A SURVEY OF THE MONGE–AMPÈRE EQUATION WITH APPLICATIONS TO OPTIMAL TRANSPORTATION

GABRIEL RIOUX

Abstract. The classical Monge–Ampère equation has a well established regularity theory. A generalized form of the Monge–Ampère equation appears in the study of optimal transportation with quadratic cost whose regularity results follow from regularity theory of the classical equation. In the study of the optimal transport problem with arbitrary cost functions, a Monge–Ampère type equation also occurs naturally. We demonstrate throughout the connection between optimal transportation theory and the Monge–Ampère equation following closely a paper of Alessio Figalli and Guido De Philippis [14].

Contents

Date: April 21, 2019.

List of Symbols

∂f	subdifferential of f Def. 1.4
μ_u	Monge–Ampère measure of u Def. 1.5
$T_\sharp \mu$	pushforward of μ through T Def. 2.1
$\Pi(\mu, \nu)$	set of all transport plans Def. 2.2
f^*	convex conjugate of f Def. 2.4
f^c	c–transform of f Def. 2.8
$(C0) - (C3)$	conditions on the cost function Rmk. 2.10
c$-\exp_x(p)$	c–exponential in x evaluated at p Def. 2.11
$\mathfrak{S}_{(x,y)}(\xi, \eta)$	Ma–Trudinger–Wang tensor evaluated at (ξ, η) Def. 3.1
MTW(K)	Ma–Trudinger–Wang condition Def. 3.2
$[y_0, y_1]_{\bar{x}}$	c–segment from y_0 to y_1 with base $\bar{x}$ Def. 3.5
$\partial_c f$	c–subdifferential of f Def. 3.6
$\partial^- f$	Fréchet subdifferential of f Def. 3.7

1. The Classical Monge–Ampère Equation

We begin by stating and motivating the classical Monge–Ampère equation.

Definition 1.1. (Classical Monge–Ampère Equation) Let $\Omega \subseteq \mathbb{R}^n$ be open and $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^+$ be given. The classical *Monge–Ampère equation* is given by

$$\det D^2 u = f(x, u, \nabla u) \quad \text{in } \Omega, \tag{1.1}$$

where $u : \Omega \to \mathbb{R}$ is taken to be convex and D^2u is the Hessian matrix of u.

Recalling that the determinant of a matrix is the product of its eigenvalues, the Monge–Ampère equation can be characterized as a prescription on the eigenvalues of the Hessian of its solutions. We note, moreover that this particular equation occurs in a variety contexts. We provide a striking case where this problem surfaces and refer the reader to [1] for other interesting applications.

Example 1.2. (Prescribed Gaussian Curvature) Let u be a solution to (1.1) with $f(x, \nabla u) = K(x)(1+|\nabla u|^2)^{(n+2)/2}$ where $K(x)$ is an arbitrary function. Then the Gaussian curvature of the graph of u at $(x, u(x))$ is equal to $K(x)$.

The prescription that u be convex in (1.1) may appear arbitrary, however this condition is natural in the study of the Monge–Ampère equation as discussed in the following remark.

Remark 1.3. (Imposed Convexity of u and Degenerate Ellipticity) Suppose u is a C^∞ solution to (1.1) for $f \in C^\infty$ strictly positive such that $f = \det D^2u > 0$. Let $e \in \mathbb{S}^{n-1}$ be arbitrary, then $\partial_e(\det D^2u) = \det(D^2u)u^{ij}\partial_{ij}u_e = f_e$ via Jacobi's identity. Here $u^{ij} = (D^2u)^{-1}_{ij}$ and repeated indices are summed over. Hence,

$$\frac{f_e}{f} = u^{ij}\partial_{ij}u_e \tag{1.2}$$

such that elliptic regularity estimates on u_e can be obtained if u^{ij} is at least positive semi–definite. Notably, if u is convex, D^2u is positive semi–definite per [16, Thm. 4.5 p.27], hence so too is u^{ij}. However, unless one can establish the bounds $\frac{\mathrm{Id}}{C} \le D^2u \le \mathrm{Id}C$, for some constant C we have no a priori estimates on the eigenvalues of u^{ij}, so they may be arbitrarily small. If the prior bounds can be established, (1.2) becomes uniformly elliptic, since $\frac{\mathrm{Id}}{C} \le u^{ij} \le \mathrm{Id}C$. We note that imposing $f > 0$ means we only need the bound $|D^2u| \le C$ as the product of the eigenvalues of D^2u will be strictly greater than zero and thus the eigenvalues need only be bounded from above.

1.1. Alexandrov solutions to the Monge–Ampère Equation.

We wish to introduce the notion of a weak solution to the Monge–Ampère equation. To this effect, we first recall the definition of the subdifferential of a function.

Definition 1.4. (Subdifferential of a Function) Given $f : \Omega \to \mathbb{R}$ convex for $\Omega \subseteq \mathbb{R}^n$ open and convex, the *subdifferential* of f at $x \in \Omega$ is given by

$$\partial f(x) = \{x^* \in \mathbb{R}^n : f(z) \ge f(x) + \langle x^*, z - x\rangle \quad \forall\, z \in \Omega\}.$$

Now, we define the Monge–Ampère measure which will lead naturally to the desired weak formulation.

Definition 1.5. (Monge–Ampère Measure) The *Monge–Ampère measure* of a convex function $u : \Omega \to \mathbb{R}$ for Ω open and convex is a Borel measure defined by $\mu_u(E) = |\partial u(E)|\ \forall\, E \subseteq \Omega$ Borel. Where $|\cdot|$ denotes the Lebesgue measure and $\partial u(E) = \cup_{x\in E}\partial u(x)$.

In light of the above definition, let $u \in C^2(\Omega)$ be as in the statement of Definition 1.5. It is a standard result that $\partial u(x) = \{\nabla u(x)\}$ for every point where ∇u is defined (c.f. [16, Thm. 25.1 p.242]). Hence, $\mu_u(E) = |\nabla u(E)| = \int_\Omega \mathbb{1}_{\nabla u(E)} \mathrm{d}x = \int_E \det D^2 u(x)\ \mathrm{d}x$ for any $E \subseteq \Omega$ Borel via a change of variables and, therefore $\mu_u = \det D^2 u(x)\ \mathrm{d}x$ in Ω as Borel measures. Notably, this term gives us the left hand side of (1.1). In the sequel, we will drop the $\mathrm{d}x$ as it is understood that the equality is in the sense of Borel measures.

As noted above, this definition leads naturally to a generalized formulation of the Monge–Ampère equation and we wish to make precise the notion of a weak solution and study their properties.

Definition 1.6. (Alexandrov Solutions) Let $\Omega \subseteq \mathbb{R}^n$ be open and convex and μ be a Borel measure on Ω. Then the convex function $u : \Omega \to \mathbb{R}$ is an *Alexandrov solution* to $\det D^2 u = \mu$ if $\mu = \mu_u$ in the sense of Borel measures. If $\mu(x) = f(x)\mathrm{d}x$ we say that u solves $\det D^2 u = f$ in the weak sense.

Of note is that the Monge–Ampère measure admits the following important properties.

Proposition 1.7. (Weak* Convergence of Monge–Ampère Measure) Let (u_k) be a sequence of convex functions converging locally uniformly to u then $\mu_{u_k} \to \mu_u$ weakly*.

The above proposition is discussed in [11, Lem. 1.2.3 p.8].

Lemma 1.8. *(Monotonicity of Monge–Ampère Measure) Let E be open and bounded, u, v be convex functions satisfying $u = v$ on ∂E and $u \leq v$ in E then $\partial u(E) \supseteq \partial v(E)$ thus $\mu_u(E) \geq \mu_v(E)$.*

Proof. Let $x \in E$ be arbitrary and let $p \in \partial v(x)$. Then $z \mapsto v(x) + \langle p, (z - x)\rangle$ is a non–vertical supporting hyperplane [16, Line 1 p.215] to v at x. Since $u \leq v$ in E, $\exists\ \alpha$ such that $y \mapsto \alpha + p \cdot (y - x)$ is a non–vertical supporting hyperplane to u at some $\bar{x} \in \bar{E}$. If $\bar{x} \in \partial E$ we have that $\alpha = v(x)$ since u and v agree on ∂E and therefore $u(x) = v(x)$. Otherwise, $p \in \partial u(E) \implies \partial v(E) \subseteq \partial u(E)$ and by the monotonicity of Lebesgue measure, $\mu_v(E) \leq \mu_u(E)$. □

We now focus on demonstrating existence and uniqueness of Alexandrov solutions to the Dirichlet problem.

Definition 1.9. (Dirichlet Problem) Let $\Omega \subseteq \mathbb{R}^n$ be a convex domain, the associated *Dirichlet problem* to the generalized Monge–Ampère equation is given by

$$\begin{cases} \det D^2 u = \mu & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{1.3}$$

To prove these properties, we will make use of a maximum principle as well as a comparison principle for Alexandrov solutions.

Theorem 1.10. *(Alexandrov Maximum Principle) Let $\Omega \subseteq \mathbb{R}^n$ be an open, bounded, and convex domain. Let $u : \Omega \to \mathbb{R}$ be convex, if $u = 0$ on $\partial\Omega$ then*

$$|u(x)|^n \leq C_n (\text{diam } \Omega)^{n-1} \operatorname{dist}(x, \partial\Omega) |\partial u(\Omega)| \quad \forall\, x \in \Omega,$$

where C_n is a constant depending only on the dimension.

Proof. Let $(x, u(x))$ be a point on the graph of u and let $C_x(y)$ be a cone with vertex $(x, u(x))$ which takes on the value 0 on its base Ω. By the convexity of u and since C_x is a cone, $u(y) \leq C_x(y)$ for any $y \in \Omega$ thus by monotonicity (Lemma 1.8), $|\partial C_x(x)| \leq |\partial C_x(\Omega)| \leq |\partial u(\Omega)|$.

Let p be a vector such that $|p| < |u(x)|/\operatorname{diam}\Omega$ and consider a plane of slope p. This plane supports C_x at some $\bar{y} \in \Omega$ and hence is also supporting at x, the vertex of C_x such that $\partial C_x(x) \supseteq B(0, |u(x)|/\operatorname{diam}\Omega)$.

Let $\bar{x} \in \partial\Omega$ be such that $\operatorname{dist}(x, \partial\Omega) = |x - \bar{x}|$ and let q be a vector having the same direction as $(\bar{x} - x)$ with $|q| < |u(x)|/\operatorname{dist}(x, \partial\omega)$. Then the plane $u(x) + q \cdot (y - x)$ supports C_x at x, i.e.

$$q = \frac{\bar{x} - x}{|\bar{x} - x|} \frac{|u(x)|}{|\operatorname{dist}(x, \partial\Omega)|} \in \partial C_x(x)$$

By convexity of $\partial C_x(x)$, the cone $\mathcal{C}$ generated by q and $B(0, |u(x)|/\operatorname{diam}(\Omega))$ is such that $\mathcal{C} \subseteq \partial C_x(x)$. Therefore by monotonicity of Lebesgue measure and an inequality on $|\mathcal{C}|$

$$|\partial u(\Omega)| \geq |\partial C_x(\Omega)| \geq |\partial C_x(x)| \geq |\mathcal{C}| \geq \frac{|u(x)|^n}{C_n (\operatorname{diam}\Omega)^{n-1} \operatorname{dist}(x, \partial\Omega)}.$$

□

Lemma 1.11. *(Comparison Principle) Let $\Omega \subseteq \mathbb{R}^n$ be open, bounded and convex. Let u, v be convex functions on Ω such that $u \geq v$ on $\partial\Omega$ and $\det D^2 u \leq \det D^2 v$ in Ω in the sense of Borel measures. Then $u \geq v$ in Ω.*

Proof. Assume without loss of generality that $\det D^2 u < \det D^2 v$ (this is true up to setting $v = v + \epsilon(|x - x_0|^2 - \operatorname{diam}(\Omega)^2)$ and taking ϵ to 0). Suppose $E = \{u < v\} \neq \emptyset$ then by monotonicity of Monge–Ampère measure, $\mu_u(E) \geq \mu_v(E)$ which contradicts the initial assumption. □

A direct consequence of this lemma is the uniqueness of Alexandrov solutions to the Dirichlet problem. Indeed given two solutions u, v to (1.3), we have that $u = v = 0$ on $\partial\Omega$ and $\det D^2 u = \det D^2 v = \mu$ in Ω. Whence by the comparison principle, $u = v$ in Ω. We now have the following stability result (see [11, Lemma 5.3.1 p.96]) which will allow us to prove existence of solutions.

