METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 28, No. 1, pp. 031–052, March 2021

003

A CLASS OF NONLINEAR OPTIMISATION AND APPLICATIONS*

JIAKUN LIU†

Dedicated to Professor John Urbas on the occasion of his 60th birthday

Abstract. In this paper, we introduce a class of nonlinear optimisation problems. Under mild assumptions, we obtain the existence of potential functions and show that the potential function is a generalised solution of a Monge-Ampère type equation.

Key words. Nonlinear optimisation, potential function, Monge-Ampère equation.

Mathematics Subject Classification. 35J60, 35B45; 49Q20, 28C99.

1. Introduction. In this paper, we introduce a class of nonlinear optimisation problems, which extends Kantorovich's linear optimisation in the optimal transport problem. Motivation and examples arise naturally from geometric optical problems such as reflector and refractor problems as summarised in §5. Let $U, V \subset \mathbb{R}^n$ be two bounded domains, and $\phi = \phi(x, y, t, s) : U \times V \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a given constraint function, which is assumed to be C^1 smooth and strictly increasing in t and s.

DEFINITION 1.1. *Let $(u, v) \in C(U) \times C(V)$ be a pair of continuous functions. We say (u, v) is a* dual pair *with respect to ϕ if it satisfies*

$$\begin{cases} u(x) = \sup\{t : \phi(x, y, t, v(y)) \le 0, \quad \forall\, y \in V\}, \\ v(y) = \sup\{s : \phi(x, y, u(x), s) \le 0, \quad \forall\, x \in U\}. \end{cases} \tag{1.1}$$

Since the function $\phi \in C^1$ and is strictly increasing in t, s, from the implicit function theorem the constraint $\phi \le 0$ can be written as

$$\phi(x, y, t, s) = t + \varphi(x, y, s) \le 0, \tag{1.2}$$

for some C^1 function $\varphi = \varphi(x, y, s)$ strictly increasing in s. We assume further that there exists a constant $\theta_0 > 0$ such that

$$\varphi_s \ge \theta_0 \quad \text{in} \;\; U \times V \times \mathbb{R}, \tag{1.3}$$

and φ satisfies the condition

(H_1) For each $x_0 \in U$, $\forall\, (p, t) \in \mathbb{R}^n \times \mathbb{R}$, there is at most one pair $(y, s) \in \mathbb{R}^n \times \mathbb{R}$ such that $(\varphi_x, \varphi)\,(x_0, y, s) = -(p, t)$; and for each $y_0 \in V$, $\forall\, (q, s) \in \mathbb{R}^n \times \mathbb{R}$, there is at most one pair $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ such that $(\varphi_y/\varphi_s, \varphi)\,(x, y_0, s) = -(q, t)$.

Denote the restriction set by K, which consists of all pairs (u, v) satisfying (1.2), namely

$$K := \{(u, v) \in C(U) \times C(V) : \phi(x, y, u(x), v(y)) \le 0 \text{ in } U \times V\}. \tag{1.4}$$

*Received April 3, 2020; accepted for publication July 15, 2020.
†School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW 2522, Australia (jiakunl@uow.edu.au). This work is supported by the Australian Research Council, DP170100929 and DP200101084.

Under above assumptions, we have the following result, whose proof is postponed in §2.

LEMMA 1.1. *For each dual pair* $(u, v) \in K$, *there exists an associated mapping* $T_{(u,v)} : U \to V$ *that solves the equation*

$$u(x) + \varphi(x, T_{(u,v)}x, v(T_{(u,v)}x)) = 0 \tag{1.5}$$

and is uniquely determined almost everywhere on U. *Meanwhile, there exists a mapping* $T^{-1}_{(u,v)} : V \to U$ *solving*

$$u(T^{-1}_{(u,v)}y) + \varphi(T^{-1}_{(u,v)}y, y, v(y)) = 0 \tag{1.6}$$

and is uniquely determined almost everywhere on V. *Moreover,* $T^{-1}_{(u,v)} \circ T_{(u,v)}(x) = x$, *a.e. on* U *and* $T_{(u,v)} \circ T^{-1}_{(u,v)}(y) = y$, *a.e. on* V.

Let f, g be two probability densities supported on U, V, respectively,

$$\int_U f(x)dx = \int_V g(y)dy = 1. \tag{1.7}$$

We also assume there exists a constant $c_0 > 0$ such that

$$c_0 \le f \le c_0^{-1} \text{ in } U, \quad c_0 \le g \le c_0^{-1} \text{ in } V. \tag{1.8}$$

DEFINITION 1.2. *A mapping* $S : U \to V$ *is called* measure preserving, *denoted by* $S_\# f = g$, *if for any* $h \in C(V)$,

$$\int_U h(Sx)f(x)\, dx = \int_V h(y)g(y)\, dy. \tag{1.9}$$

Similarly, a mapping $S^* : V \to U$ *is measure preserving, denoted by* $S^*_\# g = f$, *if for any* $\eta \in C(U)$,

$$\int_U \eta(x)f(x)\, dx = \int_V \eta(S^*y)g(y)\, dy. \tag{1.10}$$

Denote the restriction set $\mathcal{T}$ by

$$\mathcal{T} := \left\{ S : U \to V : S_\# f = g, \text{ and } S^{-1}_\# g = f \right\}. \tag{1.11}$$

Note that this condition is stronger than that in optimal transportation, since a measure reserving mapping may not be invertible in general. However, as remarked in [21], in practice the mappings like the light reflection and refraction processes are actually invertible.

In this paper, we study the nonlinear optimisation problem that maximises the functional

$$I(u, v, T) := \int_U u(x)f(x)\, dx + \int_V \varphi(T^{-1}y, y, v(y))g(y)\, dy, \tag{1.12}$$

among all pairs $(u, v) \in K$ and the mapping $T \in \mathcal{T}$, defined in (1.4) and (1.11), respectively. More generally, one could consider U, V as subsets of a Riemannian manifold $\mathcal{M}$ with general measures μ, ν, but for simplicity, here we only consider the Euclidean case with absolutely continuous measures $\mu = f\,dx,\ \nu = g\,dy$.

Note that for any $T \in \mathcal{T}$ and $(u,v) \in K$, one has

$$\int_V \varphi(T^{-1}y, y, v(y))g(y)\,dy = \int_U \varphi(x, Tx, v(Tx))f(x)\,dx,$$

and thus

$$\begin{aligned} I(u,v,T) &= \int_U u(x)f(x)\,dx + \int_U \varphi(x, Tx, v(Tx))f(x)\,dx \\ &= \int_U (u(x) + \varphi(x, Tx, v(Tx)))\,f(x)\,dx \le 0, \end{aligned}$$

where the last inequality is due to $(u,v) \in K$. Hence, $\sup_{K,\mathcal{T}} I(u,v,T) \le 0$ is bounded.

Our main result is the following solvability of the nonlinear optimisation problem (1.12).

THEOREM 1.1. *Under the hypotheses* (1.3), (1.7) *and* (H_1), *there exists a maximising dual pair* $(u,v) \in K$ *and* $T \in \mathcal{T}$ *such that* $0 = I(u,v,T) = \sup_{K,\mathcal{T}} I$. *Moreover, the mapping* $T = T_{(u,v)}$ *given in Lemma 1.1 associated to the dual pair* (u,v) *is uniquely determined almost everywhere on* U.

The functions u, v in a dual maximising pair are called *potential functions* of the nonlinear optimisation (1.12). The associated mapping T is called an *optimal mapping*. These terminologies are adopted from optimal transportation [22, 27, 28]. However, it is worth pointing out that in the nonlinear case there is generally no uniqueness for maximising pair (u,v), see Remark 2.1.

It is well known that many constrained nonlinear optimisation problems can be solved by Lagrangian dual methods (see, for example, [11]), where the convexity plays a crucial role. For the nonlinear optimisation (1.12) under conditions (1.17)–(1.18), we can show that there is no duality gap and there exists at least one Lagrange multiplier. This enables us to use the Lagrangian duality theory in §4 to study the maximisation of the functional I.

In order to state the result, let us introduce some terminology in Lagrangian duality. More details are contained in §4. Denote $X := C(U) \times C(V)$. The nonlinear optimisation (1.12) is a special case of the following *primal problem*:

$$\begin{aligned} &\text{maximise } I(u,v) = \int_{U\times V} F(x,y,u(x),v(y))\,d\gamma, \\ &\text{subject to } (u,v) \in X, \quad \psi(u,v) := \inf_{x\in U, y\in V} -\phi(x,y,u(x),v(y)) \ge 0, \end{aligned} \tag{1.13}$$

where $F = F(x,y,t,s)$ is a function in $\mathbb{R}^{2n+2}$ monotone increasing in t, s, and $d\gamma$ is a measure on $U \times V$ with dx, dy as marginals. We always assume that ϕ has the form (1.2). An element $(u,v) \in X$ is called *feasible* if $\psi(u,v) \ge 0$.

