COMPUTATIONAL NONIMAGING GEOMETRIC OPTICS: MONGE-AMPÈRE

GERARD AWANOU

1. Overview

The goal of computational nonimaging geometric optics is the efficient design of optical lenses and mirrors for the accurate control of light. Light waste in the United States is equivalent to 72.9 million mwh of unnecessary electricity generated at a cost of \$ 6.9 billion a year [10] and the amount of CO2 generated in that process is equivalent to 9.5 million cars on the roads. Light pollution also has adverse health impacts on wildlife and humans. Other examples where an accurate control of light is required include projection displays, laser weapons, concentrated solar energy and medical illuminators. Freeform illumination design, i.e. with no a priori symmetry assumption, often leads to solving numerically a nonlinear second order partial differential equation of Monge-Ampère type with non local boundary conditions. For a review of other approaches we refer to [8, section 2].

Let Ω and Ω^* be two bounded convex domains of $\mathbb{R}^d$. We are interested in the redistribution of an incoming source of light with density $f \in L^1(\Omega), f \geq 0$, by a surface defined by a function u on Ω, into a prescribed irradiance described by a density $R \in L^1(\Omega^*), R > 0$. Conservation of energy requires $\int_\Omega f(x)dx = \int_{\Omega^*} R(p)dp$. In the case the surface represents a mirror, light is reflected and we will say that we have a reflector problem. In the case of a lens, light is transmitted with a new direction of travel, i.e. the light is refracted. We will refer to this as a refractor problem. One often makes the assumption of an idealized point light source. Another design we will consider is based on the assumption that the incoming light is collimated, i.e. has parallel rays. As for the target, when it is very far from the source, the light output can be described with a set of directions on the unit sphere. This is referred to as a far-field problem. The combination of these design constraints lead to the type of problems we consider in this review, i.e. the parallel near field reflector problem, the point source far field refractor problem etc.

For the parallel far field reflector problem, $\Omega \subset \mathbb{R}^d$ and Ω^* is identified with a subset of $\mathbb{R}^d$ which is the stereographic projection of a domain of the unit sphere in $\mathbb{R}^{d+1}$. In this case, a ray originating at $x \in \Omega$ is reflected by the mirror described by the graph of u into the point $Du(x)$, the gradient of u at x. It is shown for example in [19] that one can choose u convex solving the Monge-Ampère equation

$$(1.1) \qquad R(Du(x)) \det D^2u(x) = f(x) \text{ in } \Omega,$$

with the natural boundary condition

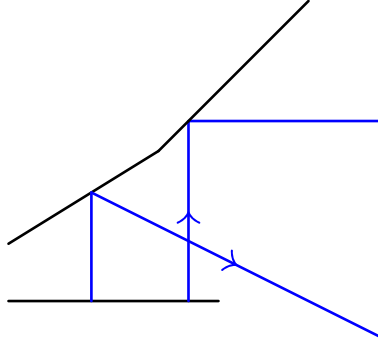$$(1.2) \qquad Du(\Omega) = \Omega^*.$$

FIGURE 1. Convex reflector for the density R discretized with two Dirac masses, c.f. section 2.5. The target Ω^* is not shown.

For a smooth function u, $D^2u = \left(\partial^2 u/(\partial x_i \partial x_j)\right)_{i,j=1,\ldots,d}$ denotes its Hessian matrix and $\det D^2u$ is its determinant. Figure 1 illustrates a convex reflector which redirects a parallel light beam into a finite number of directions of the unit sphere.

Problem (1.1)-(1.2) also appear in optimal transport problems as we discuss in section 3. In general, problems in geometric optics lead to more general Monge-Ampère equations

$$(1.3) \qquad \det[D^2u - A(.,u,Du)] = B(.,u,Du), \quad T(.,u,Du)(\Omega) = \Omega^*,$$

which may not have interpretations as optimal transport problems. The numerical resolution of (1.1) in conjunction with the Dirichlet boundary condition

$$(1.4) \qquad\qquad\qquad\qquad\qquad u = g \text{ on } \partial\Omega,$$

for a function g continuous on $\partial\Omega$, has been the subject of several reviews, c.f. [17] for the latest.

Our focus in this review is on numerical methods for (1.3). We first start with the model problem (1.1)-(1.2) in section 2. The discretizations are based on the kind of solutions for (1.1), classical solutions and various notions of weak solutions. At this point we mainly consider the recent discretization of (1.2) from [2]. We introduce the setting of generated Jacobian equations for (1.3) in section 3 where we review two notions of weak solutions for (1.3). We then discuss the convergence analysis of some of the methods in that setting in section 4. It is here that we review other discretizations of the second boundary condition. We conclude with a list of possible future directions.

This review focuses on computational aspects of generated Jacobians. For insights about the general theory, and applications beyond optics, we refer to the excellent recent review [11].