Theorem 1.12. *(Stability of Alexandrov Solutions to the Dirichlet Problem) Let $\Omega_k \subseteq \mathbb{R}^n$ be a family of convex domains with associated convex Alexandrov solutions $u_k : \Omega_k \to \mathbb{R}$ (i.e. u_k solves (1.3) on Ω_k with $\mu = \mu_k$) then if $\Omega_k \to \Omega$ in Hausdorff distance (Ω convex) and the μ_k satisfy $\sup_k \mu_k(\Omega_k) < +\infty$ with $\mu_k \to \mu$ in the weak* topology for μ a Borel measure. Then $u_k \to u$ uniformly where u is a solution to (1.3).*

The existence of Alexandrov solutions will be demonstrated in a method analogous to the method of Perron.

Theorem 1.13. *(Existence of Alexandrov Solutions) Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open convex domain and μ be a nonnegative Borel measure in Ω. Then there exists an Alexandrov solution of (1.3).*

Proof. Let $\mu_k = \sum_{i=1}^k \alpha_i \delta_{x_i}$, $\alpha_i \geq 0$ be a family of atomic measures converging weakly* to μ here δ_{x_i} refers to the Dirac delta distribution centred at x_i. Define $S[\mu_k] = \{v : \Omega \to \mathbb{R} : v \text{ is convex}, v = 0 \text{ on } \partial\Omega, \det D^2 v \geq \mu_k\}$. Note that $S[\mu_k]$ is non–empty as the function $-A \sum_{i=1}^k C_{x_i}$ for C_x the conical one–homogeneous function taking value -1 at x and 0 on $\partial\Omega$ and $A > 0$ sufficiently large is an element of $S[\mu_k]$.

This set is closed under suprema per Proposition 1.7 and we have that the maximum of any two elements of $S[\mu_k]$ is also in $S[\mu_k]$, so we define $u_k = \sup_{v \in S[\mu_k]} v \in S[\mu_k]$ in analogy with Perron's method.

It remains to be seen if u_k satisfies the Monge–Ampère equation on Ω_k with appropriate source term μ_k such that the prior stability result can be applied.

To this effect, we show that $\det D^2 u_k$ is a measure concentrated on $\{x_1, \ldots, x_k\}$. Suppose not, then $\exists\, \bar{x} \in \Omega \backslash \{x_1, \ldots, x_k\}$ and $p \in \mathbb{R}^n$ such that $p \in \partial u(\bar{x}) \backslash \partial u(\{x_1, \ldots, x_k\})$. Therefore (by definition of the subdifferential) $u_k(x_j) > u(\bar{x}) + p \cdot (x_j - \bar{x}) \quad \forall\, j \in \{1, \ldots, k\}$ and taking $\tilde{u}_k(x) = \max\{u_k(x), u_k(\bar{x}) + p \cdot (y - \bar{x}) + \delta\} \in S[\mu_k]$ with $\delta > 0$ sufficiently small yields a larger subsolution, contradicting the assumption that u_k is the supremum.

Now, we show that $\det D^2 u_k = \mu_k$. Assume this is not the case, then we have by the previous step that $\det D^2 u_k = \sum_{i=1}^{k} \beta_i \delta_{x_i}$ with $\beta_i \geq \alpha_i$ and $\beta_j > \alpha_j$ for some $j \in \{1, \ldots, k\}$. We note that $\partial u(x_j)$ is a convex set with non–zero volume, hence its interior is nonempty and $\exists\, p \in \operatorname{int}(\partial u(x_j))$. Assume without loss of generality that $p = 0$ (up to subtracting $p \cdot y$ from u_k) and define

$$\bar{u}_k(x) = \begin{cases} u_k(x) & u_k > u_k(x_j) + \delta \\ (1 - \delta) u_k(x) + \delta(u_k(x_j) + \delta) & u_k \leq u_k(x_j) + \delta \end{cases}$$

for $\delta > 0$ sufficiently small that $\bar{u}_k$ is a larger subsolution, a contradiction as before.

Finally, $u_k = 0$ on $\partial\Omega$ since $u_k(x) \geq -C \operatorname{dist}(x, \partial\Omega)^{1/n}$ as a consequence of Theorem 1.10. □

Under the assumptions of the above theorem, the existence and uniqueness of Alexandrov solutions to the Dirichlet problem have been demonstrated.

1.2. **Existence of Smooth Classical Solutions.** We now wish to show the existence of smooth solutions to the Dirichlet problem for the classical Monge–Ampère equation:

$$\begin{cases} \det D^2 u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}. \tag{1.4}$$

To this effect, we wish to establish a C^0 estimate of the Hessian of solutions to this problem. Note that not all domains will permit regularity up to the boundary, hence we establish the notion of a uniformly convex domain as it is precisely the type of domain on which this desirable property can be obtained.

Definition 1.14. (Uniformly Convex Domain) A domain Ω is called *uniformly convex* if $\exists$ $R > 0$ such that $\Omega \subseteq B_R(x_0 + R\nu_{x_0})$ $\forall\, x_0 \in \partial\Omega$ where ν_{x_0} denotes the interior normal to Ω at x_0.

We now establish out C^0 estimate on the Hessian.

Theorem 1.15. *(C^0 estimate on Hessian) Let $\Omega \subseteq \mathbb{R}^n$ be a C^3 uniformly convex domain and u be a solution of (1.4) with $f \in C^2(\bar{\Omega})$ and $\lambda \leq f \leq 1/\lambda$. Then $\exists$ a constant C depending only on Ω, λ, $||f||_{C^2(\bar{\Omega})}$ such that*

$$\left|\left|D^2 u\right|\right|_{C^0(\bar{\Omega})} \leq C$$

Proof. We obtain first a C^0 estimate on u via barrier construction. By the comparison principle, it suffices to let $v(x) = \lambda^{-1/n}(|x - x_1|^2 - R^2)$ for x_1 and R satisfying $\Omega \subseteq B_R(x_1)$ to obtain a uniform lower bound on u.

Now, a C^1 estimate. By convexity, $\sup_\Omega |\nabla u| = \sup_{\partial\Omega} |\nabla u|$, so we only need an estimate on the boundary. Recalling that $u = 0$ on $\partial\Omega$ demonstrates that any tangential derivative is zero, so only estimates of the normal derivative are necessary.

To this effect, let $x_0 \in \partial\Omega$ be arbitrary and construct the barriers $v_\pm(x) = \lambda^{\mp 1/n}(|x - x_\pm|^2 - R_\pm^2)$ for $x_\pm = x_0 + R_\pm \nu_{x_0}$ and $0 < R_- < R_+ < \infty$ chosen such that $B_{R_-}(x_-) \subseteq \Omega \subseteq B_{R_+}(x_+)$. Therefore, $v_+ \le u \le v_-$ and $-C \le \partial_\nu u(x_0) \le -1/C$ for some C

Finally, a C^2 estimate. Given a unit vector e, consider $\log\det D^2u = \log f$ and define the derivative in the direction of e as $L(u_e) = u^{ij}(u_e)_{ij} = (\log f)_e$ and the second derivative is $u^{ij}(u_{ee})_{ij} - u^{il}u^{kj}(u_e)_{ij}(u_e)_{lk} = (\log f)_{ee}$. Since u is convex, $u^{il}u^{kj}(u_e)_{ij}(u_e)_{lk} \ge 0$, so $L(u_{ee}) \ge (\log f)_{ee} \ge -C$, for C depending only on f.

Now, $L(u) = u^{ij}u_{ij} = n$ so we have $L(u_{ee} + Mu) \ge 0$ for M sufficiently large. By the maximum principle therefore, $\sup_\Omega(u_{ee} + Mu) \le \sup_{\partial\Omega}(u_{ee} + Mu)$.

Since u is bounded, we only need to estimate D^2u on the boundary. Assume $0 \in \partial\Omega$ and that $\partial\Omega = \{(x_1, \dots, x_n) \ : \ x_n = \sum_{\alpha=1}^{n-1} \frac{\kappa_\alpha}{2} x_\alpha^2 + O(|x|^3)\}$ for some $\kappa_\alpha > 0$ locally. By smoothness and uniform convexity of Ω, $1/C \le \kappa_\alpha \le C$, so $u_{\alpha\alpha}(0) = -\kappa_\alpha u_n(0)$, $u_{\alpha\beta}(0) = 0 \,\forall\, \alpha \ne \beta \in \{1, \dots, n-1\}$.

By our estimate on the normal derivative then $\mathrm{Id}_{n-1}/C \le (u_{\alpha\beta}(0))_{\alpha,\beta\in\{1,\dots,n-1\}} \le C\mathrm{Id}_{n-1}$. Then, noting that $f = \det D^2u = M^{nn}(D^2u)u_{nn} + \sum_{\alpha=1}^{n-1} M^{\alpha n}(D^2u)u_{\alpha n}$ with $M^{ij}D^2u$ denoting the cofactor of u_{ij}; this and the upper bound on $u_{\alpha\beta}(0)$ will give an upper bound on $u_{nn}(0)$ once a bound for $u_{\alpha n}(0)$ is found for $\alpha \in \{1, \dots, n-1\}$.

To this effect, consider the rotational derivative operator $R_{\alpha n} = x_\alpha\partial_n - x_n\partial_\alpha$, $\alpha \in \{1, \dots, n-1\}$. By invariance of derivatives with respect to rotation, differentiating $\log\det D^2u = \log f$ yields $L(R_{\alpha n}u) = u^{ij}(R_{\alpha n}u)_{ij} = R_{\alpha n}(\log f)$.

So, multiplying the prior equation by κ_α and using the notation $L(u_\epsilon) = u^{ij}(u_\epsilon)_{ij} = (\log f)_\epsilon$ as before yields $|L((1 - \kappa_\alpha x_n)u_\alpha + \kappa_\alpha x_\alpha u_n)| \le C$. Since $u = 0$ on $\partial\Omega$, the uniform convexity of Ω and the bound on $|\nabla u|$ give us $|(1 - \kappa_\alpha x_n)u_\alpha + \kappa_\alpha x_\alpha u_n| \le -A|x|^2 + Bx_n$ for a suitable choice of A, B, depending on Ω.

By the the AM–GM inequality, $L(-A|x|^2 + Bx_n) = -A\sum_i u^{ii} \le -\frac{nA}{(\det D^2u)^{1/n}} \le -\frac{nA}{\lambda^{1/n}}$ and choosing A large enough yields $|(1-\kappa_\alpha x_n)u_\alpha + \kappa_\alpha x_\alpha u_n| \le -A|x|^2 + Bx_n$ in Ω. Dividing through by x_n and taking $x_n \to 0$ yields $|u_{\alpha n}|(0) \le C$ for C depending only on Ω and f. □

The prior estimates yield the following existence result.

Theorem 1.16. *Let Ω be a C^3 uniformly convex domain. Then $\forall\, f \in C^2(\bar\Omega)$ with $\lambda \le f \le 1/\lambda$ there exists a $C^{2,\alpha}(\bar\Omega)$ solution of (1.4).*

Proof. The proof relies on the method of continuity. We provide an outline of this method and refer the reader to a more complete treatment in [10, Chap. 17.2 p.446]. Suppose $\bar u$ is a smooth and convex solution to (1.4) with associated source $\bar f(x)$, we wish to find from $\bar u$ a solution u with associated source $f(x)$. To this effect, define $f_t = (1-t)\bar f + tf$, $t \in [0,1]$ and consider the one–parameter family of problems

$$\begin{cases} \det D^2u_t = f_t & \text{in } \Omega \\ u_t = 0 & \text{on } \partial\Omega \end{cases}$$

Now, assume $f, \bar f$ are smooth and consider

$$\mathcal{C} = \{u \in C^{2,\alpha}(\bar\Omega) \text{ convex } : u = 0 \text{ on } \partial\Omega\}.$$

Finally, we define the map

$$\mathcal{F} : \mathcal{C} \times [0,1] \to C^{0,\alpha}(\bar\Omega) \quad (u,t) \mapsto \det D^2u - f_t$$

and wish to demonstrate that $\Gamma = \{t \in [0,1] : \exists\, u_t \in \mathcal{C} \text{ with } \mathcal{F}(u_t, t) = 0\}$ is both open and closed in $[0,1]$.

