AMERICAN UNIVERSITY OF BEIRUT

OPTIMAL MASS TRANSPORT

by
LARA KHALED BAALBAKI

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science
to the Department of Mathematics
of Faculty of Arts and Sciences
at the American University of Beirut

Beirut, Lebanon
April 2022

AMERICAN UNIVERSITY OF BEIRUT

OPTIMAL MASS TRANSPORT PROBLEM

by
LARA KHALED BAALBAKI

Approved by:

A. Sabra

Dr. Ahmad Sabra, Assistant Professor — Advisor
Department of Mathematics

Nabil Nassif

Dr. Nabil Nassif, Professor — Member of Committee
Department of Mathematics

Farouk Abi Khuzam

Dr. Farouk Abi Khuzam, Professor — Member of Committee
Department of Mathematics

Date of thesis defense: April 29, 2022

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Baalbaki Lara Khaled
Last First Middle

● Master's Thesis ○ Master's Project ○ Doctoral Dissertation

■ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☐ I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes after : **One ---- year from the date of submission of my thesis, dissertation, or project.**
Two ---- years from the date of submission of my thesis, dissertation, or project.
Three ---- years from the date of submission of my thesis, dissertation, or project.

Lara May 11, 2022

Signature Date

Acknowledgements

First and foremost, thanks to God, the Almighty, for the blessings he gave me throughout my research journey that enabled me to complete the research successfully.

It's now April 18, 2022, at 1:28 a.m. and I'm writing this acknowledgment. It was exactly eight months ago when the journey of writing my thesis started. This journey had many ups and downs, sleepless nights, and unforgettable moments. Through this journey, I challenged myself a lot and I learned even more. I knew, though, from the beginning that it would be a little bit hard for me because I did not have any similar experience from my bachelor's degree, which I successfully completed without writing a thesis, but I did it!!

This thesis would not have been possible without the inspiration and support of many wonderful people. I really appreciate them all for being a part of this journey and making this thesis possible. First of all, I would like to express my deep and sincere gratitude to my research advisor, Professor Ahmad Sabra, who generously gave time to offer me valuable comments toward improving my work. Moreover, I want to thank him for his patience, motivation, enthusiasm, and immense knowledge, which helped me a lot in writing this thesis.

Besides my advisor, I would like to thank Professor Nabil Nassif for believing in me and for being so supportive!

Moreover, I would like to thank my family. Words can not express how grateful I am for all of the sacrifices that you've made on my behalf. Your prayers for me were what sustained me thus far. A special thanks to my mother for providing me with all of the tools I needed to succeed in life and for constantly showering me with affection and love. I recognize and appreciate your immense contribution to my success, and I want you to know that I owe everything I have to you.

Last but not least, I dedicate this thesis to my sacrificing mother, Khadija Fleifel, and my brother, Ahmad Baalbaki.

Abstract
of the Thesis of

Lara Khaled Baalbaki for Master of Science
Major: Pure Mathematics

Title: Optimal Mass Transport

Mass transportation consists of optimizing the cost of transport of "goods" from a source to its target. First, it was stated by Monge in 1781 for a pile of sand, but the applications of this problem appear in different fields such as economics, optics, data science, and linear programming. In this thesis, we develop the needed theory to solve the optimal transport problem, relax the variational problem and connect it to two other problems that we'll analyze using tools from convex analysis and measure theory. We also show the existence and uniqueness of the three problems as well as the connection of their solutions.

Table of Contents

Illustrations

Chapter 1

Introduction

Gaspard Monge was a French mathematician, known as the inventor of descriptive geometry. In 1781, Monge formalized the mass transportation problem which is to move one distribution of mass onto another with a minimum average distance covered [15]. Major advances were made during World War II by the Soviet mathematician Leonid Kantorovich [11], the founder of linear programming who solved the Monge problem more than 150 years later, and was awarded for this contribution Nobel Prize in economics. Another award related to this topic is the Field medal granted to C. Villani [19] (2010), and A. Figalli (2018) [6] for their mathematical contributions.

The problem can be stated as follows: assume we have some pile of sand, and we would like to transport it into some hole on the other side. We want to come up with a good transport plan minimizing the cost of the transportation from the source to the target. Another example would be the "mines to factories" problem. Suppose we have several mines that are producing resources and now we get the resources out of the mines, and we want to transport them to our collection of factories, to which factory should I send them in a way that will minimize the cost of transporting the resources? Let $X_1, X_2, \cdots, X_m$ be m mines (sources) and $Y_1, Y_2, \cdots, Y_n$ be n factories(destinations), x_{ij} be the quantity shipped from X_i to Y_j, and c_{ij} be the unit transportation cost. The goal is to find a transport plan minimizing the cost of the transportation from the source to the destination, i.e. to find a matrix $X = (x_{ij})$ minimizing

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} c_{ij}.$$

We consider that the total capacity of the sources is equal to the total needs for the problem to be feasible, and hence in this case the number of possible transportation plans is finite then there must be a transportation plan that attains the minimum. On the other hand, finding an optimal plan is long so linear programming facilitates this issue and solves these types of equations efficiently [4].

This problem can be extended to non-discrete cases. Monge considered the cost as the distance of masses transported from the source to the target, but more costs functions are applicable. Mathematically, the problem can be stated as follows. Let X and Y be two metric spaces with measures μ and ν respectively such that $\mu(X) = \nu(Y)$ and $c : X \times Y \to \mathbb{R}$ a measurable function with respect to the product measure $\mu \otimes \nu$. Our goal is to find $T : X \mapsto Y$ such that $\mu(T^{-1}(A)) = \nu(A)$ that minimizes the transport cost

$$\int_X c(x, Tx)\, d\mu.$$

Does such T always exist, and is it unique? The quadratic cost function has been popular because it is convex and smooth, which makes evaluation of derivatives easier. This has led to its popularity stretching across linear as well as nonlinear control [17].

Well, Monge's problem is not easy to deal with because of its constraint. What shall we do? "Relax!" said Kantorovich, more than 150 years later. We'll study in Chapter 2 the relaxed formulation of Kantorovich which consists of finding a measure that minimizes $\int_{X\times Y} c(x, y)\, d\gamma$ along all measures γ with marginal μ and ν. A precise definition can be found in Section 2.2, and a solution to the Kantorovich problem is provided in Section 2.3. The support of these solutions have an interesting geometrical structure called c-cyclically monotone studied in Section 3.4. In Chapter 3, we studied the dual Kantorovich problem consisting of maximizing some integral (3.1) allowing us to understand more the structure of the support of γ. The dual problem is solved in Section 3.3. Finally both the dual and the Kantorovich problem are used in Chapter 4 to solve the Monge Problem under some conditions on the cost and the sets X and Y. As we see in Section 4.1, a solution to the Monge Problem might not always exist even in the case when the Kantorovich problem is solvable. We end the thesis with an example where we can obtain explicitly a relation between the solutions to the three problems covering a large family of costs that appear in the literature including the quadratic cost.

Optimal transport theory is used widely to solve problems in mathematics and some areas of the sciences, but it can also be used to understand a range of problems in applied economics, such as the matching between job seekers and jobs, and the determinants of real estate prices [7]. Many researchers formulated the mathematical model for transportation problem in various environments. The basic transportation model was introduced by Hitchcock [10], in 1941, in which the transportation constraints were based on crisp values. But, in the present world, the transportation parameters like demand, supply, and unit transportation cost may be uncertain due to several uncontrolled factors. In this situation, fuzzy transportation problem was formulated and solved by many researchers. The Monge-Kantorovich theory is having a growing number of applications [9] in various areas of sciences including economics [8], optic [13] (e.g. the reflector problem), meteorology [3], oceanography, kinetic theory [2], machine learning [12], partial differential equations [5] and functional analysis (e.g. geometric inequalities) [14].

Chapter 2

The Kantorovich Problem

In this chapter, we introduce the Kantorovich problem and show that it admits a solution under some conditions.

2.1 Preliminaries: Space of measures

Definition 2.1.1. *Given a metric space X, let $\mathfrak{B}(X)$ be the smallest σ-algebra that contains the open sets of X, this is known as the σ-algebra of Borel sets.*

We denote by $\mathcal{M}(X)$ the space of Borel finite signed measure that are finitely additive, that is elements μ in $\mathcal{M}(X)$ are maps $\mu : X \mapsto \mathbb{R}$ such that $|\mu(E)| < \infty$ for every $E \in \mathfrak{B}(X)$ and for every $E_1, \cdots, E_n \in \mathfrak{B}(X)$ disjoints we have

$$\mu(\cup_{i=1}^n E_i) = \sum_{i=1}^n \mu(E_i).$$

The space of $\mathcal{M}(X)$ is equipped with the variational norm

$$||\mu|| := |\mu|(X).$$

We also denote by $\mathcal{M}_+(X)$ the subspace of $\mathcal{M}(X)$ containing non-negative measures.

Definition 2.1.2 (Probability Measures)**.** *Let X be a metric space. Define $\mathcal{P}(X)$ to be the set of all probability measures, i.e. $\mu \in \mathcal{P}(X)$ if and only if $\mu \in \mathcal{M}_+(X)$ and $\mu(X) = 1$*

$$\mathcal{P}(X) := \{\mu \in \mathcal{M}_+(X) : \mu(X) = 1\}.$$

Definition 2.1.3. *Let X and Y be two metric spaces. The pushforward $T_{\#}\mu$ of a measure μ on a measurable space X along a measurable function $T : X \mapsto Y$ is defined as follows:*

$$(T_{\#}\mu)(A) = \mu(T^{-1}(A)) \qquad \forall A \subseteq Y.$$

Definition 2.1.4. *Let X be a metric space. We denote by $C(X)$ the space of all continuous functions. If X is compact then elements in $C(X)$ are bounded, in this case we equip $C(X)$ with the $\|.\|_\infty$ norm defined as follows $\|f\|_\infty = \sup_{x\in X} |f(x)|$.*

Definition 2.1.5. *Given a metric space X, and the measures $\mu_n, \mu \in \mathcal{M}(X)$, we say that μ_n weakly converges to μ and denote it by $\mu_n \rightharpoonup \mu$ if and only if $\forall f \in C(X)$ $\displaystyle\int_X f\, d\mu_n \to \int_X f\, d\mu$.*

Remark 2.1.6. By Riesz theorem [16], we know that for X compact $C(X)^* = \mathcal{M}(X)$, that is for every functional $T : C(X) \mapsto \mathbb{R}$ there exists a unique measure $\mu \in \mathcal{M}(X)$ such that

$$T(f) = \int_X f\, d\mu \quad \forall f \in C(X).$$

In this case, the weak convergence given in Definition 2.1.5 corresponds to the the weak* topology.

In the case where X is not compact, $\mathcal{M}(X)$ is not the dual of $C(X)$, but is the dual of $C_0(X)$ the space of compactly supported continuous functions.

2.2 Transport Plans

Definition 2.2.1. *Let X, Y be two metric spaces, $\mu \in \mathcal{P}(X)$, and $\nu \in \mathcal{P}(Y)$. We denote by $\Pi(\mu, \nu)$ the set of probability measures on $X \times Y$ whose first and second marginals are equal to μ and ν respectively, i.e. $(\pi_1)_\#\gamma = \mu$ and $(\pi_2)_\#\gamma = \nu$ for all $\gamma \in \Pi(\mu, \nu)$ where π_1 and π_2 denote the projections from $X \times Y$ onto the first and second factors respectively. The measures $\gamma \in \Pi(\mu, \nu)$ are called transport plans.*

Proposition 2.2.2. *Given X, Y metric spaces, $\mu \in \mathcal{P}(X)$, and $\nu \in \mathcal{P}(Y)$, then $\gamma \in \Pi(\mu, \nu)$ if and only if*

$$\gamma(A \times Y) = \mu(A)\ \forall A \in \mathfrak{B}(X), \textit{and} \qquad \gamma(X \times B) = \nu(B)\ \forall B \in \mathfrak{B}(Y).$$

Proof. Let us note that for every $A \in \mathfrak{B}(X)$, $\pi_1^{-1}(A) = A \times Y$. However, if $\gamma \in \Pi(\mu, \nu)$, then $(\pi_1)_\#\gamma = \mu$ which is $\gamma(\pi_1^{-1}(A)) = \mu(A)$ which implies that $\gamma(A \times Y) = \mu(A)$ for all $A \in \mathfrak{B}(X)$. Similar proof can be done to prove that $\gamma(X \times B) = \nu(B)$ for all $B \in \mathfrak{B}(Y)$.