2. Numerical methods for the second boundary value problem for the Monge-Ampère equation

The constraint (1.2) is referred to as the second boundary value condition for (1.1) because it was studied much later than the Dirichlet boundary condition (1.4). For
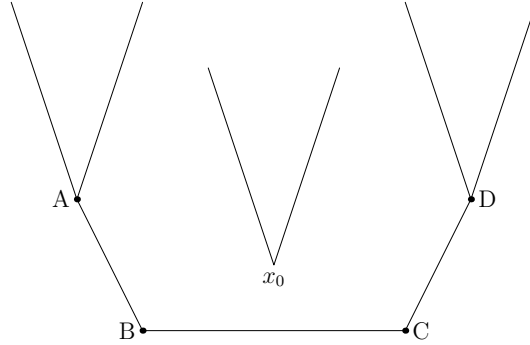
FIGURE 2. The convex hull M of $\{A, B, C, D\}$ defines a piecewise linear convex function on a finite interval. The convex hull M* of M and the cone $x_0 + K_{\Omega^*}$ with $\Omega^* = (-3, 3)$ defines a piecewise linear convex function on the real line with asymptotic cone K_{Ω^*}.

a smooth strictly convex function u, (1.2) was shown in [19] to be equivalent to

$$(2.1) \qquad Du(\partial\Omega) = \partial\Omega^*,$$

which looks more like a boundary condition and non local. Several approaches have been proposed to enforce (1.2) in a numerical scheme. We review most of them in section 4. Below we focus on discretizations of the differential operator in conjunction with the approach through asymptotic cone of [2] for enforcing (1.2). The constraint (1.2) can also be enforced directly by seeking piecewise linear functions with points in Ω^* as their piecewise gradients, c.f. section 2.5.

Next, we interpret (1.1)-(1.2) as a problem in the geometry of convex surfaces. We recall that a set $K \subset \mathbb{R}^d$ is a cone if $tx \in K$ for all $t \geq 0$ and $x \in K$. We associate to the domain Ω^* the cone

$$K_{\Omega^*} = \cap_{p \in \overline{\Omega^*}}\{\,(x, z) \in \mathbb{R}^d \times \mathbb{R}, z \geq p \cdot x\,\}.$$

Given a convex function u on Ω, recall that its epigraph is the convex set

$$M = \{\,(x, z) \in \mathbb{R}^d \times \mathbb{R}, z \geq u(x)\,\}.$$

The convex hull M_* of M and the set $(x_0, u(x_0)) + K_{\Omega^*}$, for $x_0 \in \partial\Omega$, defines an infinite convex hypersurface whose boundary defines a convex function $\tilde{u}$ on $\mathbb{R}^d$. For any $y \in M_*$, $y + K_{\Omega^*} \subset M_*$. The convex function u is said to have asymptotic cone K_{Ω^*} if $\tilde{u} = u$ on Ω. See Figure 2 for an illustration.

Problem (1.1)-(1.2) has the formulation: find a convex function u on Ω with asymptotic cone K_{Ω^*} such that (1.1) holds. If u has asymptotic cone K_{Ω^*}, it is shown in [2] that

$$(2.2) \qquad \tilde{u}(x) = \inf_{s \in \Omega} u(s) + \sup_{p \in \Omega^*} (x - s) \cdot p, x \notin \Omega.$$

For example, if $\Omega = (-1, 1)$ and $\Omega^* = (-1/2, 1/2)$, then K_{Ω^*} is the epigraph of the function $y = |x|/2$. Example of functions with asymptotic cone K_{Ω^*} are given by $u_1(x) = |x|/2$ and $u_2(x) = 0$ for $-1 \leq x \leq 1$ with $u_2(x) = |x|/2 - 1/2$ for $x \notin \Omega$.

In the sequel we will approximate the convex domain Ω^* by polygons $K^* \subset \overline{\Omega^*}$. The resulting approximate problems are shown to be convergent in [2]. Assuming now for simplicity that $K^* = \overline{\Omega^*}$, it can be shown [2] that Problem (1.1)-(1.2) is equivalent to finding a convex function u on Ω which extends to a convex function on $\mathbb{R}^d$ by (2.2) and such that (1.1) holds. Let $a_j^*, j = 1, \ldots, N$ denote the vertices of K^*. It can then be shown that

$$(2.3) \qquad \tilde{u}(x) = \inf_{s \in \Omega} u(s) + \max_{j=1,\ldots,N} (x - s) \cdot a_j^*, x \notin \Omega.$$

The infimum can be further restricted to boundary points of a computational mesh.