Define the *Lagrangian function* $L : X \times \mathbb{R} \to \mathbb{R}$ to be

$$L(u,v,\mu) = I(u,v) + \mu\psi(u,v), \tag{1.14}$$

where $\mu \in \mathbb{R}$. The *dual functional* J is defined by

$$J(\mu) = \sup_{(u,v)\in X} L(u,v,\mu), \tag{1.15}$$

and the *dual problem* is given by

$$\begin{aligned}&\text{minimise } J(\mu)\\&\text{subject to } \mu \geq 0.\end{aligned} \tag{1.16}$$

Regardless of the functional I and the constraint ϕ of the primal problem, the dual problem has a very nice convexity property, as shown in Lemma 4.2. In the language of nonlinear programming [3, 11], when $\inf_{\mu\geq 0} J(\mu) = \sup_{(u,v)\in X} I(u,v)$, we say that *there is no duality gap*, otherwise, *there is duality gap*.

THEOREM 1.2. *Assume that the function F in* (1.13) *is concave in* (t,s), *namely for any* $(x,y,t,s) \in U\times V\times\mathbb{R}^2$,

$$Hess_{t,s}F := \begin{pmatrix} \partial_{tt}F, & \partial_{ts}F \\ \partial_{st}F, & \partial_{ss}F \end{pmatrix} \leq 0, \tag{1.17}$$

and the constraint φ in (1.2) *is convex in s, namely for any* $(x,y,s)\in U\times V\times\mathbb{R}$,

$$\partial_{ss}\varphi \geq 0. \tag{1.18}$$

Suppose that there exists a pair $(\bar{u},\bar{v}) \in X$ *such that*

$$\psi(\bar{u},\bar{v}) > 0. \tag{1.19}$$

Then there is no duality gap between the primal problem (1.13) *and dual problem* (1.16), *and there exists at least one Lagrange multiplier (see Definition 4.1).*

Note that in the special case (1.12), the convexity assumptions (1.17) and (1.18) imply that $\varphi = \varphi(x,y,v)$ is linear in v, namely

$$\varphi(x,y,v) = c_0(x,y) + c_1(x,y)v,$$

for some functions c_0, c_1, where $c_1 \geq \theta_0 > 0$ in $U\times V$ by the monotonicity (1.3). When $c_1 \equiv 1$, it is optimal transportation with the associated cost function $-c_0$, see Example 5.1.

This paper is organised as follows: In §2 we prove Lemma 1.1 and Theorem 1.1. In §3 we show that the potential function u is a generalised solution of a Monge-Ampère type equation. In §4 we establish a Lagrangian duality theory, and prove Theorem 1.2. In §5 we summarise some examples and applications of the nonlinear optimisation (1.12).

2. Existence of maximisers. In this section, we obtain the solvability of the nonlinear optimisation problem (1.12). Some similar results have been obtained in [21], but for the completeness and readers' convenience we shall outline the main steps of proofs, particularly for general cases. First, we give a weak compactness of the restriction set $\mathcal{T}$.

LEMMA 2.1. *For any sequence* $\{T_k\}_{k=1}^\infty \subset \mathcal{T}$, *there exists a subsequence* $\{T_{k_j}\}_{j=1}^\infty \subset \{T_k\}_{k=1}^\infty$ *and* $T\in\mathcal{T}$ *such that* $T_{k_j} \rightharpoonup T$, *that is for any* $h \in C(V)$,

$$\int_U h(T_{k_j}x)f(x)\,dx \to \int_U h(Tx)f(x)\,dx \quad \text{as } j\to\infty. \tag{2.1}$$

Proof. Observe that for any $S\in\mathcal{T}$, S is bounded since $S(U) = V$ and V is a bounded domain. Hence, for any sequence $\{T_k\}_{k=1}^\infty \subset \mathcal{T}$, there exists a subsequence

$\{T_{k_j}\}_{j=1}^{\infty} \subset \{T_k\}_{k=1}^{\infty}$ and a limit mapping $T : U \to V$ such that (2.1) holds. For simplicity, we shall denote $\{T_{k_j}\}_{j=1}^{\infty}$ by $\{T_j\}_{j=1}^{\infty}$. It remains to show that T is a bijection from U to V a.e., and moreover, T, T^{-1} are measure preserving, so that $T \in \mathcal{T}$.

To see T is one-to-one a.e. on U, suppose to the contrary that there is a subset $E \subset U$ with $|E| \neq 0$ such that $T(E) = \{y_0\}$ is a single point $y_0 \in V$. Let $h_j \in C(V)$ be a sequence of functions such that

$$\begin{aligned} h_j(y) &= 1 \text{ for } y \in B_{\epsilon_j/2}(y_0) \cap V, \\ h_j(y) &= 0 \text{ for } y \in B^c_{\epsilon}(y_0), \end{aligned}$$

where the sequence $\epsilon_j \to 0$ as $j \to \infty$. Since T_j is measure preserving, one has

$$\int_U h_j(T_j(x)) f(x)\, dx = \int_V h_j(y) g(y)\, dy.$$

From (2.1) and (1.8), the left hand side converges to

$$\begin{aligned} \lim_{j\to\infty} \int_U h_j(T_j(x)) f(x)\, dx &= \int_{T^{-1}(y_0)} f(x)\, dx \\ &\geq \int_E f(x)\, dx \\ &\geq c_0|E| > 0. \end{aligned}$$

While by the construction of $\{h_j\}$, the right hand side converges to

$$\lim_{j\to\infty} \int_V h_j(y) g(y)\, dy = 0.$$

This contradiction implies that T is one-to-one. Since for each j, $T_j \in \mathcal{T}$, T_j^{-1} is also measure preserving. One can use a similar argument as above to show T is surjective as well. Therefore, the mapping T is a bijection from U to V.

To show T is measure preserving, for any $h \in C(V)$ one can see

$$\int_U h(T_j(x)) f(x)\, dx = \int_V h(y) g(y)\, dy$$

for each j, since T_j is measure preserving. Taking the limit by letting $j \to \infty$, from (2.1) we then obtain

$$\int_U h(T(x)) f(x)\, dx = \int_V h(y) g(y)\, dy,$$

which implies that T is measure preserving. One can similarly show that T^{-1} is measure preserving as well. Therefore, the mapping $T \in \mathcal{T}$ and the proof is done. □

We are now ready to prove the existence result in Theorem 1.1. First, let us simplify the notation a bit. As mentioned before, the functional I in (1.12) is a special case of (1.13) if the function F is chosen by

$$F_T(x, y, u(x), v(y)) = \frac{1}{|V|} u(x) f(x) + \frac{1}{|U|} \varphi(T^{-1}y, y, v(y)) g(y). \tag{2.2}$$

In the following content, we always assume F is given by (2.2). The proof of Theorem 1.1 is divided into two lemmas: In Lemma 2.2 we show there exists a maximiser $(u,v)\in K$ and $T\in\mathcal{T}$ such that $I(u,v,T)=\sup_{K,\mathcal{T}} I$, then in Lemma 2.3 we show that $0=I(u,v,T)$ and $T=T_{(u,v)}$.

LEMMA 2.2. *Under the hypotheses of Theorem 1.1, there exists a dual maximiser $(u,v)\in K$ and $T\in\mathcal{T}$ such that $I(u,v,T)=\sup_{K,\mathcal{T}} I$.*

Proof. Given any pair $(u,v)\in K$, by the monotonicity and continuity one can show [21] that $I(u,v,T)$ does not decrease if (u,v) is replaced by

$$\begin{aligned}
v^*(y) &= \sup\{s\,:\,\phi(x,y,u(x),s)\le 0,\quad \forall\, x\in U\},\\
u^*(x) &= \sup\{t\,:\,\phi(x,y,t,v^*(y))\le 0,\quad \forall\, y\in V\},
\end{aligned}$$

namely $(u^*,v^*)\in K$ and $I(u^*,v^*,T)\ge I(u,v,T)$.

Define $K_{C_0}=K\cap\{u\ge C_0\}$, where C_0 is a constant, which may be chosen negative and sufficiently small in the following context. We show that u^* and v^* are uniformly bounded if $(u,v)\in K_{C_0}$. Since $v^*\ge v, u\ge C_0$, by (1.2)–(1.3) we have for each $y\in V$, $s:=v^*(y)$,

$$C_0+\varphi(x,y,s)\le u(x)+\varphi(x,y,s)\le 0,\quad \text{for all } x\in U.$$

Then by (1.3) again, there exists a constant C_1, such that $s\le C_1$. This implies that

$$v\le v^*\le C_1, \tag{2.3}$$

we may choose C_1 such that $\sup_V v^*=C_1$. By a similar argument, there is another constant $\tilde{C}_0$ depending on φ and C_1 such that $\inf_U u^*=\tilde{C}_0$. The constant $\tilde{C}_0\ge C_0$, since $u^*\ge u$ in U, and so $(u^*,v^*)\in K_{C_0}$.

