circular, constant target intensity.*

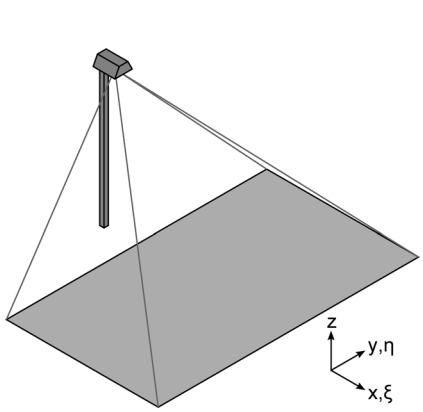
Elaborating (2.24) we find

$$\widetilde{G}(p,q) = d^2 \frac{\left(p^2+q^2+1\right)^3}{\left(p^2+q^2-1\right)^3} L_0. \tag{4.2}$$

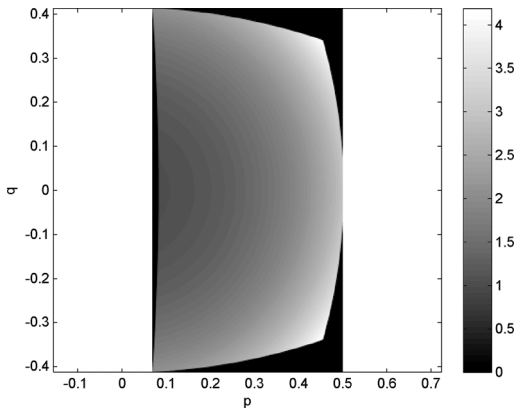
We solve the inverted problem, because $\mathcal{T}$ is not convex, as can be seen in Figure 5(b). We choose

$$f(p,q) = \frac{4\,\widetilde{G}(p,q)}{\left(p^2+q^2+1\right)^2}, \quad g(x,y) = M(x,y), \tag{4.3}$$

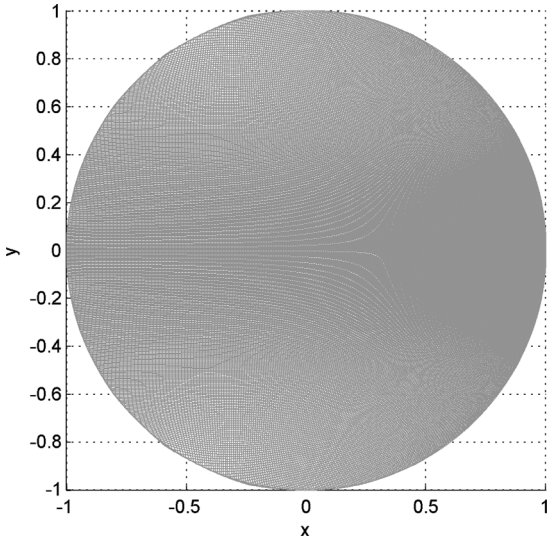
and $M_x = 200$. Because we need a square grid, we choose $M_y = 383$. In 15 Newton steps, the solver iterates to a solution with $|\boldsymbol{N}(\boldsymbol{u})|_1 / M_x M_y < 10^{-10}$. The gradient of the solution can be seen in Figure 5(c). The map is inverted using the Legendre–Fenchel transform. The inverted map is shown in Figure 5(d). The last map shows



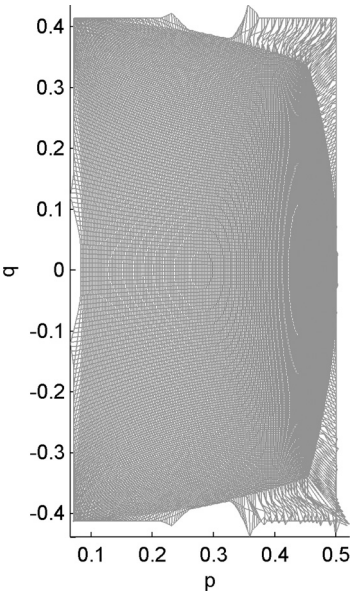
(a) Sketch of the lamppost



(b) Function $f(p,q)$, which is the converted target illuminance of the reflector