The Fréchet differential of $\mathcal{F}$ with respect to u is given by the linearized Monge–Ampère equation

$$D_u\mathcal{F}(u,t)[h] = \det(D^2u)u^{ij}h_{ij}, \quad h = 0 \text{ on } \partial\Omega, \tag{1.5}$$

We note that $\mathcal{F}$ is continuously differentiable at every point of Γ. If u is bounded in $C^{2,\alpha}$ as shown in Theorem 1.15 and f is bounded from below by λ then the smallest eigenvalue D^2u is bounded uniformly away from zero and the linearized operator becomes uniformly elliptic with $C^{0,\alpha}$ coefficients. Therefore, classical Schauder's theory yields invertibility of $D_u\mathcal{F}(u,t)$ such that the Implicit Function Theorem in Banach spaces [10, Thm. 17.6 p.447] can be implied yielding the openness of Γ.

As it pertains to the closedness of Γ, since the linearized equation is uniformly elliptic we obtain by the Evans–Krylov theorem [10, Thm. 17.26$'$ p.481] that $u \in C^{2,\alpha}(\bar{\Omega})$ and therefore by the Arzelà–Ascoli theorem Γ is also closed. In particular setting $t = 1$ in Γ demonstrates the existence of a solution that is $C^{2,\alpha}$ up to the boundary. □

1.3. **Interior Regularity for Alexandrov Solutions to the Dirichlet Problem.** We now wish to discuss results pertaining to the interior regularity of weak solutions to the Dirichlet problem.

Theorem 1.17. *(Pogorelov Interior Estimate) Let $u \in C^4(\Omega)$ be an Alexandrov solution of (1.3) with $\mu(x) = f(x)dx$, $f \in C^2(\Omega)$ and $\lambda \geq f \geq 1/\lambda$. Then $\exists\, C$ depending on λ and $||f||_{C^2}$ such that*

$$|u(x)|u_{11}(x)\exp\left(\frac{u_1^2}{2}\right) \leq C\left|\left|\exp(\frac{u_1^2}{2})(1+|u_1|+|u|)\right|\right|_{L^\infty(\Omega)} \quad \forall\, x \in \Omega.$$

Where $u_1(x) = \partial_1 u(x)$ and so on.

Proof. By convexity, we note that $u \leq 0$ in Ω thus we define $w = (-u)u_{11}e^{\frac{(u_1)^2}{2}}$ and if x_0 is a maximal point of w in $\bar{\Omega}$ we have that $x_0 \in \Omega$ since $u = 0$ on $\partial\Omega$. We now change coordinate systems to $x' = Ax$ with $\det A = 1$ leaving x_1 invariant and such that u_{ij} is diagonal at x_0.

We note that $(\log w)_i = \frac{u_i}{u} + \frac{u_{11i}}{u_{11}} + u_1u_{1i}$ and $(\log w)_{ij} = \frac{u_{ij}}{u} - \frac{u_iu_j}{u^2} + \frac{u_{11ij}}{u_{11}} - \frac{u_{11i}u_{11j}}{u_{11}^2} + u_{1j}u_{1i} + u_1u_{1ij}$. Since x_0 is a local max, at this point, $(\log w)_i = 0$ and $0 \geq (\log w)_{ij}$ and u_{ij} is nonnegative so $0 \geq u^{ij}(\log w)_{ij}$ at x_0.

Noting that $u^{ij}u_{ij} = n$, $u^{ij}(u_e)_{ij} = (\log f)_e$ and $u^{ij}(u_{ee})_{ij} - u^{il}u^{kj}(u_e)_{ij}(u_e)_{lk} = (\log f)_{ee}$ using the notation of the proof of Theorem 1.15 and noting that $u_e = u$ in our case yields $0 \geq \frac{n}{u} - \frac{u^{ij}u_iu_j}{u^2} + \frac{(\log f)_{11}}{u_{11}} + \frac{u^{il}u^{kj}u_{ij}u_{lk}}{u_{11}} - \frac{u^{ij}u_{11i}u_{11j}}{u_{11}^2} + u^{ij}u_{1j}u_{1i} + (\log f)_1u_1$ where we have multiplied through by u^{ij}.

We consider first the middle term $\frac{u^{il}u^{kj}u_{1ij}u_{1kl}}{u_{11}} - \frac{u^{ij}u_{11i}u_{11j}}{u_{11}^2} - \frac{u^{ij}u_iu_j}{u^2}$ and note that by a prior calculation, $\frac{u_i}{u} = -\frac{u_{11i}}{u_{11}} - u_1u_{1i}$ so the middle tem is equal to $\frac{u^{il}u^{kj}u_{1ij}u_{1kl}}{u_{11}} - \frac{u^{ij}u_{11i}u_{11j}}{u_{11}^2} - u^{ij}\left(\frac{u_{11i}}{u_{11}} + u_1u_{1i}\right)\left(\frac{u_{11j}}{u_{11}} + u_1u_{1j}\right)$, since both u_{ij} and u^{ij} are diagonal at x_0, this term is greater than $-u^{11}\left(\frac{u_{111}}{u_{11}} + u_1u_{11}\right)^2 = -\frac{u_1^2}{u^2u_{11}}$. Such that $0 \geq \frac{n}{u} + \frac{(\log f)_{11}}{u_{11}} + u^{ij}u_{1j}u_{1i} + (\log f)_1u_1 - \frac{u_1^2}{u^2u_{11}}$.

Since u_{ij} is diagonal we get $0 \geq \frac{n}{u} + \frac{(\log f)_{11}}{u_{11}} + u_{11} + (\log f)_1 u_1 - \frac{u_1^2}{u^2 u_{11}}$ such that multiplying by $u^2 u_{11} e^{u_1^2}$ gives us $0 \geq nuu_{11}e^{u_1^2} + (\log f)_{11}u^2 e^{u_1^2} + u^2 u_{11}^2 e^{u_1^2} + (\log f)_1 u_1 u^2 u_{11} e^{u_1^2} - u_1^2 e^{u_1^2} = -nwe^{\frac{u_1^2}{2}} + (\log f)_{11} u^2 e^{u_1^2} + w^2 - (\log f)_1 u_1 uwe^{\frac{u_1^2}{2}} - u_1^2 e^{u_1^2} = w^2 - (n + u_1 u (\log f)_1) we^{\frac{u_1^2}{2}} + e^{u_1^2}((\log f)_{11} u^2 - u_1^2) \geq w^2 - Cwe^{\frac{u_1^2}{2}}(1 + |u_1| u) - Ce^{u_1^2}(u^2 + u_1^2)$ which yields the desired bound for $C = C(f, \lambda)$. □

Definition 1.18. (Strictly Convex Functions) A convex function u is called *strictly convex* in Ω if $\forall\, x \in \Omega$ and $p \in \partial u(x)$

$$u(z) > u(x) + p \cdot (z - x) \quad \forall\, z \in \Omega \backslash \{x\}.$$

In particular stictly convex functions are characterized by the property that any of their supporting planes touches their graph precisely once.

Theorem 1.19. *(Regularity for Strictly Convex Solutions) Let $u : \Omega \to \mathbb{R}$ be a convex Alexandrov solution of $\det D^2 u = f$ for $f \in C^2(\Omega)$ and $\lambda \leq f \leq 1/\lambda$. Assume u is strictly convex in $\Omega' \subseteq \Omega$. Then $u \in C^2(\Omega')$.*

Proof. Let $x_0 \in \Omega'$, $p \in \partial u(x_0)$ and consider the section of u at height t defined via $S(x, p, t) = \{y \in \Omega \ : \ u(y) \leq u(x) + p \cdot (y - x) + t\}$. Since u is strictly convex, we choose $t > 0$ small enough that $S(x, p, t) \Subset \Omega'$. Then, consider S_ϵ, a sequence of smooth uniformly convex sets converging to $S(x_0, p, t)$ and apply Theorem 1.16 to get $v_\epsilon \in C^{2,\alpha}(S_\epsilon)$ satisfying

$$\begin{cases} \det D^2 v_\epsilon = f * \rho_\epsilon & \text{in } S_\epsilon \\ v_\epsilon = 0 & \text{on } \partial S_\epsilon \end{cases}$$

by Schauder's theory $v_\epsilon \in C^\infty(S_\epsilon)$ so by Theorem 1.17, $|D^2 v_\epsilon| \leq C$ in $S(x_0, p, t/2)$ for ϵ sufficiently small. Since $S_\epsilon \to S(x_0, p, t)$ and $u(x) = u(x_0) + p \cdot x + t$ on $\partial S(x_0, p, t)$, by uniqueness of weak solutions we have that $v_\epsilon + u(x_0) + p \cdot x + t \to u$ uniformly as $\epsilon \to 0$, hence $|D^2 u| \leq C$ in $S(x_0, p, t/2)$. This makes the linearized Monge–Ampère equation uniformly elliptic as we have discussed before such that $u \in C^2(S(x_0, p, t/4))$. Since x_0 is arbitrary we have $u \in C^2(\Omega')$ as desired. □

1.4. Regularity of Alexandrov Solutions Without Imposed Boundary Conditions.

We now wish to study regularity results for Alexandrov solutions to (1.1). The following result is due to Caffarelli (c.f. [4]).

Theorem 1.20. *(Regularity for Strictly Convex Weak Solutions) Let $u : \Omega \to \mathbb{R}$ be a strictly convex Alexandrov solution of $\det D^2 u = f$ satisfying $\lambda \leq f \leq 1/\lambda$. Then $u \in C^{1,\alpha}_{loc}(\Omega)$ for some universal α. Precisely $\forall\ \Omega' \Subset \Omega\ \exists\ C$ depending on λ, Ω' and the modulus of strict convexity of u such that*

$$\sup_{\substack{x, y \in \Omega' \\ x \neq y}} \frac{|\nabla u(x) - \nabla u(y)|}{|x - y|^\alpha} \leq C.$$

The intuition behind the proof of this theorem hinges on the following lemma.

Lemma 1.21. *(John's Lemma) Let $V \subseteq \mathbb{R}^n$ be a bounded convex set with non–empty interior. Then $\exists!$ ellipsoid E of maximal volume contained in V. Moreover,*

$$E \subseteq V \subseteq nE$$

where nE denotes dilation of E by a factor n.

We recall the notion of a normalized convex set, as solutions to the Monge–Ampère equation on such sets exhibit some useful properties.

Definition 1.22. (Normalized Convex Set) A convex set E is normalized provided $B_1 \subseteq E \subseteq nB_1$

The utility of John's lemma is that one can take any bounded open convex set Ω and transform it into a normalized convex set $A(\Omega)$ via an affine transformation A, and in particular if u is such that $\lambda \leq \det D^2u \leq 1/\lambda$ in Ω, $u = 0$ on $\partial\Omega$ then if A normalizes Ω, $v = (\det A)^{\frac{2}{n}} u \circ A^{-1}$ satisfies $\lambda \leq \det D^2 v \leq 1/\lambda$ in Ω, $v = 0$ on $\partial\Omega$.

As aforementioned solutions to the Monge–Ampère equation on normalized convex sets have some useful properties, as discussed in the following theorem.

Theorem 1.23. *Let Ω be a normalized convex set and u be a solution of $\det D^2u = f$ with $\lambda \leq f \leq 1/\lambda$. Then $\exists$ positive constants $\alpha = \alpha(n, \lambda)$ and $C = C(n, \lambda)$ such that $||u||_{C^{1,\alpha}}(B_{1/2}) \leq C$*

The proof of the above theorem hinges essentially on the ability to show that solutions of the Monge–Ampère equation in this context on normalized convex domains have a universal modulus of strict convexity. To this effect, the following proposition of Caffarelli (c.f. [2]).

Proposition 1.24. Let u be a solution of $\lambda \leq \det D^2u \leq 1/\lambda$ in a convex set Ω and $\ell : \mathbb{R}^n \to \mathbb{R}$ be a linear function supporting u at $\bar{x} \in \Omega$. If $W = \{x \in \Omega : u(x) = \ell(x)\}$ contains more than one point then it cannot have an extremal point in Ω i.e. every point of W lies on an open line segment joining two points of W.

That is, if the boundary conditions forbid u from coinciding with an affine function along a segment crossing Ω then u is strictly convex as it touches any of its supporting planes exactly once. By the weak* convergence property of the Monge–Ampère measure (Proposition 1.7) and the fact that the family of normalized convex domains is compact we have that for a normalized convex set Ω, the class of solutions is compact with respect to uniform convergence and therefore share a universal modulus of strict convexity.