We next deduce the lower bound of v^* and the upper bound of u^* by showing that u^* and v^* are locally Lipschitz. Consider two points in U, $x_1\ne x_2$ and $|x_1-x_2|<\varepsilon$ sufficiently small. There are two points $y_1,y_2\in\overline{V}$ such that

$$\begin{aligned}
&\phi(x_1,y_1,u^*(x_1),v^*(y_1))=0,\\
&\phi(x_2,y_2,u^*(x_2),v^*(y_2))=0.
\end{aligned}$$

Then by (1.2), we have

$$\begin{aligned}
0 =&\phi(x_2,y_2,u^*(x_2),v^*(y_2))-\phi(x_1,y_2,u^*(x_1),v^*(y_2))\\
&+\phi(x_1,y_2,u^*(x_1),v^*(y_2))-\phi(x_1,y_1,u^*(x_1),v^*(y_1))\\
=&u^*(x_2)-u^*(x_1)-\varphi_x(\hat{x},y_2,v^*(y_2))\cdot(x_2-x_1)\\
&+\phi(x_1,y_2,u^*(x_1),v^*(y_2)),
\end{aligned}$$

where $\hat{x}=\theta x_1+(1-\theta)x_2$ for some $\theta\in(0,1)$. Noting that $\phi(x_1,y_2,u^*(x_1),v^*(y_2))\le 0$, we have

$$u^*(x_2)-u^*(x_1)\ge -C_2|x_2-x_1|,$$

where the constant $C_2=\sup(|\varphi_x|+|\varphi_y|)$.

On the other hand, replacing $\phi(x_1,y_2,u^*(x_1),v^*(y_2))$ by $\phi(x_2,y_1,u^*(x_2),v^*(y_1))$ in the above calculation, we have

$$u^*(x_2)-u^*(x_1)\le C_2|x_2-x_1|.$$

Therefore, the Lipschitz constant of u^* on U is controlled by

$$\|u^*\|_{Lip(U)} \le C_2. \tag{2.4}$$

By switching x and y in the above argument, we can obtain the Lipschitz continuity of v^* on V,

$$\begin{aligned} |v^*(y_2) - v^*(y_1)| &\le \frac{\sup|\varphi_y|}{\inf \varphi_s}|y_2 - y_1| \\ &\le C_2\theta_0^{-1}|y_2 - y_1|, \end{aligned}$$

where θ_0 is the constant in (1.3), y_1, y_2 are two distinct points in V. This inequality implies that $\|v^*\|_{Lip(V)} \le C_2\theta_0^{-1}$. Hence, we have $u^* \le \tilde{C}_0 + C_2\text{diam}(U)$ and $v^* \ge C_1 - C_2\theta_0^{-1}\text{diam}(V)$.

We conclude, therefore, that any pair $(u, v) \in K_{C_0}$ may be replaced by a bounded, Lipschitz pair $(u^*, v^*) \in K_{C_0}$ without decreasing I. We now choose a maximising sequence $\{(u_k, v_k)\} \subset K_{C_0}$ and $\{T_k\} \subset \mathcal{T}$such that

$$I(u_k, v_k, T_k) \to \sup_{\substack{(u,v)\in K_{C_0} \\ T\in\mathcal{T}}} I(u, v, T).$$

By the above considerations we may assume that each (u_k, v_k) is a bounded, uniformly Lipschitz pair, uniformly with respect to k, so there is a subsequence converging uniformly to a bounded, Lipschitz, maximising pair $(\bar{u}, \bar{v}) \in K_{C_0}$. Thanks to Lemma 2.1, there is a subsequence of $\{T_k\}$ weakly converges to some $T \in \mathcal{T}$.

Next, letting $I_T(u, v) = I(u, v, T)$, we show that when $C_0 < 0$ sufficiently small,

$$\sup_{(u,v)\in K_{C_0}} I_T(u, v) = \sup_{(u,v)\in K} I_T(u, v), \tag{2.5}$$

or equivalently, $\sup_{K_{C_0}} I_T$ is independent of C_0. By definition, one has $\sup_{K_{C_0-1}} I_T \ge \sup_{K_{C_0}} I_T$. So it suffices to show the reverse inequality. Let $(u, v) \in K_{C_0-1}$ be a maximiser such that $I_T(u, v) = \sup_{K_{C_0-1}} I_T$, and $\{x_k\}_{k=1,\cdots,N}$ be a set of points in U. For a small constant $\varepsilon > 0$, define

$$\tilde{u} = \begin{cases} u & \text{in } U - \cup_N B_\varepsilon(x_k), \\ u + 2 & \text{in } \cup_N B_\varepsilon(x_k). \end{cases}$$

Note that we may replace $\tilde{u}$ by its mollification $\tilde{u}_h = \rho_h * \tilde{u}$, where ρ_h is the standard mollifier function [10]. For simplicity, we assume $\tilde{u}$ is continuous in the sense that for $h > 0$ sufficiently small,

$$I_T(\tilde{u}_h, v) = I_T(u, v) + O(N\varepsilon^n).$$

Define

$$\begin{aligned} \tilde{v}^*(y) &= \sup\{s : \phi(x, y, \tilde{u}(x), s) \le 0, \quad \forall\, x \in U\}, \\ \tilde{u}^*(x) &= \sup\{t : \phi(x, y, t, \tilde{v}^*(y)) \le 0, \quad \forall\, y \in V\}. \end{aligned}$$

Since the constraint function φ is C^1 smooth in s and by (1.2)–(1.3), except for a set $E \subset U$ and a set $E' \subset V$ of measure $|E| = |E'| = O(N\varepsilon^n)$,

$$\begin{aligned} \tilde{v}^* &= v - \frac{2}{\varphi_s} + O(\delta) \quad \text{in } V \setminus E', \\ \tilde{u}^* &= u + 2 + O(\delta) \quad \text{in } U \setminus E, \end{aligned}$$

where $\delta := \min_{i\neq j}\{\mathrm{dist}(x_i, x_j)\}$. Therefore, by (1.7) and the mean value theorem we have

$$\begin{aligned} I_T(\tilde{u}^*, \tilde{v}^*) &= I_T(u, v) + 2\int_{(U\setminus E)\times(V\setminus E')} \left\{F_t - \frac{F_s}{\varphi_s}\right\} d\gamma + O(\delta) + O(N\varepsilon^n) \\ &\geq I_T(u, v) - C\delta - CN\varepsilon^n. \end{aligned}$$

As $(u, v) \in K_{C_0-1}$, we may assume that $\inf_U u = C_0 - 1$. Otherwise, one has $\inf_U u = C_0 - \tau_0$ for some constant $\tau_0 < 1$. This implies that $\sup_{K_{C_0-1}} I_T = \sup_{K_{C_0}} I_T$, namely $\sup_{K_{C_0}} I_T$ is independent of C_0, and the proof is finished. By the definition, δ will become small if the number of points N is sufficiently large so that we have $(\tilde{u}^*, \tilde{v}^*) \in K_{C_0}$ and

$$\sup_{K_{C_0}} I_T \geq I_T(\tilde{u}^*, \tilde{v}^*) \geq \sup_{K_{C_0-1}} I_T - C\delta - CN\varepsilon^n.$$

Then, choosing $\varepsilon > 0$ sufficiently small, we have

$$\sup_{K_{C_0-1}} I_T \leq \sup_{K_{C_0}} I_T,$$

by letting $\delta \to 0, \varepsilon \to 0$, which implies that $\sup_{K_{C_0}} I_T$ is independent of C_0.

Note that the above argument is independent of $T \in \mathcal{T}$. Suppose that $T' \in \mathcal{T}$ satisfies

$$\sup_{K_{C_0-1}} I_{T'} = \sup_{K_{C_0-1},\mathcal{T}} I. \tag{2.6}$$

By the proof of (2.5) one has $\sup_{K_{C_0-1}} I_{T'} = \sup_{K_{C_0}} I_{T'} \leq \sup_{K_{C_0},\mathcal{T}} I = \sup_{K_{C_0}} I_T$. On the other hand, $\sup_{K_{C_0-1},\mathcal{T}} I \geq \sup_{K_{C_0},\mathcal{T}} I = \sup_{K_{C_0}} I_T$. Hence, $\sup_{K_{C_0}} I_{T'} = \sup_{K_{C_0}} I_T$. Therefore, the proof is finished. □

Remark 2.1. Note that there exist infinitely many maximising pairs. In fact, if (u, v) is a maximiser and $C_0 = \inf_U u$, then there is another maximiser $(u', v') \in K_{C_0+1}$. In the linear case, one has $(u', v') = (u+1, v-1)$. However, it is not yet clear how these maximising pairs linking with each other in the nonlinear case.

Remark 2.2. As remarked in [21], instead of assuming the condition (H_1) over the whole space $\mathbb{R}^n \times \mathbb{R}$, we can actually restrict it in some proper subsets in practice.

Now, let's prove Lemma 1.1 that defines the associated mappings $T_{(u,v)}$ and $T^{-1}_{(u,v)}$. Here we outline some main steps, see [21] for more details.