2.1. Standard discretizations.

By standard discretizations, we refer to discretizations based on the interpretation of the solution u of (1.1) as a classical $C^2(\Omega)$ solution. In that case

$$\det D^2 u = \operatorname{div} \big((\operatorname{cof} D^2 u) Du \big),$$

where $\operatorname{cof} A$ denote the cofactor matrix of the matrix A and div denotes the divergence operator. Thus (1.1) can be seen as a nonlinear Poisson equation. Pretty much methods developed for elliptic problems can be applied to (1.1)-(2.2). The resulting nonlinear discrete equations may have multiple solutions and cannot be solved by a vanilla Newton's method when the goal is to reproduce a non smooth solution. Iterative methods which preserve a notion of discrete convexity can be used. A particular solution was selected in [19] in a least squares setting with a mixed approximation, i.e. the introduction of new variables $m = Du$ and $P = Dm$. Therein, the second boundary condition was also enforced in a least squares setting. A least squares solution in $\mathbb{R}^N$ of a system of linear equations $Ax = b$, is a vector x which minimizes $||b - Ax||^2$ for the Euclidean norm $||.||$ on $\mathbb{R}^N$.

2.2. Semi-discretizations for Aleksandrov solutions.

The semi-discrete problem considered here is obtained by approximating the density f with a sum of Dirac masses $f_M = \sum_{i=1}^{M} \mu_i \delta_{x_i}$, for an integer M, weights $\mu_i \geq 0$ and $x_i \in \Omega$. We will assume that $R = 1$. For illustration, we consider a one dimensional Monge-Ampère equation, i.e. find a convex function u on $(0, 1)$ such that in a weak sense

$$u'' = \sum_{i=1}^{M} \mu_i \delta_{x_i} \text{ in } (0, 1),$$

and for all $x \in (0, 1)$ we have $u'(x) \in (-1, 2)$, i.e. $u'(0, 1) = (-1, 2)$.

For u smooth and a Borel set $B \subset (0, 1)$, the Monge-Ampère measure associated to u is defined as $M[u](B) = \int_B u''(x)\, dx$. By the change of variable $x \to \gamma(x) = u'(x) = p$ (gradient mapping) we obtain $M[u](B) = \int_{\gamma(B)} dp$. Next, we replace $\gamma(x)$ by the subgradient mapping for non smooth convex solutions

$$\partial u(x_0) = \{\, p \in \mathbb{R} : u(x) \geq u(x_0) + p(x - x_0), \text{ for all } x \in (0, 1) \,\}.$$

For $v(x) = |x - 1/2|$, we have $\partial v(x_0) = \{-1\}, x_0 < 1/2, \partial v(1/2) = [-1, 1]$ and $\partial v(x_0) = \{1\}, x_0 > 1/2$.

For a Borel set B, we have $M[u](B) = |\partial u(B)|$, where for a set S, $|S|$ denotes its Lebesgue measure. By Aleksandrov solution, we mean a convex function u such that

$\partial u(0,1) = (-1,2)$ and for all Borel sets $B \subset (0,1)$, we have $M[u](B) = \sum_{i=1}^{M} \mu_i \delta_{x_i}(B)$. We require the compatibility condition $\sum_{i=1}^{M} \mu_i = |(-1,2)| = 3$.

If we assume that the points x_i are equidistributed, i.e. $x_{i+1} - x_i = h$, then for $1 \leq i \leq M$ with $x_0 = 0$ and $x_{M+1} = 1$

$$M[u](\{x_i\}) = \frac{u_{i+1} - u_i}{h} - \frac{u_i - u_{i-1}}{h} = \frac{u_{i+1} - 2u_i + u_{i-1}}{h},$$

where $u_i \approx u(x_i)$. The main contribution in [2] is the observation that the above formula can also be used at x_1 and x_M provided that one uses the extension formula (2.3) which gives here for $i = 0, M+1$

$$u_i = \min \left\{ u_1 + \max \left(-(x_i - x_1), 2(x_i - x_1) \right), u_M + \max \left(-(x_i - x_M), 2(x_i - x_M) \right) \right\}.$$

Solutions can be shown to be unique up to a constant. One can impose the constraint $u_0 = \alpha$ for an arbitrary number α. The value at x_{M+1} is given by the above formula. The discretization just described generalizes a similar discretization proposed for the Dirichlet problem in [18]. Details for arbitrary dimensions will be reported elsewhere.

2.3. **Medius approach: lattice basis reduction.** It is possible to exploit the arithmetic of two dimensional Cartesian meshes for an efficient resolution of the non-linear discrete system obtained from a semi-discretization for Aleksandrov solutions. Though one has to relax the convexity criterion. The scheme we describe can be seen as a finite difference version of the one described in the previous section. It can be implemented through an efficient adaptive algorithm. Again, we present a variant of an existing scheme [6]. The modification was crucial for the proof of existence and uniqueness of a solution to the discrete problem [2].

The discrete operator is written as a minimization problem over subsets of the mesh. The mesh is identified with a tree and the adaptive algorithm selects subtrees [5], hence the name lattice basis reduction. By adaptivity here, we mean that the operator is evaluated in a cheap and smart way. Let h be a small positive parameter and let $\mathbb{Z}_h^2 = \{ mh, m \in \mathbb{Z}^2 \}$ denote the orthogonal lattice with mesh length h. The set of mesh points is given by $\Omega_h = \Omega \cap \mathbb{Z}_h^2$. Computations are made with a finite subset of the mesh. Let V denote a finite set of non zero elements of $\mathbb{Z}^2$ such that if $e \in V, -e \in V$. It is further assumed that elements of V have co-prime coordinates, span $\mathbb{R}^2$ and V contains the elements of the canonical basis of $\mathbb{R}^2$ and a normal to each side of the target polygonal domain K^*. We also require that V contains $\{ (a,b) \in \{-1,0,1\}^2, ab \neq 0 \}$.