Now, combining all of the prior results yields the following lemma which will allow us to prove Theorem 1.20.

Lemma 1.25. *Let Ω be a normalized convex set, v be a solution of $\det D^2v = f$ with $\lambda \leq f \leq 1/\lambda$. Let x_0 be a minimum point for $u \in \Omega$ and $\forall\, \beta \in (0, 1]$ let the cone with vertex $(x_0, v(x_0))$ and base $\{v = (1-\beta)\min v\} \times \{(1-\beta)\min v\}$ be denoted by $C_\beta \subseteq \mathbb{R}^{n+1}$. If h_β is the function whose graph is given by C_β $\exists$ a universal constant $\delta_0 > 0$ such that*

$$h_{1/2} \leq (1-\delta_0)h_1$$

Proof. (Proof of Theorem 1.20) We use throughout the notation of Lemma 1.25. Let $k \in \mathbb{N}$ be arbitrary and consider $\Omega_k = \{u \leq (1-2^{-k})\min u\}$. Renormalizing the convex set Ω_k through an affine map A_k, applying Lemma 1.25 to $v = (\det A_k)^{2/n} u \circ A_k$ and transfering the information back to u, one deduces that $h_{2^{-(k+1)}} \leq (1-\delta_0)h_{2^{-k}}$. Iterating this estimate yields $h_{2^{-k}} \leq (1-\delta_0)^k h_1 \forall k \in \mathbb{N}$. Then we get that $v \in C^{1,\alpha}$ in the sense that $u(y) - u(x_0) \leq C|y-x_0|^{1+\alpha}$. Now for every $x \in \Omega' \Subset \Omega$ and $p \in \partial u(x)$, repeat the same argument with $u(y) - p\cdot(y-x)$ in lieu of u and replacing Ω with $S(x, p, t)$ for t small satisfying $S(x, p, t) \Subset \Omega$ and taking a renormalization of S. Then $u(y) - u(x) - p\cdot(y-x) \leq C|y-x|^{1+\alpha}\,\forall\, p \in \partial u(x)$. Since $x \in \Omega'$ is arbitrary the above estimate is shown in [8, Lemma 3.1 p.4411] to yield $u \in C^{1,\alpha}_{\text{loc}}$.

□

If the source term f is assumed to be Hölder continuous, Caffarelli further improved this result in [3] to read as follows.

Theorem 1.26. *Let Ω be a normalized convex set and u be an Alexandrov solution of* $\det D^2u = f$ *with* $\lambda \leq f \leq 1/\lambda$ *and* $f \in C^{0,\alpha}(\Omega)$. *Then* $||u||_{C^{2,\alpha}(B_{1/2})} \leq C$ *for some* C *depending on* n, λ *and* $||f||_{C^{0,\alpha}(B_1)}$.

The proof of the above is provided in [3] and is based on the fact that if f is locally close to a constant, u is locally close to a solution of $\det D^2u = c$ for c constant. We have already established interior estimates for this type of solution and approximation via interpolation yields the requires bound on $||u||_{C^{2,\alpha}}(B_{1/2})$. A similar reasoning yields the following result with more details provided in [3].

Theorem 1.27. *Let Ω be a normalized convex set and u be a solution of* $\det D^2u = f$. *Then* $\forall\, p > 1$, $\exists\, \delta(p)$ *and* $C = C(p)$ *such that if* $||f - 1||_\infty \leq \delta(p)$ *then* $||u||_{W^{2,p}(B_{1/2})} \leq C$.

If one localizes this result as in the proof of Theorem 1.20 for strictly convex solutions u with f continuous one obtains that $u \in W^{2,p}_{\text{loc}}(\Omega)\, \forall\, p < \infty$. One can refine this result to the following as discussed in [15].

Theorem 1.28. *Let Ω be a normalized convex set and u be an Alexandrov solution of* $\det D^2u = f$ *such that* $\lambda \leq f \leq 1/\lambda$. *Then* $\exists\, \epsilon = \epsilon(n, \lambda)$ *and* $C = C(n, \lambda)$ *such that* $||u||_{W^{2,1+\epsilon}(B_{1/2})} \leq C$.

We have throughout this section discussed the general regularity theory of both classical and weak solutions to the Monge–Ampère equation. We now wish to establish the connection between this equation and the problems studied in optimal transportation theory.

2. Optimal Transportation and the Monge–Ampère Equation

The Monge–Ampère equation appears naturally in the context of optimal transportation theory. One can therefore derive regularity and existence results for the problems of optimal transportation using the theory established hitherto an vice–versa. We therefore provide a brief introduction of the basic concepts of optimal transportation before elaborating on these connections.

2.1. **A Brief Primer to Optimal Transportation Theory.**

Definition 2.1. (Monge Problem and Optimal Transport Map) Given two probability measures μ, ν on the measurable spaces X, Y respectively and a cost function $c : X \times Y \to [0, \infty]$. The *Monge Problem* consists of solving

$$\text{(MP)} \qquad \inf\left\{\int_X c(x, T(x))\mathrm{d}\mu(x) : T_\sharp\mu = \nu, T : X \to Y \text{ measurable}\right\}.$$

Where $T_\sharp\mu = \nu$ denotes that the pushforward of μ through T equals ν. This means that $\forall\ A \subseteq Y$ measurable, $\mu(T^{-1}(A)) = \nu(A))$. The *optimal transport map* is the arg–min of this problem provided it exists.

We now define a relaxation of the Monge problem introduced by Kantorovich.

Definition 2.2. (Kantorovich Problem and Optimal Transport Plans) Given two probability measures μ, ν on the measurable spaces X, Y respectively and a cost function $c : X \times Y \to [0, \infty]$. The *Kantorovich Problem* consists of solving

$$\text{(KP)} \qquad \inf\left\{\int_{X\times Y} c(x, y)\,\mathrm{d}\gamma : \gamma \in \Pi(\mu, \nu)\right\}.$$

Where $\Pi(\mu, \nu)$={probability measures γ on $X \times Y : (\pi_x)_\sharp\gamma = \mu, (\pi_y)_\sharp\gamma = \nu$}. Here, π_x and π_y are the projections of $X \times Y$ onto X and Y respectively. Of course the arg–min of this expression, if it exists, is referred to as the *optimal transport plan.*

Note that if γ is of the form $(\mathrm{id}, T)_\sharp\mu$ for $T : X \to Y$ measurable we get that $\forall\ A \subseteq X$ measurable, $\mu((\mathrm{id}, T^{-1})(\pi_x^{-1}(A))) = \mu(A)$ so the condition $(\pi_x)_\sharp\gamma = \mu$ is realized. Moreover, $(\pi_y)_\sharp\gamma = \nu \implies \forall\ B \subseteq Y$ measurable, $\mu((\mathrm{id}, T^{-1})(\pi_y^{-1}(B))) = \mu(T^{-1}(B)) = \nu(B)$ so, in particular, $T_\sharp\mu = \nu$ and T is an optimal transport map as the functionals being minimized in (MP) and (KP) are identical when γ takes on this particular form.

The Kantorovich problem is said to be a relaxation of the Monge problem in this respect, as if a problem is solvable in the Monge framework, one deduces a solution in the Kantorovich framework as in the above discussion. However, in some contexts, the Monge problem admits no solutions, but the Kantorovich problem does. The canonical example of this behaviour is when μ is a Dirac mass and ν is not. Thus, no measurable map can push μ onto ν, but one can form measures admitting μ and ν as marginals.

Next, we introduce the so–called dual problem which we shall show in the sequel to be equivalent to the Kantorovich problem under certain assumptions.

Definition 2.3. (Dual Problem) Given two probability measures μ, ν on the measurable spaces X, Y respectively and a cost function $c : X \times Y \to [0, \infty]$. The *Dual Problem* consists of solving

$$\text{(DP)} \qquad \sup\left\{\int_X \varphi(x)\,\mathrm{d}\mu(x) + \int_Y \psi(y)\,\mathrm{d}\nu(y) : \varphi \in C_b(X),\ \psi \in C_b(Y) \text{ and } \varphi + \psi \leq c\right\}.$$

We note that taking $\varphi \in L^1(X)$ and $\psi \in L^1(Y)$ yields an identical problem.

Another useful definition will be that of the convex conjugate or Legendre–Fenchel transformation.

Definition 2.4. (Convex Conjugate) Let X be a normed vector space and X^* be its dual space then the convex conjugate of $f : X \to \bar{\mathbb{R}}$ denoted $f^* : X^* \to \bar{\mathbb{R}}$ is given by $f^*(x^*) = \sup_{x\in X}\{\langle x^*, x\rangle - f(x)\}$ where $\langle \cdot, \cdot\rangle : X^* \times X \to \mathbb{R}$ denotes the duality pairing.

We now state a fundamental result of convex analysis which we shall use to prove the equivalence of (DP) and (KP). The proof of which can be found in [19, Cor. 2.8.5 p.125].

Theorem 2.5. *(Fenchel–Rockafellar Duality) Let X be a Fréchet space, f and g be lower semicontinuous functions from X to $\bar{\mathbb{R}}$ such that $\exists\, x \in X$ such that $f(x), g(x) < \infty$. Then* $\inf_{x\in X}\{f + g\} = \max_{x^*\in X^*}\{-f^*(-x^*) - g^*(x^*)\}$

We now have all of the machinery required to prove that the Kantorovich problem and the dual problem are equivalent following the treatment in [17, Thm. 1.3 p.19].

Theorem 2.6. *((KP)=(DP)) Suppose X and Y are compact and the cost function $c : X \times Y \to \bar{\mathbb{R}}$ is continuous then (KP)=(DP) and an optimal transport plan γ exists.*

Proof. Let $\Gamma = C_b(X \times Y)$ normed with $||\cdot||_\infty$. We have by Riesz's representation theorem that $\Gamma^* = \mathscr{R}(X \times Y)$ the space of regular Radon measures on $X \times Y$ normed by total variation. For $u \in \Gamma$ define

$$\xi(u) = \begin{cases} 0 \text{ if } u(x,y) \geq -c(x,y) \\ +\infty \text{ o.w.} \end{cases} \qquad \zeta(u) = \begin{cases} \int_X \varphi \,\mathrm{d}\mu(x) + \int_Y \psi \,\mathrm{d}\nu(y) \text{ if } u(x,y) = \varphi(x) + \psi(y) \\ +\infty \text{ o.w.} \end{cases}$$

We note that $-\inf_\Gamma\{\xi + \zeta\} = \text{(DP)}$, hence we compute the convex conjugates of ξ and ζ in order to apply Fenchel–Rockafellar duality. Let $\eta \in \mathscr{R}(X \times Y)$ be arbitrary, then

$$\begin{aligned} \xi^*(-\eta) &= \sup_{u\in\Gamma}\left\{-\int_{X\times Y} u(x,y)\,\mathrm{d}\eta : u(x,y) \geq -c(x,y)\right\} \\ &= \sup_{u\in\Gamma}\left\{\int_{X\times Y} u(x,y)\,\mathrm{d}\eta : u(x,y) \leq c(x,y)\right\}. \end{aligned}$$

Remark that if η is negative then $\exists\, u \in \Gamma$ nonpositive s.t. $\int_{X\times Y} u \,\mathrm{d}\eta > 0$ then, taking $u_n = nu$ and taking the limit as $n \to \infty$ yields $\xi^*(-\eta) = \infty$. If η is nonnegative then $\xi^*(-\eta) = \int_{X\times Y} c\,\mathrm{d}\eta$ due to the constraint on u. Hence,

$$\xi^*(-\eta) = \begin{cases} \int_{X\times Y} c\,\mathrm{d}\eta \text{ if } \eta \in \mathscr{R}_+(X \times Y) \\ +\infty \text{ o.w.} \end{cases}$$

Where $\mathscr{R}_+(X \times Y)$ denotes the set of all nonnegative Radon measures. Similarly,

$$\begin{aligned} \zeta^*(\eta) &= \sup_{u\in\Gamma}\left\{\int_{X\times Y} u(x,y)\,\mathrm{d}\eta - \int_X \varphi(x)\,\mathrm{d}\mu(x) - \int_Y \psi(y)\,\mathrm{d}\nu(y) : u(x,y) = \varphi(x) + \psi(y)\right\} \\ &= \sup_{u\in\Gamma}\left\{\int_{X\times Y} (\varphi(x) + \psi(y))\,\mathrm{d}\eta - \int_X \varphi(x)\,\mathrm{d}\mu(x) - \int_Y \psi(y)\,\mathrm{d}\nu(y)\right\} \end{aligned}$$