Proof of Lemma 1.1. For each $x \in U$, by (1.1) and the continuity there exists a point $y =: T_{(u,v)}(x) \in \overline{V}$ such that

$$\begin{aligned} u(x) + \varphi(x, y, v(y)) &= 0, \\ u(x') + \varphi(x', y, v(y)) &\leq 0, \quad \text{for any other } x' \in U. \end{aligned} \tag{2.7}$$

Since φ is C^1, one has u is locally Lipschitz and differentiable a.e. on U. Let $x \in U$ be a differentiable point of u, then

$$\varphi_x(x, y, v) + Du(x) = 0. \tag{2.8}$$

Therefore, from the condition (H_1), we obtain the mapping $y = T_{(u,v)}(x)$ that solves the equation (1.5). Since u is differentiable a.e. on U, the mapping $T_{(u,v)}$ is uniquely determined a.e. on U. Similarly, we can obtain the mapping $T^{-1}_{(u,v)}$ that solves the equation (1.6) and is uniquely determined a.e. on V.

It remains to show that $T_{(u,v)}$ and $T^{-1}_{(u,v)}$ is essentially inverse to each other. Setting $x = T^{-1}_{(u,v)}y$ in (1.6), one has

$$u(x) + \varphi(x, y, v(y)) = 0.$$

On the other hand, the above equation is uniquely solved by $y = T_{(u,v)}x$ a.e.. Hence, $T_{(u,v)} \circ T^{-1}_{(u,v)}(y) = y$ for a.e. $y \in V$. Similarly, one can see that $T^{-1}_{(u,v)} \circ T_{(u,v)}(x) = x$ for a.e. $x \in U$. □

The proof of Theorem 1.1 will be accomplished by the next lemma.

Lemma 2.3. *Let $(u, v) \in K$ and $T \in \mathcal{T}$ be the maximiser obtained in Lemma 2.2 such that $I(u, v, T) = \sup_{K,\mathcal{T}} I$. Then, one has $0 = I(u, v, T)$ and $T = T_{(u,v)}$ is the associated mapping to the dual pair (u, v), given in Lemma 1.1.*

Proof of Lemma 2.3. Let $h \in C(V)$ and ϵ is a small constant. Define

$$\begin{aligned} v_\epsilon(y) &= v(y) + \epsilon h(y), \\ u_\epsilon(x) &= \sup\{t \,:\, t + \varphi(x, y, v_\epsilon(y)) \le 0, \quad \forall y \in V\}. \end{aligned} \tag{2.9}$$

Then $(u_\epsilon, v_\epsilon) \in K$ and $(u_0, v_0) = (u, v)$. From Lemma 1.1 for $x \in U$, let $y = T_{(u,v)}x$. It was proved in [21] that at these points we have

$$u_\epsilon(x) - u(x) = -\epsilon\varphi_s(x, T_{(u,v)}x, v(T_{(u,v)}x))h(T_{(u,v)}x) + o(\epsilon). \tag{2.10}$$

Define $I(\epsilon) := I(u_\epsilon, v_\epsilon, T)$, where $T \in \mathcal{T}$ is the maximiser obtained in Lemma 2.2. Since the function $I(\epsilon)$ achieves maximum at $\epsilon = 0$, we obtain

$$0 = \lim_{\epsilon\to 0} \frac{I(u_\epsilon, v_\epsilon, T) - I(u, v, T)}{\epsilon},$$

and by (2.2),

$$\int_U f(x)\varphi_s(x, T_{(u,v)}x, v(T_{(u,v)}x))h(T_{(u,v)}x)\, dx = \int_V g(y)\varphi_s(T^{-1}y, y, v(y))h(y)\, dy.$$

Since $T \in \mathcal{T}$, T^{-1} is measure preserving, one has

$$\int_V g(y)\varphi_s(T^{-1}y, y, v(y))h(y)\, dy = \int_U f(x)\varphi_s(x, Tx, v(Tx))h(Tx)\, dx.$$

Hence, for any $h \in C(V)$, we have

$$\int_U f(x)\varphi_s(x, T_{(u,v)}x, v(T_{(u,v)}x))h(T_{(u,v)}x)\, dx = \int_U f(x)\varphi_s(x, Tx, v(Tx))h(Tx)\, dx,$$

which implies that $T = T_{(u,v)}$. Therefore, from (1.5)

$$\begin{aligned} I(u, v, T) &= \int_U u(x)f(x)\, dx + \int_U \varphi(x, Tx, v(Tx))f(x)\, dx \\ &= \int_U \big(u(x) + \varphi(x, T_{(u,v)}x, v(T_{(u,v)}x))\big)\, f(x)\, dx = 0. \end{aligned}$$

The proof is finished. □

COROLLARY 2.1. *Let $(u,v)\in K$ be a maximiser pair of I, the associated mapping $T_{(u,v)}\in\mathcal{T}$ is measure preserving.*

In the following context, we shall simply write $T_{(u,v)}$ as T with no confusion arises. As a consequence, we have the following

COROLLARY 2.2. *Assume the function F is given by* (2.2) *and the condition* (1.7) *holds. If the optimal mapping T is continuous differentiable, then*

$$|det\, DT| = \frac{f}{g\circ T}. \tag{2.11}$$

Proof. From Corollary 2.1, one has T is measure preserving in the sense of (1.9). When T is C^1 smooth, by the formula of change of coordinates,

$$\int_U f(x)h(Tx)dx = \int_U g(Tx)h(Tx)|\det DT|dx,$$

for any $h\in C(V)$. Hence the Jacobian of DT satisfies (2.11). □

As a consequence of Corollary 2.1 and Corollary 2.2, we derive the equation satisfied by the potential function u as follows. At this stage, let us assume all the functions are smooth enough, say at least C^2, so that we can do the differentiations.

Let $(u,v)\in K$ be a dual maximising pair of I, and T be the associated optimal mapping. By (2.8) we have

$$\varphi_x(x,Tx,v(Tx)) + Du(x) = 0,$$

in U. Differentiating with respect to x, we then get

$$0 = \varphi_{xx} + \varphi_{xy}DT + (\varphi_{xs}\otimes Dv)DT + D^2u, \tag{2.12}$$

where each side is regarded as an $n\times n$ matrix valued at (x,y), $y=T(x)$.

In order to eliminate Dv in (2.12), we note that for $x_0\in U$ equality (1.5) holds at $y_0=T(x_0)$, and for other $y'\in V$

$$u(x_0) + \varphi(x_0,y',v(y'))\le 0,$$

since $(u,v)\in K$. Thus, at (x_0,y_0) there holds

$$\frac{d\varphi}{dy} = \varphi_y + \varphi_s Dv = 0.$$

By the assumption (1.3), we thus get

$$Dv = -\frac{\varphi_y}{\varphi_s}. \tag{2.13}$$

Combining (2.12) and (2.13) we have the equation

$$\left|D^2u + \varphi_{xx}\right| = \left|\varphi_{xy} - \frac{1}{\varphi_s}\varphi_{xs}\otimes\varphi_y\right| |DT|\,. \tag{2.14}$$

In the case of (2.2), by Corollary 2.2 we obtain the equation

$$\left|\det\left[D^2u+\varphi_{xx}\right]\right| = \left|\det\left[\varphi_{xy}-\frac{1}{\varphi_s}\varphi_{xs}\otimes\varphi_y\right]\right|\frac{f}{g\circ T}, \tag{2.15}$$

which is a Monge-Ampère type equation [8, 10, 13]. Correspondingly, we have the natural boundary condition

$$T(U)=V. \tag{2.16}$$

Similarly, one can also derive the dual PDE for the dual potential v.

When $\varphi(x,y,v)=x\cdot y$, then (2.15) is equivalent to the standard Monge-Ampère equation

$$\det D^2u=h,$$

with the boundary condition

$$Du(U)=V,$$

where $h=f/g$. When $\varphi(x,y,v)=v-c(x,y)$ for some function $c:U\times V\to\mathbb{R}$, we have the optimal transportation equation, see Example 5.1,

$$\det\left[D^2u-D^2_xc\right]=|\det D^2_{xy}c|\frac{f}{g\circ T},$$

with the boundary condition (2.16).

It is well-known that the regularity of equation (2.15) depends crucially on the structure of constraint function φ, as in [22, 29]. In the next section, we shall introduce a notion of generalised solutions and show that if (u,v) is a dual maximising pair of I over K, then the potential u is a generalised solution of (2.15). The regularity property of u is related to that of generated Jacobian equations, which has been systematically studied by Trudinger in [24, 25]. Under some structural conditions (which are analogous to the MTW condition in optimal transportation [22]), in [24, 25, 12, 17] the authors started to develop a theory parallel to that in optimal transportation.

3. generalised solution. In this section, we introduce a notion of generalised solutions of (2.15) and show that the potential function is a generalised solution. Let φ be the constraint in (1.2). First we introduce the φ-concavity for functions, which is an extension of the c-concavity in optimal transportation, see [9, 22].