On the other hand, we have to prove that $\gamma \in \Pi(\mu, \nu)$. It is obvious that γ is a probability measure and by using the previous definition 2.2.1, it is enough to prove that $(\pi_1)_\#\gamma = \mu$ and $(\pi_2)_\#\gamma = \nu$.
By using Definition 2.1.3, $(\pi_1)_\#\gamma = \gamma(\pi_1^{-1}(A)) = \mu(A)$ for all $A \in \mathfrak{B}(X)$, so $(\pi_1)_\#\gamma = \mu$. Similarly for $(\pi_2)_\#\gamma = \nu$. $\square$

Proposition 2.2.3. *Let X, Y be metric spaces, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, then $\Pi(\mu, \nu)$ is not empty.*

Proof. Let $\gamma = \mu \otimes \nu$ be the measure on $X \times Y$. Obviously, γ is a probability measure since $\gamma(X \times Y) = \mu(X) \cdot \nu(Y) = 1$. Moreover, for all $A \in \mathfrak{B}(X)$ and $B \in \mathfrak{B}(Y)$ we have: $\gamma(A \times Y) = \mu(A) \cdot \nu(Y) = \mu(A)$ since $\nu(Y) = 1$ and $\gamma(X \times B) = \mu(X) \cdot \nu(B) = \nu(B)$ since $\mu(X) = 1$. Then, by using Proposition 2.2.2, $\gamma \in \Pi(\mu, \nu)$. Therefore, $\Pi(\mu, \nu)$ is not empty.

□

Proposition 2.2.4. *Given compact metric spaces X and Y, $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, then $\gamma \in \Pi(\mu, \nu)$ if and only if for every $\phi \in C(X)$ and $\psi \in C(Y)$*

$$\int_X \phi(x)\, d\mu = \int_{X \times Y} \phi(x)\, d\gamma, \ \textit{and} \qquad \int_Y \psi(y)\, d\nu = \int_{X \times Y} \psi(y)\, d\gamma.$$

Proof. Let $\gamma \in \Pi(\mu, \nu)$ then for all $\phi \in C(X)$ there exists $c_1, \cdots, c_n$ distinct values in $\mathbb{R}$, A_n disjoint subsets of X such that

$$\phi(x) = \sum_{i=1}^{\infty} c_i \cdot \chi_{A_i}$$

Notice that

$$\int_{X \times Y} \phi(x)\, d\gamma = \int_{X \times Y} \sum_{i=1}^{\infty} c_i \cdot \chi_{A_i}\, d\gamma = \sum_{i=1}^{\infty} c_i \int_{X \times Y} \chi_{A_i}\, d\gamma$$

(Fubini's theorem having $\sum_{i=1}^{\infty} |c_n \chi_{A_i}| \leq \sup |\phi| < \infty$). Moreover,

$$\sum_{i=1}^{\infty} c_i \int_{X \times Y} \chi_{A_i}(x)\, d\gamma = \sum_{i=1}^{\infty} c_i \int_{X \times Y} \chi_{A_i \times Y}(x, y)\, d\gamma = \sum_{i=1}^{\infty} c_i \gamma(A_i \times Y).$$

Using the above Proposition (2.2.2) we have that $\gamma(A_i \times Y) = \mu(A_i)$ so

$$\sum_{i=1}^{\infty} c_i \gamma(A_i \times Y) = \sum_{i=1}^{\infty} c_i \mu(A_i) = \int_X \sum_{i=1}^{\infty} c_i \chi_{A_i}(x)\, d\mu.$$

Therefore, $\displaystyle\int_X \phi(x)\, d\mu = \int_{X \times Y} \phi(x)\, d\gamma$.

Similarly we can prove that $\displaystyle\int_Y \psi(y)\, d\nu = \int_{X \times Y} \psi(y)\, d\gamma$. On the other hand, we have to prove that $\gamma \in \Pi(\mu, \nu)$. Let F be any closed subset of X. Let

$$f_n(x) = \max(1 - n d(x, F), 0).$$

Obviously, f_n is continuous in X and decreases to χ_F since if $x \in F$ then $f_n(x) = 1$ and if $x \notin F$ so $d(x, F) > 0$ so there exists $n \in \mathbb{N}_+$ such that $d(x, F) > \frac{1}{n}$ then $f_n(x) = 0$. We have

$$\int_X f_n(x)\, d\mu = \int_{X \times Y} f_n(x)\, d\gamma$$

but by applying Lebesgue Monotone convergence theorem we have :

$$\int_X f_n(x)\, d\mu \to \int_X \chi_F(x)\, d\mu \quad \text{and} \int_{X\times Y} f_n(x)\, d\gamma \to \int_{X\times Y} \chi_{F\times Y}(x, y)\, d\gamma$$

so $\mu(F) = \gamma(F \times Y)$. Similarly we can prove that $\nu(G) = \gamma(X \times G)$ for all closed subsets G of Y. Therefore, by Proposition 2.2.2 we can conclude that $\gamma \in \Pi(\mu, \nu)$. □

2.3 Solution of the Kantorovich Problem

Definition 2.3.1. *Let X and Y be two metric spaces , $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, and $c : X \times Y \mapsto \mathbb{R}$ a continuous function, then the Kantorovich-Problem is defined as follows*

$$(KP) : \inf \left\{ \int_{X\times Y} c(x, y)\, d\gamma : \gamma \in \Pi(\mu, \nu) \right\} \tag{2.1}$$

Proposition 2.3.2. *Let X, Y be compact metric spaces, $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, then $\Pi(\mu, \nu)$ is compact.*

Proof. Let $\gamma_n \in \Pi(\mu, \nu)$. We have to prove that γ_n has a convergent subsequence. Notice that γ_n are contained in the closed unit ball of $\mathcal{M}(X)$ which by Banach Alaoglu [1] is compact in the weak* topology then there exists a sub sequence γ_{n_k} that converges weakly to γ i.e $\gamma_{n_k} \rightharpoonup \gamma$.

It remains to show $\gamma \in \Pi(\mu, \nu)$. We have $\gamma_{n_k} \rightharpoonup \gamma$ then $\int_{X\times Y} \phi(x)\, d\gamma_{n_k} \to \int_{X\times Y} \phi(x)\, d\gamma$ for every $\phi \in C(X)$, but from Proposition 2.2.4 $\int_{X\times Y} \phi(x)\, d\gamma_{n_k} = \int_X \phi(x)\, d\mu$. Letting $k \to \infty$, we obtain that $\int_{X\times Y} \phi\, d\gamma = \int_X \phi\, d\mu$. Similarly we show that for all $\psi \in C(Y)$ we have $\int_Y \psi(y)\, d\nu = \int_{X\times Y} \psi(y)\, d\gamma$, therefore using proposition 2.2.4 we get $\gamma \in \Pi(\mu, \nu)$, and the proof follows. □

Proposition 2.3.3. *Given X and Y compact metric spaces, $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, and a continuous function $c : X \times Y \mapsto \mathbb{R}$ then the (KP) problem (2.1) admits a solution.*

Proof. To prove that (KP) admits a solution we need to show that the infimum is attained. We define $K : \Pi(\mu, \nu) \mapsto \mathbb{R}$ as follows

$$K(\gamma) = \int_{X\times Y} c(x,y)\, d\gamma.$$

Let $\gamma_n \in \Pi(\mu, \nu)$ be such that

$$K(\gamma_n) = \int_{X\times Y} c(x,y)\, d\gamma_n \to m = \inf \left\{ \int_{X\times Y} c(x,y)\, d\gamma : \gamma \in \Pi(\mu,\nu) \right\}.$$

Since $\Pi(\mu, \nu)$ is compact then γ_n has a subsequence $\gamma_{n_k} \rightharpoonup \gamma_0$ with $\gamma_0 \in \Pi(\mu, \nu)$. We have that $c \in C(X \times Y)$, we then conclude by the definition of convergence in $\Pi(\mu, \nu)$ that $\lim_{k\to\infty} K(\gamma_{n_k}) = K(\gamma_0)$, and hence

$$m = K(\gamma_0) = \int_{X\times Y} c(x,y)\, d\gamma_0.$$

$\square$

Remark 2.3.4. *A solution of the Kantorovich Problem γ is called an optimal transport plan.*

Chapter 3

The dual problem

3.1 Motivation of the Dual problem

Here is an informal way of interpreting Kantorovich duality principle. Suppose for instance an industrial is willing to transfer a huge amount of coal from his mines to his factories. He can hire trucks to do this transportation problem, but he has to pay them $c(x, y)$ for each ton of coal which is transported from position x to position y. Both the amount of coal which he can extract from each mine, and the amount which each factory will receive, are fixed. As he is trying to solve the associated Monge-Kantorovich problem in order to minimize the price he has to pay, a mathematician comes to him and tells him "I can ship all your coal with my own trucks and you won't have to care of what goes where. I will only set a price $\phi(x)$ for loading one ton of coal at position x, and a price $\psi(y)$ for unloading it at destination y. I will set the prices in such a way that your financial interest will be to let me handle all your transportation! Indeed, the industrial can check that for all x and all y, the sum $\phi(x) + \psi(y)$ will always be less than the cost $c(x, y)$. Kantorovich's dual problem presumes that if the shipper is clever enough, then he can arrange the prices in such a way that he will pay you (almost) as much as he would have been ready to spend by the other method. [19]

Mathematically, we then seek to show that

$$\inf\left\{\int_{X\times Y} c(x,y)\,d\gamma : \gamma\in\Pi(\mu,\nu)\right\}$$
$$= \sup\left\{\int_X \phi(x)\,d\mu + \int_Y \psi(y)\,d\nu : \phi\in C(X), \psi\in C(Y),\ \text{and } \phi(x)+\psi(y)\le c(x,y)\right\}.$$

We will provide in this chapter a rigorous proof of this identity however we present below a heuristic approach of why we expect these two problems to be equivalent.

$$(KP): \inf_{\gamma\in\Pi(\mu,\nu)} \left\{ \int_{X\times Y} c(x,y)\,d\gamma \right\} = \inf_{\gamma\in\mathcal{M}_+(X\times Y)} \left\{ \int_{X\times Y} c(x,y)\,d\gamma + \begin{cases} 0 & \text{if } \gamma\in\Pi(\mu,\nu) \\ +\infty & \text{otherwise} \end{cases} \right\}.$$

We also have

$$\sup_{\phi,\psi} \left\{ \int_X \phi\,d\mu + \int_Y \psi\,d\nu - \int_{X\times Y} (\phi(x)+\psi(y))\,d\gamma \right\} = \begin{cases} 0 & \text{if } \gamma\in\Pi(\mu,\nu) \\ +\infty & \text{otherwise} \end{cases}.$$

where the supremum is taken over all $\phi : X \mapsto \mathbb{R}$, $\psi : Y \mapsto \mathbb{R}$ bounded and continuous functions. We then write

$$\begin{aligned} &\inf_{\gamma\in\mathcal{M}_+(X\times Y)} \left\{ \int_{X\times Y} c(x,y)\,d\gamma + \sup_{\phi,\psi} \left\{ \int_X \phi(x)\,d\mu + \int_Y \psi(y)\,d\nu - \int_{X\times Y} (\phi(x)+\psi(y))\,d\gamma \right\} \right\} \\ &\qquad = \inf_{\gamma\in\mathcal{M}_+(X\times Y)} \sup_{\phi,\psi} \left\{ \int_{X\times Y} c(x,y)\,d\gamma + \int_X \phi(x)\,d\mu + \int_Y \psi(y)\,d\nu - \int_{X\times Y} (\phi(x)+\psi(y))\,d\gamma \right\} \end{aligned}$$

Switching "naively" the infimum and the supremum we get that the above expression is equal to

$$\begin{aligned} &\sup_{\phi,\psi} \inf_{\gamma\in\mathcal{M}_+(X\times Y)} \left\{ \int_{X\times Y} c(x,y)\,d\gamma + \int_X \phi(x)\,d\mu + \int_Y \psi(y)\,d\nu - \int_{X\times Y} (\phi(x)+\psi(y))\,d\gamma \right\} \\ &\quad = \sup_{\phi,\psi} \left\{ \int_X \phi(x)\,d\mu + \int_Y \psi(y)\,d\nu + \inf_{\gamma\in\mathcal{M}_+(X\times Y)} \left\{ \int_{X\times Y} c(x,y) - (\phi(x)+\psi(y))\,d\gamma \right\} \right\} \\ &\quad = \sup_{\phi,\psi} \left\{ \int_X \phi(x)\,d\mu + \int_Y \psi(y)\,d\nu + \begin{cases} 0 & \text{if } \phi(x)+\psi(y)\le c(x,y) \\ -\infty & \text{otherwise} \end{cases} \right\} \\ &\quad = \sup_{\phi,\psi} \left\{ \int_X \phi(x)\,d\mu + \int_Y \psi(y)\,d\nu : \phi\in C(X), \psi\in C(Y), \text{ and } \phi(x)+\psi(y)\le c(x,y) \right\}. \end{aligned}$$

Definition 3.1.1 (The Dual Problem)**.** *Let X, Y be any two metric spaces, $c : X\times Y \mapsto \mathbb{R}$ be a continuous function we then define the dual problem:*

$$(DP): \sup_{\phi,\psi} \left\{ \int_X \phi(x)\,d\mu + \int_Y \psi(y)\,d\nu : \phi\in C(X), \psi\in C(Y), \text{ and } \phi(x)+\psi(y)\le c(x,y) \right\}. \tag{3.1}$$

In this chapter, we will investigate the existence of solutions to the (DP) problem. To do this we need first to introduce the notions of $c-$transform and $\bar{c}-$transform.