Notably, if η is such that $\int_{X\times Y}(\varphi(x) + \psi(y))\,\mathrm{d}\eta \neq \int_X \varphi(x)\,\mathrm{d}\mu(x) + \int_Y \psi(y)\,\mathrm{d}\nu(y)$ for some $\varphi(x), \psi(y)$ then by rescaling, $\zeta^*(\eta) = \infty$. Hence,

$$\zeta^*(\eta) = \begin{cases} 0 \text{ if } \int_{X\times Y}(\varphi(x) + \psi(y))\,\mathrm{d}\eta = \int_X \varphi(x)\,\mathrm{d}\mu(x) + \int_Y \psi(y)\,\mathrm{d}\nu(y)\ \forall\, \varphi, \psi \\ +\infty \text{ o.w.} \end{cases}$$

and therefore by Fenchel–Rockafellar duality, $\inf_\Gamma\{\xi+\zeta\} = \max_{\mathscr{R}}\{-\xi^*(-\eta)-\zeta^*(\eta)\}$ and (DP)$=-\inf_\Gamma\{\xi+\zeta\} = -\max_{\mathscr{R}}\{-\xi^*(-\eta)-\zeta^*(\eta)\} = \min_{\mathscr{R}}\{\xi^*(-\eta)+\zeta^*(\eta)\}$ $= \min_{\Pi(\mu,\nu)}\{\int_{X\times Y} c\,\mathrm{d}\eta\} =$ (KP) where we recall that $\Pi(\mu,\nu)$ denotes the set of all probability measures admitting μ and ν as marginals on X and Y respectively and notably there exists a solution to (KP). □

A more careful approach using this argument is found in [17, Thm. 1.3 p.19] allows us t generalize this result to the case where X and Y are Polish spaces.

2.2. **Returning to the Monge–Ampère Equation.** We now derive a characterization of optimal transport maps in terms of a Jacobian equation.

Remark 2.7. (Associated Jacobian Equation) Let $X, Y \subseteq \mathbb{R}^n$, $\mu(x) = f(x)\mathrm{d}x$, $\nu(y) = g(y)\mathrm{d}y$, $T : X \to Y$ a sufficiently smooth transport map and $\chi \in \mathscr{D}(\mathbb{R}^n)$. Then the condition $T_\sharp\mu = \nu$ yields

$$\int \chi(T(x))\mathrm{d}\mu(x) = \int \chi(y)\mathrm{d}\mu(T^{-1}(x)) = \int \chi(y)\mathrm{d}\nu(y).$$

So inputting the explicit form of μ and ν, we have

$$\int \chi(T(x))f(x)\mathrm{d}x = \int \chi(y)g(y)\mathrm{d}y.$$

Taking the change of variables $y = T(x)$ on the right hand side yields

$$\int \chi(T(x))f(x)\mathrm{d}x = \int \chi(T(x))g(T(x))|\det(DT(x))|\mathrm{d}x.$$

Here D denotes the Jacobian of $T(x)$. Since χ was chosen to be arbitrary we recover the following Jacobian equation,

$$f(x) = g(T(x))|\det(DT(x))| \text{ a.e. in } X.$$

We note that the above Jacobian equation is close to the form of a Monge–Ampère equation, but we must still do some work to recover it.

Definition 2.8. (c–convexity and c–transform) $\psi : X \to \mathbb{R}\cup\{+\infty\}$ is *c–convex* if

$$\psi(x) = \sup_{y\in Y}\{\psi^c(y) - c(x,y)\}, \tag{2.1}$$

where $\psi^c : Y \to \mathbb{R}\cup\{-\infty\}$ is the c–transform of ψ given by

$$\psi^c(y) = \inf_{x\in X}\{\psi(x) + c(x,y)\}.$$

With this definition in mind we can establish an equivalent problem to (DP) in terms of a c–convex function and its c–transform.

Remark 2.9. (Alternative formulation of the Dual Problem)

The dual problem can be rewritten as a minimization problem over c–convex functions, notably.

$$\sup\left\{\int_X \varphi\,\mathrm{d}\mu(x)+\int_Y \psi\,\mathrm{d}\nu(y):\varphi\in C_b(X),\ \psi\in C_b(Y)\text{ and }\varphi+\psi\le c\right\}=$$
$$\sup\left\{-\int_X \alpha\,\mathrm{d}\mu(x)+\int_Y \beta\,\mathrm{d}\nu(y):\alpha\in C_b(X),\ \beta\in C_b(Y)\text{ and }\beta-\alpha\le c\right\}=$$
$$\sup\left\{-\int_X \alpha\,\mathrm{d}\mu(x)+\int_Y \beta\,\mathrm{d}\nu(y):\alpha\in C_b(X),\ \beta\in C_b(Y)\text{ and }\beta\le c+\alpha\right\}=$$
$$\sup\left\{-\int_X \alpha\,\mathrm{d}\mu(x)+\int_Y \alpha^c\,\mathrm{d}\nu(y):\alpha\in C_b(X)\text{ and }\alpha^c-\alpha\le c\right\}.$$

In particular, by the duality result, we have that for the optimal transport plan γ, $\int_{X\times Y} c\mathrm{d}\gamma = -\int_X \alpha\,\mathrm{d}\mu(x)+\int_Y \alpha^c\,\mathrm{d}\nu(y)=\int_{X\times Y}(\alpha^c(y)-\alpha(x))\,\mathrm{d}\gamma$ such that $c(x,y)=\alpha^c(y)-\alpha(x)$ γ–almost surely for $(x,y)\in\operatorname{supp}(\gamma)$. Thus without loss of generality, the dual problem can be taken to be the supremum over φ c–convex and its c–transform.

We now take the opportunity to introduce some conditions on the cost function that will be useful in the sequel.

Remark 2.10. (Conditions on Cost) Let $X,Y\subseteq\mathbb{R}^n$ we enumerate some conditions that will be useful in the analysis of general cost functions

- (C0) The cost function $c: X\times Y\to\mathbb{R}$ is C^4 and $||c||_{C^4(X\times Y)}<\infty$.
- (C1) $\forall\, x\in X,\ Y\ni y\mapsto D_x c(x,y)\in\mathbb{R}^n$ is injective. Also known as the twist condition in the optimal transport literature.
- (C2) $\forall\, y\in Y,\ X\ni x\mapsto D_y c(x,y)\in\mathbb{R}^n$ is injective.
- (C3) $\det(D_{xy}c)(x,y)\neq 0\,\forall\,(x,y)\in X\times Y$.

Definition 2.11. (c-exponential) If c satisfies (C0)–(C2) then $\forall\, x\in X, y\in Y, p\in\mathbb{R}^n$ we define the *c–exponential map* as

$$\text{c–}\exp_x(p)=y\iff p=-D_x c(x,y) \tag{2.2}$$

Theorem 2.12. *(Characterization of Optimal Transport Maps for General Costs) Suppose $c: X\times Y\to\mathbb{R}$ satisfies (C0)-(C1) and f, g are two positive probability densities on X and Y , two open and bounded sets respectively, then $\exists\ u: X\to\mathbb{R}$ c–convex such that the unique optimal transport map sending f onto g us given by $T(x)=c\text{–}\exp(\nabla u(x))$. Moreover, if (C2) holds then T is injective $f dx$ a.e, $|\det(DT(x))|=\frac{f(x)}{g(T(x))}$ and the inverse of T is the unique optimal transport map sending g onto f.*

Proof. From the above remark, if γ denotes the optimal transport plan then there is a c–convex function φ such that for any $(x_0,y_0)\in\operatorname{supp}(\gamma)$, the function $\varphi(x)+c(x,y_0)$ attains a minimum at x_0 (as $\varphi^c(y_0)=\varphi(x)+c(x,y_0)$). Notably, φ is the supremum of the family of uniformly Lipschitz functions $c(\cdot,y)+\lambda_y$ and hence is Lipschitz, yielding differentiability almost everywhere. Thus $\nabla\varphi(x_0)+D_x c(x_0,y_0)=0 \implies \nabla\varphi(x_0)=-D_x c(x_0,y_0)\implies$ $\text{c–}\exp_x(\nabla\varphi(x_0))=y_0$ and notably, we have constructed a representation for an optimal transport map taking x_0 to y_0. Hence, $T(x)=\text{c–}\exp_x(\nabla\varphi(x))$ is an optimal transport map as was desired. Suppose $\exists$ another optimal transport map $\tilde{T}(x)$, we noted before that $\gamma=(\mathrm{Id},T)_\sharp\mu$ and $\tilde{\gamma}=(\mathrm{Id},\tilde{T})_\sharp\mu$ are the corresponding optimal transport plans, letting μ denote $f dx$. By linearity of $\Pi(\mu,\nu)$ we have that $\frac{1}{2}(\gamma+\tilde{\gamma})$ is also optimal and is notably concentrated on the same graph as γ and $\tilde{\gamma}$ thus $\gamma=\tilde{\gamma}$ $f dx$ almost everywhere. Whence, the Jacobian equation yields $|\det(DT(x))| = \frac{f(x)}{g(T(x))}$ and T is invertible provided (C2) is satisfied due to the uniqueness of minimizers in the optimal transport problem. □

2.3. **The Quadratic Cost on $\mathbb{R}^d$.** We will now study a specific cost function that admits some straightforward regularity results, namely the case of the quadratic cost $c(x,y) = \frac{|x-y|^2}{2}$.

We first wish to write the quadratic cost in an alternative form that will be relevant in the sequel.

Remark 2.13. (Equivalent Cost) Let S denote a transport map taking μ onto ν for μ, ν probability measures on $\mathbb{R}^d$. Then, the condition $S_\sharp\mu = \nu$ yields

$$\int_{\mathbb{R}^n} \frac{|S(x)|^2}{2} \mathrm{d}\mu(x) = \int_{\mathbb{R}^n} \frac{|y|^2}{2} \mathrm{d}\nu(y).$$

We now expand the expression for quadratic cost at $y = S(x)$,

$$\int_{\mathbb{R}^n} \frac{|x - S(x)|^2}{2} \mathrm{d}\mu(x) = \int_{\mathbb{R}^n} \frac{|x|^2}{2} \mathrm{d}\mu(x) + \int_{\mathbb{R}^n} \frac{|y|^2}{2} \mathrm{d}\nu(y) - \int_{\mathbb{R}^n} (x \cdot S(x)) \mathrm{d}\mu(x).$$

We note that the only term on the right hand side that depends on the transport map chosen is the rightmose one, hence the folloring minimization problems are euquivalent:

$$\min_{S_\sharp\mu=\nu} \int_{\mathbb{R}^n} \frac{|x - S(x)|^2}{2} \mathrm{d}\mu(x) \quad \text{and} \quad \min_{S_\sharp\mu=\nu} \int_{\mathbb{R}^n} (-x \cdot S(x)) \mathrm{d}\mu(x).$$

All in all, we have that the quadratic cost $|x-y|/2$ is equivalent to the cost $c(x,y) = -x \cdot y$.

Theorem 2.14. *(Existence and Uniqueness of Optimal Transport Maps for the Quadratic Cost) Let μ, ν be compactly supported probability measures on $\mathbb{R}^n$ and $c = |x-y|^2/2$. Suppose μ is absolutely continuous with respect to the Lebesgue measure then $\exists$ a unique optimal transport map T from μ to ν of the form $T = \nabla u$ for u convex. Moreover, if $\mu(dx) = f dx$ and $\nu(dy) = g dy$ then T is differentiable μ–a.e. and*

$$|\det(DT(x))| = \frac{f(x)}{g(T(x))} \quad \text{for } \mu\text{–a.e. } x \in \mathbb{R}^n. \tag{2.3}$$

Proof. All that is required is to compute c–exp$(\nabla u(x))$ and to apply Theorem 2.12 to get the desired result. As noted above, the quadratic cost is equivalent to $-x \cdot y$ and one has that $-D_x(-x \cdot y) = y$ hence c–exp$_x$ is simply the identity map, thus $T = \nabla u$. Whence we get the desired results since c–convexity is equivalent to convexity for the quadratic cost, as can be seen from the definition (with cost $-x \cdot y$, c–convex functions are defined as a supremum of linear functions and are thus convex). □

We now recall a useful property of convex functions and refer to [18, Thm. 14.25 p.402] for the proof.