Definition 3.1. *A φ-support function of u at x_0 is a function of the form $\varphi(x,y_0,s_0)$, where $y_0\in\mathbb{R}^n$, and $s_0\in\mathbb{R}$ is a constant such that*

$$\begin{aligned} u(x_0)+\varphi(x_0,y_0,s_0)&=0,\\ u(x)+\varphi(x,y_0,s_0)&\le 0 \quad \forall x\in U. \end{aligned} \tag{3.1}$$

A continuous function u defined on $\overline{U}$ is φ-concave if for any point $x_0\in U$, there exists a φ-support function at x_0.

By definition, the potential function u is φ-concave with $y_0\in V, s_0=v(y_0)$. In the special case when $\varphi(x,y,s)=s-x\cdot y$, the notion of φ-concavity coincides with that of concavity, and the graph of a φ-support function is a support hyperplane.

Recall that φ is derived from the constraint function $\phi(x, y, u, v)$ by the strict monotonicity in u. Since ϕ is also strictly increasing in v, the constraint (1.2) can also be written as

$$\phi(x, y, u, v) = v + \varphi^*(x, y, u) \leq 0, \tag{3.2}$$

for a function $\varphi^* = \varphi^*(x, y, t)$ strictly increasing in t. The function $\varphi^* = \varphi^*(x, y, t)$ is called *dual constraint function* of φ in the sense of

$$-\varphi(x, y, -\varphi^*(x, y, t)) = t,$$
$$-\varphi^*(x, y, -\varphi(x, y, s)) = s,$$

for all $(x, y) \in U \times V$. By differentiating, we have

$$\varphi^*_t = \frac{1}{\varphi_s}, \quad \varphi^*_x = \frac{\varphi_x}{\varphi_s}, \quad \varphi^*_y = \frac{\varphi_y}{\varphi_s}. \tag{3.3}$$

For the dual constraint φ^*, from (3.3) and the condition (H_1) we have:

(H_1^*) For each $y_0 \in V$, for any $(q, s) \in \mathbb{R}^n \times \mathbb{R}$, there is at most one pair $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\left(\varphi^*_y, \varphi^*\right)(x, y_0, t) = -(q, s),$$

namely, $\varphi^*_y(x, y_0, t) = -q$ and $\varphi^*(x, y_0, t) = -s$.

The φ-concavity in Definition 3.1 and (3.1)–(3.3) are generalisations of c-concavity and c-duality in optimal transportation, where

$$\varphi(x, y, s) = s - c(x, y), \quad \varphi^*(x, y, t) = t - c(x, y),$$

for a cost function $c(x, y)$. The condition (H_1)–(H_1^*) is the counterpart of the condition (A1) assumed on the cost function $c(x, y)$ in [22]. Note that from (3.2) and (3.3), we can directly derive (2.13) for a dual pair of functions u, v.

Similarly, by switching x and y, U and V, one can also introduce the notion of φ^*-concavity for the function v. From Definition 1.1 and (3.1), when $(u, v) \in K$ is a dual pair, u is naturally φ-concave and v is φ^*-concave.

Let u be a φ-concave function in U. We define a set-valued mapping $T_u = T_{u,\varphi} : U \to V$. For any $x_0 \in U$, let $T_u(x_0)$ denote the set of points y_0 such that $\varphi(x, y_0, s_0)$ is a φ-support function of u at x_0 for some constant s_0. For any subset $E \subset U$, we denote $T_u(E) = \cup_{x \in E} T_u(x)$.

If u is C^1 smooth, by condition (H_1) (y_0, s_0) is uniquely determined by $(Du(x_0), u(x_0))$, and T_u is single valued. In this paper we call the mapping T_u the *φ-normal mapping* of u. Similarly we can define the φ^*-normal mapping for φ^*-concave functions. In particular, if $(u, v) \in K$ is a dual pair, we see that $y \in T_{u,\varphi}(x)$ if and only if $x \in T_{v,\varphi^*}(y)$.

Remark 3.1. As the constraint function φ is smooth, any φ-concave function u is semi-concave, namely there exists a constant C such that $u(x) - C|x|^2$ is concave. It follows that u is twice differentiable almost everywhere and $T_u(x)$ is a singleton for almost all $x \in U$.

Lemma 3.1. *Let $(u, v) \in K$ be a dual maximising pair of I. Assume that the constraint φ^* satisfies condition (H_1^*). Let*

$$Y = Y_u = \{y \in V \mid \exists\, x_1 \neq x_2 \in U \text{ such that } y \in T_u(x_1) \cap T_u(x_2)\}.$$

Then Y has Lebesgue measure zero.

Proof. If $y \in T_u(x_1) \cap T_u(x_2)$, we have $x_1, x_2 \in T_{v,\varphi^*}(y)$. From the proof of Theorem 1.1, v is almost everywhere differentiable. Assume y is a differentiable point, then by definition

$$\varphi^*(x_i, y, u(x_i)) = -v(y),$$
$$\varphi^*_y(x_i, y, u(x_i)) = -Dv(y),$$

for $i = 1, 2$. If $x_1 \neq x_2$, this is a contradiction to (H_1^*). □

We now define a measure $\mu = \mu_{u,g}$ in U, where $g \in L^1(V)$ is the positive measurable function in (1.7). Set $g \equiv 0$ in $\mathbb{R}^n - V$. For any Borel set $E \subset U$, define

$$\mu(E) = \int_{T_u(E)} g(y)dy. \tag{3.4}$$

It follows from Lemma 3.1 that μ is a Radon measure, and satisfies the following regularity properties:

$$\mu(E) = \inf\{\mu(D) : E \subset D \subset U,\ D \text{ open}\}$$

for all Borel sets $E \subset U$, and

$$\mu(D) = \sup\{\mu(K) : K \subset D,\ K \text{ compact}\}$$

for all open sets $D \subset U$. For further discussion of the measure μ and its stability property, see [2, 6, 8, 13, 22].

Definition 3.2. *A φ-concave function u is called a generalised solution of* (2.15) *if $\mu_{u,g} = f dx$ in the sense of measure, that is for any Borel set $E \subset U$,*

$$\int_E f = \int_{T_u(E)} g. \tag{3.5}$$

Note that since we extended $g = 0$ to $\mathbb{R} - V$, the boundary condition (2.16) is a consequence of the mass balance condition (1.7) in the sense of $|V - T_u(U)| = 0$.

The next result shows that a potential function u is a generalised solution. The existence of potentials in Theorem 1.1 then implies the existence of generalised solutions. But in general we do not have the uniqueness, see Remark 2.1.

Theorem 3.1. *Let $(u, v) \in K$ be a dual maximising pair of I. Then u is a generalised solution of* (2.15).

Proof. Let $(u, v) \in K$ be a dual maximising pair of I. Then by (1.1), u is φ-concave and v is φ^*-concave with respect to each other. By Lemma 1.1, the optimal mapping T associated to (u, v), as determined by (2.8), is equal to the mapping $T_{u,\varphi}$ almost everywhere on U. By Corollary 2.2, T is measure preserving in the sense of (1.9). Hence u is a generalised solution of (2.15). Assumption (1.7) implies that (2.16) holds. □

4. Lagrangian duality. In this section, we study the dual problem (1.16) of the constrained nonlinear optimisation (1.13), and prove Theorem 1.2. Recall that the Lagrangian function L is defined in (1.14), where the constraint ψ is given in (1.13). Denote by I^* the optimal value of the primal problem (1.13), namely

$$I^* = \sup_{(u,v)\in K} I(u,v).$$

DEFINITION 4.1. *A factor μ^* is called a* Lagrange multiplier *for the primal problem if $\mu^* \geq 0$, and*

$$I^* = \sup\{L(u,v,\mu^*) \,:\, (u,v)\in X\}.$$

LEMMA 4.1. *Let μ^* be a Lagrange multiplier. Then (u^*,v^*) is a global maximum of the primal problem if and only if (u^*,v^*) is feasible and*

$$(u^*,v^*) = \arg\max_{(u,v)\in X} L(u,v,\mu^*), \tag{4.1}$$

$$\mu^*\psi(u^*,v^*) = 0. \tag{4.2}$$

Proof. If (u^*,v^*) is a global maximum of the primal problem, then (u^*,v^*) is feasible and furthermore,

$$\begin{aligned} I^* &= I(u^*,v^*) \leq I(u^*,v^*) + \mu^*\psi(u^*,v^*) \\ &= L(u^*,v^*,\mu^*) \leq \sup\{L(u,v,\mu^*) \,:\, (u,v)\in C(U)\times C(V)\} \\ &= I^*, \end{aligned} \tag{4.3}$$

where the first inequality follows from the definition of Lagrange multiplier ($\mu^* \geq 0$) and the feasibility of (u^*,v^*) (i.e. $\psi(u^*,v^*) \geq 0$). Using again the definition of Lagrange multiplier, we have $I^* = \sup_{(u,v)\in X} L(u,v,\mu^*)$, so that equality holds throughout (4.3). This implies the equalities (4.1)–(4.2).