Let (e_1, e_2) denote the canonical basis of $\mathbb{R}^2$ and let

$$\partial\Omega_h = \{ x \in \Omega_h \text{ such that for some } i = 1,2, x + he_i \notin \Omega_h \text{ or } x - he_i \notin \Omega_h \}.$$

We define for a function v_h on $\mathbb{Z}_h^2$, $e \in \mathbb{Z}^2$ and $x \in \Omega_h$

$$\Delta_{he}v_h(x) = v_h(x + he) - 2v_h(x) + v_h(x - he).$$

We are interested in mesh functions on Ω_h which are extended to $\mathbb{Z}_h^2$ using

(2.4) $$\tilde{u}(x) = \inf_{s \in \partial\Omega_h} u(s) + \max_{j=1,\dots,N} (x - s) \cdot a_j^*, x \notin \Omega,$$

and are discrete convex in the sense that $\Delta_{he}v_h(x) \geq 0$ for all $x \in \Omega_h$ and $e \in V$. A uniform limit of mesh functions which are discrete convex in the sense above and

solve suitable discrete Monge-Ampère equations is a convex function. It is a result implicit in convergence studies of discretizations of (1.1) with $R = 1$.

Next, we consider a local version of a symmetrization of a discrete version of the subgradient

$$D_L v_h(x) = \{\, p \in \mathbb{R}^2, 2p \cdot (he) \leq \Delta_{he} v_h(x) \, \forall e \in L \,\}.$$

Define a basis of $\mathbb{Z}^2$ as a pair $(e_1, e_2) \in (\mathbb{Z}^2)^2$ such that $|\det(e_1, e_2)| = 1$. A superbase of $\mathbb{Z}^2$ is a triplet $(e_0, e_1, e_2) \in (\mathbb{Z}^2)^3$ such that $e_0 + e_1 + e_2 = 0$, and (e_1, e_2) is a basis of $\mathbb{Z}^2$. The Monge-Ampère operator with lattice basis reduction for $R = 1$ is defined as

$$MA_{LBR}\, u_h(x) = \min_{\substack{L=(e_0,e_1,e_2)\in V^3 \\ superbase}} |D_L u_h(x)|.$$

The discrete problem consists in finding a discrete convex mesh function u_h such that

$$(2.5) \qquad\qquad MA_{LBR}\, u_h(x) = \int_{E_x} f(t)dt, x \in \Omega_h,$$

where $E_x = x + [-h/2, h/2]^d$ is a cube centered at x with $E_x \cap \Omega_h = \{\, x \,\}$. The unknowns in the above equation are the mesh values $u_h(x), x \in \Omega_h$. For $x \notin \Omega_h$, the value $u_h(x)$ needed for the evaluation of $D_L v_h(x)$ is obtained from the extension formula (2.4). Here we made the simplifying assumption that $\Omega^* = K^*$ so that conservation of energy holds. Again here we impose the constraint $u_h(x_0) = \alpha$ for an arbitrary real number α and with $x_0 \in \Omega_h$ for all h. A damped Newton's method can be used for solving the nonlinear equations for $f > 0$.

2.4. **Approach through viscosity solutions.** The notion of viscosity solution for (1.1) is based on comparisons with smooth test functions. Aleksandrov solutions of (1.1) are equivalent to viscosity solutions when the right-hand side f is continuous and positive [13]. For $R = 1$, it was shown in [4] through a perturbation argument that the equivalence also holds.

For solutions of schemes to converge to a viscosity solution, it is convenient that the scheme satisfies a monotonicity property allowing comparison with smooth test functions. This often requires to write discretizations in a specific form and for schemes which violate the monotonicity condition, it is very difficult to prove convergence in the setting of viscosity solutions. For example, the scheme (2.5) may require a numerical integration leading to non monotone schemes. However, we believe that convergence can still be proven in the setting of Aleksandrov solutions through a perturbation argument taking advantage of convergence results for (2.5) which are essentially the same as the ones discussed in [2]. In fact, a non monotone approximation of $MA_{LBR}\, u_h(x)$ through a standard discretization of the gradient was actually considered in [6].

The geometric content of solutions to the Monge-Ampère equation is lost in the viscosity solution setting. It is unlikely to explain the behavior of standard discretizations for (1.1)-(1.2) and non smooth solutions. We refer to [7, 12] for explicit monotone discretizations of (1.1). We do not discuss this further since their analysis in conjunction with (2.4) is similar to the analysis of the effect of numerical integration for (2.5), a topic we wish to discuss in a separate work.