Theorem 2.15. *(Differentiability of Convex Functions) Let $\Omega \subseteq \mathbb{R}^n$ be convex and open and $u : \Omega \to \mathbb{R}$ be convex. Then for a.e. $x \in \Omega$, u is differentiable at x and $\exists$ a symmetric matrix $D^2u(x)$ such that*

$$u(y) = u(x) + \nabla u(x) \cdot (y - x) + \frac{1}{2} D^2u(x)(y - x) \cdot (y - x) + o(|y - x|^2)$$

and at such points, ∇u is differentiable

$$\nabla u(y) = \nabla u(x) + D^2u(x) \cdot (y - x) + o(|y - x|) \quad \forall\, y \in \mathrm{dom}(\nabla u)$$

Having discussed the relevant preliminary results we can now examine the Monge–Ampère equation that arises in the context of optimal transport and discuss regularity in this context.

Definition 2.16. (Brenier Solution to the Monge–Ampère Equation) Let $X, Y \subseteq \mathbb{R}^n$ be bounded, smooth and open, $\mu(x) = f(x)\mathrm{d}x$, $\nu(y) = g(y)\mathrm{d}y$ be probability measures with $\operatorname{supp} f \subseteq X$ and $\operatorname{supp} g \subseteq Y$. Moreover, let $f > 0$ on X and $g < \infty$ on Y. Then Theorem 2.14 implies $T(x) = \nabla u(x)$ is the unique optimal transport map of μ onto ν for the quadratic cost where u is convex, hence substituting $T(x)$ into (2.3) yields

$$|\det D^2 u(x)| = \frac{f(x)}{g(\nabla u(x))} \quad f\mathrm{d}x\text{–a.e.} \tag{2.4}$$

which is well defined almost everywhere since by Theorem 2.15, ∇u is differentiable a.e. and due to the bounds on f and g. Therefore this is a Monge–Ampère type equation with boundary conditions $\nabla u(X) = T(X) = Y$ induced by the condition $T_\sharp \mu = \nu$. A function u satisfying (2.4) is called a *Brenier solution.*

We note that a convex function u is an Alexandrov solution to (2.4) if $\partial u(x) = \frac{f(x)}{g(\nabla u(x))}$ in measure. In particular, $\forall\ E \subseteq X$ Borel we have that $|\partial u(E)| = \int_E \frac{f(x)}{g(\nabla u(x))}\mathrm{d}x$. Hence $\partial u(x)$ preserves Lebesgue measure up to some multiplicative constant depending on u, f, g and E and we denote this property $|E| \sim |\partial u(E)|$.

We wish to determine when Brenier solutions will exhibit the same regularity properties as their Alexandrov counterparts.

Note that a Brenier solution u is an Alexandrov solution for a.e. $x \in X$ by definition, hence the control offered by Alexandrov solutions is mimicked only by Brenier solutions at points where u is twice differentiable. It is therefore possible that the Monge–Ampère measure may still exhibit some singular behaviors. In particular, since $\partial u(X) \supseteq \nabla u(X) = Y$, we can only guarantee that $|E| \sim |\partial u(E) \cap Y|$ and thus we don't have full control of the Monge–Ampère measure of u.

However, if $\partial u(E) \subseteq Y$ we have the same type of control offered by Alexandrov solutions and the notion of solution is identical, thus we can utilize the regularity results established for Alexandrov solutions. To this effect, the following theorem.

Theorem 2.17. *(Regularity Results for Quadratic Cost) Let $X, Y \subseteq \mathbb{R}^n$ be bounded and open, let $f : X \to \mathbb{R}^+$ and $g : Y \to \mathbb{R}^+$ be probability densities which are bounded away from 0 and ∞ on X and Y respectively. Let $T = \nabla u : X \to Y$ be the unique optimal transport map sending f onto g for the quadratic cost. Let Y be convex then we have*

(i) $T \in C^{0,\alpha}_{loc}(X) \cap W^{1,1+\epsilon}_{loc}(X)$;
(ii) *If* $f \in C^{k,\beta}_{loc}(X)$ *and* $g \in C^{k,\beta}_{loc}(Y)$ *for some* $\beta \in (0,1)$ *then* $T \in C^{k+1,\beta}_{loc}(X)$;
(iii) *If* $f \in C^{k,\beta}(\bar{X}), g \in C^{k,\beta}(\bar{Y})$ *and* X, Y *are smooth and uniformly convex then* $T : \bar{X} \to \bar{Y} \in C^{K+1,\beta}(\bar{X})$ *and is a diffeomorphism.*

We precede the proof of this theorem by the following definition of a construction that will be useful only in the proof.

Definition 2.18. (Convex Hull) Given $V \subseteq \mathbb{R}^n$ the convex hull of V denoted $\operatorname{conv}(V)$ is given by $\cap\{U \text{ convex} : V \subseteq U\}$.

Proof. (Theorem 2.17) The proof hinges on the property discussed in the preliminary notes to this theorem. Notably, we will show that for Y convex we have that $\forall\ E \subseteq X\ \partial u(E) \subseteq Y$ and therefore that Brenier solutions are also Alexandrov.

By definition of the subdifferential we have that $\forall\, A \subseteq X$, $\partial u(A) \supseteq \nabla u(A \cap \mathrm{dom}(\nabla u))$. Whence, by the area formula (c.f. [9, Cor. 3.2.20 p.256]),

$$\mu_u(A) = |\partial u(A)| \geq |\nabla u(A \cap \mathrm{dom}(\nabla u))| = \int_A \mathrm{det} D^2 u(x) \mathrm{d}x = \int_A \frac{f(x)}{g(\nabla u(x))} \mathrm{d}x$$

Next, if $\partial u(A) \subseteq Y$ up to a set of measure 0, we have that $A \cap \mathrm{dom}(\nabla u) \subseteq (\nabla u)^{-1}(\partial u(A))$ since $\nabla u(A) \subseteq \partial u(A)$ by definition. Moreover,

$$\begin{aligned}
&(\nabla u)^{-1}(\partial u(A) \cap Y) \backslash A \\
&\subseteq (\nabla u)^{-1} \left(\{ y \in Y \;:\; \exists\, x_1 \neq x_2 \;:\; y \in \partial u(x_1) \cap \partial u(x_2) \}\right) \\
&\subseteq (\nabla u)^{-1} \left(\{ \text{points s.t. } u^* \text{ is not differentiable} \cap Y \}\right)
\end{aligned}$$

The first inclusion follows from the fact that if $z \in \nabla u)^{-1}(\partial u(A) \cap Y)$, then $\exists\, \bar{z}$ such that u is differentiable at $\bar{z}$ and $\nabla u(\bar{z}) \in \partial u(A) \cap Y$, so notably $(\nabla u)^{-1} u(A) \cap Y \backslash A$ is simply the set $\Gamma = \{ z \notin A : u \text{ is differentiable at } z \text{ and } \nabla u(z) \in \partial u(A) \cap Y \}$ so this set does not include any point in A where u is differentiable. Therefore, $\forall\, x_1 \in \Gamma$, $\nabla u(x_1) \in \partial u(x_1) \cap \partial u(x_2)$ for $x_2 \in A$ and hence $x_1 \neq x_2$ since $x_1 \notin A$. As such the first inclusion holds.

The second inclusion can be deduced by letting $y \in \partial u(x_1) \cap \partial u(x_2)$. Recalling the definition of the subdifferential, we have that $\partial u(x_1) = \{ p \in \mathbb{R}^n : \langle p, x_1 \rangle - u(x_1) \geq \langle p, x \rangle - u(x) \}$ for any x. We note therefore that $u^*(y) = \langle x_i, y \rangle - u(x_i)$ for $i = 1, 2$ and that $\forall\, z$, $u^*(z) \geq \langle x_i, z \rangle - u(x_i)$ by definition of the convex conjugate. So, in particular $u^*(z) \geq \langle x_i, z \rangle + \langle x_i, y - y \rangle - u(x_i) = u^*(y) + \langle x_i, z - y \rangle$, so if u^* is differentiable, its derivative is multiply defined by x_1 and x_2 which is absurd, whence the second inclusion.

Note that $|(\nabla u)^{-1}(\partial u(A) \cap Y) \backslash A| = 0$ since u^* is convex [16, Line 9 p.104] i.e. hence differentiable almost everywhere and by the fact that $(\nabla u)_\sharp (f \mathrm{d}x) = g \mathrm{d}y$ we have that the last included set is of measure 0 Thus the first set in question is a subset of a set of measure 0 and hence is of measure 0.

Assume $\partial u(A) \subseteq Y$ up to a set of measure 0 and f vanishes outside X then

$$|\partial u(A)| = \int_{\partial u(A)} \frac{g(y)}{g(y)} \mathrm{d}y = \int_{(\nabla u)^{-1}(\partial u(A))} \frac{f(x)}{g(\nabla u(x))} \mathrm{d}x \tag{2.5}$$

since $(\nabla u)_\sharp (f \mathrm{d}x) = g \mathrm{d}y$. Using the aforementioned property $A \cap \mathrm{dom}(\nabla u) \subseteq (\nabla u)^{-1}(\partial u(A))$ we have up to excision of a set of measure 0 that $(\nabla u)^{-1}(\partial u(A)) = A \cup (\nabla u)^{-1}(\partial u(A)) \backslash A$. Thus $(\nabla u)^{-1}(\partial u(A))$ is the union of a measurable set and a set of mesure 0, so

$$|\partial u(A)| = \int_A \frac{f(x)}{g(\nabla u(x))} \mathrm{d}x.$$

Note that $\{ x \in X \cap \mathrm{dom}(\nabla u) : \nabla u(x) \in Y \}$ is dense in X by the property that $\nabla u(X) = Y$. Moreover, we have the following result for convex functions per [7, Thm. 3.3.6 p.59]

$$\partial u(x) = \mathrm{Conv} \left(\{ p \;:\; \exists\, x_k \in \mathrm{dom}\, \nabla u \text{ with } x_k \to x \text{ and } \nabla u(x_k) \to p \} \right).$$

Hence, by density, $\partial u(X) \subseteq \overline{\mathrm{Conv}(Y)} = \bar{Y}$ by convexity of Y ($\mathrm{conv}(U) = U$ for U convex).

Now, since ∂Y has 0 measure we are in the scenario to apply (2.3) yielding

$$|\partial u(X)| = \int_X \frac{f(x)}{g(\nabla u(x))} \mathrm{d}x$$

Therefore, u is an Alexandrov solution. We note moreover that u is actually strictly convex in X by [5, Lem. 3 p.102] such that Theorem 1.20 and Theorem 1.28 yield (i) and Theorem

1.26 yields (ii), recalling that $T(x) = \nabla u(x)$ and that $\frac{f(x)}{g(\nabla u(x))} \in C^{k,\beta}_{\text{loc}}(X)$ if $f \in C^{k,\beta}_{\text{loc}}(X)$, $g \in C^{k,\beta}_{\text{loc}}(Y)$ and the bounds on f and g are satisfied. The boundary regularity is discussed in [6]. $\square$

The above theorem guarantees regularity of solutions provided Y is convex. This condition is however rather restrictive and we wish to discuss what results can be derived if convexity is not assumed.

Theorem 2.19. *(Regularity Without Convexity on Target) Let X, Y, f, g and T be as above save for the convexity assumption on Y. Then $\exists$ two relatively closed sets of measure zero $\Sigma_X \subseteq X, \Sigma_Y \subseteq Y$ such that $T : X\backslash\Sigma_X \to Y\backslash\Sigma_Y$ is a $C^{0,\alpha}_{loc}$ diffeomorphism for $\alpha > 0$. If $c \in C^{K+2,\alpha}_{loc}(X \times Y), f \in C^{k,\alpha}_{loc}(X)$, and $g \in C^{k,\alpha}_{loc}(Y)$ for some $k \geq 0$ and $\alpha \in (0, 1)$, then $T : X\backslash\Sigma_X \to Y\backslash\Sigma_Y$ is a $C^{k+1,\alpha}_{loc}$ diffeomorphism.*

Proof. If Y is not convex, there may exist $x \in X$ such that $\partial u(x) \not\subseteq Y$ and as mentioned before, at such points there is no control on the Monge–Ampère measure of u. We define the regular set $\text{Reg}_X = \{x \in X : \partial u(x) \subseteq Y\}$ and let $\Sigma_X = X\backslash\text{Reg}_X$. By the continuity properties of the subdifferential [7, Prop. 3.3.4 p.57] and the fact that Y is open, Reg_X is also open. Thus, the condition $(\nabla u(x))_\sharp(f\mathrm{d}x) = g\mathrm{d}y$ yields $\nabla u(x) \in Y$ for a.e. $x \in X$, that is $|\Sigma_X| = 0$. Then, following the proof of the previous theorem, u is a strictly convex solution on Reg_X and the previous regularity results apply. $\square$

Throughout this section, we have discussed the basics of optimal transportation theory and have made the connection between the case of quadratic cost and the Monge–Ampère equation.