Conversely, if (u^*,v^*) is feasible and (4.1)–(4.2) hold, we have from the definition of Lagrange multiplier,

$$\begin{aligned} I(u^*,v^*) &= I(u^*,v^*) + \mu^*\psi(u^*,v^*) \\ &= L(u^*,v^*,\mu^*) = \max_{(u,v)\in X} L(u,v,\mu^*) = I^*, \end{aligned}$$

so (u^*,v^*) is a global maximum. □

Recall the definitions of the dual functional J in (1.15) and the dual problem in (1.16). Note that $J(\mu)$ may be equal to $+\infty$ for some μ. In this case, we define the domain of J to be the set of μ for which $J(\mu)$ is finite:

$$D = \{\mu\in\mathbb{R} \,:\, J(\mu) < +\infty\}.$$

Regardless of the functional I and the constraint ϕ of the primal problem, the dual problem (1.16) has a nice convexity property, as shown in the following lemma.

LEMMA 4.2. *The domain D of the dual functional J is convex and J is convex over D.*

Proof. For any $u, v, \mu, \bar{\mu}$, and $\alpha \in [0,1]$, we have

$$L(u, v, \alpha\mu + (1-\alpha)\bar{\mu}) = \alpha L(u, v, \mu) + (1-\alpha) L(u, v, \bar{\mu}).$$

Taking the supremum over all $(u, v) \in X$, we obtain

$$\sup L(u, v, \alpha\mu + (1-\alpha)\bar{\mu}) \leq \alpha \sup L(u, v, \mu) + (1-\alpha) \sup L(u, v, \bar{\mu}),$$

or equivalently

$$J\left(\alpha\mu + (1-\alpha)\bar{\mu}\right) \leq \alpha J(\mu) + (1-\alpha) J(\bar{\mu}).$$

Therefore if μ and $\bar{\mu}$ belong to D, the same is true for $\alpha\mu + (1-\alpha)\bar{\mu}$, so D is convex. Furthermore, J is convex over D. □

Another important property is that the optimal dual value

$$J^* = \inf_{\mu \geq 0} J(\mu)$$

is always an upper bound of the optimal primal value, as shown in the next lemma.

LEMMA 4.3. *We have*

$$I^* \leq J^*.$$

Proof. For all $\mu \geq 0$, and $(u, v) \in X$ with $\psi(u, v) \geq 0$, we have

$$\begin{aligned} J(\mu) &= \sup_{(\tilde{u}, \tilde{v}) \in X} L(\tilde{u}, \tilde{v}, \mu) \\ &\geq I(u, v) + \mu\psi(u, v) \geq I(u, v), \end{aligned}$$

and therefore,

$$J^* = \inf_{\mu \geq 0} J(\mu) \geq \sup_{(u,v) \in K} I(u, v) = I^*.$$

□

In the language of nonlinear programming [3, Chapter 5] and [11, Chapter 4], if $J^* = I^*$ we say that *there is no duality gap*; if $J^* > I^*$ *there is duality gap.* Note that if there exists a Lagrange multiplier μ^*, the above lemma ($J^* \geq I^*$) and the definition of Lagrange multiplier ($I^* = J(\mu^*) \geq J^*$) imply that there is no duality gap.

The following is a sufficient condition for the existence of Lagrange multiplier, which is also a proof of Theorem 1.2.

LEMMA 4.4. *Let the assumptions* (1.17)–(1.19) *hold for the primal problem* (1.13). *Then there is no duality gap and there exists at least one Lagrange multiplier.*

Proof. Consider the subset of $\mathbb{R}^2$ given by

$$\begin{aligned} A = \{(z, w) \ : \ \exists\, (u, v) \in X \text{ such that} \\ \psi(u, v) \geq z, \quad I(u, v) \geq w\}. \end{aligned}$$

We first show that A is convex. Let $(z, w) \in A$ and $(\tilde{z}, \tilde{w}) \in A$ be two different elements, we show that their convex combinations belong to A.

The definition of A implies that for some $(u,v) \in X$ and $(\tilde{u},\tilde{v}) \in X$, we have

$$I(u,v) \geq w, \quad \psi(u,v) \geq z,$$
$$I(\tilde{u},\tilde{v}) \geq \tilde{w}, \quad \psi(\tilde{u},\tilde{v}) \geq \tilde{z}.$$

For any $\alpha \in [0,1]$, by the concavity of F in (1.17), we obtain

$$\begin{aligned} I\left(\alpha u + (1-\alpha)\tilde{u}, \alpha v + (1-\alpha)\tilde{v}\right) &\geq \alpha I(u,v) + (1-\alpha) I(\tilde{u},\tilde{v}) \\ &\geq \alpha w + (1-\alpha)\tilde{w}. \end{aligned}$$

By the convexity of φ in (1.18) and noting that $\inf_\Omega (f+h) \geq \inf_\Omega f + \inf_\Omega h$ for any $f,h \in C^0(\Omega)$, we obtain

$$\begin{aligned} \psi\left(\alpha u + (1-\alpha)\tilde{u}, \alpha v + (1-\alpha)\tilde{v}\right) &\geq \alpha \psi(u,v) + (1-\alpha) \psi(\tilde{u},\tilde{v}) \\ &\geq \alpha z + (1-\alpha)\tilde{z}. \end{aligned}$$

Since the combination $(\alpha u + (1-\alpha)\tilde{u}, \alpha v + (1-\alpha)\tilde{v}) \in X$, the above equations imply that the convex combination of (z,w) and $(\tilde{z},\tilde{w})$, i.e.

$$\left(\alpha z + (1-\alpha)\tilde{z}, \alpha w + (1-\alpha)\tilde{w}\right),$$

belongs to A. This proves the convexity of A.

We next observe that $(0, I^*)$ is not an interior point of A; otherwise, the point $(0, I^* + \varepsilon)$ would belong to A for some small $\varepsilon > 0$, contradicting the definition of I^* as the optimal primal value.

Therefore, there exists a supporting hyperplane passing through $(0, I^*)$ and containing A in one side. In particular, there exists a vector $(\mu, \beta) \neq (0,0)$ such that

$$\beta I^* \geq \mu z + \beta w, \qquad \forall\, (z,w) \in A. \tag{4.4}$$

We observe that if $(z,w) \in A$, then $(z, w-\gamma) \in A$ and $(z-\gamma, w) \in A$ for all $\gamma > 0$. The inequality (4.4) thus implies that

$$\mu \geq 0, \qquad \beta \geq 0. \tag{4.5}$$

We now claim that $\beta > 0$. If not, $\beta = 0$ and from (4.4)

$$0 \geq \mu z, \qquad \forall\, (z,w) \in A.$$

By the assumption (1.19), there exists a pair $(\bar{u},\bar{v}) \in X$ such that

$$\psi(\bar{u},\bar{v}) > 0.$$

Since $(\psi(\bar{u},\bar{v}), I(\bar{u},\bar{v})) \in A$, we have

$$0 \geq \mu \psi(\bar{u},\bar{v}),$$

which in view of $\mu \geq 0$ in (4.5) implies that $\mu = 0$. This means, however, that $(\mu,\beta) = (0,0)$ arriving at a contradiction. Thus, we must have $\beta > 0$ and by dividing if necessary the vector (μ,β) by β, we may assume that $\beta = 1$. Note that

$$(\psi(u,v), I(u,v)) \in A, \qquad \forall\, (u,v) \in X.$$

Equation (4.4) implies that

$$I^* \geq I(u,v) + \mu\psi(u,v), \qquad \forall\,(u,v) \in X. \tag{4.6}$$

Taking the supremum over $(u,v) \in X$ and using the fact $\mu \geq 0$, we obtain

$$\begin{aligned} I^* &\geq \sup_{(u,v)\in X} L(u,v,\mu) \\ &= J(\mu) \geq J^*, \end{aligned}$$

where J^* is the optimal dual value. By Lemma 4.3 we have the equalities hold above, namely μ is a Lagrange multiplier and there is no duality gap. □

5. Examples and applications. In this section we summarise some examples and applications of the nonlinear optimisation (1.12) from optimal transportation and geometric optics as in [19, 21].

5.1. Optimal transportation. Let $c \in C^4(U \times V)$ be a cost function. Kantorovich introduced a linear optimisation as a dual functional in optimal transportation

$$I(u,v) = \int_U u(x)f(x)dx + \int_V v(y)g(y)dy \tag{5.1}$$

over the constraint set

$$K = \{(u,v) \in C(U) \times C(V)\,:\, u(x) + v(y) \leq c(x,y), \quad \forall\, x \in U, y \in V\}. \tag{5.2}$$

The reader is referred to [1, 5, 7, 9, 22, 27, 28] for further discussion on the optimal transport problem. Note that (5.1)–(5.2) is a linear case of (1.12). In the form of (1.13), one can set

$$\begin{aligned} F(x,y,u,v) &= u(x)f(x) + v(y)g(y), \quad d\gamma = dx \otimes dy, \\ \phi(x,y,u,v) &= u(x) + v(y) - c(x,y), \end{aligned}$$

or equivalently, $\varphi(x,y,v) = v - c(x,y)$ in (1.2) and $\varphi^*(x,y,u) = u - c(x,y)$ in (3.2). All the hypotheses in Theorem 1.1 are satisfied when the cost function $c(x,y)$ satisfies [22]:

(A1) For any $x, p \in \mathbb{R}^n$, there is a unique $y \in \mathbb{R}^n$ such that $D_x c(x,y) = p$; and for any $y, q \in \mathbb{R}^n$, there is a unique $x \in \mathbb{R}^n$ such that $D_y c(x,y) = q$.