2.5. Semi-discretizations for Brenier solutions. Here we assume that the density R is approximated by a sum of Dirac masses $\sum_{i=1}^{M} r_i \delta_{P_i}$ for $P_i \in \Omega^*$ and $r_i \in \mathbb{R}$ for all i. Energy conservation reads $\sum_{i=1}^{M} r_i = \int_{\Omega} f(x)\, dx$. In the case of one Dirac mass $r_i \delta_{P_i}$ the surface which reflects all rays with direction $(0, \ldots, 0, 1) \in \mathbb{R}^{d+1}$ from Ω into a direction of the unit sphere with stereographic projection the point P_i is, by Snell's law, given by a plane $x \cdot P_i - b_i$ for a parameter b_i. The reflector is then given by the graph of the convex function

$$u_M(x) = \max_{i=1,\ldots,M} x \cdot P_i - b_i,$$

with rays in the region

$$W_i(b) = \{\, x \in \Omega, x \cdot P_i - b_i \geq x \cdot P_j - b_j \text{ for all } j = 1, \ldots, M \,\},$$

reflected in the direction P_i. We thus need

$$\int_{W_i(b)} f(x)\, dx = r_i, i = 1, \ldots, M,$$

which is the nonlinear equations to be solved for $b_i, i = 1, \ldots, M$. The constraint (1.2) is enforced implicitly in the sense that by construction $\partial u_M(\Omega) \subset \Omega^*$ when Ω^* is convex.

3. Generated Jacobian equations

Generated Jacobian equations are a class of prescribed Jacobian equations, i.e. one seeks a mapping T between two bounded domains Ω and Ω^* of $\mathbb{R}^d$ whose Jacobian $\det DT(x)$ is prescribed by an equation

$$\det DT(x) = \psi(x, T(x)),$$

for a given function ψ on $\Omega \times \Omega^*$. An example of such a mapping is the optimal transport map, discussed above, between two measures supported respectively on Ω and Ω^*. In that case, one requires $\psi \geq 0$ and the mapping T is generated by a convex function u on Ω in the sense that $T(x) = Du(x)$. In geometric optics problems, $\psi \geq 0$ as well and the mapping T_u taking a light described by $x \in \Omega$ into a point $T_u(x)$ on the target, is also generated by a scalar function u on Ω which describes the optical surface and solves the generated Jacobian equation

(3.1) $$\det DT_u(x) = \psi(x, u(x), T_u(x)), \quad T_u(x) = T(x, u(x), Du(x)),$$

where now ψ and T are functions on $\Omega \times \mathbb{R} \times \Omega^*$ which take values in $\mathbb{R}$ and $\mathbb{R}^d$ respectively. We assume in this paper that ψ is separable in the sense that

$$\psi(x, u, p) = \frac{f(x)}{R(T(x, u, p))},$$

for positive functions $f \in L^1(\Omega)$ and $g \in L^1(\Omega^*)$. This structural assumption encompasses applications in geometric optics. The second boundary value problem for the generated Jacobian equation (3.1) is to prescribe in addition the image of Ω by T_u, i.e.

(3.2)
$$R(T_u(x)) \det DT_u(x) = f(x), x \in \Omega$$
$$T_u(\Omega) = \Omega^*.$$

The transformation T and the "potential" u are now related through a generating function $G : \overline{\Omega} \times \overline{\Omega^*} \times \mathbb{R}^+ \mapsto \mathbb{R}$ and $T(x, u, p)$ is obtained by solving the system

$$D_x G(x, T, Z) = p \qquad G(x, T, Z) = u,$$

where Z is an additional unknown. It is assumed that the above system has a unique solution, that G is sufficiently smooth and strictly decreasing with respect to z. When $G(x, y, z) = x \cdot y + \log z$, we obtain $T(x, y, p) = p$, i.e. $T_u(x) = Du(x)$ and ψ does not depend on $u(x)$. The same holds for optimal transport problems with a general cost function $c(x, y)$, with compatible assumptions on $c : \Omega \times \Omega^* \to \mathbb{R}$, in which case $G(x, y, z) = c(x, y) + \log z$. The mapping T_u is then the optimal transport map, i.e.

$$T_u(x) = \arg\min_T \int_\Omega c(x, T(x)) f(x) dx,$$

where the infimum is taken over mappings T which pushforward the measure with density f onto the measure with density g, that is, mappings T which satisfy

$$\int_\Omega \phi(T(x)) f(x) dx = \int_{\Omega^*} \phi(y) g(y) dy.$$

for all continuous functions $\phi : \Omega^* \to \mathbb{R}$.

For weak solutions of (3.2) one has, as in the case of (1.1)-(1.2) the Aleksandrov theory and the Brenier formulation.