(c) Gradient map of the inverse problem

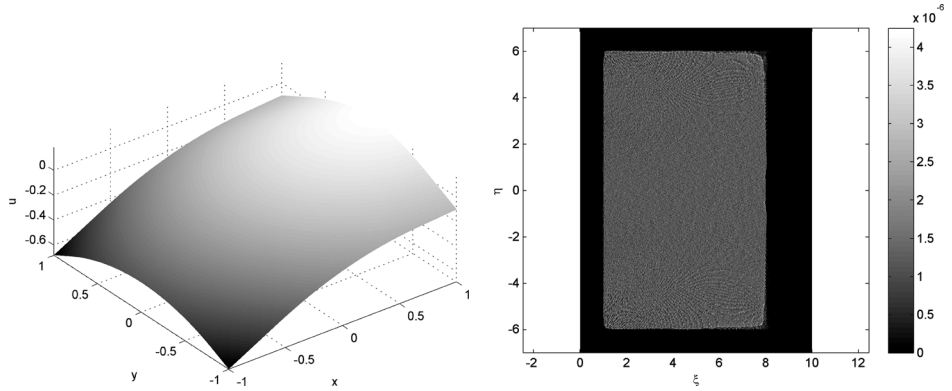


(d) Gradient of the reflector

FIG. 5. *The target function and gradient maps of the concave reflector for a lamppost with uniform target illuminance.*

some irregularities at the boundary. These irregularities correspond to parts of the inverse reflector where $\widetilde{G}(p,q)$ is zero, and therefore the maximum in (3.24) is not well-defined. These irregularities are not a problem in practice, as little or no light is involved. The final reflector can be seen in Figure 6(a). A constant is added to $u(x,y)$ to lift it above $z = 0$. The surface is exported to LightTools and verified using raytracing. The raytracing result is shown in Figure 6(b). The result shows the expected rectangular uniform illumination pattern.

4.3. A painting. For the last test, we use the Dutch painting "Girl with a Pearl Earring," shown in Figure 7(a). The painting is a work by Johannes Vermeer from the 17th century. In 2006 it was elected as the most beautiful Dutch painting in an



(a) Reflector surface (b) Target illuminance in the simulation

FIG. 6. *Reflector and simulation result for the lamppost.*

election organized by the newspaper Trouw. The painting was also subject of a novel and a movie with the same name [37].

We construct a convex reflector that illuminates a vertical wall with an illuminance $L(\xi, \eta)$ given by an interpolation of the summed RGB values of an image file of the painting. The light source is a square with constant emittance $M(x, y) = 1$. Again the illuminance is scaled to satisfy (2.26). The vertical wall is located parallel to the x-z plane at distance $d = 1$. The painting is projected on the rectangle $(-0.665, 0.6650) \times \{1\} \times (-0.7875, 0.7875)$.

We use (2.19) and (2.24) with $d = 1$, $\hat{\boldsymbol{n}}_P = (0, -1, 0)^T$, $\hat{\boldsymbol{e}}_\xi = (1, 0, 0)$, and $\hat{\boldsymbol{e}}_\eta = (0, 0, 1)$. We find

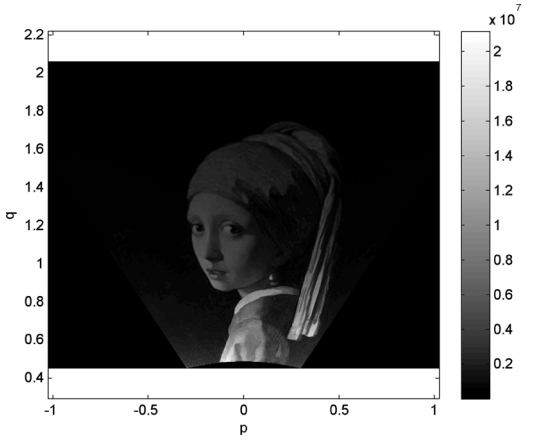
$$\xi = -\frac{p}{q}, \quad \eta = \frac{p^2 + q^2 - 1}{2\,q}, \tag{4.4a}$$

$$\widetilde{G}(p, q) = \frac{\left(p^2 + q^2 + 1\right)^3}{8q^3} L(\xi(p, q), \eta(p, q)). \tag{4.4b}$$