In fact, the condition (A1) implies both (H_1) and (H_1^*).

The existence of potentials in optimal transportation was previously proved in [4, 5, 9]. By directly applying the formula (2.15), one obtains the optimal transport equation

$$\left|\det\left[D^2u - D^2_{xx}c\right]\right| = |\det c_{x,y}|\frac{f}{g}. \tag{5.3}$$

We remark that in the linear case (5.1)–(5.2), both F and ϕ in (1.13) are linear in t, s variables, that is a border situation of F being concave and ϕ being convex in t, s, simultaneously.

There are numerous applications of optimal transportation including two important ones in geometric optics: In [30], Wang showed that the far field reflector problem is an optimal transport problem, and so is a linear optimisation problem. The associated cost function $c(x, y) = -\log(1 - x \cdot y)$, where x, y are points on the unit sphere $\mathbb{S}^2$. Later on in [14] Gutiérrez and Huang showed that the far field refractor problem is also an optimal transport problem. Let κ be the refractor index from the initial media to the target media. Then the associated cost function $c(x, y) = -\log(1-\kappa x\cdot y)$ when $\kappa < 1$; and $c(x, y) = \log(\kappa x \cdot y - 1)$ when $\kappa > 1$, where x, y are points on the unit sphere $\mathbb{S}^n$.

5.2. Near field reflector problem with point source. In [19] it is proved that the near field reflector problem is a nonlinear optimisation. For the convenience of the reader, we summaries the result as follows. Assume that the light emits from the origin O and passes through $\Omega \subset \mathbb{S}^n$ with a positive density $f \in L^1(\Omega)$. After being reflected from a surface Γ, the light will illuminate the target surface Ω^* in $\mathbb{R}^{n+1}$ with a prescribed positive density $g \in L^1(\Omega^*)$. Assume the energy conservation condition $\int_\Omega f = \int_{\Omega^*} g$.

Represent the reflector Γ in polar coordinate system as

$$\Gamma_\rho = \{X\rho(X) \,:\, X \in \Omega\},$$

where ρ is a positive function. From the optical property of ellipsoid, the reflector Γ_ρ is an envelope of a family of ellipsoids [19], with a focal function η satisfying the dual relation

$$\begin{aligned} \rho(X) &= \inf_{Y\in\Omega^*} \frac{1}{\eta(Y)\left(1 - \epsilon(\eta(Y))\langle X, \frac{Y}{|Y|}\rangle\right)}, \\ \eta(Y) &= \inf_{X\in\Omega} \frac{1}{\rho(X)\left(1 - \epsilon(\eta(Y))\langle X, \frac{Y}{|Y|}\rangle\right)}, \end{aligned} \tag{5.4}$$

where $\epsilon(\eta) = \sqrt{1 + 1/\eta^2|Y|^2} - 1/\eta|Y|$ is the eccentricity. We remark that the above dual relation is analogous to the classical relation between the radial and support functions for convex bodies, for example, see [26].

Inspired by [30], the reflector problem can be formulated to a nonlinear optimisation (1.12) as follows. Let $u = \log \rho$ and $v = \log \eta$. Set the functional

$$I(u, v) = \int_\Omega f(X)u + \int_{\Omega^*} g(Y)\left(v + \log(1 - \frac{\langle T^{-1}Y, Y\rangle}{e^{-v} + \sqrt{|Y|^2 + e^{-2v}}})\right), \tag{5.5}$$

and the constraint set

$$K = \{(u, v) \in C(\Omega) \times C(\Omega^*) \,:\, \phi(X, Y, u, v) \le 0\},$$

with the constraint function

$$\phi(X, Y, u, v) = u + v + \log\left(1 - \frac{\langle X, Y\rangle}{e^{-v} + \sqrt{|Y|^2 + e^{-2v}}}\right). \tag{5.6}$$

Under certain physical conditions, in [19] we obtain the existence of a maximiser $(u, v) \in K$ and $\rho = e^u$ is a solution of the reflector problem with given densities (Ω, f) and (Ω^*, g). Moreover, by directly applying the formula (2.15), we also obtain the PDE in the near field case in [19], which was previously obtained by Karakhanyan and Wang in [18].

5.3. Near field reflector problem with parallel source. The parallel case can be described as follows [20]: a parallel light emits from $\Omega \subset \mathbb{R}^n \times \{0\}$ along $e_{n+1} = (0, \cdots, 0, 1)$ with a positive density $f \in L^1(\Omega)$. After being reflected by the surface Γ in $\mathbb{R}^{n+1}$, it will illuminate the target domain $\Omega^* \subset \mathbb{R}^n \times \{0\}$ with the prescribed density $g \in L^1(\Omega^*)$. Assume the energy conservation condition $\int_\Omega f = \int_{\Omega^*} g$. We represent the reflector Γ as the graph of a positive function u,

$$\Gamma_u = \{(x, u(x)) \,:\, x \in \Omega\}.$$

Considering the optical property of the *"inverse" paraboloid* $\Gamma_p = \{(x, p(x)) \,:\, x \in \mathbb{R}^n\}$,

$$p(x) = p_y(x) = \frac{1}{2v} - \frac{v}{2}|x-y|^2, \quad v > 0, \tag{5.7}$$

one has the reflector Γ_u is the envelope of the family of paraboloids $\{\Gamma_{p_y} \,:\, y \in \Omega^*\}$, and the defining function u and the dual function v satisfy the dual relation:

$$u(x) = \inf_{y\in\Omega^*} \left\{ \frac{1}{2v(y)} - \frac{v(y)}{2}|x-y|^2 \right\}, \quad x \in \Omega, \tag{5.8}$$

$$v(y) = \inf_{x\in\Omega} \left\{ \frac{-u + \sqrt{u^2 + |x-y|^2}}{|x-y|^2} \right\}, \quad y \in \Omega^*. \tag{5.9}$$

Based on (5.8)–(5.9) we can formulate this problem to a nonlinear optimisation problem (1.12). Set the functional

$$I(u,v) = \int_\Omega f(x)u(x) + \int_{\Omega^*} g(y) \left(\frac{v}{2}|T^{-1}y - y|^2 - \frac{1}{2v} \right), \tag{5.10}$$

and the constraint set

$$\mathcal{K} = \{(u,v) \in C(\Omega) \times C(\Omega^*) \,:\, \phi(x,y,u,v) \le 0\},$$

with the constraint function

$$\phi(x,y,u,v) = u - \frac{1}{2v} + \frac{v}{2}|x-y|^2. \tag{5.11}$$

Under certain physical condition ensuring the hypotheses of Theorem 1.1, we can obtain a dual maximising pair $(u,v) \in K$ of I with $|Du| < 1$, which gives a solution of the reflector problem with given data (Ω, f) and (Ω^*, g). Moreover, by using the formula (2.15) we can derive the equation

$$\det \left[D^2 u + \frac{1 - |Du|^2}{2u} I \right] = \frac{(1 - |Du|^2)^{n+1}}{(2u)^n (1 + |Du|^2)} \frac{f}{g}, \tag{5.12}$$

which was previously obtained in [20] by directly computing the Jacobian of the reflection mapping.

5.4. Near field refractor problem with point source. Suppose the light emits from the origin surrounded by medium I with positive intensity $f(X)$ for $X \in \Omega$, where $\Omega \subset \mathbb{S}^n$. There is a surface $\mathcal{R}$, separates two homogeneous and isotropic media I and II, such that all rays refracted by $\mathcal{R}$ into medium II illuminate a target hypersurface Ω^* in $\mathbb{R}^{n+1}$ with positive intensity g on Ω^*. Assume that f, g satisfy the energy conservation condition $\int_\Omega f = \int_{\Omega^*} g$.

Let n_1, n_2 be the indices of refraction of media I, II, respectively, and $\kappa = n_2/n_1$. When $\kappa < 1$, the refracted rays tend to bent away from the normal, when $\kappa > 1$, the refracted rays tend to bent towards the normal.