The functions $x \mapsto G(x, ., .)$ play the role hyperplanes play as support functions in the theory of convex functions. Given $y_0 \in \Omega^*$ and $\lambda_0 \in \mathbb{R}^+$, the function $x \mapsto G(x, y_0, \lambda_0)$ is said to be a G-support to a function $u : \Omega \to \mathbb{R}$ at $x = x_0 \in \Omega$ if $u(x) \geq G(x, y_0, \lambda_0) \forall x \in \Omega$ with equality at $x = x_0$. The function u is said to be G-convex if it has a G-support at all points $x \in \Omega$. Equivalently u is G-convex if and only if there exists a set $\mathcal{A} \subset \Omega^* \times \mathbb{R}^+$ such that

$$u(x) = \sup_{(y, \lambda) \in \Omega^* \times \mathbb{R}^+} G(x, y, \lambda).$$

The G-subdifferential of u at $x_0 \in \Omega$ is defined as the set-valued function

$$\partial_G u(x_0) = \{\, y \in \Omega^*, \exists \lambda_0 \in \mathbb{R}^+ \text{ such that } G(x, y, \lambda_0) \text{ is a G-support to } u \text{ at } x_0 \,\}.$$

It is known that for a G-convex function u, the set $\partial_G u(E) = \cup_{x \in E} \partial_G u(x)$ is measurable when E is measurable [1, Lemma 2.1] and [20, p. 12–13]. Moreover the set function

$$M[u](E) = \int_{\partial_G u(E)} g(p) dp,$$

is a Radon mesure. A weak solution of (3.2) in the sense of Aleksandrov, is a G-convex function u such that

$$M[u](E) = \int_E f(x) dx \text{ for all Borel sets } E \subset \Omega.$$

Thus we have the necessary condition
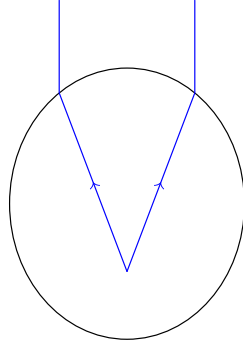
$$\int_\Omega f(x) dx = \int_{\Omega^*} g(p) dp.$$

FIGURE 3. An example of surface with a uniform refracting property

The tracing map of u for $y_0 \in \Omega^*$ is defined as

$$\tau_G u(y_0) = \{\, x_0 \in \Omega, \exists \lambda_0 \in \mathbb{R}^+ \text{ such that } G(x, y_0, \lambda_0) \text{ supports } \phi \text{ at } x_0 \,\}.$$

Note that τ_G is the inverse of the G-subdifferential, i.e. $\tau_G u(y_0) = (\partial_G u)^{-1}(y_0)$ and can be interpreted as the set of directions from which light emanating from the origin is redirected in the direction y.

For a subset $F \subset \Omega^*$ we define $\tau_G u(F) = \cup_{y \in F} \tau_G u(y)$ and for a G-convex function u, we define the set function

$$(3.3) \qquad \eta_u(F) = \int_{\tau_G u(F)} f(x) dx.$$

A weak solution of (3.2) in the sense of Brenier, is a G-convex function u such that

$$(3.4) \qquad \eta_u(F) = \int_F g(p) dp \text{ for all Borel sets } F \subset \Omega^*.$$

Explicit expressions of the generating function G and the terms for the differential equation (1.3) can be found in [20, 8]. For the far field reflector problem with an incoming parallel light beam, as in section 2.5, the surface $x \mapsto x \cdot P_i - b_i$ gives the generating function $G(x, P_i, \lambda_0) = x \cdot P_i + \log \lambda_0$. In general, the generating function describes a basic optical surface which converts light from $x \in \Omega$ into $y \in \Omega^*$. The optical surface can be made of ellipses, parabolas, hyperbolas, Cartesian ovals, etc. Figure 3 shows an ellipse refracting light from a point light source into a uniform direction.

4. CONVERGENCE OF NUMERICAL METHODS FOR GENERATED JACOBIAN EQUATIONS

There are two type of convergence to be adressed. Convergence of an iterative method for solving the discrete nonlinear system resulting from a discretization and convergence of the numerical solution to the exact solution.

Many of the developments have taken place with discretizations of (1.2) different from the approach through asymptotic cone taken in section 2. We review them below. It is our goal to systematically extend the Aleksandrov solution approach and the lattice basis reduction approach to generated Jacobian equations.

4.1. **Other discretizations of the second boundary condition.**

4.1.1. *Defining function of the target domain.* Let H be a defining function of Ω^*, i.e. $\Omega^* = \{ x \in \mathbb{R}^d, H(x) < 0 \}$. The boundary condition $Du(\Omega) = \Omega^*$ can be shown to be equivalent to

$$H(Du) = 0 \text{ on } \partial\Omega,$$

for a defining function H of Ω^*, c.f. for example [7].

Let $d_{\partial\Omega^*}$ denote the distance function to the boundary $\partial\Omega^*$ of Ω^*, i.e. $d_{\partial\Omega*}(x) = \inf_{z \in \partial\Omega^*} ||x - z||$. An example of defining function is given by the signed distance-to-the-boundary defined as

$$\delta_{\partial\Omega^*}(x) = -d_{\partial\Omega^*}(x),\ x \in \Omega^* \text{ and } \delta_{\partial\Omega^*}(x) = d_{\partial\Omega^*}(x),\ x \notin \Omega^*.$$

If the goal is to prove convergence to a viscosity solution, one choses a monotone discretization of $H(Du) = 0$ as in [7, 12].