3.2 c and $\bar{c}$ transforms

Definition 3.2.1. *Let X and Y be two metric spaces, $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, given $\phi \in C(X)$, $\psi \in C(Y)$, we define $I(\phi, \psi)$ as follows*

$$I(\phi, \psi) = \int_X \phi \, d\mu + \int_Y \psi \, d\nu$$

Definition 3.2.2. *Let X and Y be two metric spaces, given a function $f : X \mapsto \overline{\mathbb{R}}$, we define its c-transform $f^c : Y \mapsto \overline{\mathbb{R}}$ as follows*

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

We also define the $\bar{c}$-transform of $g : Y \mapsto \overline{\mathbb{R}}$ by $g^{\bar{c}} : X \mapsto \overline{\mathbb{R}}$ given as follows

$$g^{\bar{c}}(x) = \inf_{y \in Y} \{c(x, y) - g(y)\}.$$

Remark 3.2.3. Notice that if $f \leq m$ then $f^c(y) \geq m^c(y)$ for every $y \in Y$. Similarly, if $g \leq n$ then $g^{\bar{c}}(x) \geq n^{\bar{c}}(x)$ for every $x \in X$.

Proposition 3.2.4. *Given two metric spaces X and Y, and the measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, let $c : X \times Y \mapsto \mathbb{R}$ be a continuous function, $\phi \in C(X)$ and $\psi \in C(Y)$ then the followings are true:*

1. *$\phi(x) + \phi^c(y) \leq c(x, y)$, and similarly $\psi(y) + \psi^{\bar{c}}(x) \leq c(x, y)$ for all $(x, y) \in X \times Y$.*
2. *If for every $(x, y) \in X \times Y$, $\phi(x) + \psi(y) \leq c(x, y)$ then $I(\phi, \phi^c) \geq I(\phi, \psi)$ and $I(\psi^{\bar{c}}, \psi) \geq I(\phi, \psi)$.*
3. *$I(\phi - \epsilon, \psi + \epsilon) = I(\phi, \psi)$.*
4. *$(\phi - \epsilon)^c = \phi^c + \epsilon$, and $(\psi + \epsilon)^{\bar{c}} = \psi^{\bar{c}} - \epsilon$.*
5. *If $\phi_n \to \phi$ uniformly on X then $\phi_n^c \to \phi^c$ uniformly on Y. Similarly, if $\psi_n \to \psi$ uniformly on Y then $\psi_n^{\bar{c}} \to \psi^{\bar{c}}$ uniformly on X.*

Proof. The proof of these claims follow directly from the definitions of the c and $\bar{c}$ transform and of $I(\phi, \psi)$, we show (2), (4), and (5).

Proof of (2). We have $\psi(y) \leq c(x, y) - \phi(x)$ for every $(x, y) \in X \times Y$, and so taking the infimum over X we get

$$\psi(y) \leq \inf_x \{c(x, y) - \phi(x)\} = \phi^c(y).$$

Hence,

$$I(\phi, \phi^c) = \int_X \phi \, d\mu + \int_Y \phi^c \, d\nu \geq \int_X \phi \, d\mu + \int_Y \psi \, d\nu = I(\phi, \psi).$$

The other inequality follows similarly. $\square$

Proof of (4). We have $(\phi-\epsilon)^c(y) = \inf_x\{c(x,y)-(\phi-\epsilon)(x)\} = \inf_x\{(c(x,y)-\phi(x))+\epsilon\} = \phi^c(y)+\epsilon$. The other equality follows similarly. $\square$

Proof of (5). For $y \in Y$,

$$|\phi_n^c(y)-\phi^c(y)| = \left|\inf_x\{c(x,y)-\phi_n(x)\} - \inf_x\{c(x,y)-\phi(x)\}\right| \le \sup_{x\in X}|\phi_n(x)-\phi(x)|\,[1].$$

Hence, $\sup_{y\in Y}|\phi_n^c(y)-\phi^c(y)| \le \sup_{x\in X}|\phi_n(x)-\phi(x)|$, and the proof follows. $\square$

$\square$

Definition 3.2.5. *Given X and Y any two metric spaces, a function $\psi : Y \mapsto \overline{\mathbb{R}}$ is $\overline{c}-$concave if there exists $f : X \mapsto \overline{\mathbb{R}}$ such that $\psi = f^c$. We denote by $\overline{c}$-conc(Y) the set of $\overline{c}-$concave functions.*

A function $\phi : X \mapsto \overline{\mathbb{R}}$ is $c-$concave if there exists $g : Y \mapsto \overline{\mathbb{R}}$ such that $\phi = g^{\overline{c}}$. We denote by c-conc(X) the sets of $c-$concave functions.

Remark 3.2.6. Notice that if X and Y are compact metric spaces and $\phi \in C(X)$ and $\psi \in C(Y)$, then ϕ and ψ are bounded and so is their c and $\bar{c}$ transforms, we can then in this case consider the functions to have range in $\mathbb{R}$ as opposed to $\bar{\mathbb{R}}$.

Proposition 3.2.7. *Let X and Y be two metric spaces, $c : X \times Y \mapsto \mathbb{R}$ continuous. For $\phi : X \mapsto \mathbb{R}\cup\{-\infty\}$ we have $\phi^{c\overline{c}} \ge \phi$. Moreover, we get equality $\phi^{c\overline{c}} = \phi$ if and only if ϕ is $c-$concave.*

Similarly, for $\psi : Y \mapsto \mathbb{R}\cup\{-\infty\}$ we have $\psi^{\overline{c}c} \ge \psi$. Moreover, we get $\psi^{\overline{c}c} = \psi$ if and only if ψ is $\overline{c}-$concave.

Proof. Given $x \in X$, we realize that for every $y \in Y$,

$$\begin{aligned} c(x,y)-\phi^c(y) &= c(x,y) - \inf_{x'}\{c(x',y)-\phi(x')\} \\ &\ge c(x,y)-c(x,y)+\phi(x) = \phi(x). \end{aligned}$$

Then taking the infimum over y, we get that for every $x \in X$

$$\phi^{c\overline{c}}(x) = \inf_y\{c(x,y)-\phi^c(y)\} \ge \phi(x).$$

Similarly, it follows that $\psi^{\overline{c}c}(y) \ge \psi(y)$, $y \in Y$.

Next, we prove that $\phi^{c\overline{c}} = \phi$ if and only if ϕ is $c-$concave. The forward implication follows from the definition of $c-$concave function. To show the converse, assume ϕ is $c-$concave then there exists ψ such that $\phi = \psi^{\overline{c}}$, then $\phi^c = \psi^{\overline{c}c} \ge \psi$ and so from Remark 3.2.3 $\phi^{c\overline{c}} \le \psi^{\overline{c}} = \phi$.

$\square$

Remark 3.2.8. Above proposition implies $\phi^{c\overline{c}}$ is the smallest $c-$concave function larger than ϕ. In fact, take $\tilde{\phi}$ any $c-$concave function $\tilde{\phi} \ge \phi$, we get from Remark 3.2.3 that $\tilde{\phi}^c \le \phi^c$ and hence by the same remark and Proposition 3.2.7

$$\tilde{\phi} = \tilde{\phi}^{c\overline{c}} \ge \phi^{c\overline{c}}.$$

[1] The inequality follows from the fact that $|\inf_x F(x) - \inf_x G(x)| \le \sup_x |F(x)-G(x)|$

3.3 Existence of solution to the DP problem

In this section, we investigate the existence of solutions of the (DP) problem (3.1).

Lemma 3.3.1. *Given two metric spaces X and Y, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, and $c : X \times Y \mapsto \mathbb{R}$ continuous then there exists a sequence $\phi_n \in C(X)$ such that*

1. $\phi_n(x_0) = 0$ *for some* $x_0 \in X$, *and all* $n \in \mathbb{N}$.
2. $\phi_n \in c-$*conc*(X).
3. $\lim_{n\to\infty} I(\phi_n, \phi_n^c) = \sup \{I(\phi, \psi) : \phi \in C(X), \psi \in C(Y), \phi(x) + \psi(y) \le c(x, y)\}$.

Proof. Let f_n be a sequence of functions in $C(X)$ and g_n be a sequence of functions in $C(Y)$ such that $f_n(x) + g_n(y) \le c(x, y)$ for all $(x, y) \in X \times Y$ and

$$I(f_n, g_n) \to M := \sup \{I(\phi, \psi) : \phi \in C(X), \psi \in C(Y), \text{ and } \phi(x) + \psi(y) \le c(x, y)\}.$$

By Proposition 3.2.4, we get that

$$I(f_n, g_n) \le I(g_n^{\bar{c}}, g_n) \le I(g_n^{\bar{c}}, g_n^{\bar{c}c}) \le M.$$

Hence, by Squeeze theorem $I(g_n^{\bar{c}}, g_n^{\bar{c}c}) \to M$.

Given $x_0 \in X$, let $\phi_n(x) = g_n^{\bar{c}}(x) - g_n^{\bar{c}}(x_0)$, we have $\phi_n(x_0) = 0$. Notice that from Proposition 3.2.4

$$\phi_n = (g_n + g_n^{\bar{c}}(x_0))^{\bar{c}}, \text{ and} \qquad \phi_n^c(y) = (g_n^{\bar{c}} - g_n^{\bar{c}}(x_0))^c = g_n^{\bar{c}c}(y) + g_n^{\bar{c}}(x_0).$$

Hence, $\phi_n \in c-\text{conc}(X)$ and

$$I(\phi_n, \phi_n^c) = I\left(g_n^{\bar{c}} - g_n^{\bar{c}}(x_0), g_n^{\bar{c}c} + g_n^{\bar{c}}(x_0)\right) = I(g_n^{\bar{c}}, g_n^{\bar{c}c}) \to M.$$

□

Lemma 3.3.2. *Given compact metric spaces X and Y, $\phi : X \mapsto \mathbb{R}$, $\psi : Y \mapsto \mathbb{R}$ and $c : X \times Y \mapsto \mathbb{R}$ continuous. If $\phi \in c-conc(X)$ then*

$$|\phi(x_1) - \phi(x_2)| \le \omega_c(d_X(x_1, x_2)) \qquad \forall x_1, x_2 \in X,$$

where ω_c denotes the modulus of continuity of the function c.

Similarly, if $\psi \in \bar{c} - conc(Y)$ then

$$|\psi(y_1) - \psi(y_2)| \le \omega_c(d_Y(y_1, y_2)) \qquad \forall y_1, y_2 \in Y.$$

Proof. c is continuous and finite on compact set then it is uniformly continuous, and so there exists an increasing continuous function $\omega_c : \mathbb{R}^+ \mapsto \mathbb{R}^+$ such that $\omega_c(0^+) = 0$, and

$$|c(x_1, y_1) - c(x_2, y_2)| \leq \omega_c \left(d_X(x_1, x_2) + d_Y(y_1, y_2)\right) \qquad \forall (x_1, y_1), (x_2, y_2) \in X \times Y.$$

We have $\phi \in c-conc(X)$ then there exists $g : Y \mapsto \mathbb{R}$ such that $\phi = g^{\bar{c}}$, therefore

$$\begin{aligned} |\phi(x_1) - \phi(x_2)| = |g^{\bar{c}}(x_1) - g^{\bar{c}}(x_2)| &= \left|\inf_y \{c(x_1, y) - g(y)\} - \inf_y \{c(x_2, y) - g(y)\}\right| \\ &\leq \sup_y |c(x_1, y) - c(x_2, y)| \\ &\leq \omega_c(d_X(x_1, x_2)). \end{aligned}$$

The proof for ψ follows similarly. □

Theorem 3.3.3. *Given X and Y two compact spaces, and $c : X \times Y \mapsto \mathbb{R}$ continuous function then there exists a solution (ϕ_0, ψ_0) to the (DP) problem (3.1) such that $\phi_0 \in c - conc(X), \psi_0 \in \bar{c} - conc(Y)$ and $\psi_0 = \phi_0^c$.*

Proof. Let $M := \sup\{I(\phi, \psi) : \phi \in C(X), \psi \in C(Y), \text{ and } \phi(x) + \psi(y) \leq c(x, y)\}$. From Lemma 3.3.1, there exists a sequence $\phi_n \in c-conc(X)$ such that $\phi_n(x_0) = 0$ for some $x_0 \in X$, and $I(\phi_n, \phi_n^c) \to M$.

From Lemma 3.3.2, ϕ_n is equicontinuous, and for every $x \in X$ we have

$$|\phi_n(x)| = |\phi_n(x) - \phi_n(x_0)| \leq \omega_c(d_X(x, x_0)) \leq \omega_c(diamX)$$

so ϕ_n is uniformly bounded. Arzelà-Ascoli theorem hence implies that ϕ_n has a subsequence ϕ_{n_k} that converges uniformly on X.

Set $\phi_0 = \lim_{k\to\infty} \phi_{n_k}$. By Proposition 3.2.4, $\phi_{n_k}^c$ converges uniformly to ϕ_0^c, and

$$M = \lim_{k\to\infty} I(\phi_{n_k}, \phi_{n_k}^c) = \lim_{k\to\infty} \int_X \phi_{n_k}\, d\mu + \int_Y \phi_{n_k}^c\, d\nu = \int_X \phi_0\, d\mu + \int_Y \phi_0^c\, d\nu = I(\phi_0, \phi_0^c).$$

Therefore, there exists (ϕ_0, ψ_0) solution to the (DP) problem such that $\psi_0 = \phi_0^c$ and $\psi_0 \in \bar{c} - conc(Y)$.