3. A Class of Monge–Ampère Type Equations

In the previous section, we determined that in the case of the quadratic cost, the Monge–Ampère equation occurs naturally and guarantees some regularity properties on the corresponding optimal transport maps. This result was demonstrated in Theorem 2.14 and we have discussed a similar result in Theorem 2.12 for the case of more general cost functions $c(x,y)$ satisfying (C0)– (C2). Hence, for such a cost function we have that

$$|\det(DT(x))| = \frac{f(x)}{g(T(x))} \quad \text{a.e.},$$

and that $T(x) = \text{c–exp}_x(\nabla u(x))$. By the definition of $c\text{–exp}_x$, $z \mapsto u(z)+c(z, \text{c–exp}_x(\nabla u(x)))$ attains a minimum at $z = x$ for every point of x where u is differentiable. Hence, if u is twice differentiable at x,

$$D^2u(x) + D_{xx}c(x, \text{c–exp}_x(\nabla u(x))) \geq 0.$$

Therefore, recalling that $T(x) = \text{c–exp}_x(\nabla u(x)) \implies -D_x c(x, T(x)) = \nabla u(x)$ yields $-D_{xx}c(x, T(x)) - D_{yx}c(x, T(x))DT(x) = D^2u(x)$ and taking the determinant on both sides, we get

$$\begin{aligned} &\det(D^2u(x) + D_{xx}c(x, \text{c–exp}_x(\nabla u(x)))) \\ &\quad = |\det(D_{xy}c(x, \text{c–exp}_x(\nabla u(x))))| \frac{f(x)}{g(\text{c–exp}_x(\nabla u(x)))}. \end{aligned}$$

which we recognize as a Monge–Ampère type equation of the form

$$\det(D^2u - \mathcal{A}(x, \nabla u)) = f(x, \nabla u) \tag{3.1}$$

with $\mathcal{A}(x, \nabla u(x)) = -D_{xx}c(x, \text{c–exp}_x(\nabla u(x)))$.

The regularity of optimal transport maps of this form is dependent on the properties of $\mathcal{A}$.

3.1. The MTW Condition and Smooth Solutions.

A major breakthrough in regularity theory for this type of problem was demonstrated in [13] and appears in the form of the following tensor.

Definition 3.1. (Ma–Trudinger–Wang (MTW) Tensor) The *MTW tensor* $\mathfrak{S}_{(x,y)}(\xi, \eta)$ is defined as

$$\begin{aligned} \mathfrak{S}_{(x,y)}(\xi, \eta) &= D^2_{p_\eta p_\eta} \mathcal{A}(x, p)[\xi, \xi] \\ &= \sum_{i,j,k,l,p,q,r,s} (c_{ij,p}c^{p,q}c_{q,rs} - c_{ij,rs})c^{r,k}c^{s,l}\xi^i\xi^j\eta^k\eta^l \quad \xi, \eta \in \mathbb{R}^n. \end{aligned}$$

Here, $\mathcal{A}$ is identical to that in (3.1). Moreover, the cost function c in all of this expression is evaluated at $(x, y) = (x, \text{c–}\exp_x(p))$ and we have used the following convention for brevity, $c_j = \partial_{x^j}c$, $c_{jk} = \partial_{x^jx^k}c$, $_{i,j} = \partial_{x^iy^j}c$, $c^{i,j} = \partial_{x^iy^j}c$, $c^{i,j} = (c_{i,j})^{-1}$.

The following condition will play a pivotal role in the regularity theory for these general costs.

Definition 3.2. (MTW Condition) Given $K \geq 0$, the cost function c satisfies the MTW(K) condition provided $\forall\,(x, y) \in (X \times Y)$ and $\forall\,\xi, \eta \in \mathbb{R}^n$,

$$\mathfrak{S}_{(x,y)}(\xi, \eta) \geq K|\xi|^2|\eta|^2 \quad \text{whenever } \xi \perp \eta \tag{3.2}$$

The MTW condition may appear unremarkable for the moment, however the following result is a testament to its utility.

Theorem 3.3. *(Smoothness from the MTW Condition) Let $c : X \times Y \to \mathbb{R}$ satisfy (C0)–(C3) and MTW(K) holds for $K > 0$. Moreover, let f, g be smooth and bounded away from $0, \infty$ on their respective supports X, Y. Also, suppose*

(i) *X and Y are smooth;*
(ii) *$D_x c(x, Y)$ is uniformly convex $\forall\, x \in X$;*
(iii) *$D_y c(X, y)$ is uniformly convex $\forall\, y \in Y$;*

Then the optimal transport map sending f onto g is of the form $T(x) = c\text{–}\exp_x(\nabla u(x))$ for $u \in C^\infty(\bar{X})$ and $T : \bar{X} \to \bar{Y}$ is a smooth diffeomorphism.

Proof. Assume $u \in C^4(X)$ is a solution of (3.1) with $T(x) = \text{c–}\exp_x(\nabla u(x))$ and the natural boundary conditions $T(X) = Y$. Hence, $|\nabla u(x)| = |D_x c(x, T(x))| \le C$, so u is globally Lipschitz. Let $w_{ij} = D_{x^i x^j} u + D_{x^i x^j} c(x, \text{c–}\exp_x(\nabla u(x)))$ which is positive definite by c–convexity of u. Thus, (3.1) simplifies to $\det(w_{ij}) = f(x, \nabla u(x))$. It will be simpler to work with $\log\det(w_{ij})$ as notably $\partial_{x^k} \log\det(w_{ij}) = \frac{\text{adj}(w)_{ji}}{\det(w_{ij})} w_{ij,k} = w^{ij} w_{ij,k}$ via the Jacobi identity and we have again used the convention $w^{ij} = (w_{ij})^{-1}$ and $w_{ij,k} = \partial_{x^k} w_{ij}$. Hence, letting $\varphi(x) = f(x, \nabla u(x))$ yields $w^{ij} w_{ij,k} = \varphi_k$. Taking another derivative yields $w^{ij} w_{ij,kk} - w^{is} w^{jt} w_{ij,k} w_{st,k} = \varphi_{kk}$ or $w^{ij} w_{ij,kk} = \varphi_{kk} + w^{st} w^{ij} w_{ij,k} w_{st,k} \ge \varphi_{kk}$ so, inputting the explicit form of w_{ij} yields

$$w^{ij}(u_{ijk} + c_{ijk} + c_{ij,s} T_{s,k}) = \varphi_k$$
$$w^{ij}(u_{ijkk} + c_{ijkk} + 2c_{ijk,s} T_{s,k} + c_{ij,s} T_{s,kk} + c_{ij,st} T_{s,k} T_{t,k}) \ge \varphi_{kk}$$

Now, we take $\bar{x} \in X$ and let η be a cut–off function around $\bar{x}$, define $G : X \times \mathbb{S}^{n-1} \to \mathbb{R}$ $(x, \xi) \mapsto \eta(x)^2 w_{\xi\xi}$ for $w_{\xi\xi} = w_{ij}\xi^i \xi^j$. Note that w_{ij} is positive definite hence $G > 0$. We now wish to demonstrate an upper bound on G and employ the same strategy as in Theorem 1.15.

Let $x_0 \in X$ and $\xi_0 \in \mathbb{S}^{n-1}$ be a point where G is maximal. We take rotation of coordinates with $\xi_0 = e_1$ such that at x_0, $0 = \log(G)_i = \frac{w_{11,i}}{w_{11}} + 2\frac{\eta_i}{\eta}$ and $\log(G)_{ij} = \frac{w_{11,ij}}{w_{11}} + 2\frac{\eta_{ij}}{\eta} - 6\frac{\eta_i \eta_j}{\eta}$. Note that $\log G$ is nonpositive, hence $0 \ge w_{11} w^{ij} (\log G)_{ij}$ and differentiating $\nabla u = -D_x c(x, T(x))$ yields $w_{ij} c^{k,i} = T_{k,j}$ notably $|\nabla T| \le C w_{11}$.

Combining these equations yields $0 \ge w^{ij}[c^{k,l} c_{ij,k} c_{l,st} - c_{ij,st}] c^{s,p} c^{t,q} w_{p1} w_{q1} - C$. One can take a rotation of coordinates that leaves e_1 invariant and assume that w_{ij} is diagonal at x_0 to get $w^{ii}[c^{k,l} c_{ii,k} c_{l,st} - c_{ii,st}] c^{s,1} c^{t,1} w_{11} w_{11} \le C$. Now, apply MTW(K) to $\xi_1 = (0, \sqrt{w^{22}}, \ldots, \sqrt{w^{nn}})$ and $\xi_2 = (w_{11}, 0, \ldots, 0)$ yielding $K w_{11}^2 \sum_{i=2}^n w^{ii} \le C + w^{11}[c^{k,l} c_{11,k} c_{l,st} - c_{11,st}] c^{s,1} c^{t,1} w_{11} w_{11}$, so using (C0) and the fact that $w^{ij} = w_{ij}^{-1}$ gives $w_{11}^2 \sum_{i=2}^n w^{ii} \le C(1 + w_{11})$ such that the AM-GM inequality yields

$$\frac{1}{n-1}\sum_{i=2}^n w^{ii} \ge \left(\prod_{i=2}^n w^{ii}\right) \ge c_0 (w^{11})^{-1/(n-1)} = c_0 (w_{11})^{1/(n-1)}$$

for $c_0 = \inf_{x \in X} h(x, \nabla u(x))^{-1/(n-1)} > 0$. We therefore obtain $w_{11}(x_0)^{2+1/(n-1)} \le C(1 + w_{11}(x_0))$ hence $G(x, \xi) \le G(x_0, \xi_0) \le C$ for any $(x, \xi) \in X \times \mathbb{S}^{n-1}$. □

We note that this result can be used to proved boundary regularity for diffent Monge–Ampère type equations.

Remark 3.4. (Boundary Regularity Results for MTW Condition) It has been shown in [13] that the MTW condition is coordinate invariant. Hence if u solves $\det(D^2 u - \mathcal{A}(x, \nabla u)) =$

$f(x, u, \nabla u)$ with $\mathcal{A}$ satisfying the MTW condition then taking Φ to be a smooth diffeomorphism, $u \circ \Phi$ satisfies the same equation with $\tilde{\mathcal{A}}$ satisfying the MTW condition in place of $\mathcal{A}$. Therefore to prove boundary regularity it is possible to simplify the problem using such a diffeomorphism. This result also holds for a solution to the classical Monge–Ampère equation ($\mathcal{A} = 0$ is MTW(0)) say u, then $u \circ \Phi$ satisfies the a similar equation with $\tilde{\mathcal{A}}$ satisfying MTW(0) in lieu of $\mathcal{A}$.

3.2. **Geometry, Regularity and the MTW condition.** The geometric interpretation of the MTW condition stems from the work of Loeper (c.f. [12]). Recall that the convexity of the subdifferential of a convex function was one of the main results used to prove the regularity results for the classical Monge–Ampère equation. Moreover, recall that convex sets are connected and thus a natural extension in this case is to study whether a modification of the subdifferential for c–convex function is connected. To this effect, the following definitions.