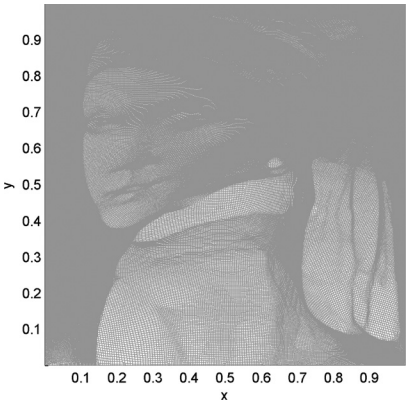
The target function is plotted in Figure 7(b). Again the converted target is not convex, so we choose f as the target and g as the source in the Monge–Ampère solver. We solved the equation on a 766×601 grid. In 21 Newton steps, the solver iterated to a solution with $\frac{|\boldsymbol{N}(\boldsymbol{u})|_1}{M_x M_y} < 10^{-7}$. The solution was converted with a Legendre–Fenchel transform to a 601×601 grid. The resulting gradient maps are shown in Figures 7(c) and 7(d), and the resulting reflector is shown in Figure 7(e). Also this reflector was evaluated using LightTools. The illuminance on the wall from the evaluation can be seen in Figure 7(f). The details of the original picture 7(a) are clearly visible. Also, the boundary of the image is close to rectangular.

5. Summary, discussion, and conclusions. We developed a new method to design reflectors for lighting applications where the light source is a parallel beam. We derived the differential equation and boundary condition governing the reflector surface and applied state of the art numerical methods to find solutions. This is the first time a numerical method has been published for free-form reflector design that can be used with the complicated boundary condition for this problem and that is efficient and stable enough to handle detailed features and large contrasts in the target

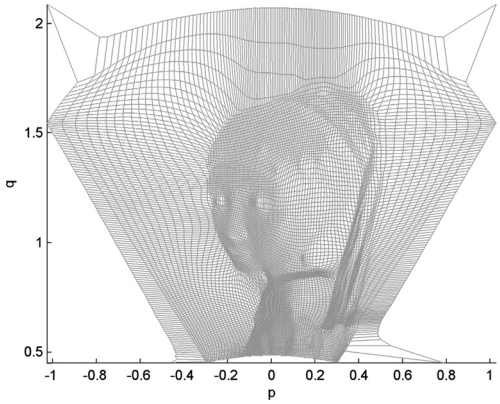
(a) The painting "Girl with a Pearl Earring" by Johannes Vermeer



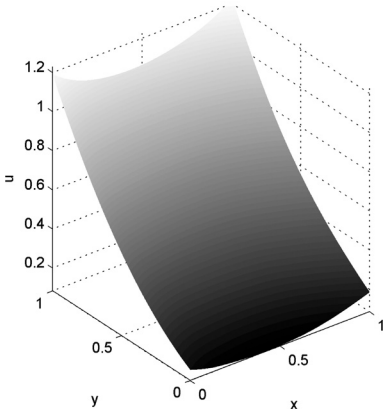
(b) Function $f(p, q)$, which is the converted target illuminance of the reflector