It remains to show that $\phi_0 \in c-conc(X)$. In fact, since $\phi_{n_k} \in$ c$-conc(X)$ then by Proposition 3.2.7 we have $\phi_{n_k}^{c\bar{c}} = \phi_{n_k}$. Moreover, $\phi_{n_k}^c \to \phi_0^c$ uniformly hence, by Proposition 3.2.4 we get $\phi_{n_k} = \phi_{n_k}^{c\bar{c}} \to \phi_0^{c\bar{c}}$ uniformly. By uniqueness of limit, we conclude that $\phi_0 = \phi_0^{c\bar{c}}$ and our proof follows using Proposition 3.2.7. □

Remark 3.3.4. *The solution ϕ_0 of the (DP) problem 3.1 is called a Kantorovich Potential.*

3.4 c-Cyclically Monotone

Definition 3.4.1. *Given metric spaces X and Y, and $c : X \times Y \mapsto \mathbb{R}$. We say that a set $\Gamma \subseteq X \times Y$ is c-cyclically monotone (briefly c-CM) if for every $k \in \mathbb{N}$, permutation σ of $\{1, 2, \cdots, k\}$ we have*

$$\sum_{i=1}^{k} c(x_i, y_i) \leq \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}).$$

for all $(x_1, y_1), \cdots, (x_k, y_k) \in \Gamma$.

Remark 3.4.2. The word "cyclical" refers to the fact that, since every permutation is the disjoint composition of cycles, it is enough to check the inequality for cyclical permutations, i.e. replacing $\sum_{i=1}^{k} c(x_i, y_{\sigma(i)})$ by $\sum_{i=1}^{k} c(x_i, y_{i+1})$ in the definition (with the convention $y_{k+1} = y_1$).

The word "monotone" refers instead to the behavior of those sets when $X = Y = \mathbb{R}^d$ and $c(x, y) = -x \cdot y$, in this case the set $\Gamma \subseteq \mathbb{R}^d \times \mathbb{R}^d$ is simply called cyclically monotone if and only if for every $k \in \mathbb{N}$ and permutation σ of $\{1, 2, \cdots, k\}$ we have

$$\sum_{i=1}^{k} x_i \cdot y_i \geq \sum_{i=1}^{k} x_i \cdot y_{\sigma(i)}$$

for every $(x_1, y_1), \cdots, (x_k, y_k) \in \Gamma$.

Theorem 3.4.3. *Let $c : X \times Y \mapsto \mathbb{R}$ continuous. If $\Gamma \neq \emptyset$ is a c-CM set in $X \times Y$, then there exists a $c-$concave function $\phi : X \mapsto \mathbb{R} \cup \{-\infty\}$ such that*

$$\Gamma \subseteq \{(x, y) \in X \times Y : \phi(x) + \phi^c(y) = c(x, y)\}. \tag{3.2}$$

Proof. Fix $(x_0, y_0) \in \Gamma$. Define the following function on X

$$\phi(x) = \inf\{c(x, y_n) - c(x_n, y_n) + c(x_n, y_{n-1}) - c(x_{n-1}, y_{n-1}) + \cdots + c(x_1, y_0)$$
$$-c(x_0, y_0) : n \in \mathbb{N}, (x_i, y_i) \in \Gamma \text{ for all } i = 1, \cdots, n\}.$$

$\phi(x)$ never takes the value $+\infty$ since c is real valued and Γ is non-empty. Notice also that since Γ is $c-$CM then for every $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n) \in \Gamma$

$$\begin{aligned}
&c(x_0, y_n) - c(x_n, y_n) + c(x_n, y_{n-1}) - c(x_{n-1}, y_{n-1}) + \cdots + c(x_1, y_0) - c(x_0, y_0) \\
&= (c(x_0, y_n) + c(x_1, y_0) + \cdots + c(x_n, y_{n-1})) - (c(x_0, y_0) + c(x_1, y_1) + \cdots + c(x_n, y_n)) \\
&\geq 0
\end{aligned}$$

Therefore $\phi(x_0) \geq 0$ implying that ϕ is not identically $-\infty$.

Define also the following function on Y

$$\psi(y) = -\inf\{-c(x_n, y) + c(x_n, y_{n-1}) - c(x_{n-1}, y_{n-1}) + ... + c(x_1, y_0) - c(x_0, y_0);$$

$$n \in \mathbb{N}, (x_i, y_i) \in \Gamma \text{ for all } i = 1, ..., n\,; y_n = y\}.$$

Notice that $\psi(y) \neq -\infty$ if and only if there exist $x \in X$ such that $(x, y) \in \Gamma$ which is equivalent to say that $y \in \pi_2(\Gamma)$.

The goal is to first show that $\phi = \psi^{\bar{c}}$. For $y \in \pi_2(\Gamma)$, we have that for every $n \in \mathbb{N}$ and $(x_1, y_1), (x_2, , y_2), \cdots, (x_{n-1}, y_{n-1}), (x_n, y_n) \in \Gamma$, with $y_n = y$,

$$\begin{aligned}\phi(x) &\leq c(x, y) - c(x_n, y) + c(x_n, y_{n-1}) + \cdots + c(x_1, y_0) - c(x_0, y_0) \\ &\leq c(x, y) + (-c(x_n, y) + c(x_n, y_{n-1}) + \cdots + c(x_1, y_0) - c(x_0, y_0))\end{aligned}$$

so we get $\phi(x) \leq c(x, y) - \psi(y)$ for every $y \in \pi_2(\Gamma)$. The inequality also holds for $y \notin \pi_2(\Gamma)$ since in this case the right hand side is equal to $+\infty$. Taking the infimum over all $y \in Y$ we get that $\phi(x) \leq \psi^{\bar{c}}(x)$ for every $x \in X$. Moreover applying the inequality for $x = x_0$ we get that for every $y \in Y$,

$$\psi(y) \leq c(x_0, y) - \phi(x_0) \leq c(x_0, y) < +\infty.$$

Next we prove that $\psi^{\bar{c}}(x) \leq \phi(x)$. Notice that for every $n \in \mathbb{N}$ and for all $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n) \in \Gamma$

$$-\psi(y_n) \leq -c(x_n, y_n) + c(x_n, y_{n-1}) + \cdots + c(x_1, y_0) - c(x_0, y_0),$$

and so for every $x \in X$

$$\begin{aligned}\psi^{\bar{c}}(x) &= \inf_{y \in Y} \{c(x, y) - \psi(y)\} \\ &\leq c(x, y_n) - \psi(y_n) \\ &\leq c(x, y_n) + (-c(x_n, y_n) + c(x_n, y_{n-1}) + \cdots c(x_1, y_0) - c(x_0, y_0)).\end{aligned}$$

Taking the infimum over $(x_1, y_1), \cdots, (x_n, y_n) \in \Gamma$ yields to $\phi(x) \geq \psi^{\bar{c}}(x)$. Hence, $\phi(x) = \psi^{\bar{c}}(x)$. This shows that ϕ is $c-$concave.

It remains to show the inclusion (3.2). Let $(x, y) \in \Gamma$ from Proposition 3.2.4 it is enough to show that we have $\phi(x) + \phi^c(y) \geq c(x, y)$. For $n \in \mathbb{N}$ and $(x_1, y_1), \cdots, (x_n, y_n) \in \Gamma$ we write

$$\begin{aligned}&c(x, y_n) - c(x_n, y_n) + c(x_n, y_{n-1}) - c(x_{n-1}, y_{n-1}) + \cdots + c(x_1, y_0) - c(x_0, y_0) \\ &= c(x, y) + [-c(x, y) + c(x, y_n) - c(x_n, y_n) + c(x_n, y_{n-1}) - c(x_{n-1}, y_{n-1}) + \\ &\qquad\qquad \cdots + (c(x_1, y_0) - c(x_0, y_0))] \\ &\geq c(x, y) - \psi(y),\end{aligned}$$

where we used the fact that $(x_1, y_1), \cdots, (x_n, y_n), (x, y) \in \Gamma$. Taking the infimum over all $n \in \mathbb{N}$ and $(x_1, y_1), \cdots, (x_n, y_n) \in \Gamma$, we get that $\phi(x) \geq c(x, y) - \psi(y)$. Therefore for every $(x, y) \in \Gamma$, from Proposition 3.2.7 and the fact that $\phi = \psi^{\bar{c}}$ we conclude that

$$\phi(x) + \phi^c(y) = \phi(x) + \psi^{\bar{c}c}(y) \geq \phi(x) + \psi(y) \geq c(x, y).$$

□

3.5 $(KP) = (DP)$

In this section, we show equality of the Kantorovich Problem and its dual. The setting is as follows, we are given X, Y metric spaces, $c : X \times Y \mapsto \mathbb{R}$ continuous, $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$.

Notice that for every $\phi \in C(X)$ and $\psi \in C(Y)$ with $\phi(x) + \psi(y) \leq c(x, y)$ and for every measure $\gamma \in \Pi(\mu, \nu)$ we have from Propositions 2.2.4, and 3.2.4

$$I(\phi, \psi) = \int_X \phi(x)\, d\mu + \int_Y \psi(y)\, d\nu \int_{X \times Y} \phi(x) + \psi(y)\, d\gamma \leq \int_{X \times Y} c(x, y)\, d\gamma.$$

and hence

$$(DP) \leq (KP). \tag{3.3}$$

Definition 3.5.1. *Let $\mathcal{X}$ be a metric space and $\eta \in \mathcal{P}(\mathcal{X})$ the support of η is defined as follows*

$$spt(\eta) = \{x \in \mathcal{X} : \eta(B(x, r)) > 0 \, for \; all \; r > 0\}.$$

Theorem 3.5.2. *Let $\gamma \in \Pi(\mu, \nu)$ be an optimal transport plan, i.e. a solution to the corresponding (KP) problem, then $spt(\gamma)$ is a c-CM set.*

Proof. Suppose that there exist $k \in \mathbb{N}$, a permutation σ, and $(x_1, y_1), \cdots, (x_k, y_k) \in spt(\gamma)$ such that

$$\sum_{i=1}^{k} c(x_i, y_i) > \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}).$$

Take now $0 < \epsilon < \frac{1}{2k} \left(\sum_{i=1}^{k} c(x_i, y_i) - \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}) \right)$. Since c is continuous then there exists $r > 0$ such that for all $i = 1, \cdots, k$ we have and for all $(x, y) \in B(x_i, r) \times B(y_i, r)$ we have

$$c(x, y) > c(x_i, y_i) - \epsilon \qquad \forall (x, y) \in B(x_i, r) \times B(y_i, r), \tag{3.4}$$

$$c(x, y) < c(x_i, y_{\sigma(i)}) + \epsilon \quad \forall (x, y) \in B(x_i, r) \times B(y_{\sigma(i)}, r). \tag{3.5}$$

We will construct a measure $\tilde{\gamma} \in \Pi(\mu, \nu)$ such that $\displaystyle\int_{X \times Y} c\, d\tilde{\gamma} < \int_{X \times Y} c\, d\gamma$, obtaining then a contradiction since γ is given to be an optimal transport plan. Denote $V_i := B(x_i, r) \times B(y_i, r)$, notice that $\gamma(V_i) > 0$ for every i because $(x_i, y_i) \in spt(\gamma)$. Define the measures $\gamma_i \in \mathcal{P}(X \times Y)$ to be the normalized restriction of γ on V_i, that is,

$$\gamma_i(A) = \frac{\gamma(A \cap V_i)}{\gamma(V_i)} \qquad \forall A \in \mathfrak{B}(X \times Y),$$

and let $\mu_i = (\pi_1)_{\#}\gamma_i$ and $\nu_i = (\pi_2)_{\#}\gamma_i$. Take also the product measures $\tilde{\gamma}_i = \mu_i \otimes \nu_{\sigma(i)}$. We then take

$$\tilde{\gamma} = \gamma - \frac{\min_i \gamma(V_i)}{2k} \sum_{i=1}^{k} (\gamma_i - \tilde{\gamma}_i).$$

Notice that for $A \in \mathfrak{B}(X \times Y)$ we have

$$\tilde{\gamma}(A) \geq \gamma(A) - \frac{\min_i \gamma(V_i)}{2k} \sum_{i=1}^{k} \frac{\gamma(A \cap V_i)}{\gamma(V_i)} \geq \gamma(A) - \frac{1}{2k} \sum_{i=1}^{k} \gamma(A) = \frac{1}{2}\gamma(A) \geq 0,$$

and $\gamma(X \times Y) = 1$, then $\gamma \in \mathcal{P}(X \times Y)$. Moreover, since $\gamma \in \Pi(\mu, \nu)$, $\gamma_i \in \Pi(\mu, \nu)$, $\tilde{\gamma}_i \in \Pi(\mu_i, \nu_{\sigma(i)})$ then

$$(\pi_1)_{\#}\tilde{\gamma} = \mu - \frac{\min \gamma(V_i)}{2k} \sum_{i=1}^{k} (\mu_i - \mu_i) = \mu$$
$$(\pi_2)_{\#}\tilde{\gamma} = \nu - \frac{\min \gamma(V_i)}{2k} \sum_{i=1}^{k} (\nu_i - \nu_{\sigma(i)}) = \nu - \frac{\min \gamma(V_i)}{2k} \left(\sum_{i=1}^{k} \nu_i - \sum_{i=1}^{k} \nu_{\sigma(i)} \right) = \nu$$

where the last equality is due to the fact that σ is a permutation.