Definition 3.5. (c–segment) Let $\bar{x} \in X$, y_0, $y_1 \in Y$. Then the *c–segment* from y_0 to y_1 with base $\bar{x}$ is given by

$$[y_0, y_1]_{\bar{x}} = \{y_t = \text{c–exp}_{\bar{x}}((1-t)(\text{c–exp}_{\bar{x}})^{-1}(y_0) + t(\text{c–exp}_{\bar{x}})^{-1}(y_1)) \mid t \in [0, 1]\} \tag{3.3}$$

Definition 3.6. (c–subdifferential) For a c–convex function ψ its *c–subdifferential* at x is

$$\partial_c \psi(x) = \{y \in Y : \psi(x) = \psi^c(y) - c(x, y)\}$$

or equivalently

$$\partial_c \psi(x) = \left\{y \in \bar{Y} : \psi(z) \geq -c(z, y) + c(x, y) + \psi(x) \quad \forall\, x \in X\right\}$$

Definition 3.7. (Fréchet Subdifferential) The *Fréchet subdifferential* of ψ at x is given by

$$\partial^- \psi(x) = \{p \in \mathbb{R}^n : \psi(z) \geq u(x) + p \cdot (z - x) + o(|z - x|)\} \tag{3.4}$$

Notably, for $\psi \in C^1$ if $y \in \partial_c \psi(x)$ then $-D_x c(x, y) \in \partial^- c(x, y)$ and moreover $\partial_c \psi(x) \subseteq \text{c–}\exp_x(\partial^- \psi(x))$.

With these definitions in mind and recalling the question of connectedness of the c–subdifferential, the following theorem confirms that it is a necessary condition for smoothness of optimal transport (see [12, Rmk. 3 p.257]).

Theorem 3.8. *(Discontinuous Optimal Transport Maps) Suppose $\exists\, \bar{x} \in X$ and $\psi : X \to \mathbb{R}$ c–convex such that $\partial_c \psi(\bar{x})$ is not connected. Then one can find f, g smooth positive probability densities such that the optimal map is discontinuous.*

With this negative result in mind, it remains to be seen if connectedness of the c–subdifferential is sufficent to obtain smoothness. We therefore introduce some alternate characterizations of this condition to see if we can establish a connection with the MTW condition.

Theorem 3.9. *(Characterizations of Connectedness of the c-subdifferential) The following are equivalent*

(i) *For any c–convex ψ, $\forall\, \bar{x} \in X$, $\partial_c \psi(\bar{x})$ is connected*
(ii) *For any c–convex ψ, $\forall\, \bar{x} \in X$, $(\text{c–exp}_{\bar{x}})^{-1}(\partial_c \psi(\bar{x}))$ is convex and coincides with $\partial^- \psi(\bar{x})$*
(iii) $\forall\, \bar{x} \in X, \forall\, y_0, y_1 \in Y$, *if* $[y_0, y_1]_{\bar{x}} = (y_t)_{t \in [0,1]} \subseteq Y$ *then*

$$c(x, y_t) - c(\bar{x}, y_t) \geq \min\{c(x, y_0) - c(\bar{x}, y_0), c(x, y_1) - c(\bar{x}, y_1)\}$$

$\forall\, x \in X, t \in [0, 1]$

(iv) $\forall\, \bar{x} \in X, y \in Y, \eta, \xi \in \mathbb{R}^n$ *with* $\xi \perp \eta$

$$\frac{d^2}{ds^2}\bigg|_{s=0} \frac{d^2}{dt^2}\bigg|_{t=0} c(c\text{–}\exp_{\bar{x}}(t\xi), c\text{–}\exp_{\bar{x}}(p+s\eta)) \leq 0 \tag{3.5}$$

for $p = (c\text{–}\exp_{\bar{x}})^{-1}(y)$

and if any of these conditions is not satisfied then C^1 *c–convex functions are not dense in the set of Lipschitz c–convex functions.*

Proof. $(ii) \implies (i)$ Since $(\text{c–}\exp_x)^{-1}(\partial_c \psi(x))$ is convex, it is also connected, whence $\partial_c\psi(x)$ is connected as well.
$(i) \implies (ii)$ Let $\psi_{\bar{x},y_0,y_1} = \max\{-c(\cdot, y_0) + c(\bar{x}, y_0), -c(\cdot, y_1) + c(\bar{x}, y_1)\}$. Then $(\text{c–}\exp_{\bar{x}})^{-1}(\partial_c\psi_{\bar{x},y_0,y_1}(\bar{x})) \subseteq [(\text{c–}\exp_{\bar{x}})^{-1}(y_0), (\text{c–}\exp_{\bar{x}})^{-1}(y_1)]$ which is a segment. Here, connectedness and convexity are equivalent, so if (i) holds $\partial_c\psi_{\bar{x},y_0,y_1}(\bar{x}) = [y_0, y_1]_{\bar{x}} = \text{c–}\exp_{\bar{x}}(\partial^-\psi_{\bar{x},y_0,y_1}(\bar{x}))$. Generally, taking $y_0, y_1 \in \partial_c\psi(\bar{x})$ one sees that $\partial_c\psi(\bar{x}) \supseteq \partial_c\psi_{\bar{x},y_0,y_1}(\bar{x}) = [y_0, y_1]_{\bar{x}}$
$(ii) \iff (iii)$ Note that the equation in (iii) is equivalent to $\partial_c\psi_{\bar{x},y_0,y_1} = [y_0, y_1]_{\bar{x}}$ thus the same arguments above can be used.
$(iii) \implies (iv)$ Take $\bar{x} \in X$, $y \in Y$ such that $y = \text{c–}\exp_{\bar{x}}(p)$. Take ξ, η perpendicular and of norm 1 and define $y_0 = \text{c–}\exp_{\bar{x}}(p - \epsilon\eta)$ and $y_1 = \text{c–}\exp_{\bar{x}}(p + \epsilon\eta)$ for some $\epsilon > 0$ and $h_0(x) = c(\bar{x}, y_0) - c(x, y_0)$, $h_1(x) = c(\bar{x}, y_1) - c(x, y_1)$ and finally $\psi = \max\{h_0, h_1\} = \psi_{\bar{x},y_0,y_1}$. Let $\gamma(t)$ be a curve contained in the set $\{h_0 = h_1\}$ for which $\gamma(0) = \bar{x}$, $\dot{\gamma}(0) = \xi$.

Since $y \in [y_0, y_1]_{\bar{x}}$ we get by (iii) that $y \in \partial_c\psi(\bar{x})$ and therefore $\frac{1}{2}[h_0(\bar{x}) + h_1(\bar{x})] + c(\bar{x}, y) = \psi(\bar{x}) + c(\bar{x}, y) \leq \psi(\gamma(t)) + c(\gamma(t), y) = \frac{1}{2}[h_0(\gamma(t)) + h_1(\gamma(t))] + c(\gamma(t), y)$ as $h_0 = h_1$ along γ. Thus, $\frac{1}{2}[c(\gamma(t), y_0) + c(\gamma(t), y_1)] - c(\gamma(t), y) \leq \frac{1}{2}[c(\bar{x}, y_0) + c(\bar{x}, y_1)] - c(\bar{x}, y)$ and the function on the left hand side attains its minimum at $t = 0$, so $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big|_{t=0} \left(\frac{1}{2}[c(\gamma(t), y_0) + c(\gamma(t), y_1)] - c(\gamma(t), y)\right) \leq 0$. Since $D_x c(\bar{x}, y) = \frac{1}{2}[D_x c(\bar{x}, y_0) + D_x c(\bar{x}, y_1)]$ one obtains

$$\left\langle \left[\frac{1}{2}(D_{xx}c(\bar{x}, y_0) + D_{xx}c(\bar{x}, y_1)) - D_{xx}c(\bar{x}, y)\right] \cdot \xi, \xi \right\rangle \leq 0.$$

Therefore $\langle D_{xx}c(\bar{x}, \text{c–}\exp_{\bar{x}}(p + s\eta)) \cdot \xi, \xi\rangle$ is concave as a function of s, proving (iv). □

Remark 3.10. (Equivalence of Connectedness and MTW(0) Condition) By direct computation, one obtains that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\bigg|_{s=0} \frac{\mathrm{d}^2}{\mathrm{d}t^2}\bigg|_{t=0} c(\text{c–}\exp_{\bar{x}}(t\xi), \text{c–}\exp_{\bar{x}}(p + s\eta)) = \mathfrak{S}_{(x,y)}(\xi, \eta)$$

Therefore by Theorem 3.9 we get that the MTW(0) condition is equivalent to connectedness of the c–subdifferential of any c–convex function. Thus, since we have already shown in by the results of Theorems 3.3 and 3.8; the MTW condition is necessary and sufficient for smoothness of the optimal transport map.

Loeper further proved in [12] the following regularity results for cost functions satisfying the MTW(K) condition. Notably, this gives us regularity of the optimal transport map as well as the regularity of solutions to 3.1.

Theorem 3.11. *(Regularity from MTW(K)) Let* $c : X \times Y \to \mathbb{R}$ *satisfy (C0)-(C3) and MTW(K) for* $K > 0$*. Let* f *be bounded from above on* X *and* g *be bounded away from* 0 *on* Y *and denote the optimal transport map* T *sending* f *onto* g*. Suppose* $D_x c(x, y)$ *is convex* $\forall\, x \in X$*. Then* $u \in C^{1,\alpha}(X)$ *with* $\alpha = 1/(4n - 1)$ *hence* $T_u \in C^{0,\alpha}(X)$

This theorem was generalized in [14] to the following.

Theorem 3.12. *(Regularity from MTW(0)) Let* $c : X \times Y \to \mathbb{R}$ *satisfy (C0)-(C3) and MTW(0). Let* f *be bounded from above on* X *and* g *be bounded away from* 0 *and* ∞ *on* Y*. Also, assume* $D_x c(x, Y)$ *and* $D_y c(X, y)$ *are uniformly convex* $\forall\ x \in X, y \in Y$*. Then* $u \in C^{1,\alpha}_{loc}(\bar{X}') \ \forall\ X' \subseteq X$ *where* f *is uniformly bounded away from* 0*.*

Throughout this section we have discussed the regularity of optimal transport maps arising with more general cost functions. We have equally connected the regularity of these maps to the regularity of solutions to a general family of Monge–Ampère type equations.

References

[1] Luigi Ambrosio, Luis A. Caffarelli, Yann Brenier, Giuseppe Buttazzo, and Cédric Villani, *Optimal Transportation and Applications*, Springer, 2003.

[2] L. A. Caffarelli, *A Localization Property of Viscosity Solutions to the Monge-Ampere Equation and their Strict Convexity*, Annals of Mathematics **131** (1990), no. 1, 129–134.

[3] Luis A. Caffarelli, *Interior $W^{2,p}$ Estimates for Solutions of the Monge-Ampere Equation*, Annals of Mathematics **131** (1990), no. 1, 135–150.

[4] ———, *Some Regularity Properties of Solutions of Monge–Ampre Equations*, Communications on Pure and Applied Mathematics **44** (1991), no. 89, 965–969.

[5] ———, *The Regularity of Mappings with a Convex Potential*, Journal of the American Mathematical Society **5** (1992), no. 1, 99–104.

[6] ———, *Boundary Regularity of Maps with Convex Potentials*, Annals of Mathematics **144** (1996), no. 3, 453–496.

[7] Piermarco Cannarsa and Carlo Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Birkhuser, 2004.

[8] Guido De Philippis and Alessio Figalli, *Optimal Regularity of the Convex Envelope*, Transactions of the American Mathematical Society **367** (2014), 1.

[9] Herbert Federer, *Geometric Measure Theory*, Springer, 1996.

[10] David Gilbarg and Neil S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer Berlin Heidelberg, 1983.

[11] Cristian Gutiérrez, *The Monge–Ampère Equation*, Birkhäuser, 2001.

[12] Grégoire Loeper, *On the Regularity of Solutions of Optimal Transportation Problems*, Acta Mathematica **202** (2009), no. 2, 241–283.

[13] Xi-Nan Ma, Neil S. Trudinger, and Xu-Jia Wang, *Regularity of Potential Functions of the Optimal Transportation Problem*, Archive for Rational Mechanics and Analysis **177** (2005), no. 2, 151–183.

[14] Guido De Philippis and Alessio Figalli, *The Monge–Ampère Equation and its Link to Optimal Transportation*, Bulletin of the American Mathematical Society **51** (2014), no. 4, 527–580.

[15] Guido De Philippis, Alessio Figalli, and Ovidiu Savin, *A Note on Interior $W^{2,1+\varepsilon}$ Estimates for the Monge–Ampère Equation*, Mathematische Annalen **357** (2013), no. 1, 11–22.

[16] Ralph T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[17] Cédric Villani, *Topics in Optimal Transportation*, American Mathematical Society, 2003.

[18] ———, *Optimal Transport Old and New*, Springer Verlag, 2009.

[19] Constantin Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, 2002.