Represent the surface $\mathcal{R}$ in polar coordinate system as

$$\mathcal{R}_\rho = \{X\rho(X) \,:\, X \in \Omega\},$$

where ρ is a positive function. When $\kappa < 1$, by considering the refraction property of Cartesian oval that refracts all rays emitting from the origin O into a single point Y, together with a dual function η, the pair (ρ, η) satisfies

$$\rho(X) = \inf_{Y\in\Omega^*} \frac{1-\eta(\kappa^2 X\cdot Y + \sqrt{\Delta(X\cdot Y)})}{\eta(1-\kappa^2)}, \tag{5.13}$$

$$\eta(Y) = \inf_{X\in\Omega} \frac{1-\eta(\kappa^2 X\cdot Y + \sqrt{\Delta(X\cdot Y)})}{\rho(1-\kappa^2)}. \tag{5.14}$$

Recently in [21] we formulate the refractor problem to a nonlinear optimisation as follows. Let $u = \log\rho$ and $v = \log\eta$. Set the functional

$$I(u,v) = \int_\Omega f(X)u + \int_{\Omega^*} g(Y)\left(v + \log(\frac{1-\kappa^2}{1-e^v(\kappa^2 T^{-1}Y\cdot Y + \sqrt{\Delta(T^{-1}Y\cdot Y)})})\right), \tag{5.15}$$

and the constraint set

$$K = \{(u,v) \in C(\Omega)\times C(\Omega^*) \,:\, \phi(X,Y,u,v) \le 0\},$$

with the constraint function

$$\phi(X,Y,u,v) = u + v + \log\left(\frac{1-\kappa^2}{1-e^v(\kappa^2 X\cdot Y + \sqrt{\Delta(X\cdot Y)})}\right). \tag{5.16}$$

Under some physical conditions for refraction as in [15], we obtain a dual maximising pair $(u,v) \in K$, and $\rho = e^u$ is a solution of the refractor problem in [21]. A similar formulation for the case $\kappa > 1$ was also obtained in [21] under different physical constraints.

5.5. Near field refractor problem with parallel source. The parallel case can be described as follows. Suppose that a parallel light emits from $\Omega \subset \mathbb{R}^n \times \{0\}$ along $e_{n+1} = (0, \cdots, 0, 1)$ with positive intensity $f \in L^1(\Omega)$, Ω^* is a hypersurface in $\mathbb{R}^{n+1}$, which is referred to as the target domain. Suppose that Ω and Ω^* are surrounded by two homogeneous and isotropic media I and II, respectively. One seeks an optical surface $\mathcal{R}$ interface between media I and II, such that all rays refracted by $\mathcal{R}$ into medium II are received at the surface Ω^*, and the prescribed radiation intensity received at each point $Y \in \Omega^*$ is $g(Y)$. Assume the energy conservation condition $\int_\Omega f = \int_{\Omega^*} g$.

Similarly as before, the refraction property depends on the ratio of indice $\kappa = n_1/n_2$. We may assume $\kappa < 1$. The case $\kappa > 1$ can be treated similarly. For simplicity, we also assume that $\Omega^* \subset \{y_{n+1} = h\}$ for a constant $h > 0$, and denote $Y = (y, h)$

for points on Ω^* and $X = (x, 0)$ for points on Ω. Represent the refractor $\mathcal{R}$ as graph $u|_\Omega$ for $u > 0$, namely

$$\mathcal{R}_u = \{(x, u(x)) : x \in \Omega\}.$$

Recently in [21], we formulate this problem to a nonlinear optimisation problem (1.12) as follows. Set the functional

$$I(u,v) = \int_\Omega f(x)u + \int_{\Omega^*} g(y) \left(\frac{\kappa v(y)}{1-\kappa^2} + \sqrt{\frac{v(y)^2}{(1-\kappa^2)^2} - \frac{|T^{-1}y - y|^2}{1-\kappa^2}} \right), \tag{5.17}$$

and the constraint set

$$K = \{(u,v) \in C(\Omega) \times C(\Omega^*) \, : \, \phi(X, Y, u, v) \le 0\},$$

with the constraint function

$$\phi(X, Y, u, v) = u + \frac{\kappa v(y)}{1-\kappa^2} + \sqrt{\frac{v(y)^2}{(1-\kappa^2)^2} - \frac{|x-y|^2}{1-\kappa^2}} - h. \tag{5.18}$$

Under some physical conditions as in [16] we then obtain a maximising dual pair $(u, v) \in K$ and thus a parallel refractor u satisfying the given data (Ω, f) and (Ω^*, g).

Remark 5.1. In addition to the examples arising in reflectors and refractors, there are many other nonlinear optimisation problems with potentials. For example, one can perturb the linear optimisation problem, such as the optimal transportation, to get a nonlinear one. Moreover, similarly to [23] one can show that the objective functional of any solvable linear optimisation problem can be perturbed by a differentiable, convex or Lipschitz continuous nonlinear functional in such a way that (i) a solution of the original linear problem is a local or global solution of the perturbed nonlinear problem; (ii) each global solution of the perturbed nonlinear problem is also a solution of the linear problem.

REFERENCES

[1] L. Ambrosio, *Lecture notes on optimal transport problems*, Mathematical aspects of evolving interfaces (Funchal, 2000), pp. 1–52, Lecture Notes in Math., 1812, Springer, Berlin, 2003.

[2] I. Bakelman, *Convex analysis and nonlinear geometric elliptic equations.* Springer, Berlin, Heidelberg, 1994.

[3] D. P. Bertsekas, *Nonlinear programming*, 2nd edition. Athena Scientific, 1999.

[4] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 44 (1991), pp. 375–417.

[5] L. Caffarelli, *Allocation maps with general cost functions*, in "Partial Differential Equations and Applications" (P. Marcellini, G. Talenti, and E. Vesinitini eds). Lecture Notes in Pure and Appl. Math., 177 (1996), pp. 29–35.

[6] S. Y. Cheng and S.-T. Yau, *On the regularity of the Monge-Ampère equation* $det(\partial^2 u/\partial x_i \partial x_j) = F(x, u)$, Comm. Pure Appl. Math., 30 (1977), pp. 41–68.

[7] L. C. Evans, *Partial differential equations and Monge-Kantorovich mass transfer*, in "Current developments in mathematics", Int. Press, 1997 pp. 65–126.

[8] A. Figalli, *The Monge-Ampère equation and its applications*, Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), 2017.

[9] W. Gangbo and R. J. McCann, *Optimal maps in Monge's transport problem*, C. R. Acad. Sci. Paris Sér. I. Math., 321 (1995), pp. 1653–1658.

[10] D. Gilbarg and N. Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag, Berlin, 1983.

[11] C. J. Goh and X. Q. Yang, *Duality in optimisation and variational inequalities*, optimisation Theory and Applications, 2. Taylor & Francis, Ltd., London, 2002.

[12] N. Guillen and J. Kitagawa, *Pointwise estimates and regularity in geometric optics and other generated Jacobian equations*, Comm. Pure Appl. Math., 70 (2017), pp. 1146–1220.

[13] C. Gutiérrez, *The Monge-Ampère equation*, Progress in Nonlinear Differential Equations and their Applications, 44. Birkhäuser Boston, Inc., Boston, MA, 2001.

[14] C. E. Gutiérrez and Q. Huang, *The refractor problem in reshaping light beams*, Arch. Rational Mech. Anal., 193 (2009), pp. 423–443.

[15] C. E. Gutiérrez and Q. Huang, *The near field refractor*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 31 (2014), pp. 655–684.

[16] C. E. Gutiérrez and F. Tournier, *The parallel refractor*, From Fourier analysis and number theory to Radon transforms and geometry, pp. 325–334. Developments in Mathematics 28, Springer, 2013.

[17] F. Jiang and N. S. Trudinger, *On Pogorelov estimates in optimal transportation and geometric optics*, Bull. Math. Sci., 4 (2014), pp. 407–431.

[18] A. Karakhanyan and X.-J. Wang, *On the reflector shape design*, J. Diff. Geom., 84 (2010), pp. 561–610.

[19] J. Liu, *Light reflection is nonlinear optimization*, Calc. Var. and PDEs, 46 (2013), pp. 861–878.

[20] J. Liu and N. S. Trudinger, *On the classical solvability of near field reflector problems*, Discrete Contin. Dyn. Syst., 36 (2016), pp. 895–916.

[21] J. Liu and X.-J. Wang, *Light refraction is nonlinear optimisation*, J. Math. Study, to appear.

[22] X. N. Ma, N. S. Trudinger and X.-J. Wang, *Regularity of potential functions of the optimal transportation problem*, Arch. Rat. Mech. Anal., 177 (2005), pp. 151–183.

[23] O. L. Mangasarian and R. R. Meyer, *Nonlinear perturbation of linear programs*, SIAM J. Control Optim., 17 (1979), pp. 745–752.

[24] N. S. Trudinger, *On the prescribed Jacobian equation*, Proceedings of International Conference for the 25th Anniversary of Viscosity Solutions, pp. 243–255, GAKUTO Internat. Ser. Math. Sci. Appl. 30, Gakkotosho, Tokyo, 2008.

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

[26] R. Schneider, *Convex Bodies*, The Brunn-Minkowski Theory, Cambridge University Press, Cambridge, 1993.

[27] J. Urbas, *Mass transfer problems*, Lecture Notes, Univ. of Bonn, 1998.

[28] C. Villani, *Optimal transport. Old and new*, Grundlehren Math. Wiss., Vol. 338, Springer-Verlag, Berlin, 2009.

[29] X.-J. Wang, *On the design of a reflector antenna*, Inverse problems, 12 (1996), pp. 351–375.

[30] X.-J. Wang, *On the design of a reflector antenna II*, Calc. Var. and PDEs, 20 (2004), pp. 329–341.