Finally, since γ_i is supported on $B(x_i, r) \times B(y_i, r)$ and $\tilde{\gamma}_i$ on $B(x_i, r) \times B(y_{\sigma(i)}, r)$, it follows from inequalities (3.4) and (3.5)

$$\begin{aligned}
\int_{X \times Y} c\, d\gamma - \int_{X \times Y} c\, d\tilde{\gamma} &= \frac{\min_i \gamma(V_i)}{2k} \sum_{i=1}^{k} \left(\int_{X \times Y} c\, d\gamma_i - \int_{X \times Y} c\, d\tilde{\gamma}_i \right) \\
&= \frac{\min_i \gamma(V_i)}{2k} \sum_{i=1}^{k} \left(\int_{B(x_i, r) \times B(y_i, r)} c\, d\gamma_i - \int_{B(x_i, r) \times B(y_{\sigma(i)}, r)} c\, d\tilde{\gamma}_i \right) \\
&\geq \frac{\min_i \gamma(V_i)}{2k} \left(\sum_{i=1}^{k} (c(x_i, y_i) - \epsilon) - \sum_{i=1}^{k} (c(x_i, y_{\sigma(i)}) + \epsilon) \right) \\
&= \frac{\min_i \gamma(V_i)}{2k} \Big(\sum_{i=1}^{k} c(x_i, y_i) - \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}) - 2k\epsilon \Big) > 0.
\end{aligned}$$

Hence, a contradiction so *spt*(γ) is a c-CM set. □

Theorem 3.5.3. *Let X and Y be compact metric spaces, $c : X \times Y \mapsto \mathbb{R}$ continuous, $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, then*

$$(KP) = (DP).$$

Proof. From Theorem 2.3.3, there exists γ_0 solution to the (KP) problem. By Theorem 3.5.2, *spt*(γ_0) is $c-$CM then from Theorem 3.4.3 there exists $\phi_0 \in c-\text{conc}(X)$ such that

$$spt(\gamma_0) \subseteq \{(x, y) \in X \times Y : \phi_0(x) + \phi_0^c(y) = c(x, y)\} \tag{3.6}$$

By Lemma 3.3.2, ϕ_0 and ϕ_0^c are continuous and from Proposition 3.2.4 $\phi_0(x) + \phi_0^c(y) \leq c(x,y)$ for all $(x,y) \in X \times Y$ therefore from Proposition 2.2.4

$$
\begin{aligned}
(KP) &= \inf\left\{\int_{X\times Y} c(x,y)\, d\gamma : \gamma \in \Pi(\mu,\nu)\right\} \\
&= \int_{X\times Y} c(x,y)\, d\gamma_0 = \int_{X\times Y} \phi_0(x) + \phi_0^c(y)\, d\gamma_0 = \int_X \phi_0(x)\, d\mu + \int_Y \phi_0^c(y)\, d\nu \\
&\leq \sup\{I(\phi,\psi) : \phi \in C(X), \psi \in C(Y) \text{ and } \phi(x) + \psi(y) \leq c(x,y)\} = (DP).
\end{aligned}
$$

Hence by inequality (3.3), we get

$$(KP) = \int_{X\times Y} c(x,y)\, d\gamma_0 = I(\phi_0, \phi_0^c) = (DP).$$

□

Remark 3.5.4. Notice that above proof can be used to show that every measure $\gamma \in P(X \times Y)$ with a $c-$CM support is an optimal transport plan with respect to the measures $\mu = (\pi_1)_{\#}\gamma$ and $\nu = (\pi_2)_{\#}\gamma$, this corresponds to the converse of Theorem 3.5.2.

We can also infer that every transport plan γ_0 solution to the (KP) problem induces a Kantorovich potential $\phi_0 \in c\text{-conc}(X)$ solution to (DP), and satisfying (3.6). We end this chapter by showing that (3.6) is true for every Kantorovich potential ϕ.

Proposition 3.5.5. *Given the setting of Theorem 3.5.3, let $\gamma \in \Pi(\mu,\nu)$ be an optimal transport plan and ϕ be a Kantorovish potential then*

$$\phi(x) + \phi^c(y) = c(x,y) \qquad\qquad a.e\, \gamma$$

Proof. Since $(KP) = (DP)$ then

$$\int_{X\times Y} c(x,y)\, d\gamma = I(\phi,\phi^c) = \int_X \phi(x)\, d\mu + \int_Y \phi^c(y)\, d\nu = \int_{X\times Y} (\phi(x) + \phi^c(y))\, d\gamma,$$

and so

$$\int_{X\times Y} c(x,y) - (\phi(x) + \phi^c(y))\, d\gamma = 0$$

The results hence follows from Proposition 3.2.4. □

Chapter 4

The Monge Problem

We are now ready to study the Monge Problem presented in the introduction. In other words, given an optimal transport plan corresponding to a cost function $c : X \times Y \mapsto \mathbb{R}$, does it induce a map $T : X \mapsto Y$ minimizing the transport cost? The problem can be formulated as follows, let X and Y be two metric spaces with corresponding probability measures μ and ν and $c : X \times Y \mapsto \mathbb{R}$ continuous, we define the Monge problem as follows

$$(MP) : \inf \left\{ \int_X c(x, T(x))\, d\mu(x), T : X \mapsto Y \text{ and } T_{\#}\mu = \nu \right\}. \tag{4.1}$$

Solutions to the (MP) problems are called optimal transport maps. As in the previous chapters, we shall find sufficient conditions for existence of these maps, and for obtaining $(MP) = (KP) = (DP)$.

We make the following observation, that justifies why the Monge Problem is considered to be a relaxation of the Kantorovich problem.

Theorem 4.0.1. *Given X and Y two compact metric spaces with corresponding probability measures μ and ν, we are given a continuous function $c : X \times Y \mapsto \mathbb{R}$, and a measurable map $T : X \mapsto Y$ with $T_{\#}\mu = \nu$. Define $\gamma_T = (id, T)_{\#}\mu$, that is*

$$\gamma_T(A \times B) = \mu(A \cap T^{-1}B) \qquad \forall A \times B \subseteq X \times Y.$$

then $\gamma_T \in \Pi(\mu, \nu)$ and

$$\int_X c(x, Tx)\, d\mu = \int_{X \times Y} c(x, y)\, d\gamma. \tag{4.2}$$

Proof. Let $A \subseteq X$ and $B \subseteq Y$ then

$$\gamma_T(A \times Y) = \mu(T^{-1}Y \cap A) = \mu(X \cap A) = \mu(A)$$
$$\gamma_T(X \times B) = \mu(T^{-1}B \cap X) = \mu(T^{-1}B) = \nu(B).$$

Hence, by Proposition 2.2.2 $\gamma_T \in \Pi(\mu,\nu)$. Now let's prove that

$$\int_X c(x,Tx)\,d\mu = \int_{X\times Y} c(x,y)\,d\gamma_T.$$

Since X and Y are two compact spaces and c is continuous then there exists c_n distincts values in $\mathbb{R}$, $A_i \times B_i \in \mathfrak{B}(X\times Y)$ disjoints such that

$$c(x,y) = \sum_{i=1}^{\infty} c_i \cdot \chi_{A_i\times B_i}(x,y)$$

$$\int_{X\times Y} c(x,y)\,d\gamma_T = \int_{X\times Y} \sum_{i=1}^{\infty} c_i \cdot \chi_{A_i\times B_i}(x,y)\,d\gamma_T = \sum_{i=1}^{\infty} c_i \cdot \int_{X\times Y} \chi_{A_i\times B_i}(x,y)\,d\gamma_T$$

(Fubini's theorem having $\sum_{i=1}^{\infty} |c_i\chi_{A_i\times B_i}| \leq \sup|c| < \infty$) Then,

$$\int_{X\times Y} c(x,y)\,d\gamma_T = \sum_{i=1}^{\infty} c_i\gamma_T(A_i\times B_i) = \sum_{i=1}^{\infty} c_i\mu(A_i\cap T^{-1}B_i).$$

While,

$$\int_X c(x,Tx)\,d\mu = \int_X \sum_{i=1}^{\infty} c_i\chi_{A_i\times B_i}(x,Tx)\,d\mu = \sum_{i=1}^{\infty} c_i \int_X \chi_{A_i\cap T^{-1}B_i}(x)\,d\mu = \sum_{i=1}^{\infty} c_i\mu(A_i\cap T^{-1}B_i).$$

We hence conclude (4.2). □

This imply for every T such that $T_{\#}\mu = \nu$ we can associate a $\gamma_T \in \Pi(\mu,\nu)$ such that $\gamma_T = (id,T)_{\#}\mu$ and satisfying (4.2). In other words, the space of transport plan is larger than the space of transport maps concluding that

$$(KP) \leq (MP).$$

Notice that the result in Theorem 4.0.1 is still valid for X and Y not necessarily compact as long as the function c is continuous and bounded.

4.1 Preliminary examples for inexistence

4.1.1 *Example 1. Nonexistence Case of atomic measure*

Let X and Y be metric spaces and $a \in X$. Take μ the Dirac mass at a that is such that for $A \subseteq X$

$$\mu(A) = \delta_a(A) = \begin{cases} 1 & \text{if } a \in A \\ 0 & \text{if } a \notin A, \end{cases}.$$

Notice that for every Borel map $T : X \mapsto Y$ we have $T_{\#}\delta_a = \delta_{T(a)}$. Let $B \subseteq Y$ in fact for every $B \subseteq Y$

$$(T_{\#}\delta_a)(B) = \delta_a(T^{-1}(B)) = \begin{cases} 1 & \text{if } a \in T^{-1}(B) \\ 0 & \text{if } a \notin T^{-1}(B) \end{cases} = \begin{cases} 1 & \text{if } T(a) \in B \\ 0 & \text{if } T(a) \notin B \end{cases} = \delta_{T(a)}(B)$$

Therefore if ν is not a dirac mass then one can't find a transport map such that $T_{\#}\mu = \nu$, and hence no optimal transport map exist in this case.

However, an optimal transport plan exists for any Borel measure ν and continuous cost c, since in this case we'll prove that $\Pi(\mu, \nu)$ contains only the canonical product measure. Take $\gamma \in \Pi(\delta_a, \nu)$, $A \in \mathfrak{B}(X)$ and $B \in \mathfrak{B}(Y)$. If $a \notin A$, then from Proposition 2.2.2

$$\gamma(A \times B) \leq \gamma(A \times Y) = \delta_a(A) = 0,$$

and so $\gamma(A \times B) = 0 = \delta_a(A)\nu(B)$. On the other hand, if $a \in A$, then using again Proposition 2.2.2 and the above result

$$\gamma(A \times B) = \gamma(X \times B) - \gamma(A^c \times B) = \nu(B) = \delta_a(A)\nu(B).$$

4.1.2 *Example 2. Nonexistence case of quadratic cost*

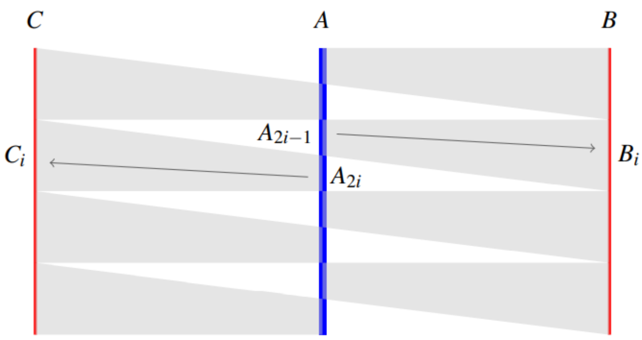

Figure 4.1: Example 2 figure [18]

We consider the quadratic cost $c(x, y) = |x - y|^2$, and $X = A$, $Y = B \cup C$ where A, B and C are three vertical parallel segments in $\mathbb{R}^2$ whose vertices lie on the two lines $y = 0$ and $y = 1$ and the abscissas are $0, 1$ and -1 respectively as shown in Figure 4.1.2. We define the following two measures on X and Y

$$\mu(E) = \mathcal{H}^1(E \cap A) \qquad \forall E \in \mathfrak{B}(A) \qquad \text{and}$$
$$\nu(F) = \frac{\mathcal{H}^1(F \cap B) + \mathcal{H}^1(F \cap C)}{2} \qquad \forall F \in \mathfrak{B}(B \cup C)$$

where $\mathcal{H}^1$ corresponds to the one-dimensional Hausdorff measure.