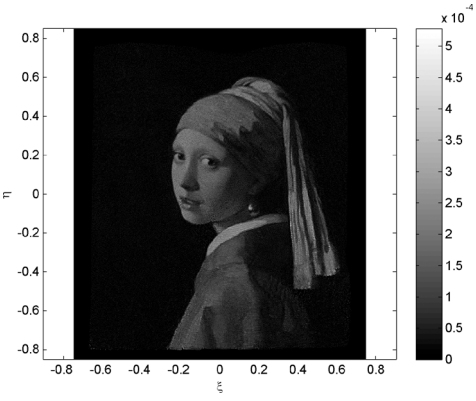
(c) Gradient map from f to g



(d) Gradient of the reflector. Because of the large number of grid lines, only 1/5 of the lines is plotted



(e) The reflector surface



(f) The target illuminance in the simulation

FIG. 7. *Gradient maps of the reflector for the painting.*

distribution. The method is tested for a circular beam of light with uniform intensity, a street light, and a projection of a painting on a wall. The reflectors are tested using commercial ray tracing software, with promising results.

A convergence proof exists for the numerical method for the Monge–Ampère equation described in this article. However, this proof does not fully apply to our implementation, because the implementation of the boundary condition is not monotone. A similar issue arises in the discretization of the interior near the boundary as given in (3.14). In practice, we did not encounter any convergence problems. It is probably possible to combine a high accuracy boundary implementation with provable convergence using a filtered scheme, such as described in [17]. This may be a topic of future research.

A next step in our research is an extension of this method to lens design. This would make the method applicable to the practical problem of the design of microlens arrays for illumination. Furthermore, we would like to extend the method to point light sources and more general types of light sources. This last problem is more challenging, because the Monge–Ampère equation governing these type of problems contains additional second order derivatives and is thus not of the same form as in this paper. This equation has a different condition for uniqueness of the solution, equivalent to, but different from, convexity of the surface. We are very interested in a Monge–Ampère solver, suitable for reflector and lens surfaces for point light sources and more general light sources.

REFERENCES

[1] S. Angenent, S. Haker, and A. Tannenbaum, *Minimizing flows for the Monge–Kantorovich problem*, SIAM J. Math. Anal., 35 (2003), pp. 61–97.

[2] G.B. Barles and P.E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptot. Anal., 4 (1991), pp. 271–283.

[3] J. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393.

[4] J. Benamou, B.D. Froese, and A.M. Oberman, *Numerical solution of the optimal transportation problem using the Monge–Ampère equation*, J. Comput. Phys., 260 (2014), pp. 107–126.

[5] J. Benamou, B.D. Froese, and A.M. Oberman, *Numerical solution of the optimal transportation problem via viscosity solutions for the Monge–Ampère equation*, preprint, arXiv:1208.4873, 2012.

[6] M. Born and E. Wolf, *Principles of Optics*, 7th ed., Cambridge University Press, Cambridge, 1999.

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004; also available from http://www.stanford.edu/∼boyd/cvxbook/.

[8] T.A. Davis, *A column pre-ordering strategy for the unsymmetric-pattern multifrontal method*, ACM Trans. Math. Softw., 30 (2004), pp. 165–195.

[9] T.A. Davis, *Algorithm* 832: *UMFPACK, an unsymmetric-pattern multifrontal method*, ACM Trans. Math. Software, 30 (2004), pp. 196–199.

[10] T.A. Davis and I.S. Duff, *An unsymmetric-pattern multifrontal method for sparse LU factorization*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 140–158.

[11] T.A. Davis and I.S. Duff, *A combined unifrontal/multifrontal method for unsymmetric sparse matrices*, ACM Trans. Math. Software, 25 (1999), pp. 1–19.

[12] F. Fournier, *Freeform Reflector Design with Extended Sources*, Ph.D. thesis, University of Central Florida, Orlando, FL, 2010; also available online from http://etd.fcla.edu/CF/CFE0003311/Fournier_Florian_R_201008_PhD.pdf.

[13] F.R. Fournier, W.J. Cassarly, and J.P. Rolland, *Fast freeform reflector generation using source-target maps*, Opt. Express, 18 (2010), pp. 5295–5304.

[14] B.D. Froese, *A numerical method for the elliptic Monge–Ampere equation with transport boundary conditions*, SIAM J. Sci. Comput., 34 (2012), pp. A1432–A1459.