4.1.2. *Iterated projection algorithm.* Numerical experiments in [9] suggested that the following iterative method converges to a solution of (1.1)-(1.2). Let n denote the outward normal to $\partial\Omega$. We consider the sequence u^k defined by

$$R(Du^k) \det D^2 u^{k+1} = f \text{ in } \Omega$$
$$\frac{\partial u^{k+1}}{\partial n} = \left(\text{Proj}_{\partial\Omega^*} Du^k\right) \cdot n \text{ on } \partial\Omega$$
$$u^{k+1}(x_0) = \alpha,$$

for $x_0 \in \partial\Omega$ and α a real number. Here for a vector $v \in \mathbb{R}^d$, we define

$$\text{Proj}_{\partial\Omega^*} v = \inf_{y \in \partial\Omega^*} ||y - v||.$$

4.1.3. *Enforced in a least squares sense.* Introducing new variables $m = Du$ and $P = Dm$, (2.2) becomes $m(\partial\Omega) = \partial\Omega^*$. In [19] (1.1) is written in terms of the new variables and solved along with the constraint $m(\partial\Omega) = \partial\Omega^*$ in a least squares setting.

4.1.4. *Enforced throughout the source domain.* It has been suggested in [12] that instead of enforcing $H(Du) = 0$ on $\partial\Omega$, one can enforce $H(Du) = 0$ on Ω. The motivation was to get convergence results to a viscosity solution of (1.1)-(1.2). A similar idea was previously used in [15] where the authors sought a piecewise linear convex approximation with the requirement that its piecewise gradients be vectors in Ω^* enforced as a constraint in an optimization scheme.

Although we have assumed in this paper that Ω^* is convex, several of the discretizations proposed for (1.2) should work for non convex domains, an exception being the approach based on a defining function of Ω^*.

4.2. **Numerical methods for generated Jacobian equations.** A general approach for handling (1.3) was initiated in [8] for the point source near field reflector and refractor problems. It is based on the iterative projection algorithm for handling the second boundary condition.

A priori the semi-discrete approximations with Brenier solutions can be applied to generated Jacobian equations. But this raises the practical issue of how to compute the analogues of the sets $W_i(b)$ described in section 2.5. For far field problems a computational geometry approach was used in [16].

4.3. **Convergence of iterative methods.** No convergence analysis has been reported for the iterated method based on projections proposed in [9]. Damped Newton methods have also been used. A damped Newton's method is a variant of Newton's method for which the Jacobian matrix is multiplied by a damping factor. With the goal of having convergence of the iterates independent of the closeness to the solution of an initial guess.

For the semi-discrete problems with Brenier solutions, a convergence analysis for a damped Newton's method was given in [14]. It does not cover for example the far field refractor problem which is included in the class of generated Jacobian equations for which convergence of an iterative method is proven in [1].

4.4. **Convergence of discretizations.** Not much is known about convergence of standard discretizations for (1.1)-(1.2). For smooth solutions and the Dirichlet problem, existence of a solution can be proven for various discretizations but for h sufficiently small. The least squares method for the discretization of (1.1)-(1.4) has not been analyzed for smooth solutions. We gave a theory of convergence of standard discretizations for Aleksandrov solutions of the Dirichlet problem for the Monge-Ampère equation [3]. It is based on the assumption that computers do not see the difference between a computational domain and a fictitious subdomain arbitrarily close. It would be interesting to have a theory without that assumption.

The convergence of semi-discrete approximations for Aleksandrov and Brenier solutions is central to the theory of generated Jacobian equations [20]. For the medius approach convergence of the discretization was proven in [6, 2]. We note that convergence of the discretization was also proven for both approaches [15, 12] where the second boundary condition is enforced throughout the domain.

4.5. **Performance of the numerical methods.** There have not been a comparative numerical study of discretizations for (1.1)-(1.2). No numerical experiments were reported in [2, 12]. The approach through asymptotic cones of [2] should yield results similar to the ones reported in [6] for the medius approach. Methods based on standard discretizations may not be very efficient. A possible exception is the least squares approach which has been applied to a variety of optics problems. In [15], it was reported that a method based on standard discretizations of the gradient and Hessian vastly outperforms an analogue which is provably competent. It is reasonable to expect that methods based on the iterative projection algorithm would be less efficient than a more direct approach. Understanding the mechanisms of the not so

efficient methods could give a better understanding of the computational process and may lead to more efficient numerical methods.

5. Possible future directions

Computational nonimaging geometric optics based on Monge-Ampère equations is like a painting which is largely incomplete. The main open issue is to what extent existing methods can be made more efficient. Several possible moves or combinations of the ideas discussed above are possible. For example, one can adapt to the second boundary value problem methods proposed for the Dirichlet problem [17]. A possible direction is to adapt advances in computational optimal transport to generated Jacobian equations. Also, the state of the art in computational mathematics such as fast solvers and adaptive methods have not been applied to geometric optics problems. There are also many unanswered questions which deal with the analysis of several of the numerical methods that have been proposed. In addition to several of the issues mentioned above, we give several more examples.