Notice that for all $T : A \mapsto B \cup C$ we have $|x - Tx|^2 \geq 1$. This is clear since $d(A, B \cup C) = 1$ then

$$\int_A c(x, Tx)\, d\mu = \int_A |x - Tx|^2\, d\mu \geq \int_A 1\, d\mu = \mu(A) = 1.$$

then

$$\inf \left\{ \int_A c(x, Tx)\, d\mu : T_{\#}\mu = \nu \right\} \geq 1.$$

We prove that the infimum is actually equal to 1.

We construct a sequence $T_n : A \mapsto B \cup C$ as follows. Divide A into $2n$ equal segments $(A_i)_{i=1,\cdots,2n}$, B into $n-$segments $(B_i)_{i=1,\cdots,n}$, and C into $n-$segments $(C_i)_{i=1,\cdots,n}$, T_n is an affine map such that

$$A_{2i-1} \mapsto B_i \text{ and } A_{2i} \mapsto C_i$$

Notice from the geometry of the figure for all $x \in A$

$$1 \leq |x - T_n(x)|^2 \leq |(0,0) - (-1, 1/n)|^2 = 1 + \frac{1}{n^2},$$

letting $n \to \infty$ we get that $\displaystyle\int_A c(x, T_n x)\, d\mu \to 1$.

Notice that the length of the subintervals A_i is doubled when mapped by T_n, so using the definition of μ and ν we get that for every $F \subseteq (B \cup C)$

$$\begin{aligned}
\mu(T_n^{-1}F) &= \mu\left(T_n^{-1}(F \cap B) \cup (F \cap C)\right) \\
&= \mu\left(T_n^{-1}(F \cap B)\right) + \mu\left(T_n^{-1}(F \cap C)\right) \\
&= \mathcal{H}^1(T_n^{-1}(F \cap B)) + \mathcal{H}^1(T_n^{-1}(F \cap C)) \\
&= \frac{1}{2}\mathcal{H}^1(F \cap B) + \frac{1}{2}\mathcal{H}^1(F \cap C) \\
&= \nu(F)
\end{aligned}$$

As a result, we conclude that

$$\inf \left\{ \int_A c(x, Tx)\, d\mu, T_{\#}\mu = \nu \right\} = 1.$$

However, there is no optimal transport map for which the infimum is attained. In fact, if $T : A \mapsto B \cup C$ is such that $\int_A |x - Tx|^2\, d\mu = 1$ then $\int_A |x - Tx|^2 - 1\, d\mu = 0$ but the integrand is non negative then $|x - Tx|^2 = 1$ a.e μ, and so T moves horizontally. Hence for every $E \subseteq B$

$$2\nu(E) = \mathcal{H}^1(E) = \mathcal{H}^1(T^{-1}E) = \mu(T^{-1}E),$$

and so $T_{\#}\mu \neq \nu$, that is the map T does not belong to the class of admissible maps.

We prove however that an optimal transport plan, i.e. a solution to the corresponding (KP) problem exists. As above we have for every $\gamma \in \Pi(\mu, \nu)$, $\displaystyle\int_{A\times(B\cup C)} |x-y|^2\, d\gamma \geq 1$, and so $(KP) \geq 1$. Construct the optimal transport plan as follows. Let $T_+ : A \mapsto B$ moving horizontally to the right i.e. $T_+x = x + (1,0)$. Notice that for every $U \subseteq B$

$$\mathcal{H}^1(U) = \mathcal{H}^1(T_+^{-1}U) = \mu(T_+^{-1}U)$$

Then, $(T_+)_{\#}\mu = \mathcal{H}^1|_B$. We then define as in Theorem 4.0.1 the associated probability measure on $A \times B$, $\gamma_{T_+} \in \Pi(\mu, \mathcal{H}^1|_B)$.

Similarly, let $T_- : A \mapsto C$ moving horizontally to the left i.e. $T_-x = x + (-1, 0)$. Notice that for every $V \subseteq C$

$$\mathcal{H}^1(V) = \mathcal{H}^1(T_-^{-1}V) = \mu(T_-^{-1}V)$$

Then, $(T_-)_{\#}\nu = \mathcal{H}^1|_C$ and the associated measure $\gamma_{T_-} \in \Pi(\mu, \mathcal{H}^1|_C)$. Let $\gamma = \frac{1}{2}\gamma_{T_+} + \frac{1}{2}\gamma_{T-}$ a measure on $A \times (B \cup C)$ that is

$$\gamma(E\times F) = \frac{1}{2}\gamma_{T_+}(E\times(F\cap B)) + \frac{1}{2}\gamma_{T_-}(E\times(F\cap C), \qquad \forall E\times F \subseteq A\times(B\cup C).$$

Let's check that $\gamma \in \Pi(\mu, \nu)$ using Proposition 2.2.2. For $F \subseteq B \cup C$,

$$\begin{aligned}\gamma(A\times F) &= \frac{1}{2}\gamma_{T_+}(A\times(F\cap B)) + \frac{1}{2}\gamma_{T_-}(A\times(F\cap C)) \\ &= \frac{1}{2}\mathcal{H}^1(F\cap B) + \frac{1}{2}\mathcal{H}^1(F\cap C) \\ &= \nu(F)\end{aligned}$$

For $E \subseteq A$,

$$\begin{aligned}\gamma(E\times(B\cup C)) &= \frac{1}{2}\gamma_{T_+}(E\times B) + \frac{1}{2}\gamma_{T_-}(A\times C) \\ &= \frac{1}{2}\mu(E) + \frac{1}{2}\mu(E) \\ &= \mu(E)\end{aligned}$$

Finally, we show that γ is optimal. From (4.2) applied to γ_{T_+} and γ_{T_-}, we have

$$\begin{aligned}\int_{A\times(B\cup C)} |x-y|^2\, d\gamma &= \frac{1}{2}\int_{A\times B} |x-y|^2\, d\gamma_{T_+} + \frac{1}{2}\int_{A\times C} |x-y|^2\, d\gamma_{T_-} \\ &= \frac{1}{2}\int_A |x-T_+x|^2\, d\mu + \frac{1}{2}\int_A |x-T_-x|^2\, d\mu \\ &= \frac{1}{2}\int_A 1\, d\mu + \frac{1}{2}\int_A 1\, d\mu \\ &= \mu(A) = 1.\end{aligned}$$

Hence, we're done.

4.2 Measure preserving maps

Definition 4.2.1. *Let X and Y be measure spaces, let $N : X \mapsto 2^Y$ be a multivalued map such that $N(x) \subseteq Y$ for each $x \in X$. For $F \subseteq Y$ we denote*

$$N^{-1}(F) = \{x \in X : N(x) \cap F \neq \emptyset\}$$

Let $\mu \in \mathcal{P}(X)$. Assume N satisfies the following

1. *N is measurable i.e $N^{-1}(F) \in \mathfrak{B}(X)$ for every $F \in \mathfrak{B}(Y)$*
2. *The set $\{x \in X : N(x)$ is not a singleton$\}$ has μ-measure zero.*

Given $\nu \in \mathcal{P}(Y)$, we say that N is measure preserving from μ to ν if and only if $N_{\#}\mu = \nu$.

Example 4.2.2. *Given $\Omega \subseteq \mathbb{R}^n$ open, let $\phi : \Omega \mapsto \mathbb{R}$ be a strictly convex C^1 function. In this case, we have that $\nabla\phi$ is injective. In fact, suppose that there exist $x, y \in \Omega$ such that $\nabla\phi(x) = \nabla\phi(y)$ where $x \neq y$. We have by strict convexity*

$$\nabla\phi(x) \cdot (y - x) < \phi(y) - \phi(x) \tag{4.3}$$

Similarly, switching the roles of y and x, we get $\nabla\phi(y) \cdot (x - y) < \phi(x) - \phi(y)$. Having $\nabla\phi(x) = \nabla\phi(y)$ we get

$$\nabla\phi(x) \cdot (y - x) > \phi(y) - \phi(x)$$

which contradicts (4.3) hence $x = y$. Now, let $\Omega^* = \nabla\phi(\Omega)$. We have $\nabla\phi : \Omega \mapsto \Omega^*$ is bijective, we then define $N = (\nabla\phi)^{-1} : \Omega^* \mapsto \Omega$. N is single valued, and $N^{-1} = \nabla\phi$ is measurable as $\nabla\phi$ is continuous and injective and hence open.

Take a probability measure μ on Ω^*, and define the Borel measure ν on Ω as follows

$$\nu(E) = \mu(\nabla\phi(E)),$$

then N is a measure preserving map from from μ to ν.

Proposition 4.2.3. *Given two compact metric spaces X and Y with corresponding probability measures μ and ν, let $N : X \mapsto Y$ be a map satisfying the above conditions (1) and (2). N is measure preserving from μ to ν if and only if*

$$\int_X v(N(x))\, d\mu = \int_Y v(y)\, d\nu \tag{4.4}$$

for each $v \in C(Y)$.

Proof. Assume N is measure preserving. Let $A \in \mathfrak{B}(Y)$, notice that from (2) $\chi_A(N(x)) = \chi_{N^{-1}(A)}(x)$ for $\mu-$a.e x. Therefore, since N is measure preserving

$$\int_X \chi_A(N(x))\, d\mu = \int_X \chi_{N^{-1}A}(x)\, d\mu = \mu(N^{-1}A) = \nu(A) = \int_Y \chi_A(y)\, d\nu$$

obtaining hence (4.4) for characteristic functions.

Consider $v \in C(Y)$, since Y is compact then there exists Borel disjoint sets A_n and real numbers a_n such that

$$v(y) = \sum_{n=1}^{\infty} a_n \chi_{A_n}(y),$$

Moreover, by Fubini's theorem having $\sum_{n=1}^{\infty} |a_n \chi_{A_n}| \le \sup |v| < \infty$ and by the above

$$\begin{aligned}\int_Y v(y)\, d\nu = \int_Y \sum_{n=1}^{\infty} a_n \chi_{A_n}(y)\, d\nu &= \sum_{n=1}^{\infty} a_n \int_Y \chi_{A_n}(y)\, d\nu \\ &= \sum_{n=1}^{\infty} a_n \int_X \chi_A(N(x))\, d\nu(x) = \int_X v(N(x))\, d\mu.\end{aligned}$$

To prove the converse, let's prove for a borel set E the following inequality

$$N_{\#}\mu(E) \le \nu(E) \tag{4.5}$$

Let G be an open set in Y and $K \subseteq G$ compact then by Urysohn's lemma there exist a function $v \in C(X)$ such that $0 \le v \le 1, v = 0$ outside G and $v = 1$ on K. Notice that from the construction of v we have $\int_X \chi_K(N(x))\, d\mu \le \int_X v(N(x))\, d\mu$ Hence, using the fact that $\chi_K(N(x)) = \chi_{N^{-1}(K)}(x)$, and the above inequality we have

$$N_{\#}\mu(K) = \int_X \chi_{N^{-1}(K)}(x)\, d\mu = \int_X \chi_K(N(x))\, d\mu \le \int_X v(N(x))\, d\mu$$

Moreover, using the construction of v and (4.4) we have

$$\int_X v(N(x))\, d\mu = \int_Y v(y)\, d\nu \le \nu(G).$$

Therefore, $N_{\#}\mu(K) \le \nu(G)$. From condition 1, $N_{\#}\mu$ is a probability measure on the compact space Y then by regularity

$$N_{\#}\mu(G) = \sup_{K \subseteq G, K compact} N_{\#}\mu(K) \le \nu(G).$$

Now, let E be any Borel set, ν is a probability measure on the compact space Y then ν is regular so there exists an open set U where $E \subseteq U$ and $\nu(U \setminus E) < \epsilon$. Using the inequality on U we get

$$N_{\#}\mu(E) \leq N_{\#}\mu(U) \leq \nu(U) < \nu(E) + \epsilon$$

Let $\epsilon \to 0$ then $N_{\#}\mu(E) \leq \nu(E)$.

Now we prove the reverse inequality. First notice that for $F \subseteq Y$

$$\{x \in X : N(x) \neq \emptyset\} \cap (N^{-1}(F))^c \subseteq N^{-1}(F^c)$$

and

$$(N^{-1}(F))^c =$$
$$\big((N^{-1}(F))^c \cap \{x \in X : N(x) \neq \emptyset\}\big) \bigcup \big((N^{-1}(F))^c \cap \{x \in X : N(x) = \emptyset\}\big)$$

From condition (2), $\mu(\{x \in X : N(x) = \emptyset\}) = 0$ then using (4.5) on F^c we obtain

$$\mu((N^{-1}(F))^c) = \mu((N^{-1}(F))^c \cap \{x \in X : N(x) \neq \emptyset\}) \leq \mu(N^{-1}(F^c)) \leq \nu(F^c)$$

then $\mu(N^{-1}(F)) \geq \nu(F)$ for any F borel set in Y. Therefore, N is a measure preserving from μ to ν. □

4.3 (MP)=(DP)=(KP)

Definition 4.3.1. *Consider X and Y two metric spaces and $\phi : X \mapsto \mathbb{R}$. Let $x \in X$, we define the $c-$normal mapping of ϕ to be*

$$N_{c,\phi}(x) = \{y \in Y : \phi(x) + \phi^c(y) = c(x, y)\} \tag{4.6}$$

Example 4.3.2. *Let $X = Y = B(0, 1)$ be the unit ball in $\mathbb{R}^n$, $c(x, y) = |x - y|$, and $\phi(x) = -|x|$. We have to find ϕ^c, notice that, by definition, we have for $y \in Y$,*

$$\phi^c(y) = \inf_{x \in B(0,1)} \{c(x, y) + |x|\}.$$

From the triangular inequality, we have $|x - y| + |x| \geq |y|$, and we have equality when $x = y$, then $\phi^c(y) = |y|$, concluding that

$$N_{c,\phi}(x) = \{y \in B(0, 1) : |y| = |x - y| + |x|\}.$$

For $x = 0$, we have $N_{c,\phi}(0) = B(0, 1)$. For $x \neq 0$ in $B(0, 1)$, we have that $y \in N_{c,\phi}(x)$ if and only if $y = \lambda x$ for some $\lambda > 1$. Therefore

$$N_{c,\phi}(x) = \{\lambda x : \lambda > 1, \text{and } |\lambda x| < 1\} = \left\{\lambda x : 1 \leq \lambda \leq \frac{1}{|x|}\right\}.$$

Notice that $N_{c,\phi}(x)$ is not single valued for any $x \in B(0, 1)$ so it violates (1) for any Borel measure μ.