[15] B.D. Froese and A.M. Oberman, *Convergent finite difference solvers for viscosity solutions of the elliptic Monge–Ampére equation in dimensions two and higher*, SIAM J. Numer. Anal., 49 (2011), pp. 1692–1714.
[16] B.D. Froese and A.M. Oberman, *Fast finite difference solvers for singular solutions of the elliptic Monge–Ampère equation*, J. Comput. Phys., 230 (2011), pp. 818–834.
[17] B.D. Froese and A.M. Oberman, *Convergent filtered schemes for the Monge–Ampère partial differential equation*, SIAM J. Numer. Anal., 51 (2013), pp. 423–444.
[18] A.S. Glassner, *An Introduction to Ray Tracing*, Academic Press, New York, 1991.
[19] T. Glimm and V.I. Oliker, *Optical design of single reflector systems and the Monge–Kantorovich mass transfer problem*, J. Math. Sci., 117 (2003), pp. 47–66.
[20] P. Guan and X.J. Wang, *On a Monge–Ampère equation arising in geometric optics*, J. Differential Geom., 48 (1998), pp. 205–223.
[21] E. Haber, T. Rehman, and A. Tannenbaum, *An efficient numerical method for the solution of the l_2 optimal mass transfer problem*, SIAM J. Sci. Comput., 32 (2010), pp. 197–211.
[22] E. Hecht, *Optics*, 4th ed., Addison-Wesley, Reading, MA, 2002.
[23] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[24] S.A. Kochengin and V.I. Oliker, *Determination of reflector surfaces from near-field scattering data*, Inverse Problems, 13 (1996), pp. 363–373.
[25] S.A. Kochengin and V.I. Oliker, *Determination of reflector surfaces from near-field scattering data* II. *Numerical solution*, Numer. Math., 79 (1998), pp. 553–568.
[26] S.A. Kochengin and V.I. Oliker, *Computational algorithms for constructing reflectors*, Comput. Vis. Sci., 6 (2003), pp. 15–21.
[27] *Ora Lighttools*, http://www.opticalres.com/lt/ltprodds_f.html (31 May 2013).
[28] A.P. Norris and B.S. Westcott, *Realisation of generalised far fields by reflector synthesis*, Electron. Lett., 10 (1974), pp. 322–323.
[29] A.M. Oberman, *Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton–Jacobi equations and free boundary problems*, SIAM J. Numer. Anal., 44 (2006), pp. 879–895.
[30] A.M. Oberman, *Wide stencil finite difference schemes for the elliptic Monge–Ampère equation and functions of the eigenvalues of the Hessian*, Discrete Contin. Dyn. Syst. Ser. B, 10 (2008), pp. 221–238.
[31] W. Rengmao, X. Liang, L. Peng, Z. Yaqin, Z. Zhenrong, L. Haifeng, and L. Xu, *Freeform illumination design: A nonlinear boundary problem for the elliptic Monge–Ampère equation*, Optics Lett., 38 (2013), pp. 229–231.
[32] W. Rengmao, L. Peng, Z. Yaqin, Z. Zhenrong, L. Haifeng, and L. Xu, *A mathematical model of the single freeform surface design for collimated beam shaping*, Optics Express, 21 (2013), pp. 20974–20989.
[33] J.S. Schruben, *Formulation of a reflector-design problem for a lighting fixture*, J. Optical Soc. Amer., 62 (1972), pp. 1498–1501.
[34] J. Urbas, *On the second boundary value problem for equations of Monge–Ampère type*, J. Reine Angew. Math., 487 (1997), pp. 115–124.
[35] C. Villani, *Topics in Optimal Transportation*, AMS, Providence, RI, 2003.
[36] X.J. Wang, *On the design of reflector antenna*, Inverse Problems, 12 (1996), pp. 351–375.
[37] *Meisje met de parel*, http://nl.wikipedia.org/wiki/Meisje_met_de_parel (30 May 2013).