(1) Problems with loss of energy, i.e. when only part of the radiation is transmitted and problems with multiple source or extended fields and systems with two lenses could be addressed with recent advances on proven convergent numerical methods.
(2) The semigeostrophic flow equations were formulated as a coupled system consisting of the Monge-Ampère equation with the second boundary condition and a transport equation. It would be interesting to see how recent advances can be used for its numerical resolution.
(3) There is no theory for viscosity solutions of generated Jacobian equations. The special case (1.1)-(1.2) has been recently solved in [12].

Acknowledgments

References

[1] F. Abedin and C. E. Gutiérrez. An iterative method for generated Jacobian equations. *Calc. Var. Partial Differential Equations*, 56(4):Art. 101, 14 pp, 2017.
[2] G. Awanou. The second boundary value problem for a discrete Monge-Ampère equation with symmetrization. https://arxiv.org/abs/1910.14376.
[3] G. Awanou. Standard finite elements for the numerical resolution of the elliptic Monge-Ampère equation: Aleksandrov solutions. *ESAIM Math. Model. Numer. Anal.*, 51(2):707–725, 2017.
[4] G. Awanou and R. Awi. Convergence of finite difference schemes to the Aleksandrov solution of the Monge-Ampère equation. *Acta Applicandae Mathematicae*, 144(1):87–98, 2016.

[5] J.-D. Benamou, F. Collino, and J.-M. Mirebeau. Monotone and consistent discretization of the Monge-Ampère operator. *Math. Comp.*, 85(302):2743–2775, 2016.

[6] J.-D. Benamou and V. Duval. Minimal convex extensions and finite difference discretisation of the quadratic Monge–Kantorovich problem. *European J. Appl. Math.*, 30(6):1041–1078, 2019.

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

[8] K. Brix, Y. Hafizogullari, and A. Platen. Designing illumination lenses and mirrors by the numerical solution of Monge–Ampère equations. *JOSA A*, 32(11):2227–2236, 2015.

[9] B. D. Froese. A numerical method for the elliptic Monge-Ampère equation with transport boundary conditions. *SIAM J. Sci. Comput.*, 34(3):A1432–A1459, 2012.

[10] T. Gallaway, R. N. Olsen, and D. M. Mitchell. The economics of global light pollution. *Ecological Economics*, 69(3):658–665, 2010.

[11] N. Guillen. A primer on generated Jacobian equations: geometry, optics, economics. *Notices Amer. Math. Soc.*, 66(9):1401–1411, 2019.

[12] B. F. Hamfeldt. Convergence Framework for the Second Boundary Value Problem for the Monge–Ampère Equation. *SIAM J. Numer. Anal.*, 57(2):945–971, 2019.

[13] F. Jiang and X. Yang. Weak solutions of Monge-Ampère type equations in optimal transportation. *Acta Math. Sci. Ser. B (Engl. Ed.)*, 33(4):950–962, 2013.

[14] J. Kitagawa, Q. Mérigot, and B. Thibert. Convergence of a Newton algorithm for semi-discrete optimal transport. *J. Eur. Math. Soc. (JEMS)*, 21(9):2603–2651, 2019.

[15] M. Lindsey and Y. A. Rubinstein. Optimal transport via a Monge-Ampère optimization problem. *SIAM J. Math. Anal.*, 49(4):3073–3124, 2017.

[16] J. Meyron, Q. Mérigot, and B. Thibert. Light in power: a general and parameter-free algorithm for caustic design. In *SIGGRAPH Asia 2018 Technical Papers*, page 224. ACM, 2018.

[17] M. Neilan, A. J. Salgado, and W. Zhang. The Monge-Ampère equation. *arXiv preprint arXiv:1901.05108*, 2019.

[18] V. I. Oliker and L. D. Prussner. On the numerical solution of the equation $(\partial^2 z/\partial x^2)(\partial^2 z/\partial y^2) - ((\partial^2 z/\partial x \partial y))^2 = f$ and its discretizations. I. *Numer. Math.*, 54(3):271–293, 1988.

[19] C. Prins. *Inverse methods for illumination optics*. Ph.D. Dissertation, Technische Univeristeit Eindhoven. 2014.

[20] N. S. Trudinger. On the local theory of prescribed Jacobian equations. *Discrete Contin. Dyn. Syst.*, 34(4):1663–1681, 2014.

DEPARTMENT OF MATHEMATICS, STATISTICS, AND COMPUTER SCIENCE, M/C 249. UNIVERSITY OF ILLINOIS AT CHICAGO, CHICAGO, IL 60607-7045, USA

Email address: `awanou@uic.edu`

URL: `http://www.math.uic.edu/~awanou`