Example 4.3.3. *Given $X = Y = \mathbb{R}^n, \phi : X \mapsto \mathbb{R}$ and $c : X \times Y \mapsto \mathbb{R}$ such that $c(x, y) = x \cdot y$ then $N_{c,\phi}(x) = \partial^*\phi(x)$ where $\partial^*\phi$ is the super-differential of ϕ*

$$\partial^*\phi(x) = \{m \in \mathbb{R}^n : \phi(y) \leq \phi(x) + m \cdot (y - x)\, \forall y \in Y\}.$$

In fact, $m \in \partial^*\phi(x)$, if and only if for every $y \in Y$ we have

$$\phi(y) \leq \phi(x) + m \cdot (y - x),$$

i.e. $m \cdot x - \phi(x) \leq m \cdot y - \phi(y)$. Equivalently,

$$\phi^c(m) = \inf_{y \in \mathbb{R}^n} \{m \cdot y - \phi(y)\} = m \cdot x - \phi(x),$$

and our claim follows.

One can show that if ϕ is a $c-$concave function then $\partial^*\phi$ satisfies Condition (1), and (2) with μ being the Lebesgue measure on $\mathbb{R}^n$. [9]

Lemma 4.3.4. *Given X, Y two compact metric spaces, $c : X \times Y \mapsto \mathbb{R}$ such that $N_{c,\phi}$ satisfies (1) and (2) for every ϕ $c-$concave then*

1. *If ϕ is $c-$concave and $N_{c,\phi}$ is measure preserving from μ to ν, then ϕ is a Kantorovich potential.*

2. *If ϕ is $c-$concave and ϕ is a Kantorovich potential then $N_{c,\phi}$ is measure preserving from μ to ν.*

Proof of (1). Let $u \in C(X)$ and $v \in C(Y)$ such that $u(x)+v(y) \leq c(x, y)\, \forall (x, y) \in X \times Y$. From (2), we have that for $\mu-$a.e $x \in X$

$$u(x) + v(N_{c,\phi}(x)) \leq c(x, N_{c,\phi}(x)).$$

By using the Definition 3.2.1 of $I(u, v)$ and Proposition 4.2.3 we have

$$I(u, v) = \int_X u(x)\, d\mu + \int_Y v(y)\, d\nu = \int_X u(x) + v(N_{c,\phi}(x))\, d\mu \leq \int_X c(x, N_{c,\phi}(x))\, d\mu.$$

We also have from the definition of $N_{c,\phi}$ that $c(x, N_{c,\phi}(x)) = \phi(x) + \phi^c(N_{c,\phi}(x))$ for $\mu-$almost every $x \in X$, then similarly by Proposition 4.2.3

$$I(\phi, \phi^c) = \int_X c(x, N_{c,\phi}(x))\, d\mu \geq I(u, v).$$

Hence ϕ is a Kantorovich potential. □

Proof of (2). Let $\psi = \phi^c$, for $v \in C(Y)$ and $\theta \in \mathbb{R}$, we denote

$$\psi_\theta(y) = \psi(y) + \theta v(y),$$

and $\phi_\theta = (\psi_\theta)^{\bar{c}}$.

We first prove that

$$\lim_{\theta\to 0} \frac{I(\phi_\theta,\psi_\theta)-I(\phi,\psi)}{\theta} = \int_X -v(N_{c,\phi}(x))\,d\mu + \int_Y v(y)\,d\nu \tag{4.7}$$

In fact,

$$\frac{I(\phi_\theta,\psi_\theta)-I(\phi,\psi)}{\theta} = \int_X \frac{\phi_\theta-\phi}{\theta}\,d\mu + \int_Y \frac{\psi_\theta-\psi}{\theta}\,d\nu = \int_X \frac{\phi_\theta-\phi}{\theta}\,d\mu + \int_Y v(y)\,d\nu.$$

To apply Lebesgue Dominated Convergence Theorem we have to show that $\dfrac{\phi_\theta-\phi}{\theta}$ is uniformly bounded in X for $|\theta|$ close to 0 and $\dfrac{\phi_\theta(x)-\phi(x)}{\theta} \to -v(N_{c,\phi}(x))$ $\mu-$a.e $x\in X$ as $\theta\to 0$.

Let S the set of singular point that is $S=\{x\in X : N_{c,\phi}(x) \text{ is not a singleton}\}$. We take $x\in X\setminus S$. Denote $y_1 = N_{c,\phi}(x)$, then $\phi(x) = c(x,y_1)-\psi(y_1)$, and

$$\phi_\theta(x) = \inf_y\{c(x,y)-\psi_\theta(y)\} \le c(x,y_1)-\psi_\theta(y_1) = c(x,y_1)-\psi(y_1)-\theta v(y_1) = \phi(x)-\theta v(y_1).$$

On the other hand, we have $\phi_\theta(x) = (\psi_\theta)^{\bar{c}}(x) = \inf_y\{c(x,y)-\psi_\theta(y)\}$. By compactness of Y and continuity of c and ϕ there exists $y_\theta\in Y$ such that

$$\phi_\theta(x) = c(x,y_\theta)-\psi_\theta(y_\theta) = c(x,y_\theta)-\psi(y_\theta)-\theta v(y_\theta) \ge \psi^{\bar{c}}(x)-\theta v(y_\theta) = \phi(x)-\theta v(y_\theta),$$

where for the last equality we used by Proposition 3.2.7. We conclude

$$-\theta v(y_\theta) \le \phi_\theta(x)-\phi(x) \le -\theta v(y_1).$$

We deduce that $\left|\dfrac{\phi_\theta-\phi}{\theta}\right|$ is uniformly bounded and $\lim_{\theta\to 0}\phi_\theta(x) = \phi(x)$. It remains to show that $y_\theta \to y_1 = N_{c,\phi}(x)$.

Suppose there exists a subsequence y_{θ_k} such that $y_{\theta_k}\to y_0 \neq y_1$, as $k\to\infty$. We have that

$$\phi(x) = \lim_{k\to\infty}\phi_{\theta_k}(x) = c(x,y_0)-\psi(y_0)$$

which implies that $y_0\in N_{c,\phi}(x)$, but $x\notin S$ then $N_{c,\phi}(x)$ is a singleton hence $y_0=y_1$ which is a contradiction. We obtain then (4.7).

Since ϕ is a Kantorovich Potential then $I(\phi_\theta,\psi_\theta)\le I(\phi,\psi)$, and so

$$\lim_{\theta\to 0^+}\frac{I(\phi_\theta,\psi_\theta)-I(\phi,\psi)}{\theta}\le 0 \qquad \lim_{\theta\to 0^-}\frac{I(\phi_\theta,\psi_\theta)-I(\phi,\psi)}{\theta}\ge 0,$$

but since we proved that the limit exists we obtain

$$\lim_{\theta\to 0}\frac{I(\phi_\theta,\psi_\theta)-I(\phi,\psi)}{\theta} = 0$$

which implies that

$$\int_X v(N_{c,\phi}(x))\,d\mu = \int_Y v(y)\,d\nu$$

Hence, by Proposition 4.2.3 $N_{c,\phi}$ is measure preserving from μ to ν. □

Theorem 4.3.5. *Let X and Y be two compact metric spaces, $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$. Assume $c : X \times Y \mapsto \mathbb{R}$ continuous such that $N_{c,\phi}$ satisfies (1) and (2) for every $c-$concave function ϕ.*

Let ϕ_0 be a Kantorovich potential, then $T_0(x) = N_{c,\phi_0}(x)$ is a solution to the corresponding (MP) problem, and we have

$$(MP) = (DP).$$

Proof. ϕ_0 is a Kantorovich potential then from Lemma 4.3.4 $T_0 = N_{c,\phi_0}$ is measure preserving from μ to ν.

Let $T : X \mapsto Y$ be such that $T_{\#}\mu = \nu$ then

$$\begin{aligned}
\int_X c(x, T(x))\, d\mu &\geq \int_X \phi_0(x) + \phi_0^c(T(x))\, d\mu \\
&= \int_X \phi_0(x)\, d\mu + \int_X \phi_0^c(T(x))\, d\mu \\
&= \int_X \phi_0(x)\, d\mu + \int_Y \phi_0^c(y)\, d\nu \text{ (by Proposition 4.2.3)} \\
&= I(\phi_0, \phi_0^c) \\
&= \int_X \phi_0(x) + \phi_0^c(N_{c,\phi_0}(x))\, d\mu \text{ (by Proposition 4.2.3)} \\
&= \int_X c(x, N_{c,\phi_0}(x))\, d\mu
\end{aligned}$$

Then,

$$(MP) = \int_X c(x, T_0(x))\, d\mu(x) = I(\phi_0, \phi_0^c) = (DP).$$

□

Remark 4.3.6. Let X and Y be two compact metric spaces, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ and $c : X \times Y \mapsto \mathbb{R}$ be continuous function. We obtain that

$$\begin{aligned}
\min &\left\{ \int_{X\times Y} c(x, y)\, d\gamma : \gamma \in \Pi(\mu, \nu) \right\} \\
&= \max \{ I(u, v) : u(x) + v(y) \leq c(x, y) \} \\
&= \inf \left\{ \int_X c(x, Tx)\, d\mu : T_{\#}\mu = \nu \right\}
\end{aligned}$$

which is $(KP) = (DP) = (MP)$.

Notice that if T is an optimal transport map solving the (MP) problem then the associated measure $\gamma_T = (id, T)_{\#}\mu$ solves the (KP) problem. Now from Chapter 3 γ_T induces a $c-$concave function ϕ such that

$$spt(\gamma_T) \subseteq \{(x, y) \in X \times Y : \phi(x) + \phi^c(y) = c(x, y)\},$$

such ϕ is a Kantorovich potential that is (ϕ, ϕ^c) solves the (DP) problem.

Proposition 4.3.7 (Uniqueness)**.** *Let X and Y be two compact metric spaces, $\mu \in \mathcal{P}(X)$, and $\nu \in \mathcal{P}(Y)$. Let $c : X \times Y \to \mathbb{R}$ continuous be such that $N_{c,\phi}$ satisfies conditions (1) and (2) for every $c-$concave function ϕ, then the solution of the (MP) problem found in Theorem 4.3.5 is unique.*

Proof. From Theorem 4.3.5 then $T = N_{c,\phi}$ is the solution to the (MP) problem where ϕ is a Kantorovich potential. Let T_0 be another solution to (MP) problem, we have

$$(DP) = (MP) = \int_X c(x, T_0 x)\, d\mu.$$

Therefore, since T_0 is measure preserving then from Proposition 4.2.3 we get

$$0 = \int_X c(x, T_0 x) - \phi(x) - \phi^c(T_0 x)\, d\mu.$$

Since, the integrand is nonegative (from Proposition 3.2.4) then

$$\phi(x) + \phi^c(T_0 x) = c(x, T_0 x) \qquad\qquad \mu - a.e.$$

We hence obtain that $T_0 x \in N_{c,\phi}(x)\, \mu - a.e.$ But $N_{c,\phi}$ is a singleton $\mu - a.e$ concluding that $T_0 x = N_{c,\phi}(x)$ for μ almost every x. □

Remark 4.3.8. Notice that with the setting of Proposition 4.3.7, if γ is a transport plan, ϕ is a corresponding Kantorovich potential, and T an optimal transport map then for $(x, y) \in spt(\gamma)$ then $\phi(x) + \phi^c(y) = c(x, y)$ and so $y = N_{c,\phi}(x) = Tx$. Let $A \times B \subseteq \mathfrak{B}(X \times Y)$, then

$$\begin{aligned}
\gamma(A \times B) &= \gamma\left(\{(x, y) : x \in A, y \in B\}\right) \\
&= \gamma\left(\{(x, Tx) : x \in A, Tx \in B\}\right) \\
&= \gamma\left(\left\{(x, Tx) : x \in A \cap T^{-1}B\right\}\right)
\end{aligned}$$

But also since $\gamma \in \Pi(\mu, \nu)$, then from Lemma 2.2.2 and similarly as above

$$\begin{aligned}
\mu(A \cap T^{-1}B) = \gamma((A \cap T^{-1}B) \times Y) &= \gamma\left(\left\{(x, y) : x \in A \cap T^{-1}B, y \in Y\right\}\right) \\
&= \gamma\left(\left\{(x, y) : x \in A \cap T^{-1}B, Tx \in Y\right\}\right) \\
&= \gamma\left(\left\{(x, Tx) : x \in A \cap T^{-1}B\right\}\right)
\end{aligned}$$

Concluding that $\gamma(A \times B) = \mu(A \cap T^{-1}B) = \gamma_T(A \times B)$. Concluding, hence, that in this case the transport plan is unique.

4.3.1 *The case* $c(x, y) = h(x - y)$ *with* h *is strictly convex*

Let $X \subseteq \mathbb{R}^d$ be a compact set with $\lambda(\partial X) = 0$ where λ is the Lebesgue measure. Let $\mu, \nu \in \mathcal{P}(X)$, $c : X \times X \mapsto \mathbb{R}$ such that $c(x, y) = h(x - y)$ where h is a C^1 strictly convex function. We assume that μ is absolutely continuous to λ.

Let $\gamma \in \Pi(\mu, \nu)$ a transport plan, i.e. solution to the corresponding (KP) problem, then there exists a Kantorovich potential ϕ such that

$$spt(\gamma) \subseteq \{(x, y) \in X \times X : \phi(x) + \phi^c(y) = c(x, y).$$

From Lemma 3.3.2, ϕ is C^1.

Let $(x_0, y_0) \in spt(\gamma)$, we have for every $x \in X$

$$\phi^c(y_0) \leq c(x, y_0) - \phi(x),$$

with equality at $x = x_0$, this imply that the function $c(x, y_0) - \phi(x)$ attains a minimum at $x = x_0$. Assume moreover that x_0 is an interior point of X (as $\mu(\partial X) = 0$) we get that

$$\nabla_x c(x_0, y_0) - \nabla_x \phi(x_0) = 0,$$

obtaining then

$$\nabla_x \phi(x_0) = \nabla_x c(x_0, y_0) = \nabla_x h(x_0 - y_0).$$

Since h is strictly convex then $\nabla_x h$ is injective and so

$$y_0 = x_0 - (\nabla_x h)^{-1}(\nabla_x \phi(x_0)).$$

From Theorem 4.3.5, the optimal transport map is

$$Tx = x - (\nabla_x h)^{-1}(\nabla_x \phi(x)).$$

The C^1 assumption on h can be dropped as from [18] strictly convex function are differentiable a.e. with respect to the Lebesgue measure and hence to μ, since $\mu \ll \lambda$.

References

[1] H. Brezis, *Analyse fonctionnelle*, ser. Collection Mathématiques Appliquées pour la Maîtrise [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983, pp. xiv+234, Théorie et applications. [Theory and applications], ISBN: 2-225-77198-7.

[2] J. A. Carrillo, R. J. McCann, and C. Villani, "Kinetic equilibration rates for granular media and related equations: Entropy dissipation and mass transportation estimates," *Rev. Mat. Iberoamericana*, vol. 19, no. 3, pp. 971–1018, 2003, ISSN: 0213-2230. DOI: 10.4171/RMI/376. [Online]. Available: https://doi.org/10.4171/RMI/376.

[3] M. Cullen and W. Gangbo, "A variational approach for the 2-dimensional semi-geostrophic shallow water equations," *Arch. Ration. Mech. Anal.*, vol. 156, no. 3, pp. 241–273, 2001, ISSN: 0003-9527. DOI: 10.1007/s002050000124. [Online]. Available: https://doi.org/10.1007/s002050000124.

[4] G. B. Dantzig, *Linear programming and extensions*, corrected, ser. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1998, pp. xviii+627, ISBN: 0-691-05913-6.

[5] L. C. Evans, "Partial differential equations and Monge-Kantorovich mass transfer," in *Current developments in mathematics, 1997 (Cambridge, MA)*, Int. Press, Boston, MA, 1999, pp. 65–126.

[6] A. Figalli and F. Glaudo, *An invitation to optimal transport, Wasserstein distances, and gradient flows*, ser. EMS Textbooks in Mathematics. EMS Press, Berlin, [2021] ©2021, pp. vi+136, ISBN: 978-3-98547-010-5. DOI: 10.4171/ETB/22. [Online]. Available: https://doi.org/10.4171/ETB/22.

[7] A. Galichon, *Optimal transport methods in economics*. Princeton University Press, Princeton, NJ, 2016, pp. xii+170, ISBN: 978-0-691-17276-7. DOI: 10.1515/9781400883592. [Online]. Available: https://doi.org/10.1515/9781400883592.

[8] A. Galichon and R. McCann, "Special Issue: Optimal transportation, equilibrium, and applications to economics [Editorial]," *Econom. Theory*, vol. 67, no. 2, pp. 345–347, 2019, ISSN: 0938-2259. DOI: 10.1007/s00199-019-01182-y. [Online]. Available: https://doi.org/10.1007/s00199-019-01182-y.

[9] C. E. Gutierrez, "Optimal transport and applications to geometric optics preliminary version," 2021.

[10] F. L. Hitchcock, "The distribution of a product from several sources to numerous localities," *J. Math. Phys. Mass. Inst. Tech.*, vol. 20, pp. 224–230, 1941, ISSN: 0097-1421. DOI: 10.1002/sapm1941201224. [Online]. Available: https://doi.org/10.1002/sapm1941201224.

[11] L. V. Kantorovich, "On a problem of Monge," *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, vol. 312, no. Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 11, pp. 15–16, 2004, ISSN: 0373-2703. DOI: 10.1007/s10958-006-0050-9. [Online]. Available: https://doi.org/10.1007/s10958-006-0050-9.

[12] S. Kolouri, S. Park, M. Thorpe, D. Slepcev, and G. K. Rohde, *Optimal mass transport: Signal processing and machine-learning applications*. IEEE signal processing magazine,34(4), 2017.

[13] R. K. Luneburg, *Mathematical theory of optics*. University of California Press, Berkeley, Calif., 1964, pp. xxx+448, Foreword by Emil Wolf; supplementary notes by M. Herzberger.

[14] N. P. Miranda, *Optimal mass transport and functional inequalities*. [Online]. Available: https://upcommons.upc.edu/bitstream/handle/2117/106637/memoria.pdf.

[15] G. Monge, *Mémoire sur la théorie des déblais et des remblais*. De l'Imprimerie Royale, 1781. [Online]. Available: https://books.google.com.lb/books?id=IG7CGwAACAAJ.

[16] H. Royden and P. Fitzpatrick, *Real Analysis*, ser. Pearson Modern Classics for Advanced Mathematics Series. Pearson, 2017, ISBN: 9780134689494. [Online]. Available: https://books.google.com.lb/books?id=7GP3MAAACAAJ.

[17] R. Sanjoy, *What is the reason for using quadratic term in optimal control?* [Online]. Available: https://www.researchgate.net/post/What-is-the-reason-for-using-quadratic-term-in-optimal-control/576696bb217e20a294217cab/citation/download.

[18] F. Santambrogio, *Optimal transport for applied mathematicians*, ser. Progress in Nonlinear Differential Equations and their Applications. Birkhäuser/Springer, Cham, 2015, vol. 87, pp. xxvii+353, Calculus of variations, PDEs, and modeling, ISBN: 978-3-319-20827-5; 978-3-319-20828-2. DOI: 10.1007/978-3-319-20828-2. [Online]. Available: https://doi.org/10.1007/978-3-319-20828-2.

[19] C. Villani, *Optimal transport*, ser. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2009, vol. 338, pp. xxii+973, Old and new, ISBN: 978-3-540-71049-3. DOI: 10.1007/978-3-540-71050-9. [Online]. Available: https://doi.org/10.1007/978-3-540-71050-9.

Bibliography

[1] H. Brezis, *Analyse fonctionnelle*, ser. Collection Mathématiques Appliquées pour la Maîtrise [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983, pp. xiv+234, Théorie et applications. [Theory and applications], ISBN: 2-225-77198-7.

[2] J. A. Carrillo, R. J. McCann, and C. Villani, "Kinetic equilibration rates for granular media and related equations: Entropy dissipation and mass transportation estimates," *Rev. Mat. Iberoamericana*, vol. 19, no. 3, pp. 971–1018, 2003, ISSN: 0213-2230. DOI: 10.4171/RMI/376. [Online]. Available: https://doi.org/10.4171/RMI/376.

[3] M. Cullen and W. Gangbo, "A variational approach for the 2-dimensional semi-geostrophic shallow water equations," *Arch. Ration. Mech. Anal.*, vol. 156, no. 3, pp. 241–273, 2001, ISSN: 0003-9527. DOI: 10.1007/s002050000124. [Online]. Available: https://doi.org/10.1007/s002050000124.

[4] G. B. Dantzig, *Linear programming and extensions*, corrected, ser. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1998, pp. xviii+627, ISBN: 0-691-05913-6.

[5] L. C. Evans, "Partial differential equations and Monge-Kantorovich mass transfer," in *Current developments in mathematics, 1997 (Cambridge, MA)*, Int. Press, Boston, MA, 1999, pp. 65–126.

[6] A. Figalli and F. Glaudo, *An invitation to optimal transport, Wasserstein distances, and gradient flows*, ser. EMS Textbooks in Mathematics. EMS Press, Berlin, [2021] ©2021, pp. vi+136, ISBN: 978-3-98547-010-5. DOI: 10.4171/ETB/22. [Online]. Available: https://doi.org/10.4171/ETB/22.

[7] A. Galichon, *Optimal transport methods in economics*. Princeton University Press, Princeton, NJ, 2016, pp. xii+170, ISBN: 978-0-691-17276-7. DOI: 10.1515/9781400883592. [Online]. Available: https://doi.org/10.1515/9781400883592.

[8] A. Galichon and R. McCann, "Special Issue: Optimal transportation, equilibrium, and applications to economics [Editorial]," *Econom. Theory*, vol. 67, no. 2, pp. 345–347, 2019, ISSN: 0938-2259. DOI: 10.1007/s00199-019-01182-y. [Online]. Available: https://doi.org/10.1007/s00199-019-01182-y.

[9] C. E. Gutierrez, "Optimal transport and applications to geometric optics preliminary version," 2021.

[10] F. L. Hitchcock, "The distribution of a product from several sources to numerous localities," *J. Math. Phys. Mass. Inst. Tech.*, vol. 20, pp. 224–230, 1941, ISSN: 0097-1421. DOI: 10.1002/sapm1941201224. [Online]. Available: https://doi.org/10.1002/sapm1941201224.

[11] L. V. Kantorovich, "On a problem of Monge," *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, vol. 312, no. Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 11, pp. 15–16, 2004, ISSN: 0373-2703. DOI: 10.1007/s10958-006-0050-9. [Online]. Available: https://doi.org/10.1007/s10958-006-0050-9.

[12] S. Kolouri, S. Park, M. Thorpe, D. Slepcev, and G. K. Rohde, *Optimal mass transport: Signal processing and machine-learning applications.* IEEE signal processing magazine,34(4), 2017.

[13] R. K. Luneburg, *Mathematical theory of optics.* University of California Press, Berkeley, Calif., 1964, pp. xxx+448, Foreword by Emil Wolf; supplementary notes by M. Herzberger.

[14] N. P. Miranda, *Optimal mass transport and functional inequalities.* [Online]. Available: https://upcommons.upc.edu/bitstream/handle/2117/106637/memoria.pdf.

[15] G. Monge, *Mémoire sur la théorie des déblais et des remblais.* De l'Imprimerie Royale, 1781. [Online]. Available: https://books.google.com.lb/books?id=IG7CGwAACAAJ.

[16] H. Royden and P. Fitzpatrick, *Real Analysis*, ser. Pearson Modern Classics for Advanced Mathematics Series. Pearson, 2017, ISBN: 9780134689494. [Online]. Available: https://books.google.com.lb/books?id=7GP3MAAACAAJ.

[17] R. Sanjoy, *What is the reason for using quadratic term in optimal control?* [Online]. Available: https://www.researchgate.net/post/What-is-the-reason-for-using-quadratic-term-in-optimal-control/576696bb217e20a294217cab/citation/download.

[18] F. Santambrogio, *Optimal transport for applied mathematicians*, ser. Progress in Nonlinear Differential Equations and their Applications. Birkhäuser/Springer, Cham, 2015, vol. 87, pp. xxvii+353, Calculus of variations, PDEs, and modeling, ISBN: 978-3-319-20827-5; 978-3-319-20828-2. DOI: 10.1007/978-3-319-20828-2. [Online]. Available: https://doi.org/10.1007/978-3-319-20828-2.

[19] C. Villani, *Optimal transport*, ser. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2009, vol. 338, pp. xxii+973, Old and new, ISBN: 978-3-540-71049-3. DOI: 10.1007/978-3-540-71050-9. [Online]. Available: https://doi.org/10.1007/978-3-540-71050-9.